Н. Алон
Дж. Спенсер

Вероятностный метод

МАТЕМАТИКА

Лаборатория ЗНАНИЙ

Вероятностный метод

Wiley-Interscience Series in Discrete Mathematics and Optimization

The Probabilistic Method

Second Edition

Noga Alon

Joel H. Spencer

A Wiley-Interscience Publication
JOHN WILEY & SONS, Inc.
New York Chichester Weinheim Brisbane Singapore Toronto

Н. Алон, Дж. Спенсер

Вероятностный метод

Перевод 2-го английского издания
под редакцией **А. А. Сапоженко**

5-е издание, электронное

Допущено
учебно-методическим советом
по прикладной математике и информатике УМО
по классическому университетскому образованию
в качестве учебного пособия для студентов
высших учебных заведений, обучающихся
по специальности и направлению
«Прикладная математика и информатика»
и по направлению «Информационные технологии»

Москва
Лаборатория знаний
2024

УДК 519.1
ББК 22.176
А45

Алон Н.

А45 Вероятностный метод : учебное пособие / Н. Алон, Дж. Спенсер ; пер. 2-го англ. изд. — 5-е изд., электрон. — М. : Лаборатория знаний, 2024. — 323 с. — Систем. требования: Adobe Reader XI ; экран 10". — Загл. с титул. экрана. — Текст : электронный.

ISBN 978-5-93208-688-9

Одна из самых известных зарубежных книг в области применения вероятностных методов в комбинаторике. В книге содержатся основные элементы методологии. Строгие обоснования и доказательства сопровождаются ясными и неформальными обсуждениями задач, методов и их приложений. Каждый метод иллюстрируется целым рядом точно подобранных примеров.

Для специалистов в области дискретной математики и теории случайных графов, студентов, аспирантов и преподавателей соответствующих дисциплин.

УДК 519.1
ББК 22.176

Деривативное издание на основе печатного аналога: Вероятностный метод : учебное пособие / Н. Алон, Дж. Спенсер ; пер. 2-го англ. изд. — М. : БИНОМ. Лаборатория знаний, 2007. — 320 с. : ил.
ISBN 978-5-94774-556-6

Первый тираж книги выпущен при финансовой поддержке РФФИ
в рамках издательского проекта № 05-01-14048

ISBN 978-5-93208-688-9

Оглавление

Предисловие редактора перевода 9
Предисловие авторов к русскому изданию 11
Предисловие 13
Благодарности 15

Часть I. МЕТОДЫ

Глава 1. **Основы** 18
1.1. Вероятностный метод 18
1.2. Теория графов 20
1.3. Комбинаторика 24
1.4. Комбинаторная теория чисел 27
1.5. Пары с пустым пересечением 28
1.6. Упражнения 29
Вероятностный взгляд: *Теорема Эрдёша—Ко—Радо* 31

Глава 2. **Линейность математического ожидания** 32
2.1. Основы 32
2.2. Разбиение графов 33
2.3. Два быстрых результата 35
2.4. Балансировка векторов 36
2.5. Разбалансировка лампочек 38
2.6. Без подбрасывания монет 39
2.7. Упражнения 40
Вероятностный взгляд: *Теорема Брегмана* 42

Глава 3. **Малые вариации** 44
3.1. Числа Рамсея 44
3.2. Независимые множества 46
3.3. Комбинаторная геометрия 47
3.4. Упаковка 48
3.5. Перекраска 49
3.6. Непрерывное время 53
3.7. Упражнения 58
Вероятностный взгляд: *Большой обхват и большое хроматическое число* 59

Глава 4. **Второй момент** 60
4.1. Основы 60
4.2. Теория чисел 61

4.3. Дополнительные теоретические сведения 64
4.4. Случайные графы 66
4.5. Максимальный размер клики 70
4.6. Различные суммы 71
4.7. Подход Рёдля 73
4.8. Упражнения 78
Вероятностный взгляд: *Гамильтоновы пути* 80

Глава 5. **Локальная лемма** 83
5.1. Лемма 83
5.2. Свойство B и разноцветные множества действительных чисел 86
5.3. Нижние оценки для чисел Рамсея 87
5.4. Геометрический результат 89
5.5. Линейная древесность графов 90
5.6. Латинские трансверсали 95
5.7. Алгоритмический аспект 96
5.8. Упражнения 100
Вероятностный взгляд: *Ориентированные циклы* 101

Глава 6. **Корреляционные неравенства** 102
6.1. Теорема о четырех функциях Альсведе и Дайкина 103
6.2. FKG-неравенство 106
6.3. Монотонные свойства 107
6.4. Линейные расширения частично упорядоченных множеств 109
6.5. Упражнения 112
Вероятностный взгляд: *Теорема Турана* 113

Глава 7. **Мартингалы и плотная концентрация** 115
7.1. Определения 115
7.2. Большие уклонения 117
7.3. Хроматическое число 118
7.4. Два обобщения 121
7.5. Четыре примера 125
7.6. Неравенство Талаграна 127
7.7. Приложения неравенства Талаграна 130
7.8. Полиномиальная концентрация Кима—Ву 133
7.9. Упражнения 135
Вероятностный взгляд: *Теорема Вейерштрасса о приближении* 136

Глава 8. **Парадигма Пуассона** 137
8.1. Неравенства Янсона 137
8.2. Доказательства 139
8.3. Решето Бруна 142
8.4. Большие уклонения 145
8.5. Оценка числа расширений 146
8.6. Число представлений 148

8.7. Дальнейшие обобщения 151
8.8. Упражнения 153
Вероятностный взгляд: *Локальная раскраска* 154

Глава 9. **Псевдослучайность** 156
9.1. Турниры квадратичных вычетов 157
9.2. Собственные значения и расширители 160
9.3. Квазислучайные графы 167
9.4. Упражнения 173
Вероятностный взгляд: *Случайные блуждания* 174

Часть II. Приложения

Глава 10. **Случайные графы** 178
10.1. Подграфы 179
10.2. Размер максимальной клики 181
10.3. Хроматическое число 183
10.4. Ветвящиеся процессы 184
10.5. Гигантская компонента 188
10.6. Фазовый переход изнутри 192
10.7. Законы «нуля или единицы» 195
10.8. Упражнения 204
Вероятностный взгляд: *Число подграфов* 205

Глава 11. **Сложность схем** 207
11.1. Предварительные замечания 207
11.2. Случайные ограничения и схемы ограниченной глубины 209
11.3. Еще о схемах ограниченной глубины 213
11.4. Монотонные схемы 216
11.5. Формулы 219
11.6. Упражнения 221
Вероятностный взгляд: *Максимальные антицепи* 222

Глава 12. **Разброс** 223
12.1. Основы 223
12.2. Достаточность шести стандартных отклонений 224
12.3. Линейный и наследственный разброс 228
12.4. Нижние оценки 231
12.5. Теорема Бека—Фиала 233
12.6. Упражнения 235
Вероятностный взгляд: *Несбалансированные матрицы* 237

Глава 13. **Геометрия** 238
13.1. Наибольший угол между точками в евклидовом пространстве 239
13.2. Пустые треугольники, определяемые точками плоскости 240
13.3. Геометрическая реализация ±1-матриц 242

13.4. ε-сети и VC-размерности ранжированных пространств 244
13.5. Двойственная функция расщепления и разброс 250
13.6. Упражнения 253
Вероятностный взгляд: *Эффективная упаковка* 254

Глава 14. **Коды, игры и энтропия** 256
14.1. Коды 256
14.2. Игра лжеца 259
14.3. Игра «постоянная должность» 261
14.4. Игра «балансировка векторов» 263
14.5. Неадаптивные алгоритмы 265
14.6. Энтропия 266
14.7. Упражнения 272
Вероятностный взгляд: *Экстремальные графы* 273

Глава 15. **Дерандомизация** 275
15.1. Метод условных вероятностей 275
15.2. d-независимые случайные величины в пространствах малого размера 280
15.3. Упражнения 284
Вероятностный взгляд: *Число пересечений, инцидентности, суммы и произведения* 285

Приложение A. Оценки для больших уклонений 287
A.1. Оценки для больших уклонений 287
A.2. Упражнения 295
Вероятностный взгляд: *Графы, свободные от треугольников, содержат большие независимые множества* 296

Приложение B. Пол Эрдёш 298
B.1. Труды 298
B.2. Гипотезы 300
B.3. Об Эрдёше 301
B.4. Дядюшка Пол 302

Литература 305
Предметный указатель 314
Именной указатель 319

Предисловие редактора перевода

Мне приятно представить читателю эту замечательную книгу двух выдающихся специалистов в области дискретной математики. Нога Алон — член Национальной Академии наук Израиля, автор более чем трехсот работ по комбинаторике и теории сложности, обладатель премий Эрдёша (1989), Фейера (1991), Пойа (2000), Бруно (2001), Ландау (2005), Гёделя (2005). Джоэл Спенсер — профессор Института Куранта Нью-Йоркского университета, автор около двухсот работ по теории случайных графов, комбинаторике и теории сложности, соавтор Пола Ердёша по книге «Случайные графы», один из основателей и главных редакторов журнала «Случайные структуры и алгоритмы». Авторы являются членами редакций многих математических журналов. Они неоднократно приглашались в качестве пленарных докладчиков на международные конференции и конгрессы. Не последнее место в ряду их достижений занимает монография «Вероятностный метод»

Первое издание книги, вышедшее в свет в 1991 г., стало одной из самых цитируемых книг в сообществе математиков, специализирующихся в области дискретной математики, информатики и применения вероятностных методов. Идея перевода ее на русский язык возникла еще в 1993 г., когда Джоэл Спенсер подарил мне экземпляр «Вероятностного метода» и еще более окрепла, когда я получил второе издание книги от Н. Алона. Благодаря Российскому фонду фундаментальных исследований идея перевода книги стала реальностью.

Главная цель монографии — изложение идей вероятностного подхода к решению задач дискретной математики. Авторы явно придерживаются известного тезиса о том, что пример учит лучше, чем теория. Подбор примеров внутри глав отвечает самым высоким требованиям целесообразности и вкуса, а примеры, помещенные в промежутках между главами, являются избранными шедеврами. По существу, это — мастер-класс двух маэстро для лиц, заинтересованных в освоении вероятностных методов.

В книге делается акцент на методологии. При относительно небольшом объеме она обладает высокой плотностью идей, приходящихся на страницу текста. Авторы часто жертвуют законченностью результатов в пользу ясности и краткости изложения. Строгое изложение утверждений как правило предваряется содержательным обсуждением метода. Наличие упражнений способствует более продуктивному восприятию материала и приобретению навыков в применении методов.

Книга Н. Алона и Д. Спенсера удачно дополняет монографии отечественных авторов В. Ф. Колчина, В. Н. Сачкова, Б. А. Севастьянова, В. П. Чистякова и др. по аналогичной тематике.

Книга будет полезна специалистам в области дискретной математики (комбинаторики, теории сложности, приложений теории вероятностей) как краткая энциклопедия приемов, связанных с применением вероятностных методов. Преподаватели вузов найдут в ней обширный материал для спецкурсов и аспирантских экзаменов. Известно, что спецкурсы по материалам книги читаются во многих университетах мира, в том числе и российских. Книга будет полезна студентам и аспирантам математических специальностей для первоначального ознакомления с предметом. Она доступна читателям, знакомым с университетскими курсами математического анализа и теории вероятностей. Специалисты в области теории вероятностей найдут много замечательных примеров применения вероятностных методов в комбинаторике и тории чисел.

Авторы иногда используют устоявшиеся понятия без определений. Для читателей, знакомых с университетским курсом дискретной математики, это не доставляет каких-либо неудобств. В книгах, добавленных при переводе, можно найти все используемые здесь понятия.

Над переводом книги работали Ф. Ю. Воробьев, К. Г. Омельянов, Т. Г. Петросян и автор этих строк. Общее редактирование осуществлялось мною. Т. В. Андреева много сделала для улучшения стиля перевода. А. Б. Дайняк принял участие в подготовке оригинал-макета. Весьма ценные замечания сделаны Н. Н. Кузюриным, Д. С. Романовым и Б. С. Стечкиным. Г. А. Махина оказала нам помощь при переводе эпиграфов к главам.

Авторы книги любезно предоставили электронный вариант рукописи и список замечаний от читателей, что предотвратило внесение опечаток при наборе формул и позволило исправить некоторые имеющиеся.

Редактор берет на себя ответственность за качество перевода и будет признателен всем, кто укажет на его возможные недостатки.

А. А. Сапоженко

Предисловие авторов к русскому изданию

Написание предисловия к книге вещь всегда приятная, поскольку на самом деле является завершением долгой работы над проектом. Нам доставляет особое удовольствие написать предисловие к русскому изданию *Вероятностного метода* поскольку в данном случае большая, тщательная и профессиональная работа по переводу выполнена А. А. Сапоженко и его аспирантами К. Омельяновым, Т. Петросяном и Ф. Воробьевым, в то время как наша задача состояла в основном в добавлении этих коротких комментариев.

Открытие того, что детерминированные утверждения могут быть доказаны с помощью вероятностных соображений, позволило уже в первой половине XX в. доказать ряд замечательных утверждений из анализа, теории чисел, комбинаторики и теории информации. Вскоре стало ясно, что метод, который сейчас называется *вероятностным*, является весьма мощным инструментом получения результатов в дискретной математике. Ранние результаты такого сорта сочетали комбинаторные соображения с элементарной вероятностной техникой, однако развитие метода в последние годы потребовало применения все более изощренных инструментов теории вероятности.

Применение вероятностного метода в дискретной математике было инициировано Полом Эрдёшем, который сделал для его развития больше, чем кто-либо другой. Полученные им результаты можно разбить на три группы. К первой относится изучение определенных классов комбинаторных объектов, таких как случайные графы или случайные матрицы. Эти результаты по существу являются теоретико-вероятностными, хотя большинство из них мотивировано задачами из комбинаторики. Вторая группа состоит из примеров применения вероятностных соображений для доказательства существования комбинаторных структур, обладающих рядом предписанных свойств. Доказательства существования такого типа часто приводят к экстремальным решениям различных задач дискретной математики. Третья группа состоит из самых поразительных примеров, фокусирующих внимание на применении вероятностных соображений к доказательству тех утверждений, формулировка которых не дает каких-либо указаний на то, что вероятность может быть полезна при их исследовании. Книга содержит результаты каждой из этих трех групп.

Многие фундаментальные и наиболее важные элементы теории вероятностей были получены русскими математиками. Такие исследователи как П. Л. Чебышёв, А. А. Марков, А. Я. Хинчин и А. Н. Колмогоров заложили основы теории вероятностей, обеспечив тем самым математические основы для последующего развития вероятностного метода. Таким образом, русское изда-

ние данной книги основано в значительной мере на достижениях российской математики.

Нам очень приятно выразить свою признательность нашему другу и коллеге Александру Сапоженко за идею перевода нашей книги, а также всем, кто способствовал успешному осуществлению этой идеи. Мы убеждены, что этот перевод сделает книгу доступной для широкого круга российских исследователей.

Нога Алон
Джоэл Спенсер

Посвящается Нурит и Мэри Энн

Предисловие

В последнее время наблюдается весьма интенсивное развитие вероятностного метода. Он стал самым мощным и широко применяемым инструментом в комбинаторике. Одной из главных причин столь быстрого развития является важная роль случайности в теоретической информатике — области, которая является в настоящее время источником многих интересных комбинаторных задач.

Взаимосвязь между дискретной математикой и информатикой подразумевает алгоритмическую точку зрения при изучении вероятностного метода в комбинаторике, и именно такой подход мы пытались проводить в этой книге. Поэтому монография включает обсуждение алгоритмических аспектов наряду с изучением классического метода и современных приемов, применяемых в ней. Первая часть книги содержит описание методов, применяемых в вероятностных доказательствах, включая традиционную технику, использующую математическое ожидание и дисперсию, а также более современные методы, связанные с применением мартингалов и корреляционных неравенств. Вторая часть включает изучение различных тем, в которых вероятностная техника оказалась весьма успешной. Эта часть содержит главы, посвященные разбросу и случайным графам, а также некоторым областям теоретической информатики: сложности логических схем, вычислительной геометрии и дерандомизации вероятностных алгоритмов. Промежутки между главами заполнены жемчужинами под общим заголовком «Вероятностный взгляд». Они представляют собой элегантные доказательства, относящиеся к главам, после которых они появляются и, как правило, могут читаться отдельно.

Основная идея вероятностного метода может быть описана следующим образом: чтобы доказать существование комбинаторной структуры с определенными свойствами, мы конструируем соответствующее вероятностное пространство и показываем, что случайно выбранный элемент этого пространства обладает данными свойствами с положительной вероятностью. Этот метод был предложен Полом Эрдёшем, который за последние пятьдесят лет внес в его развитие столь значительный вклад, что представляется справедливым назвать его «методом Эрдёша». Его вклад состоит не только в огромном числе глубоких результатов, касающихся вероятностного подхода, но также во многих интересных проблемах и гипотезах, которые в значительной степени стимулировали исследования в этой области.

Написание энциклопедической книги по вероятностному методу представляется невозможным. Слишком много новых интересных результатов получено с помощью вероятностных соображений, и мы не пытаемся даже упомянуть каждый из них. В книге делается акцент на методологию. Главной

целью является изложение идей. Поэтому мы не всегда приводим наилучший известный результат, если он технически слишком сложен для того, чтобы ясно изложить его. Многие результаты являются асимптотическими. При их изложении используются стандартные асимптотические обозначения. Пусть f и g — две функции, мы пишем $f = O(g)$, если $f \leqslant c_1 g + c_2$ для всех значений переменных, где c_1, c_2 — абсолютные константы. Мы пишем $f = \Omega(g)$, если $g = O(f)$, и $f = \Theta(g)$, если $f = O(g)$ и $f = O(g)$. Запись $f = o(g)$ означает, что предел отношения f/g стремится к нулю при стремлении переменных этих функций к бесконечности. Наконец, $f \sim g$ означает, что $f = (1 + o(1))g$, т. е. что f/g стремится к 1, когда переменные стремятся к бесконечности.

Каждая глава заканчивается упражнениями. Наиболее трудные помечены символом *. Упражнения, добавленные в этом издании книги, позволяют читателю проверить понимание материала, а также дают возможность использовать книгу как учебник.

Кроме этих упражнений второе издание содержит некоторые улучшенные результаты и развивает различные темы, обсуждавшиеся в первом издании. Среди добавлений упомянем следующие. В гл. 3 описан непрерывный подход к дискретным вероятностным задачам, в гл. 7 излагаются некоторые новые неравенства для концентрации, в гл. 13 обсуждаются отношения между разбросом и VC-размерностью, в гл. 14 описываются некоторые приложения функции энтропии и ее свойства. Имеются также добавления в последних двух «Вероятностных взглядах» и новое приложение о Поле Эрдёше, в котором приведены список его статей, его гипотезы и воспоминания о нем.

С особым удовольствием мы благодарим наших жен Нурит и Мэри Энн. Их терпение, понимание и поддержка явились ключевым моментом в успехе данного предприятия.

Нога Алон
Джоэл Спенсер

Благодарности

Мы очень признательны нашим студентам и коллегам, принявшим участие в создании второго издания книги путем совместных публикаций, полезных обсуждений и ценных замечаний. В их числе: Грег Бачелис, Амир Дембо, Эхуд Фрейтгут, Марк Фоссорье, Донг Фу, Сванте Янсон, Гай Котцерс, Михаил Кривелевич, Альберт Ли, Боян Моар, Янош Пач, Ювал Перес, Аравинд Шринивасан, Бенни Судаков, Тибор Сабо, Грег Соркин, Джон Тромп, Дэвид Уилсон, Ник Уормалд и Ури Цвик, которые указали на различные небрежности и ошибки. Их предложения позволили улучшить изложение и результаты. Необходимо отметить, что ответственность за остающиеся ошибки, так же как и ответственность за новые ошибки (надеемся, что их немного) лежит целиком на нас.

С удовольствием благодарим Орена Нечуштана за большую техническую работу, проделанную при окончательной подготовке рукописи.

Часть I

МЕТОДЫ

1

Основы

Все, что нужно, — это чтобы ваш разум был открыт.

Пол Эрдёш

1.1. ВЕРОЯТНОСТНЫЙ МЕТОД

Вероятностный метод является мощным инструментом для решения многих задач дискретной математики. Грубо говоря, этот метод работает следующим образом: пытаясь доказать, что структура с некоторыми искомыми свойствами существует, мы определяем подходящее вероятностное пространство структур, а затем показываем, что искомые свойства выполняются для случайно выбранного элемента в этом пространстве с положительной вероятностью.

Метод лучше всего проиллюстрировать примерами. Ниже — один из них. Число Рамсея $R(k,l)$ есть наименьшее целое n, такое, что при любой раскраске ребер полного n-вершинного графа в синий и красный цвета либо существует красный подграф K_k (т. е. полный подграф на k вершинах, каждое ребро которого раскрашено в красный цвет), либо существует синий подграф K_l. В 1929 г. Рамсей показал, что число $R(k,l)$ конечно для любых k и l. Мы найдем нижнюю оценку для диагонального числа Рамсея $R(k,k)$.

Предложение 1.1.1. *Если* $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$, *то* $R(k,k) > n$. *Таким образом,* $R(k,k) > \lfloor 2^{k/2} \rfloor$ *для всех* $k \geqslant 3$.

Доказательство. Рассмотрим случайную раскраску ребер графа K_n, полученную раскраской каждого из ребер независимо в красный или синий цвет, причем каждый цвет появляется с равной вероятностью. Определим для каждого фиксированного k-вершинного множества R событие A_R, состоящее в том, что подграф графа K_n, порожденный подмножеством R, является *монохроматическим* (т.е. либо все его ребра являются красными, либо все являются синими. Очевидно, что $\Pr[A_R] = 2^{1-\binom{k}{2}}$. Так как существует $\binom{n}{k}$ возможностей для выбора R, вероятность того, что по крайней мере одно из событий A_R произойдет, не больше чем $\binom{n}{k} 2^{1-\binom{k}{2}} < 1$. Таким образом, с положительной вероятностью ни одно из событий A_R не произойдет, а значит, существует 2-раскраска графа K_n без монохроматических подграфов K_k, т. е. $R(k,k) > n$.

Заметим, что если $k \geqslant 3$ и $n = \lfloor 2^{k/2} \rfloor$, то $\binom{n}{k} 2^{1-\binom{k}{2}} < \frac{2^{1+\frac{k}{2}}}{k!} \cdot \frac{n^k}{2^{k^2/2}} < 1$, а, следовательно, $R(k,k) > \lfloor 2^{k/2} \rfloor$ для всех $k \geqslant 3$. ■

Этот простой пример демонстрирует суть вероятностного метода. Чтобы доказать существование хорошей раскраски, мы вовсе не должны предъявить ее явно. Достаточно неконструктивным образом показать, что она существует. Этот пример появился в статье П. Эрдёша в 1947 г. Хотя уже в 1943 г. Селе применил вероятностный метод к другой комбинаторной задаче, упомянутой в гл. 2, Эрдёш был определенно первым, кто вполне осознал силу этого метода и на протяжении многих лет успешно применял его для решения большого числа задач.[1)] Можно, конечно, сказать, что в рассмотренном выше примере легко обойтись и без вероятности. Столь же простое доказательство можно получить чисто комбинаторным путем. Достаточно лишь убедиться в том, что общее число 2-раскрасок графа K_n больше числа тех, что содержат монохроматический подграф K_k.

Кроме того, поскольку большинство вероятностных пространств, рассматривавшихся при изучении комбинаторных задач, являются конечными, это утверждение справедливо в отношении большинства приложений вероятностного метода в дискретной математике. Теоретически это так. Однако на практике использование понятия вероятности является существенным. Было бы безнадежно пытаться заменить применение многих изложенных в книге приемов — например, метод второго момента, локальную лемму Ловаса и доказательство концентрации с помощью мартингалов — комбинаторными соображениями даже в случае их применения к конечным вероятностным пространствам.

Вероятностный метод имеет интересный алгоритмический аспект. Рассмотрим, например, доказательство предложения 1.1.1, которое показывает, что существует реберная 2-раскраска графа K_n без монохроматических клик $K_{2\log_2 n}$. Можем ли мы представить такую раскраску явно? Такой вопрос, как уже говорилось, может показаться нелепым. Ведь общее число возможных раскрасок конечно, поэтому можно проверить их все, пока не найдем подходящую. Однако такая процедура может потребовать $2^{\binom{n}{2}}$ проверок. Это количество экспоненциально относительно размера входа задачи (равного $\binom{n}{2}$). Алгоритмы, время работы которых больше полинома от размера (входа) задачи, обычно рассматриваются как практически неприменимые. Класс задач, которые решаются за полиномиальное время, обычно обозначается через $\mathbf{P}$ (см., например, [Aho, Hopcroft and Ullman (1974)]) и рассматривается как класс разрешимых задач. В этом смысле подход, связанный с полным перебором, в применении к поиску хорошей раскраски графа K_n, неприемлем, и в этом причина нашего замечания о неконструктивности доказательства предложения 1.1.1. Оно не дает конструктивного, эффективного и детерминированного алгоритма

[1)] Отметим, что в 1942 г. появилась статья [Гончаров (1942)], в которой с помощью теоретико-вероятностных методов решались комбинаторные задачи о числе циклов в перестановках (см. также [Гончаров (1944)]). — *Прим. ред.*

построения раскраски с требуемыми свойствами. Однако, более внимательный взгляд на доказательство показывает, что оно может быть использовано для эффективного построения раскраски, которая с большой вероятностью является хорошей. В самом деле, для больших k при $n = \lfloor 2^{k/2} \rfloor$ выполнено $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < \frac{2^{1+\frac{k}{2}}}{k!}\left(\frac{n}{2^{k/2}}\right)^k \leqslant \frac{2^{1+\frac{k}{2}}}{k!} \ll 1$. Следовательно, случайная раскраска графа K_n с большой вероятностью не содержит монохроматических подграфов $K_{2\log n}$. Это означает, что если по некоторой причине *необходимо* представить явно 2-раскраску ребер графа K_{1024} без монохроматических подграфов K_{20}, то можно просто подбросить правильную монету $\binom{1024}{2}$ раз и тем самым получить требуемую раскраску с достаточной степенью уверенности. Вероятность того, что раскрашенный полный граф содержит монохроматический подграф K_{20}, меньше $\frac{2^{11}}{20!}$, что, наверное, гораздо меньше, чем шанс совершить ошибку в любом строгом доказательстве того, что некоторая раскраска является хорошей! Следовательно, в некоторых случаях вероятностный, неконструктивный метод дает эффективные вероятностные алгоритмы. Эта тема достаточно подробно обсуждается в гл. 15.

Вероятностный метод является мощным инструментом в комбинаторике и теории графов. Он оказывается также крайне полезным в теории чисел и вычислительной геометрии. С недавних пор он стал применяться для разработки эффективной алгоритмической техники и при изучении различных вычислительных задач. В конце этой главы мы приведем простые примеры, демонстрирующие широкий спектр вопросов, для решения которых полезен этот метод. Более сложные примеры, включающие более тонкие вероятностные соображения, появятся в дальнейшем.

1.2. ТЕОРИЯ ГРАФОВ

Турнир $T = (V, E)$ на множестве V из n игроков есть результат ориентации ребер полного графа со множеством вершин V. Таким образом, для любых двух элементов x и y множества V либо пара (x, y), либо пара (y, x) принадлежат множеству E, но не обе вместе. Название «турнир» вполне естественно, поскольку его можно представлять себе как множество игроков V, в котором каждая пара участников проводит одну встречу, а присутствие дуги (x, y) означает, что x *побеждает* y. Скажем, что T обладает свойством S_k, если в каждом множестве из k игроков найдется хотя бы один, который побеждает их всех. Например, ориентированный треугольник $T_3 = (V, E)$, в котором $V = \{1, 2, 3\}$ и $E = \big\{(1, 2), (2, 3), (3, 1)\big\}$, обладает свойством S_1. Верно ли, что для любого конечного k существует турнир T (с более чем k вершинами), обладающий свойством S_k? Как показал Эрдёш [Erdős (1963b)], эта проблема, поставленная Шютте, может быть решена почти тривиально с помощью вероятностных соображений. Более того, эти соображения позволяют получить довольно точную оценку минимально возможного числа вершин в таком турнире. Основная (и естественная) идея заключается в том, что если n достаточно велико как функция аргумента k, то *случайный турнир* на множестве $V = \{1, \ldots, n\}$ из

n игроков с большой вероятностью обладает свойством S_k. Под случайным турниром мы подразумеваем турнир T на множестве вершин V, полученный выбором для каждой пары $1 \leqslant i < j \leqslant n$ независимо или дуги (i,j), или дуги (j,i), причем каждая из этих возможностей равновероятна. Заметим, что при этом все $2^{\binom{n}{2}}$ возможных турниров на множестве V равновероятны, т. е. рассматриваемое вероятностное пространство симметрично. Мы часто используем симметричные вероятностные пространства в приложениях. В этих случаях мы будем иногда говорить об элементе пространства как о *случайном элементе*, без точного описания распределения вероятностей. Так, например, в доказательстве предложения 1.1.1 рассматривались случайные 2-раскраски, т. е. все возможные раскраски были равновероятны. Аналогично, в доказательстве следующего простого результата мы изучаем случайные турниры на множестве V.

Теорема 1.2.1. *Если* $\binom{n}{k}(1-2^{-k})^{n-k} < 1$, *то существует турнир на* n *вершинах, обладающий свойством* S_k.

Доказательство. Рассмотрим случайный турнир на множестве $V = \{1, \ldots, n\}$. Для каждого фиксированного k-элементного подмножества K множества V определим событие A_K, состоящее в том, что не существует вершины, которая побеждает все вершины из K. Ясно, что $\Pr[A_K] = (1-2^{-k})^{n-k}$, поскольку для каждой фиксированной вершины $v \in V - K$ вероятность того, что v не побеждает всех игроков из K, равна $1-2^{-k}$, и все эти $n-k$ событий, соответствующих различным выборам вершин v, независимы. Отсюда следует, что

$$\Pr\left[\bigvee_{\substack{K\subset V\\ |K|=k}} A_K\right] \leqslant \sum_{\substack{K\subset V\\ |K|=k}} \Pr[A_K] = \binom{n}{k}(1-2^{-k})^{n-k} < 1.$$

Следовательно, с положительной вероятностью ни одно из событий A_K не происходит, т. е. существует турнир на n вершинах, обладающий свойством S_k. ■

Пусть $f(k)$ — минимальное число вершин в турнире, обладающем свойством S_k. Поскольку $\binom{n}{k} < \left(\frac{en}{k}\right)^k$ и $(1-2^{-k})^{n-k} < e^{-(n-k)/2^k}$, из теоремы 1.2.1 следует, что $f(k) \leqslant k^2 \cdot 2^k \cdot (\ln 2)\big(1+o(1)\big)$. Нетрудно проверить, что $f(1) = 3$ и $f(2) = 7$. Как доказал Секереш (см. книгу [Moon (1968)]), $f(k) \geqslant c_1 \cdot k \cdot 2^k$.

Существует ли явная конструкция турнира с не более чем c_2^k вершинами, обладающего свойством S_k? Такая конструкция известна, но нетривиальна. Мы опишем ее в гл. 9.

Доминирующим называется такое множество $U \subseteq V$ неориентированного графа $G = (V, E)$, что всякая вершина $v \in V - U$ имеет хотя бы одного соседа в U.

Теорема 1.2.2. *Пусть* $G = (V, E)$ — *граф с* n *вершинами и минимальной степенью* $\delta > 1$. *Тогда* G *имеет доминирующее множество с не более чем* $n\frac{1+\ln(\delta+1)}{\delta+1}$ *вершинами.*

Доказательство. Пусть p — некоторое число из отрезка $[0,1]$. Будем выбирать случайно и независимо каждую вершину из V с вероятностью p. Пусть X — (случайное) множество всех выбранных таким образом вершин, а $Y = Y_X$ — множество тех вершин из $V - X$, которые не имеют соседей в X. Математическое ожидание величины $|X|$ равно, очевидно, np. Для каждой фиксированной вершины $v \in V$, по определению

$$\Pr[v \in Y] = \Pr[v \text{ и все ее соседи не принадлежат } X] \leqslant (1-p)^{\delta+1}.$$

Так как математическое ожидание суммы случайных величин равно сумме их математических ожиданий (даже если они не являются независимыми), и поскольку случайные переменные $|Y|$ могут быть представлены в виде сумм n индикаторных случайных величин X_i (короче, индикаторов) χ_v ($v \in V$), где $\chi_v = 1$, если $v \in Y$, и $\chi_v = 0$ в противном случае, мы заключаем, что математическое ожидание величины $|X| + |Y|$ не превышает $np + n(1-p)^{\delta+1} = n\frac{\ln(\delta+1)}{\delta+1} + n\bigl(1 - \frac{\ln(\delta+1)}{\delta+1}\bigr)^{\delta+1} \leqslant n\frac{1+\ln(\delta+1)}{\delta+1}$. Следовательно, существует по крайней мере одно множество $X \subseteq V$, такое, что $|X| + |Y_X| \leqslant n\frac{1+\ln(\delta+1)}{\delta+1}$. Множество $U = X \cup Y_X$ является, очевидно, доминирующим множеством в графе G, и его мощность не превосходит $n\frac{1+\ln(\delta+1)}{\delta+1}$.

Проведем некоторые элементарные вычисления с целью оптимизации результата. Для удобства мы оценим сверху $1 - p \leqslant e^{-p}$ (это справедливо для всех действительных p), чтобы вывести более простую оценку

$$|U| \leqslant np + ne^{-p(\delta+1)}.$$

Возьмем производную правой части по p и приравняем ее нулю. Правая часть минимальна при

$$p = \frac{\ln(\delta+1)}{\delta+1}.$$

При таком значении p имеем $|U| \leqslant n\frac{1+\ln(\delta+1)}{\delta+1}$, что и требовалось. ■

В последнем доказательстве содержатся три простые, но важные идеи. Первая — это линейность математического ожидания. Многие применения этого простого, но важного принципа встретятся нам в гл. 2. Вторая, быть может, более тонкая идея является примером принципа «малых вариаций». Он будет обсуждаться в гл. 3. Случайный выбор не позволяет получить требуемое доминирующее множество сразу. Он дает только множество X, которое должно быть слегка изменено (путем добавления множества Y_X) для получения требуемого доминирующего множества. Третья идея заключается в выборе оптимального p. Часто хочется сделать случайный выбор, не фиксируя p заранее. Идея состоит в том, чтобы провести доказательство, имея p в качестве *параметра*, и получить результат как функцию от p. Значение параметра p выбирается в конце с целью получения оптимального результата.

Имеется и четвертая идея, которую можно назвать асимптотическим вычислением. Мы хотели получить асимптотику величины $\min(np + n(1-p)^{\delta+1})$, где p принимает значения из отрезка $[0,1]$. Иметь дело с точным минимумом величины $p = 1 - (\delta+1)^{-1/\delta}$ затруднительно, и во многих подобных случаях

этот точный минимум невозможно представить в явном виде. Предпочтительнее потерять немного в точности, заменив $1-p$ на e^{-p} для получения «легко воспринимаемой» оценки. Значительная часть *искусства* вероятностного метода лежит в поиске близкого к оптимальному, но ясного результата. Много ли мы теряем в данном случае? Ответ зависит от конкретной ситуации. При $\delta = 3$ наша грубая оценка дает $|U| \leqslant 0.596n$, в то время как уточненный подсчет приводит к $|U| \leqslant 0.496n$, что, возможно, существенно. Для больших δ оба метода дают асимптотически $n\frac{\ln\delta}{\delta}$.

Из результатов Алона [Alon (1990b)] легко выводится, что оценка в теореме 1.2.2 почти оптимальна. *Не*вероятностное, алгоритмическое доказательство этой теоремы может быть получено выбором вершин доминирующего множества одной за другой, так, чтобы на каждом шаге покрывалось максимальное число еще не покрытых вершин.[2)] В самом деле, обозначим через $C(v)$ множество вершин, состоящее из v и всех ее соседей. Пусть в процессе выбора вершин число тех вершин u, которые не лежат в объединении множеств $C(v)$, взятом по всем вершинам v, выбранным к этому моменту, равно r. По предположению сумма мощностей множеств $C(u)$ по всем непокрытым вершинам u равна по меньшей мере $r(\delta+1)$. Тогда, вычисляя среднее, получаем, что существует вершина v, принадлежащая по меньшей мере $r(\delta+1)/n$ таким множествам $C(u)$. Добавляя такую вершину к уже выбранным, мы замечаем, что число непокрытых вершин не превосходит $r(1-\frac{\delta+1}{n})$. Отсюда следует, что на каждом шаге нашей процедуры число непокрытых вершин уменьшается в $1-(\delta+1)/n$ раз. Следовательно, после $\frac{n}{\delta+1}\ln(\delta+1)$ шагов останутся непокрытыми $n/(\delta+1)$ вершин. Последние могут быть добавлены в число выбранных с целью получения доминирующего множества. Размер последнего превышает размер множества, полученного в теореме 1.2.2, не более чем на единицу.

Сочетая сказанное с некоторыми соображениями Поддерюгина и Матулы, мы можем получить весьма эффективный алгоритм для решения вопроса о том, является ли n-вершинный граф, скажем, $\frac{n}{2}$-реберно связным. *Разрез* в графе $G=(V,E)$ — это разбиение множества вершин V на две непустые и непересекающиеся части $V = V_1 \cup V_2$. Будем говорить, что разрез *разделяет* v_1 и v_2, если $v_1 \in V_1$ и $v_2 \in V_2$. *Размер* разреза определяется как число ребер графа, у которых один конец лежит в V_1, а другой в V_2. В действительности, мы иногда отождествляем разрез с множеством этих ребер. *Реберная связность* графа G — это минимальный размер разреза графа G. Следующая лемма принадлежит Поддерюгину и Матуле (независимо).

Лемма 1.2.3. *Пусть $G=(V,E)$ — граф с минимальной степенью δ, и пусть $V = V_1 \cup V_2$ — разрез в графе G размера, меньшего, чем δ. Тогда каждое доминирующее множество U графа G содержит вершины как из V_1, так и из V_2.*

[2)] Подобного типа процедуры часто называют термином «жадный алгоритм» (см., например, разд. «Без подбрасывания монет»). В отечественной литературе используются также названия: «алгоритм наискорейшего спуска» (см. [Нигматуллин (1974)]) и «градиентный алгоритм»(см. [Сапоженко (1972)]). — *Прим. ред.*

Доказательство. Предположим, это неверно и $U \subseteq V_1$. Выберем произвольно вершину $v \in V_2$, и пусть $v_1, v_2, \ldots, v_\delta$ — некоторые δ ее соседей. Для каждого i, $1 \leqslant i \leqslant \delta$, определим ребро e_i данного разреза следующим образом. Если $v_i \in V_1$, то $e_i = \{v, v_i\}$, иначе $v_i \in V_2$ и, поскольку множество U является доминирующим, существует по крайней мере одна вершина $u \in U$, такая, что $\{u, v_i\}$ является ребром. Возьмем такую вершину u и положим $e_i = \{u, v_i\}$. Тогда все δ ребер $e_1, \ldots, e_\delta$ различны и принадлежат данному разрезу, что противоречит предположению о том, что размер разреза меньше δ. Тем самым утверждение доказано. ■

Пусть $G = (V, E)$ — граф на n вершинах. Предположим, мы хотим выяснить, является ли G $n/2$-реберно-связным, т. е. верно ли, что его реберная связность не меньше $n/2$. Матула показал с помощью леммы 1.2.3, что это может быть сделано за время $O(n^3)$. В силу замечания к теореме 1.2.2, можно слегка улучшить это утверждение и получить оценку вида $O(n^{8/3} \log n)$ следующим образом. Сначала мы проверим, что минимальная степень δ вершины в графе G не меньше $n/2$. Если это не так, то G не является $n/2$-реберно-связным, и алгоритм завершает работу. В противном случае, по теореме 1.2.2 существует доминирующее множество $U = \{u_1, \ldots, u_k\}$ графа G с $k = O(\log n)$. Оно может быть найдено за время $O(n^2)$. Далее укажем для каждого i, $2 \leqslant i \leqslant k$, минимальный размер s_i разреза, разделяющего u_1 и u_i. Каждая из этих задач может быть сведена к решению стандартной задачи построения минимального потока в сети за время $O(n^{8/3})$ (см. например, [Tarjan (1983)]). По лемме 1.2.3 реберная связность графа G — это просто минимум из δ и $\min\limits_{2 \leqslant i \leqslant k} s_i$. Общее время работы алгоритма не превышает $O(n^{8/3} \log n)$, что и требовалось.

1.3. КОМБИНАТОРИКА

Пара $H = (V, E)$, где V — конечное множество, элементы которого называются *вершинами*, а E — некоторое семейство подмножеств множества V, называемых *ребрами*, называется *гиперграфом*. Гиперграф называется *n-однородным*, если каждое его ребро содержит ровно n вершин. Говорят, что гиперграф H обладает свойством B или, что его можно раскрасить в два цвета, если существует раскраска множества V в два цвета, такая, что никакое ребро не является монохроматическим. Через $m(n)$ обозначим минимальное число ребер n-однородного гиперграфа, который не обладает свойством B.

Предложение 1.3.1 [Erdős (1963a)]. *Каждый n-однородный гиперграф с числом ребер, меньшим 2^{n-1}, обладает свойством B, т. е. $m(n) \geqslant 2^{n-1}$.*

Доказательство. Пусть $H = (V, E)$ является n-однородным гиперграфом с числом ребер, меньшим 2^{n-1}. Раскрасим множество V в два цвета случайным образом. Пусть для каждого $e \in E$ событие A_e заключается в том, что ребро e

является монохроматическим. Ясно, что $\Pr[A_e] = 2^{1-n}$. Следовательно,

$$\Pr\left[\bigvee_{e\in E} A_e\right] \leqslant \sum_{e\in E} \Pr[A_e] < 1.$$

Таким образом, существует 2-раскраска без монохроматических ребер. ■

В разд. 3.5 гл. 3, будет представлено более тонкое рассуждение, принадлежащее Радхакришнану и Сринивасану и основанное на идее Бека, из которого следует, что $m(n) \geqslant \Omega\big((\frac{n}{\ln n})^{1/2}2^n\big)$.

Лучшая известная оценка сверху для числа $m(n)$ получена путем «переворачивания вероятностного доказательства с ног на голову». Идея заключается в следующем: каждое множество выбирается случайным образом, а каждая раскраска определяет некоторое событие. Зафиксируем множество V с v точками (в дальнейшем мы оптимизируем значение этого v). Пусть раскраска χ множества V окрашивает a точек в один цвет и, соответственно, $b = v - a$ точек — в другой. Пусть $S \subset V$ — равновероятно выбранное n-множество. Тогда

$$\Pr[S \text{ является монохроматическим относительно } \chi] = \frac{\binom{a}{n} + \binom{b}{n}}{\binom{v}{n}}.$$

Предположим для удобства, что v четно. Из выпуклости функции $\binom{y}{n}$ следует, что данное выраженное достигает минимума при $a = b$. Получаем, что

$$\Pr[S \text{ является монохроматическим относительно } \chi] \geqslant p,$$

где

$$p = \frac{2\binom{v/2}{n}}{\binom{v}{n}}.$$

Пусть $S_1, \ldots, S_m$ — равновероятно и независимо выбранные n-множества; m будет определено позже. Для каждой раскраски χ событие A_χ заключается в том, что ни одно из множеств S_i не является монохроматическим. Из независимости выбора S_i следует, что

$$\Pr[A_\chi] \leqslant (1-p)^m.$$

Число раскрасок равно 2^v, поэтому

$$\Pr\left[\bigvee_{\chi} A_\chi\right] \leqslant 2^v(1-p)^m.$$

Если эта величина меньше 1, существуют $S_1, \ldots, S_m$, такие, что ни одно из событий A_χ не выполняется, т. е. $S_1, \ldots, S_m$ не могут быть раскрашены в два цвета и, следовательно, $m(n) \leqslant m$.

Доказанная асимптотическая оценка является характерным примером оценок, возникающих при использовании вероятностного метода. Мы пользуемся неравенством $1-p \leqslant e^{-p}$. Оно верно для всех действительных p, причем левая

и правая части неравенства близки, когда p мало. Если[3)]

$$m = \left\lceil \frac{v \ln 2}{p} \right\rceil,$$

то $2^v(1-p)^m < 2^v e^{-pm} \leqslant 1$. Откуда следует, что $m(n) \leqslant m$. Теперь нужно найти v, минимизирующее значение v/p. Величину p можно интерпретировать как удвоенную вероятность извлечения n белых шаров из урны с $v/2$ белыми и $v/2$ черными шарами (выбор производится без возвращения шаров). Возникает желание оценить p величиной 2^{-n+1}, вероятностью для выборки с возвращением. Эта аппроксимация дала бы оценку $m \sim v2^{n-1}\ln 2$. Однако, когда v уменьшается, аппроксимация становится менее точной и, поскольку наша цель состоит в минимизации m, возникает необходимость в более точном приближении. Мы используем аппроксимацию второго порядка

$$p = \frac{2\binom{v/2}{n}}{\binom{v}{n}} = 2^{1-n}\prod_{i=0}^{n-1}\frac{v-2i}{v-i} \sim 2^{1-n}e^{-n^2/2v}$$

до тех пор, пока $v \gg n^{3/2}$. Мы использовали оценку $\frac{v-2i}{v-i} = 1 - \frac{i}{v} + O(\frac{i^2}{v^2}) = e^{-\frac{i}{v}+O(\frac{i^2}{v^2})}$. Элементарный анализ дает оптимальное значение $v = n^2/2$. Четность v может потребовать изменений не более чем на 2, что в данном случае асимптотически пренебрежимо. Это дает следующий результат [Erdős (1964)].

Теорема 1.3.2. *Имеет место соотношение*

$$m(n) < (1+o(1))\frac{e\ln 2}{4}n^2 2^n.$$

Пусть $\mathcal{F} = \{(A_i, B_i)\}_{i=1}^h$ — семейство пар подмножеств произвольного множества. Семейство $\mathcal{F}$ называется (k,ℓ)-*системой*, если $|A_i| = k$ и $|B_i| = \ell$ при всех $1 \leqslant i \leqslant h$, $A_i \cap B_i = \varnothing$ и $A_i \cap B_j \neq \varnothing$ для всех различных i, j, $1 \leqslant i, j \leqslant h$. Боллобаш в [B(1965)] доказал следующий результат, который имеет множество различных обобщений и применений.

Теорема 1.3.3. *Если семейство* $\mathcal{F} = \{(A_i, B_i)\}_{i=1}^h$ *является* (k,ℓ)-*системой, то* $h \leqslant \binom{k+\ell}{k}$.

Доказательство. Пусть $X = \bigcup_{i=1}^{h}(A_i \cup B_i)$. Рассмотрим случайный порядок π на множестве X. Пусть для каждого i, $1 \leqslant i \leqslant k$, событие X_i состоит в том, что все элементы подмножества A_i предшествуют всем элементам подмножества B_i относительно данного порядка π. Ясно, что $\Pr[X_i] = 1/\binom{k+\ell}{k}$. Легко проверить, что события X_i попарно несовместны. В самом деле, предположим,

[3)] Везде в дальнейшем $\lceil a \rceil$ есть наименьшее целое, которое не меньше a, $\lfloor a \rfloor$ — целая часть числа a. — *Прим. ред.*

что это не так, и пусть порядок π таков, что все элементы множества A_i предшествуют элементам множества B_i, и все элементы из A_j предшествуют элементам из B_j. Без ограничения общности можно считать, что последний элемент множества A_i не следует за последним элементом A_j. Но тогда все элементы из A_i предшествуют элементам из B_j, что противоречит условию $A_i \cap B_j \neq \varnothing$. Поэтому, все события X_i попарно несовместны. Отсюда следует, что $1 \geqslant \Pr\left[\bigvee_{i=1}^{h} X_i\right] = \sum_{i=1}^{h} \Pr[X_i] = h \cdot 1/\binom{k+\ell}{k}$. ■

Пример семейства $\mathcal{F} = \big\{(A, X \setminus A) : A \subset X, |A| = k\big\}$, где множество $X = \{1, 2, \ldots, k+\ell\}$, показывает, что теорема 1.3.3 дает точную оценку.

1.4. КОМБИНАТОРНАЯ ТЕОРИЯ ЧИСЕЛ

Подмножество A абелевой группы G называется *свободным от сумм*, если $(A + A) \cap A = \varnothing$, т. е. не существует трех элементов $a_1, a_2, a_3 \in A$, таких, что $a_1 + a_2 = a_3$.

Теорема 1.4.1 [Erdős (1965a)]. *Каждое множество $B = \{b_1, \ldots, b_n\}$, состоящее из n целых чисел, отличных от нуля, содержит множество A, свободное от сумм, такое, что $|A| > \frac{1}{3}n$.*

Доказательство. Пусть $p = 3k + 2$ — простое число, удовлетворяющее неравенству $p > 2 \max\limits_{1 \leqslant i \leqslant n} |b_i|$. Пусть $C = \{k+1, k+2, \ldots, 2k+1\}$. Заметим, что C является подмножеством циклической группы $\mathbb{Z}_p$, свободным от сумм, и что $\frac{|C|}{p-1} = \frac{k+1}{3k+1} > \frac{1}{3}$. Выберем случайным образом натуральное число x, $1 \leqslant x < p$, согласно равномерному распределению на множестве $\{1, 2, \ldots, p-1\}$. Пусть числа $d_1, \ldots, d_n$ удовлетворяют сравнениям $d_i \equiv xb_i (\mathrm{mod}\, p)$, $0 \leqslant d_i < p$. Очевидно, что если x пробегает числа $1, 2, \ldots, p-1$, то для любого фиксированного i, $1 \leqslant i \leqslant n$, величина d_i пробегает все ненулевые элементы группы $\mathbb{Z}_p$. Отсюда получаем $\Pr[d_i \in C] = \frac{|C|}{p-1} > \frac{1}{3}$. Таким образом, математическое ожидание числа элементов b_i, таких, что $d_i \in C$, превосходит $\frac{n}{3}$. Следовательно, существуют x, $1 \leqslant x < p$, и подпоследовательность A последовательности B, такая, что $|A| > \frac{n}{3}$ и $xa (\mathrm{mod} p) \in C$ для всех $a \in A$. Ясно, что множество A свободно от сумм. Действительно, если равенство $a_1 + a_2 = a_3$ выполнялось бы для некоторых $a_1, a_2, a_3 \in A$, то тогда выполнялось бы сравнение $xa_1 + xa_2 \equiv xa_3 (\mathrm{mod}\, p)$, а это противоречило бы тому, что C является подмножеством группы Z_p, свободным от сумм. ■

В статье [Alon and Kleitman (1990)] показано, что в любом подмножестве мощности n абелевой группы, не содержащем нейтрального элемента, существует более чем $2n/7$ элементов, образующих множество, свободное от сумм. Причем константа $2/7$ неулучшаема. Наилучшая константа в утверждении теоремы 1.4.1 неизвестна.

1.5. ПАРЫ С ПУСТЫМ ПЕРЕСЕЧЕНИЕМ

Наиболее поразительные результаты, полученные вероятностным методом, связаны с его применением при доказательстве теорем, условие которых казалось бы совсем не требует такого подхода. Большинство примеров из предыдущих разделов являются простыми случаями таких теорем. В этом разделе мы приведем доказательство несколько более сложной теоремы, полученной в статье [Alon and Frankl (1985)] и дающей решение гипотезы Дайкина и Эрдёша.

Пусть $\mathcal{F}$ является семейством из m различных подмножеств множества $X = \{1, 2, \ldots, n\}$. Обозначим через $d(\mathcal{F})$ количество непересекающихся пар в семействе $\mathcal{F}$, т. е.

$$d(\mathcal{F}) = \left|\left\{(F, F') : F, F' \in \mathcal{F}, F \cap F' = \varnothing\right\}\right|.$$

Дайкин и Эрдёш предположили, что если $m = 2^{\left(\frac{1}{2}+\delta\right)n}$, то $d(\mathcal{F}) = o(m^2)$ для любого фиксированного $\delta > 0$ и n, стремящемся к бесконечности. Это утверждение следует из следующей теоремы, которая, в свою очередь, является частным случаем более общего результата.

Теорема 1.5.1. *Пусть $\mathcal{F}$ — семейство состоящее из $m = 2^{\left(\frac{1}{2}+\delta\right)n}$ подмножеств множества $X = \{1, 2, \ldots, n\}$, где $\delta > 0$. Тогда*

$$d(\mathcal{F}) < m^{2-\frac{\delta^2}{2}}. \tag{1.1}$$

Доказательство. Предположим, что неравенство (1.1) неверно. Выберем (в качестве элементов семейства $\mathcal{F}$) случайно и независимо t множеств $A_1, A_2, \ldots, A_t$ где t — большое натуральное число, которое будет определено позже. Мы покажем, что с положительной вероятностью выполняется неравенство $|A_1 \cup A_2 \cup \cdots \cup A_t| > n/2$, и что данное объединение не пересекается более чем с $2^{n/2}$ различными подмножествами множества X. Это противоречие покажет, что неравенство (1.1) верно. В самом деле,

$$\begin{aligned}\Pr\left[|A_1 \cup A_2 \cup \cdots \cup A_t| \leqslant n/2\right] &\leqslant \sum_{S \subset X, |S| \leqslant n/2} \Pr[A_i \subset S, i = 1, \ldots, t] \leqslant \\ &\leqslant 2^n \left(2^{n/2}/2^{((1/2)+\delta)n}\right)^t = 2^{n(1-\delta t)}. \end{aligned} \tag{1.2}$$

Введем обозначение

$$v(B) = \left|\{A \in \mathcal{F} : B \cap A = \varnothing\}\right|.$$

Ясно, что

$$\sum_{B \in \mathcal{F}} v(B) = 2d(\mathcal{F}) \geqslant 2m^{2-\delta^2/2}.$$

Пусть Y — случайная величина, значение которой равно числу множеств $B \in \mathcal{F}$, которые не пересекаются с каждым из A_i $(1 \leqslant i \leqslant t)$. В силу выпуклости функции z^t математическое ожидание величины Y удовлетворяет

условиям

$$\mathbf{E}[Y] = \sum_{B\in\mathcal{F}} \bigl(v(B)/m\bigr)^t = \frac{1}{m^t}\cdot m\left(\frac{\sum v(B)^t}{m}\right) \geqslant$$

$$\geqslant \frac{1}{m^t}\cdot m\left(\frac{2d(\mathcal{F})}{m}\right)^t \geqslant 2m^{1-t\delta^2/2}. \tag{1.3}$$

Так как $Y \leqslant m$, то

$$\Pr[Y \geqslant m^{1-t\delta^2/2}] \geqslant m^{-t\delta^2/2}. \tag{1.4}$$

Легко проверить, что при $t = \lceil 1 + 1/\delta \rceil$ выполняется $m^{1-t\delta^2/2} > 2^{n/2}$, и правая часть неравенства (1.4) больше правой части неравенства (1.2). Итак, с положительной вероятностью выполняется неравенство $|A_1 \cup A_2 \cup \cdots \cup A_t| > n/2$ и, к тому же, это объединение не пересекается с более чем $2^{n/2}$ множествами из семейства $\mathcal{F}$. Это противоречие доказывает неравенство (1.1). ■

1.6. УПРАЖНЕНИЯ

1. Доказать, что если существует действительное число p, $0 \leqslant p \leqslant 1$, такое, что

$$\binom{n}{k} p^{\binom{k}{2}} + \binom{n}{t}(1-p)^{\binom{t}{2}} < 1,$$

то число Рамсея $R(k,t)$ удовлетворяет неравенству $R(k,t) > n$. С использованием этого доказать неравенство

$$r(4,t) \geqslant \Omega(t^{3/2}/(\ln t)^{3/2}).$$

2. Пусть $n \geqslant 4$, и H является n-однородным гиперграфом с не более чем $\frac{4^{n-1}}{3^n}$ ребрами. Доказать, что существует раскраска вершин графа H в четыре цвета, такая, что в каждом ребре представлены все эти четыре цвета.

3*. Доказать, что для любых двух независимых, одинаково распределенных действительных случайных величин X и Y

$$\Pr[|X - Y| \leqslant 2] \leqslant 3\Pr[|X - Y| \leqslant 1].$$

4*. Пусть $G = (V, E)$ — граф с n вершинами, причем минимальная степень δ его вершины больше 10. Доказать, что существует разбиение множества V на два непересекающихся подмножества A и B, такое, что $|A| \leqslant O(\frac{n\ln\delta}{\delta})$ и каждая вершина из множества B имеет по крайней мере одного соседа в каждом из множеств A и B.

5*. Рассмотрим граф $G = (V, E)$ на $n \geqslant 10$ вершинах. Допустим, что при добавлении любого ребра число копий полных графов на десяти вершинах в нем увеличивается. Доказать, что число ребер в графе G не меньше, чем $8n - 36$.

6*. Теорема 1.2.1 утверждает, что для любого натурального числа k существует турнир $T_k = (V, E)$, $|V| > k$, в котором для любого множества U, состоящего не более чем из k вершин, существует вершина v, такая, что все ориентированные ребра $\{(v, u) : u \in U\}$ содержатся в E. Доказать, что каждый такой турнир содержит не меньше чем $\Omega(k2^k)$ вершин.

7. Пусть $\{(A_i, B_i), 1 \leqslant i \leqslant h\}$ — семейство пар подмножеств множества целых чисел, такое, что $|A_i| = k$ для всех i, а $|B_i| = l$ для всех i. Кроме того, пусть $A_i \cap B_i = \varnothing$ и $(A_i \cap B_j) \cup (A_j \cap B_i) \neq \varnothing$ для всех $i \neq j$. Доказать, что $h \leqslant \frac{(k+l)^{k+l}}{k^k l^l}$.

8. (*Коды, обладающие свойством префикса; неравенство Крафта.*) Пусть F — конечное множество наборов из нулей и единиц конечной длины. Пусть никакой набор из F не является префиксом другого набора из F. Обозначим через N_i число наборов длины i в F. Доказать, что

$$\sum_i \frac{N_i}{2^i} \leqslant 1.$$

9*. (*Однозначно декодируемые коды; неравенство Крафта—Макмиллана.*) Пусть F — конечное множество наборов из нулей и единиц конечной длины. Допустим, что никакие две различные конкатенации двух конечных последовательностей кодовых слов не дают в результате один и тот же набор. Обозначим через N_i число наборов длины i в F. Доказать, что

$$\sum_i \frac{N_i}{2^i} \leqslant 1.$$

10. Доказать, что существует абсолютная константа $c > 0$ со следующим свойством. Пусть A — квадратная матрица размера $n \times n$ с попарно различными элементами. Тогда существует перестановка строк матрицы A, такая, что ни один из столбцов полученной матрицы не содержит возрастающую подпоследовательность длины, не меньшей, чем $c\sqrt{n}$.

ВЕРОЯТНОСТНЫЙ ВЗГЛЯД:

Теорема Эрдёша—Ко—Радо

Семейство $\mathcal{F}$ множеств называется *пересекающимся*, если для любых двух множеств $A, B \in \mathcal{F}$ выполняется условие $A \cap B \neq \varnothing$. Пусть $n \geqslant 2k$ и семейство $\mathcal{F}$ является пересекающимся семейством k-элементных подмножеств множества мощности n. Для определенности будем считать, что элементы $\mathcal{F}$ являются подмножествами множества $\{0, \ldots, n-1\}$. Теорема Эрдёша—Ко—Радо утверждает, что $|\mathcal{F}| \leqslant \binom{n-1}{k-1}$. Данная оценка достижима на семействе k-множеств, содержащих какую-то определенную точку. Мы приводим короткое доказательство, полученное в [Katona (1972)].

Лемма 1.6.1. *Для каждого $s, 0 \leqslant s \leqslant n-1$, рассмотрим множество $A_s = \{s, s+1, \ldots, s+k-1\}$, где сумма берется по модулю n. Тогда $\mathcal{F}$ может содержать не более k множеств A_s.*

Доказательство. Зафиксируем некоторое $A_s \in \mathcal{F}$. Из всех остальных множеств A_t, пересекающих A_s, составим $k-1$ пар $\{A_{s-i}, A_{s+k-i}\}$, $1 \leqslant i \leqslant k-1$. Элементы каждой такой пары являются непересекающимися. Утверждение леммы теперь вытекает из того, что $\mathcal{F}$ не может содержать более одного элемента из каждой пары. ■

Теперь мы докажем теорему Эрдёша—Ко—Радо. Выберем случайным образом перестановку σ на множестве $\{0, \ldots, n-1\}$ и число $i \in \{0, \ldots, n-1\}$. Пусть $A = \{\sigma(i), \sigma(i+1), \ldots, \sigma(i+k-1)\}$, где сумма, как и раньше, берется по модулю n. В силу произвольности выбора σ из леммы вытекает, что $\Pr[A \in \mathcal{F}] \leqslant k/n$. Следовательно, $\Pr[A \in \mathcal{F}] \leqslant k/n$. Но A выбирается равновероятно из всех k-множеств, поэтому

$$\frac{k}{n} \geqslant \Pr[A \in \mathcal{F}] = \frac{|\mathcal{F}|}{\binom{n}{k}}$$

и

$$|\mathcal{F}| \leqslant \frac{k}{n}\binom{n}{k} = \binom{n-1}{k-1}. \qquad ■$$

2

Линейность математического ожидания

Поиск истины важнее обладания ею.

Альберт Эйнштейн

2.1. ОСНОВЫ

Пусть $X_1, \ldots, X_n$ — случайные величины, $X = c_1X_1 + \ldots + c_nX_n$. Линейность математического ожидания означает, что

$$\mathbf{E}[X] = c_1\mathbf{E}[X_1] + \ldots + c_n\mathbf{E}[X_n].$$

Сила этого принципа в том, что не требуется никаких ограничений на зависимость или независимость слагаемых X_i. Во многих случаях $\mathbf{E}[X]$ может быть легко вычислено путем разложения на простые случайные величины X_i (чаще всего индикаторы).

Пусть σ — случайная перестановка на множестве $\{1, \ldots, n\}$, выбранная равновероятно. Обозначим через $X(\sigma)$ количество неподвижных точек перестановки σ. Чтобы найти $\mathbf{E}[X]$, рассмотрим разложение $X = X_1 + \ldots + X_n$, где X_i — индикатор события $\sigma(i) = i$. Тогда

$$\mathbf{E}[X_i] = \Pr[\sigma(i) = i] = \frac{1}{n},$$

и

$$\mathbf{E}[X] = \frac{1}{n} + \ldots + \frac{1}{n} = 1.$$

В приложениях мы часто пользуемся тем, что существует точка вероятностного пространства, для которой $X \geqslant \mathbf{E}[X]$ и точка, для которой $X \leqslant \mathbf{E}[X]$. Результаты, представленные в этой главе, подобраны так, чтобы проиллюстрировать эту базовую методологию. Следующая теорема Селе [Szele (1943)] многократно упоминается как первое использование вероятностного метода.

Теорема 2.1.1. *Существует турнир T с n игроками и по меньшей мере с $n!2^{-(n-1)}$ гамильтоновыми путями.*[1)]

[1)] Гамильтоновым называется остовный путь, несамопересекающийся по вершинам; см. Харри Ф. Теория графов. — М.: Мир, 1973. — *Прим. ред.*

Доказательство. Обозначим через X число гамильтоновых путей в случайном турнире. Для каждой перестановки σ обозначим через X_σ индикатор того, что σ порождает гамильтонов путь, т. е. выполняется условие $(\sigma(i), \sigma(i+1)) \in T$ для $1 \leqslant i < n$. Тогда $X = \sum X_\sigma$ и

$$\mathbf{E}[X] = \sum \mathbf{E}[X_\sigma] = n! 2^{-(n-1)}.$$

Таким образом, существует турнир, имеющий по крайней мере $\mathbf{E}[X]$ гамильтоновых путей. ■

Селе предположил, что максимально возможное число гамильтоновых путей в турнире из n игроков не превосходит $\frac{n!}{(2-o(1))^n}$. Это было доказано в [Alon (1990a)] и представлено в разд. «Вероятностный взгляд: гамильтоновы пути» (гл. 4).

2.2. РАЗБИЕНИЕ ГРАФОВ

Теорема 2.2.1. *Пусть $G = (V, E)$ — n-вершинный граф с e ребрами. Тогда G содержит двудольный подграф с не менее чем $e/2$ ребрами.*

Доказательство. Пусть $T \subseteq V$ — случайное подмножество, заданное распределением $\Pr[x \in T] = 1/2$, причем элементы подмножества выбираются независимо друг от друга. Положим $B = V - T$. Назовем ребро $\{x, y\}$ *соединяющим*, если ровно одна из вершин x, y принадлежит T. Через X обозначим число соединяющих ребер. Разложим

$$X = \sum_{\{x,y\} \in E} X_{xy},$$

где X_{xy} — индикатор того, что ребро $\{x, y\}$ является соединяющим. Тогда

$$\mathbf{E}[X_{xy}] = 1/2,$$

так как вероятность того, что результаты двух подбрасываний «правильной» монеты будут различными, равна $1/2$. Следовательно,

$$\mathbf{E}[X] = \sum_{\{x,y\} \in E} \mathbf{E}[X_{xy}] = \frac{e}{2}.$$

Таким образом, $X \geqslant e/2$ для некоторого T, а двудольный граф определяется множеством соединяющих ребер. ■

Более тонко построенное вероятностное пространство позволяет немного улучшить результат.

Теорема 2.2.2. *Если граф G содержит $2n$ вершин и e ребер, то в нем найдется двудольный подграф с не менее чем $\frac{en}{2n-1}$ ребрами. Если граф G содержит $2n+1$ вершин и e ребер, тогда в нем найдется двудольный подграф с не менее чем $\frac{e(n+1)}{2n+1}$ ребрами.*

Доказательство. Пусть граф G имеет $2n$ вершин, выберем T случайно среди всех n-элементных подмножеств V. Каждое ребро $\{x, y\}$ является соединяющим с вероятностью $\frac{n}{2n-1}$, доказательство проводится аналогично предыдущему. Пусть G имеет $2n+1$ вершин, выберем T случайно из множества всех n-элементных подмножеств V, дальнейшее доказательство аналогично предыдущему. ■

Рассмотрим теперь более сложный пример, в котором выбор распределения требует предварительной леммы. Положим $V = V_1 \cup \ldots \cup V_k$, где V_i представляют собой непересекающиеся множества мощности n. Рассмотрим отображение $h : [V]^k \to \{-1, +1\}$ как раскраску k-элементного множества в два цвета. Множество E, состоящее из k элементов (в дальнейшем, k-множество) является соединяющим, если оно содержит в точности один элемент из каждого V_i. Для $S \subseteq V$ положим $h(S) = \sum h(E)$, где сумма берется по всем k-множествам $E \subseteq S$.

Теорема 2.2.3. *Предположим, что $h(E) = +1$ для всех соединяющих k-множеств E. Тогда найдется множество $S \subseteq V$, для которого выполнено неравенство*

$$|h(S)| \geqslant c_k n^k,$$

где c_k — положительная константа, не зависящая от n.

Лемма 2.2.4. *Обозначим через P_k множество всех однородных полиномов $f(p_1, \ldots, p_k)$ степени k с коэффициентами, не превосходящими по абсолютному значению единицы, у которых коэффициент при $p_1 p_2 \cdots p_k$ равен единице. Тогда для всех $f \in P_k$ существует набор $p_1, \ldots, p_k \in [0, 1]$ такой, что*

$$|f(p_1, \ldots, p_k)| \geqslant c_k,$$

где c_k — положительная константа, не зависящая от f.

Доказательство. Положим

$$M(f) = \max_{p_1, \ldots, p_k \in [0,1]} |f(p_1, \ldots, p_k)|.$$

Для каждого $f \in P_k$ выполнено неравенство $M(f) > 0$, так как f является ненулевым полиномом. Из того, что P_k является компактом, а отображение $M : P_k \to R$ непрерывно, следует, что M достигает своего минимума, равного c_k. ■

Доказательство теоремы 2.2.3. Выберем случайное множество $S \subseteq V$, полагая

$$\Pr[x \in S] = p_i, \quad x \in V_i,$$

где выбор каждого элемента происходит независимо от других. Определим вероятности p_i. Положим $X = h(S)$. Для каждого k-множества E положим

$$X_E = \begin{cases} h(E), & \text{если } E \subseteq S, \\ 0 & \text{иначе.} \end{cases}$$

Будем говорить, что множество E имеет тип $(a_1, \ldots, a_k)$, если $|E \cap V_i| = a_i$, $1 \leqslant i \leqslant k$. Для таких E справедливо равенство

$$\mathbf{E}[X_E] = h(E)\Pr[E \subseteq S] = h(E)p_1^{a_1} \cdots p_k^{a_k}.$$

Группируя слагаемые по их типу, получаем

$$\mathbf{E}[X] = \sum_{a_1+\ldots+a_k=k} p_1^{a_1} \cdots p_k^{a_k} \sum_{E \text{ имеет тип } (a_1,\ldots,a_k)} h(E).$$

Если $a_1 = \ldots = a_k = 1$, то все $h(E) = 1$ по условию теоремы. Таким образом,

$$\sum_{E \text{ имеет тип } (1,\ldots,1)} h(E) = n^k.$$

Для каждого другого типа существует менее n^k слагаемых, каждое из которых имеет значение ± 1, тогда

$$\left| \sum_{E \text{ имеет тип } (a_1,\ldots,a_k)} h(E) \right| \leqslant n^k.$$

Следовательно,

$$\mathbf{E}[X] = n^k f(p_1, \ldots, p_k),$$

причем $f \in P_k$, где P_k определено в лемме 2.2.4.

Выберем теперь $p_1, \ldots, p_k \in [0, 1]$ так, что $|f(p_1, \ldots, p_k)| \geqslant c_k$. Тогда

$$\mathbf{E}[|X|] \geqslant |\mathbf{E}[X]| \geqslant c_k n^k.$$

Некоторое значение $|X|$ должно превосходить или быть равным своему математическому ожиданию. Следовательно, найдется множество $S \subseteq V$, такое, что

$$|X| = |h(S)| \geqslant c_k n^k. \qquad \blacksquare$$

Теорема 2.2.3 имеет интересное применение в теории Рамсея. Известно (см. [Erdős (1965b)]), что при заданной раскраске в два цвета k-множеств, являющихся подмножествами некоторого n-множества, существуют такие k непересекающихся m-множеств, $m = \Theta((\ln n)^{1/(k-1)})$, что все соединяющие k-множества одноцветны. Из теоремы 2.2.3 следует, что существует множество размера $\Theta((\ln n)^{1/(k-1)})$, в котором по крайней мере $\frac{1}{2} + \varepsilon_k$ множеств размера k одноцветны. Это довольно неожиданно, так как известно, что существуют раскраски, в которых размер наибольшего монохроматического множества не превосходит $(k-2)$-кратного логарифма n.

2.3. ДВА БЫСТРЫХ РЕЗУЛЬТАТА

Линейность математического ожидания позволяет иногда быстро получать интересные результаты.

Теорема 2.3.1. *Существует раскраска в два цвета графа K_n, при которой число монохроматических подграфов K_a не превосходит*

$$\binom{n}{a} 2^{1-\binom{a}{2}}.$$

Доказательство [набросок]. Рассмотрим некоторую случайную 2-раскраску. Обозначим через X число монохроматических подграфов K_a и найдем $\mathbf{E}[X]$. Для некоторой раскраски значение X не превосходит значения математического ожидания. ■

В гл. 15 показано, как такая раскраска может быть найдена с помощью детерминированного и эффективного алгоритма.

Теорема 2.3.2. *Существует раскраска в два цвета графа $K_{m,n}$, при которой число монохроматических подграфов $K_{a,b}$ не превосходит*

$$\binom{m}{a}\binom{n}{b} 2^{1-ab}.$$

Доказательство [набросок]. Рассмотрим произвольную случайную 2-раскраску. Обозначим через X число монохроматических подграфов $K_{a,b}$ и найдем $\mathbf{E}[X]$. Для некоторой раскраски значение X не превосходит значения математического ожидания. ■

2.4. БАЛАНСИРОВКА ВЕКТОРОВ

У следующего результата имеется элегантное *не*вероятностное доказательство, которое мы приводим в конце этой главы. Здесь через $|v|$ обозначена обычная евклидова норма вектора v.

Теорема 2.4.1. *Пусть векторы $v_1, \ldots, v_n \in \mathbb{R}^n$ таковы, что все $|v_i| = 1$. Тогда существует набор $\varepsilon_1, \ldots, \varepsilon_n = \pm 1$, такой, что*

$$|\varepsilon_1 v_1 + \ldots + \varepsilon_n v_n| \leqslant \sqrt{n}.$$

Кроме того существует набор $\varepsilon_1, \ldots, \varepsilon_n = \pm 1$, такой, что

$$|\varepsilon_1 v_1 + \ldots + \varepsilon_n v_n| \geqslant \sqrt{n}.$$

Доказательство. Выберем элементы набора $\varepsilon_1, \ldots, \varepsilon_n$ равновероятно и независимо из множества $\{-1, +1\}$. Положим

$$X = |\varepsilon_1 v_1 + \ldots + \varepsilon_n v_n|^2.$$

Тогда

$$X = \sum_{i=1}^{n} \sum_{j=1}^{n} \varepsilon_i \varepsilon_j v_i \cdot v_j.$$

Следовательно,

$$\mathbf{E}[X] = \sum_{i=1}^{n}\sum_{j=1}^{n} v_i \cdot v_j \mathbf{E}[\varepsilon_i \varepsilon_j].$$

Если $i \neq j$, то $\mathbf{E}[\varepsilon_i\varepsilon_j] = \mathbf{E}[\varepsilon_i]\mathbf{E}[\varepsilon_j] = 0$. Если же $i = j$, то $\varepsilon_i^2 = 1$ и $\mathbf{E}[\varepsilon_i^2] = 1$. Тогда

$$\mathbf{E}[X] = \sum_{i=1}^{n} v_i \cdot v_i = n.$$

Таким образом, найдутся наборы $\varepsilon_1, \ldots, \varepsilon_n = \pm 1$, такие, что $X \geqslant n$, и такие, что $X \leqslant n$. Извлекая квадратные корни, получаем требуемые утверждения. ■

Следующий результат включает в себя теорему 2.4.1 как частный случай (при $p_1 = \ldots = p_n = 1/2$).

Теорема 2.4.2. *Пусть векторы $v_1, \ldots, v_n \in \mathbb{R}^n$ таковы, что все $|v_i| \leqslant 1$, а значения $p_1, \ldots, p_n \in [0,1]$ произвольны. Положим $w = p_1v_1 + \ldots + p_nv_n$. Тогда существует набор $\varepsilon_1, \ldots, \varepsilon_n \in \{0,1\}$, такой, что при $v = \varepsilon_1 v_1 + \ldots + \varepsilon_n v_n$ выполняется неравенство*

$$|w - v| \leqslant \frac{\sqrt{n}}{2}.$$

Доказательство. Выберем величины ε_i независимо друг от друга с вероятностями

$$\Pr[\varepsilon_i = 1] = p_i, \Pr[\varepsilon_i = 0] = 1 - p_i.$$

Случайный выбор чисел ε_i порождает случайный вектор v и случайную величину

$$X = |w - v|^2.$$

Заметим, что

$$X = \left|\sum_{i=1}^{n}(p_i - \varepsilon_i)v_i\right|^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} v_i \cdot v_j (p_i - \varepsilon_i)(p_j - \varepsilon_j).$$

Тогда

$$\mathbf{E}[X] = \sum_{i=1}^{n}\sum_{j=1}^{n} v_i \cdot v_j \mathbf{E}[(p_i - \varepsilon_i)(p_j - \varepsilon_j)].$$

Для $i \neq j$ имеем

$$\mathbf{E}[(p_i - \varepsilon_i)(p_j - \varepsilon_j)] = \mathbf{E}[p_i - \varepsilon_i]\mathbf{E}[p_j - \varepsilon_j] = 0.$$

Для $i = j$ получаем

$$\mathbf{E}[(p_i - \varepsilon_i)^2] = p_i(p_i - 1)^2 + (1 - p_i)p_i^2 = p_i(1 - p_i) \leqslant \frac{1}{4},$$

(величина $\mathbf{E}[(p_i - \varepsilon_i)^2] = \mathrm{Var}[\varepsilon_i]$, называемая *дисперсией*, будет рассмотрена в гл. 4). Таким образом,

$$\mathbf{E}[X] = \sum_{i=1}^{n} p_i(1 - p_i)|v_i|^2 \leqslant \frac{1}{4}\sum_{i=1}^{n}|v_i|^2 \leqslant \frac{n}{4},$$

и доказательство завершается так же, как и в теореме 2.4.1. ■

2.5. РАЗБАЛАНСИРОВКА ЛАМПОЧЕК

Теорема 2.5.1. *Пусть $a_{ij} = \pm 1$ при $1 \leqslant i, j \leqslant n$. Тогда существуют $x_i, y_j = \pm 1$, $1 \leqslant i, j \leqslant n$, такие, что*

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_iy_j \geqslant \left(\sqrt{\frac{2}{\pi}} + o(1)\right) n^{3/2}.$$

У этого результата есть забавная интерпретация. Рассмотрим матрицу размера $n \times n$, составленную из лампочек, каждая из которых либо включена ($a_{ij} = +1$), либо выключена ($a_{ij} = -1$). Предположим, что для каждой строки и каждого столбца имеется переключатель, поворот которого ($x_i = -1$ для строки i и $y_j = -1$ для столбца j) переключает все лампочки в соответствующей линии: с «вкл.» на «выкл.» и с «выкл.» на «вкл.». Тогда для любой начальной конфигурации можно установить такое положение переключателей, что число выключенных лампочек будет не меньше $\left(\sqrt{\frac{2}{\pi}} + o(1)\right) n^{3/2}$.

Доказательство теоремы 2.5.1. Забудем пока о числах x_i. Пусть величины $y_1, \ldots, y_n = \pm 1$ выбираются независимо и равновероятно. Положим

$$R_i = \sum_{j=1}^{n} a_{ij}y_j, \quad R = \sum_{i=1}^{n} |R_i|.$$

Зафиксируем i. Независимо от a_{ij} величины $a_{ij}y_j$ принимают значения $+1$ или -1 с вероятностями $1/2$, кроме того, они независимы (по j). (То есть, независимо от того, что представляла из себя i-я строка изначально, после случайного переключения она становится равномерно распределенной, и все 2^n возможных значений становятся равновероятными.) Таким образом, случайная величина R_i распределена так же, как величина $\mathbf{S}_n$, являющаяся суммой n независимых, одинаково распределенных случайных величин, принимающих значения $\{\pm 1\}$. Поэтому

$$\mathbf{E}[|R_i|] = \mathbf{E}[|\mathbf{S}_n|] = \left(\sqrt{\frac{2}{\pi}} + o(1)\right)\sqrt{n}.$$

После некоторых элементарных преобразований эта асимптотика может быть получена при оценке $\mathbf{S}_n$ величиной $\sqrt{n}N$, где N — случайная величина со стандартным нормальным распределением. С другой стороны, точную формулу

$$\mathbf{E}[|\mathbf{S}_n|] = n2^{1-n}\binom{n-1}{\lfloor (n-1)/2 \rfloor}$$

можно вывести комбинаторными методами (задача из олимпиады Путнама 1974 г.), а асимптотика следует из формулы Стирлинга.

Воспользуемся теперь линейностью математического ожидания

$$\mathbf{E}[R] = \sum_{i=1}^{n} \mathbf{E}[|R_i|] = \left(\sqrt{\frac{2}{\pi}} + o(1)\right) n^{3/2}.$$

Существует набор $y_1, \ldots, y_n = \pm 1$, при котором R принимает не меньшее значение. В заключение выберем числа x_i так, чтобы знак каждого из них

совпадал со знаком соответствующей величины R_i. Тогда

$$\sum_{i=1}^{n} x_i \sum_{j=1}^{n} a_{ij} y_j = \sum_{i=1}^{n} x_i R_i = \sum_{i=1}^{n} |R_i| = R \geqslant \left(\sqrt{\frac{2}{\pi}} + o(1) \right) n^{3/2}. \qquad \blacksquare$$

Другой результат по разбалансировке лампочек приводится в разд. «Вероятностный взгляд: несбалансированные матрицы» (гл. 12).

2.6. БЕЗ ПОДБРАСЫВАНИЯ МОНЕТ

*Не*вероятностное доказательство теоремы 2.2.1 может быть получено путем последовательного включения каждой вершины в множества T или B. На каждом шаге поместим x либо в T, либо в B так, что бы по крайней мере половина ребер из x, инцидентных предыдущим вершинам, были соединяющими. При использовании этого эффективного алгоритма по крайней мере половина ребер будут соединяющими.

Существует простой итерационный алгоритм для выбора знаков в теореме 2.4.1. Для выбора знака v_i вычислим частичную сумму $w = \varepsilon_1 v_1 + \ldots + \varepsilon_{i-1} v_{i-1}$. Теперь, если требуется получить малое значение суммы, следует выбрать значение $\varepsilon_i = \pm 1$ так, чтобы вектор $\varepsilon_i v_i$ составлял тупой (или прямой) угол с w. Если же желательно получить большое значение суммы, то следует сделать угол острым или прямым. В вырожденном случае, когда все углы прямые, с помощью теоремы Пифагора и индукции можно показать, что конечный вектор w будет иметь норму $\sqrt{n}$, в остальных случаях норма будет либо меньше $\sqrt{n}$, либо больше $\sqrt{n}$, как нам того и хотелось.

В теореме 2.4.2 требуемые ε_i можно получить с помощью так называемого «жадного» алгоритма. Пусть нам даны $v_1, \ldots, v_n \in \mathbb{R}^n$, $p_1, \ldots, p_n \in [0, 1]$. Предположим, что величины $\varepsilon_1, \ldots, \varepsilon_{s-1} \in \{0, 1\}$ уже выбраны. Рассмотрим частичную сумму $w_{s-1} = \sum_{i=1}^{s-1} (p_i - \varepsilon_i) v_i$. Выберем ε_s так, чтобы вектор

$$w_s = w_{s-1} + (p_s - \varepsilon_s) v_s = \sum_{i=1}^{s} (p_i - \varepsilon_i) v_i$$

имел минимальную норму. Случайное число $\varepsilon_s \in \{0, 1\}$, выбранное с вероятностью $\Pr[\varepsilon_s = 1] = p_s$, дает нам

$$\begin{aligned} \mathbf{E}[|w_s|^2] &= |w_{s-1}|^2 + 2 w_{s-1} \cdot v_s \mathbf{E}[p_s - \varepsilon_s] + |v_s|^2 \mathbf{E}[p_s - \varepsilon_s]^2 = \\ &= |w_{s-1}|^2 + p_s (1 - p_s) |v_s|^2. \end{aligned} \tag{2.1}$$

Таким образом, для некоторого выбора $\varepsilon_s \in \{0, 1\}$ выполняется неравенство

$$|w_s|^2 \leqslant |w_{s-1}|^2 + p_s (1 - p_s) |v_s|^2.$$

Поскольку это верно для всех $1 \leqslant s \leqslant n$ (при $w_0 = 0$), для вектора, получаемого жадным алгоритмом, будет выполняться неравенство

$$|w_n|^2 \leqslant \sum_{i=1}^{n} p_i (1 - p_i) |v_i|^2.$$

Несмотря на то, что эти доказательства схожи, прямое применение доказательства теоремы 2.4.2 для нахождения $\varepsilon_1, \ldots, \varepsilon_n$ может привести к перебору, требующему экспоненциального времени. При применении жадного алгоритма на стадии с номером s производится два вычисления значения $|w_s|^2$: для $\varepsilon_s = 0$ и 1, с последующим выбором того значения ε_s, которое дает наименьшее значение. Таким образом, мы проделываем лишь линейное количество вычислений норм, а алгоритм в целом занимает лишь квадратичное время. В гл. 15 мы обсудим несколько схожих примеров при более общей постановке задач.

2.7. УПРАЖНЕНИЯ

1. Пусть $n \geqslant 2$, и $H = (V, E)$ — n-однородный гиперграф с числом ребер, равным $|E| = 4^{n-1}$. Показать, что существует такая раскраска множества вершин V в четыре цвета, что ни одно ребро не является монохроматическим.

2. Доказать, что существует положительная константа c, такая, что в каждом множестве A, состоящем из n отличных от нуля действительных чисел, содержится подмножество $B \subset A$ мощности $|B| \geqslant cn$, в котором нет четверок $b_1, b_2, b_3, b_4 \in B$, удовлетворяющих условию
$$b_1 + 2b_2 = 2b_3 + 2b_4.$$

3. Доказать, что каждое множество, состоящее из n отличных от нуля **действительных** чисел, содержит подмножество A мощности, **строго** большей чем $n/3$, в котором нет троек $a_1, a_2, a_3 \in A$, удовлетворяющих равенству $a_1 + a_2 = a_3$.

4. Предположим, что $p > n > 10m^2$, p — простое, и пусть $0 < a_1 < a_2 < \ldots < a_m < p$ — целые. Доказать, что существует целое x, $0 < x < p$, для которого m чисел
$$\big((xa_i) \pmod p\big) \bmod n, \quad 1 \leqslant i \leqslant m,$$
попарно различны.

5. Рассмотрим граф H и целое число $n > |V(H)|$. Предположим, что существует граф с n вершинами и t ребрами, не содержащий копий графа H, и пусть $tk > n^2 \ln n$. Показать, что для полного графа на n вершинах существует раскраска его ребер в k цветов без монохроматических копий графа H.

6*. С использованием соображений вероятностного взгляда на гамильтоновы пути, доказать, что существует константа $c > 0$, такая, что для любого четного $n \geqslant 4$ верно следующее утверждение: для всякого неориентированного полного графа K с n вершинами, ребра которого раскрашены в красный и синий цвета, число альтернирующих гамильтоновых циклов в K (т. е.

правильно раскрашенных по ребрам циклов длины n) не больше

$$n^c \frac{n!}{2^n}.$$

7. Пусть $\mathcal{F}$ — семейство подмножеств множества $N = \{1, 2, \ldots, n\}$, в котором нет пар $A, B \in \mathcal{F}$, удовлетворяющих отношению $A \subset B$. Пусть $\sigma \in \mathbb{S}_n$ — случайная перестановка элементов множества N. Рассмотрим случайную величину
$$X = |\{i : \{\sigma(1), \sigma(2), \ldots, \sigma(i)\} \in \mathcal{F}\ \}|.$$
Путем вычисления математического ожидания величины X доказать, что $|\mathcal{F}| \leqslant \binom{n}{\lfloor n/2 \rfloor}$.

8*. Пусть X — множество попарно ортогональных единичных векторов из $\mathbb{R}^n$. Предположим, что проекция каждого из этих векторов на первые k координат имеет евклидову норму, не меньшую ε. Показать, что $|X| \leqslant k/\varepsilon^2$, и что эта оценка является точной для всех $\varepsilon = k/2^r < 1$.

9. Рассмотрим двудольный граф $G = (V, E)$ с n вершинами. Пусть для каждой вершины $v \in V$ задан список $S(v)$ из более чем $\log_2 n$ цветов. Доказать, что существует правильная раскраска графа G, приписывающая каждой вершине v цвет из ее списка $S(v)$.

ВЕРОЯТНОСТНЫЙ ВЗГЛЯД:

Теорема Брегмана

Пусть $A = [a_{ij}]$ — матрица размера $n \times n$ с элементами $a_{ij} \in \{0, 1\}$, а $r_i = \sum_{1 \leqslant j \leqslant n} a_{ij}$ — количество единиц в i-й строке. Обозначим через S множество таких перестановок $\sigma \in \mathbb{S}_n$, для которых $a_{i,\sigma i} = 1$ при всех i, $1 \leqslant i \leqslant n$. Тогда перманент[2)] $\operatorname{per}(A)$ есть просто $|S|$. Следующий результат был сформулирован Минком и доказан Брегманом [Брегман (1973)]. Представленное здесь доказательство аналогично тому, которое предложил Схрийвер [Schrijver (1978).]

Теорема (Брегман). *Имеет место неравенство*

$$\operatorname{per}(A) \leqslant \prod_{1 \leqslant i \leqslant n} (r_i!)^{1/r_i}.$$

Выберем $\sigma \in S$ и $\tau \in \mathbb{S}_n$ независимо и равновероятно. Положим $A^1 = A$. Пусть $R_{\tau 1}$ — количество единиц в строке $\tau 1$ матрицы A^1. Удалив строку $\tau 1$ и столбец $\sigma\tau 1$ из A^1, получим матрицу A^2. В общем случае, пусть матрица A^i получена из A удалением строк $\tau 1, \ldots, \tau(i-1)$ и столбцов $\sigma\tau 1, \ldots, \sigma\tau(i-1)$, и пусть $R_{\tau i}$ — число единиц в строке τi из A^i. (Оно не равно 0, поскольку $\sigma\tau i$-й столбец содержит 1.) Пусть

$$L = L(\sigma, \tau) = \prod_{1 \leqslant i \leqslant n} R_{\tau i}.$$

Грубо говоря, можно рассматривать L как вычисление перманента Лентяем. Существуют $R_{\tau 1}$ возможностей выбора единицы в строке $\tau 1$, каждая из которых ведет к различным вычислениям субперманента. Вместо этого Лентяй берет множитель $R_{\tau 1}$, берет одну из перестановок σ и рассматривает A^2. Поскольку $\sigma \in S$ выбирается равновероятно, Лентяй стремится к максимальным субперманентам и поэтому не удивительно, что это приводит к завышенной оценке перманента. Чтобы сделать ее точной, мы определим геометрическое среднее $G[Y]$. Если $Y > 0$ принимает значения $a_1, \ldots, a_s$ с вероятностями $p_1, \ldots, p_s$ соответственно, то $G[Y] = \prod a_i^{p_i}$. Или, что то же, $G[Y] = e^{\mathbf{E}[\ln Y]}$. Линейность математического ожидания переводит произведение средних геометрических в среднее геометрическое произведений.

Утверждение. *Имеет место неравенство* $\operatorname{per}(A) \leqslant G[L]$.

Доказательство. Докажем это для произвольного фиксированного τ. Для краткости положим $\tau 1 = 1$. Проведем индукцию по размеру матрицы. Для удобства переупорядочим столбцы так, чтобы первая строка начиналась с r единиц, где $r = r_1$. Для $1 \leqslant j \leqslant r$ пусть t_j является перманентом матрицы,

[2)] Для произвольной матрицы $A = [a_{ij}]$ ее перманентом называется число $\operatorname{per}(A) = \sum_{\sigma \in \mathbb{S}_n} \prod_i a_{i,\sigma i}$. — *Прим. ред.*

полученной из A удалением первой строки и j-го столбца. Этот перманент равен числу перестановок $\sigma \in S$, таких, что $\sigma 1 = j$. Положим

$$t = \frac{t_1 + \ldots + t_r}{r},$$

так что $\mathrm{per}(A) = rt$ при условии, что $\sigma 1 = j$. Произведение $R_2 \cdots R_n$ представляет собой вычисление Лентяем величины $\mathrm{per}(A^2)$, где матрица A^2 получена из матрицы A удалением первой строки и j-го столбца. По индукции

$$G[R_2 \cdots R_n | \sigma 1 = j] \geqslant t_j$$

и, следовательно,

$$G[L] \geqslant \prod_{j=1}^{r} (rt_j)^{t_j/\mathrm{per}(A)} = r \prod_{j=1}^{r} t_j^{t_j/rt}.$$

Для продолжения доказательства утверждения воспользуемся следующей леммой.

Лемма. *Справедливо неравенство*

$$\left(\prod_{j=1}^{r} t_j^{t_j} \right)^{1/r} \geqslant t^t.$$

Доказательство. Логарифмируя, приходим к эквивалентному неравенству

$$\frac{1}{r} \sum_{j=1}^{r} t_j \ln t_j \geqslant t \ln t,$$

которое вытекает из выпуклости функции $f(x) = x \ln x$. ■

Применяя лемму, имеем

$$G[L] \geqslant r \prod_{j=1}^{r} t_j^{t_j/rt} \geqslant r(t^t)^{1/t} = rt = \mathrm{per}(A).$$ ■

Теперь вычислим $G[L]$ для фиксированной перестановки σ. Для удобства проведем переупорядочение так, что $\sigma i = i$ для всех i, и предположим, что первая строка имеет единицы ровно в первых r_1 столбцах. Столбцы $1, \ldots, r_1$ удаляются в порядке τ, выбранном случайным образом из всех $r_1!$ возможностей. R_1 — число столбцов, остающихся в момент, когда первый столбец должен быть удален. Поскольку первый столбец может оказаться в любой из r_1 позиций, то R_1 равномерно распределен на множестве от 1 до r_1, и мы имеем $G[R_1] = (r_1!)^{1/r_1}$. «Линейность» дает теперь

$$G[L] = G\left[\prod_{i=1}^{n} R_i \right] = \prod_{i=1}^{n} G[R_i] = \prod_{i=1}^{n} (r_i!)^{1/r_i}.$$

Величина $G[L]$ является средним геометрическим условного $G[L]$ и, следовательно, имеет то же значение. То есть

$$\mathrm{per}(A) \leqslant G[L] = \prod_{i=1}^{n} (r_i!)^{1/r_i}.$$ ■

3

Малые вариации

Красота — важнейший тест: в мире нет места для некрасивой математики.

Г. Х. Харди

Идея вероятностного метода сформулирована в гл. 1следующим образом: пытаясь доказать, что некоторая структура с определенными свойствами существует, мы определяем подходящее вероятностное пространство структур, а затем показываем, что требуемые свойства выполняются в этом пространстве с положительной вероятностью. В этой главе рассматриваются ситуации, в которых «случайная» структура не обладает всеми требуемыми свойствами, а может иметь несколько изъянов. С помощью небольшой переделки мы удаляем эти изъяны, получая требуемую структуру.

3.1. ЧИСЛА РАМСЕЯ

Напомним (см. гл. 1, разд. 1.1), что неравенство $R(k,l) > n$ означает существование 2-раскраски ребер полного графа K_n в красный и синий цвета, при которой не существует ни красных подграфов K_k, ни синих подграфов K_l.

Теорема 3.1.1. *Для любого натурального n*

$$R(k,k) > n - \binom{n}{k} 2^{1-\binom{k}{2}}.$$

Доказательство. Рассмотрим случайную 2-раскраску ребер графа K_n, полученную раскраской каждого ребра независимо и равновероятно либо красным, либо синим цветом. Для всякого множества R из k вершин пусть X_R является индикатором следующего события: «индуцированный подграф K_n на множестве R является монохроматическим». Положим $X = \sum X_R$, где сумма берется по всем R. Из линейности математического ожидания вытекает, что

$$\mathbf{E}[X] = \sum \mathbf{E}[X_R] = m, \text{ где} m = \binom{n}{k} 2^{1-\binom{k}{2}}.$$

Таким образом, существует 2-раскраска, для которой $X \leqslant m$. Зафиксируем такую раскраску. Удалим по одной вершине из каждого такого монохроматического k-множества. Понадобится удалить не более m вершин (возможно, что некоторая вершина удаляется более одного раза, но это ничему не мешает). Поэтому остается $s \geqslant n-m$ вершин. Полученная раскраска на этих s вершинах не имеет монохроматических k-множеств. ■

В заключение рассмотрим «оптимизационную» задачу получения наиболее сильной формы доказанного неравенства путем выбора подходящего значения n. Анализ показывает, что следует взять $n \sim e^{-1}k2^{k/2}(1-o(1))$, при котором

$$R(k,k) > \frac{1}{e}(1+o(1))k2^{k/2}.$$

Внимательный анализ предложения 1.1.1 (см. гл. 1) дает нижнюю оценку

$$R(k,k) > \frac{1}{e\sqrt{2}}(1+o(1))k2^{k/2}.$$

Более мощная локальная лемма Ловаса (см. гл. 5) позволяет получить

$$R(k,k) > \frac{\sqrt{2}}{e}(1+o(1))k2^{k/2}.$$

Исследование различий между этими границами может показаться нелогичным, поскольку наилучшая известная верхняя оценка для $R(k,k)$ имеет вид $(4+o(1))^k$. Для получения верхних оценок вероятностные методы не используются. Эти оценки можно найти, например, в [Graham, Rothschild and Spencer (1990)]. Мы даем все три нижние оценки, следуя философии нашего подхода, состоящего в главенстве методологии над результатами.

Что касается недиагональных чисел Рамсея, то различия между базовым методом и перестройкой иллюстрируются следующими двумя результатами.

Теорема 3.1.2. *Если существует такое $p \in [0,1]$, что*

$$\binom{n}{k}p^{\binom{k}{2}} + \binom{n}{l}(1-p)^{\binom{l}{2}} < 1,$$

то $R(k,l) > n$.

Теорема 3.1.3. *Для любых целых n и $p \in [0,1]$ верно неравенство*

$$R(k,l) > n - \binom{n}{k}p^{\binom{k}{2}} - \binom{n}{l}(1-p)^{\binom{l}{2}}.$$

Доказательство. В обоих случаях рассматривается случайная 2-раскраска графа K_n, полученная независимым раскрашиванием каждого ребра в красный или синий цвет, при котором ребро окрашивается красным цветом с вероятностью p. Пусть X — сумма числа красных k-множеств и числа синих l-множеств. Из линейности математического ожидания следует, что

$$\mathbf{E}[X] = \binom{n}{k}p^{\binom{k}{2}} + \binom{n}{l}(1-p)^{\binom{l}{2}}.$$

В теореме 3.1.2 $\mathbf{E}[X] < 1$, поэтому существует 2-раскраска с $X = 0$. В теореме 3.1.3 существует 2-раскраска с s «плохими» множествами (красными k-множествами или синими l-множествами), $s \leqslant \mathbf{E}[X]$. Удаление по одной вершине из каждого плохого множества позволяет получить раскраску по крайней мере $n - s$ вершин без плохих множеств. ■

Получение асимптотик для теорем 3.1.2, 3.1.3 может оказаться весьма сложным. Часто теорема 3.1.3 дает существенное улучшение по сравнению с теоремой 3.1.2. Дальнейшие продвижения могут быть получены с использованием локальной леммы Ловаса. Эти оценки проанализированы Спенсером; см. [Spencer (1977)].

3.2. НЕЗАВИСИМЫЕ МНОЖЕСТВА

Здесь приводятся короткие и изящные соображения, позволяющие, грубо говоря, доказать половину знаменитой теоремы Турана.. Пусть $\alpha(G)$ — максимальный размер независимого множества в графе G. Неравенство $\alpha(G) \geqslant t$ означает, что существует t вершин без ребер между ними.

Теорема 3.2.1. *Пусть граф $G = (V, E)$ имеет n вершин и $nd/2$ ребер, $d \geqslant 1$. Тогда $\alpha(G) \geqslant n/2d$.*

Доказательство. Пусть $S \subseteq V$ — случайное подмножество, определяемое равенством

$$\Pr[v \in S] = p,$$

где p будет определено позднее. События $v \in S$ являются взаимно независимыми. Пусть $X = |S|$ и пусть Y — число ребер в $G|_S$. Для каждого $e = \{i, j\} \in E$ обозначим через Y_e индикатор события «$i, j \in S$». Тогда $Y = \sum_{e \in E} Y_e$. Для каждого такого e имеем

$$\mathbf{E}[Y_e] = \Pr[i, j \in S] = p^2.$$

В силу линейности математического ожидания,

$$\mathbf{E}[Y] = \sum_{e \in E} \mathbf{E}[Y_e] = \frac{nd}{2} p^2.$$

Ясно, что $\mathbf{E}[X] = np$. Поэтому вновь в силу линейности математического ожидания,

$$\mathbf{E}[X - Y] = np - \frac{nd}{2} p^2.$$

Чтобы максимизировать эту величину, положим $p = 1/d$ (здесь используется, что $d \geqslant 1$) и получим

$$\mathbf{E}[X - Y] = \frac{n}{2d}.$$

Таким образом, существует множество S, для которого число вершин множества S минус число ребер в S не меньше $n/2d$. Выберем по вершине в

каждом ребре множества S и удалим их. Это приводит к множеству S^* с не менее чем $n/2d$ вершинами. Все ребра разрушены, поэтому S^* — независимое множество. ■

Полное доказательство результата Турана дано в разд. «Вероятностный взгляд: теорема Турана» (после гл. 6).

3.3. КОМБИНАТОРНАЯ ГЕОМЕТРИЯ

Рассмотрим множество S из n точек в единичном квадрате U. Обозначим через $T(S)$ минимальную из площадей треугольников с вершинами в трех различных точках из S. Положим $T(n) = \max T(S)$, где максимум берется по семейству всех n-точечных подмножеств S множества U. Хейлбронн предположил, что $T(n) = O(1/n^2)$. Эта гипотеза опровергнута в работе [Komlós, Pintz and Szemerédi (1982)], где с помощью довольно сложной конструкции показано, что существует множество S из n точек множества U, такое, что $T(S) = \Omega(\log n/n^2)$. Поскольку это доказательство довольно громоздкое, мы приведем здесь более простое доказательство, показывающее, что $T(n) = \Omega(1/n^2)$.

Теорема 3.3.1. *Существует множество S из n точек в единичном квадрате U, такое, что $T(S) \geqslant 1/(100n^2)$.*

Доказательство. Сначала проведем некоторые вычисления. Пусть точки P, Q, R выбираются независимо и равновероятно из U. Обозначим через $\mu = \mu(PQR)$ площадь треугольника PQR. Мы оценим величину $\Pr[\mu \leqslant \varepsilon]$ следующим образом. Пусть x — расстояние между P и Q. При этом $\Pr[b \leqslant x \leqslant b + \Delta b] \leqslant \pi(b + \Delta b)^2 - \pi b^2$, а в пределе $\Pr[b \leqslant x \leqslant b + db] \leqslant 2\pi b db$. Если расстояние между точками P, Q равно b, то высота, опущенная из R на прямую PQ, должна иметь величину $h \leqslant 2\varepsilon/b$, и, таким образом, точка R должна лежать в полосе ширины $4\varepsilon/b$ и длины не больше чем $\sqrt{2}$. Это происходит с вероятностью, не превышающей $4\sqrt{2}\varepsilon/b$. Так как $0 \leqslant b \leqslant \sqrt{2}$, искомая вероятность ограничивается сверху величиной

$$\int_0^{\sqrt{2}} (2\pi b)(4\sqrt{2}\varepsilon/b) db = 16\pi\varepsilon.$$

Пусть теперь точки $P_1, \ldots, P_{2n}$ выбираются независимо и равновероятно из U, и пусть X обозначает число треугольников $P_i P_j P_k$ площади, меньшей чем $1/(140n^2)$. Вероятность появления произвольной тройки i, j, k меньше чем $0.6n^{-2}$, и поэтому

$$\mathbf{E}[X] \leqslant \binom{2n}{3}(0.6n^{-2}) < n.$$

Таким образом, существует множество из $2n$ вершин, в котором менее n треугольников имеют площадь меньше, чем $1/(100n^2)$. Удалим по вершине

из каждого такого треугольника. При этом остается не менее n вершин, но теперь уже нет треугольников, имеющих площадь, меньшую чем $1/(100n^2)$. ■

Мы опишем следующую конструкцию Эрдёша, показывающую, что $T(n) \geqslant 1/(6(n-1)^2)$ при простом n. На квадрате $[0, n-1] \times [0, n-1]$ рассмотрим n точек (x, x^2), где x^2 берется по $\bmod n$. (Более формально, (x, y), где $y \equiv x^2 \bmod n$ и $0 \leqslant y < n$.) Если через некоторые три точки этого множества можно провести прямую, то она будет задаваться уравнением $y = mx + b$, где m — рациональное число со знаменателем, не превосходящим n. Но тогда в Z_n^4 парабола $y = x^2$ должна пересекать прямую $y = mx + b$ в трех точках, и, тем самым, квадратный трехчлен $x^2 - mx - b$ имеет три различных корня, что невозможно. Площадь каждого треугольника с вершинами в целых точках плоскости является либо целым числом, либо половиной целого. Следовательно, эти площади должны быть не меньше $1/2$. Сжимая плоскость путем деления обеих координат на $n-1$, получаем требуемое множество. Хотя этот изящный результат лучше теоремы 3.3.1, он все еще слабее результата, полученного в [Komlós et al. (1982)].

3.4. УПАКОВКА

Пусть J — ограниченное измеримое подмножество пространства $\mathbb{R}^d$, а через $B(x)$ обозначен куб $[0, x]^d$ со стороной x. Упаковкой множества C в $B(x)$ называется семейство взаимно непересекающихся копий множества C, каждая из которых лежит в $B(x)$. Обозначим через $f(x)$ наибольшую меру такого семейства. *Константа упаковки* определяется равенством

$$\delta(C) = \mu(C) \lim_{x \to \infty} f(x) x^{-d},$$

где $\mu(C)$ — мера множества C. Таким образом, $\delta(C)$ — это максимальная доля пространства, которая может быть заполнена непересекающимися копиями множества C. (Можно доказать, что предел в определении $\delta(C)$ всегда существует. Но даже без этого следующий результат имеет место с заменой lim на $\liminf$.)

Теорема 3.4.1. *Пусть множество C ограничено, выпукло и центрально симметрично относительно начала координат. Тогда*

$$\delta(C) \geqslant 2^{-d-1}.$$

Доказательство. Пусть точки P, Q выбираются независимо и равномерно из $B(x)$. Рассмотрим событие «$(C + P) \cap (C + Q) \neq \varnothing$». Для того чтобы оно произошло, мы должны иметь для некоторых $c_1, c_2 \in C$

$$P - Q = c_1 - c_2 = 2\frac{c_1 - c_2}{2} \in 2C$$

в силу центральной симметрии и выпуклости. Событие $Q \in Q + 2C$ имеет вероятность, не превышающую $\mu(2C)x^{-d}$ для любого заданного Q, следовательно,

$$\Pr[(C + P) \cap (C + Q) \neq \varnothing] \leqslant \mu(2C)x^{-d} = 2^d x^{-d} \mu(C).$$

Теперь пусть $P_1, \ldots, P_n$ выбираются независимо и равномерно из $B(x)$, а через X обозначено число тех $i < j$, для которых $(C + P_i) \cap (C + P_j) \neq \varnothing$. В силу линейности математического ожидания имеем

$$\mathbf{E}[X] \leqslant \frac{n^2}{2} 2^d x^{-d} \mu(C).$$

Следовательно, существует выборка из n точек с меньшим, чем указанное, числом пересекающихся пар копий множества C. Для каждой пары P_i, P_j с $(C + P_i) \cap (C + P_j) \neq \varnothing$ удалим или P_i, или P_j из множества. Остаются по меньшей мере $n - \frac{n^2}{2} 2^d x^{-d} \mu(C)$ непересекающихся копий множества C. Положим $n = x^d 2^{-d}/\mu(C)$, чтобы максимизировать это количество. Отсюда вытекает существование по меньшей мере $x^d 7^{-d-1}/\mu(C)$ попарно непересекающихся копий множества C. Не все они лежат внутри $B(x)$, но, обозначив через w верхнюю оценку абсолютных величин координат точек множества C, мы обнаружим, что они все лежат внутри куба со стороной $x+2w$. Следовательно,

$$f(x + 2w) \geqslant x^d 2^{-d-1}/\mu(C),$$

а значит,

$$\delta(C) \geqslant \lim_{x \to \infty} \mu(C) f(x + 2w)(x + 2w)^{-d} \geqslant 2^{-d-1}. \qquad \blacksquare$$

Обычный жадный алгоритм дает несколько лучший результат. Пусть $P_1, \ldots, P_m$ — какое-нибудь максимальное подмножество множества $[0, x]^d$, обладающее тем свойством, что множества $C+P_i$ не пересекаются. Мы заметили, что $C+P_i$ перекрывает $C+P$ тогда и только тогда, когда $P \in 2C+P_i$. Следовательно, множества $2C+P_i$ должны покрывать $[0, x]^d$. Поскольку каждое такое множество имеет меру $\mu(2C) = 2^d \mu(C)$, необходимо чтобы $m \geqslant x^d 2^{-d}/\mu(C)$. Как и прежде, все множества $C + P_i$ лежат в кубе со стороной $x + 2w$, где w — константа. Поэтому

$$f(x + 2w) \geqslant m \geqslant x^d 2^{-d}/\mu(C)$$

и, значит,

$$\delta(C) \geqslant 2^{-d}.$$

Дальнейшие улучшения еще появятся в разд. «Вероятностный взгляд: эффективная упаковка» (после гл. 13).

3.5. ПЕРЕКРАСКА

Предположим, при случайной раскраске возникает некоторое множество дефектов. Тогда мы применяем случайную перекраску к дефектам с целью их удаления. Если перекраска является слишком слабой, то не все дефекты устраняются. Если перекраска является слишком сильной, то создаются новые дефекты. Интенсивность перекраски задается параметром p, а указанные две возможности соответствуют убыванию и возрастанию некоторых функций параметра p. Последующие вычисления дают нам оптимальное значение p. Мы используем определение свойства B из разд. 1.3. Свойство B: $m(n) > n$

означает, что для данного n-однородного гиперграфа $H = (V, E)$ с m ребрами существует 2-раскраска множества V такая, что никакое ребро не является монохроматическим. Бек в [Beck (1978)] улучшил оценку Эрдёша (1938), доказав, что $m(n) = \Omega(2^n n^{1/3})$. С использованием его методов доказано (см. [Radhakrishnan and Srinivasan (2000)]), что $m(n) = \Omega(2^n (n/\ln n)^{1/2})$. Именно это доказательство и будет представлено. Оно не является ни длинным, ни технически сложным и разбито на ряд тонких и красивых шагов. Не удивительно, что понадобилось тридцать пять лет, чтобы его найти. Как было сказано, верхняя и нижняя оценки для $m(n)$ по-прежнему далеки друг от друга.

Теорема 3.5.1. *Если существует $p \in [0, 1]$, удовлетворяющее условию*

$$k(1-p)^n + k^2 p < 1,$$

то $m(n) > 2^{n-1}k$.

Следствие 3.5.2. *Справедливо равенство $m(n) = \Omega(2^n (n/\ln n)^{1/2})$.*

Доказательство следствия 3.5.2. Используем неравенство $1 - p \leqslant e^{-p}$. Функция $ke^{-pn} + k^2 p$ принимает наименьшее значение при $p = \ln(n/k)/n$. Подставляя это значение, получаем

$$\frac{k^2}{n}\,[1 + \ln(n/k)] < 1.$$

Тем самым условие теоремы 3.5.1 выполнено. Это неравенство справедливо при $k = c(n/\ln n)^{1/2}$ для всех $c < \sqrt{2}$ и достаточно больших n. ■

Условие теоремы 3.5.1 в некотором смысле типично. Необходимо, чтобы общая вероятность ошибки была меньше 1, причем имеется два типа ошибок. Часто удается найти разумные границы наложением более сильных требований, чем те, при которых ошибка каждого типа имеет вероятность меньше $1/2$. Здесь $k^9 p \leqslant \frac{3}{2}$ дает $p \leqslant \frac{1}{2}k^{-2}$. Подставляя максимально возможное p во второе неравенство $k(1-p)^n \leqslant \frac{1}{2}$, получаем $2k^2 \ln(2k) \leqslant n$. Это справедливо также и при $k = c(n/\ln n)^{1/2}$, хотя при этом мы имеем более слабое условие $c < 9$. Можно рекомендовать этот грубый подход в качестве первой попытки решения задачи, когда приблизительный разброс параметров еще под сомнением. Тонкости вычислений можно оставить до публикации работы!

Доказательство теоремы 3.5.1. Зафиксируем гиперграф $H = (V, E)$ с $m = 2^{n-1}k$ ребрами и число p, удовлетворяющее условию теоремы. Мы опишем рандомизированный алгоритм, приводящий к раскраске множества V. Проведем предварительную обработку случайности: для каждой вершины $v \in V$ бросим сначала первую монету, при этом герб выпадает с вероятностью $\frac{1}{2}$, а затем — вторую монету, при этом герб выпадает с вероятностью p (представляя потенциальную перекраску). Вдобавок (это важно) заметим, что вершины из множества V упорядочены случайно.

Шаг 1. Раскрасим вершину $v \in V$ в красный цвет, если при соответствующем

подбрасывании первой монеты выпал герб. В противном случае раскрасим ее синим. Назовем это первой раскраской. Обозначим через D (в честь dangerous) множество вершин $v \in V$, которые лежат в некотором (возможно, в нескольких) монохроматическом ребре $e \in E$.

Шаг 2. Рассмотрим элементы множества D последовательно в (случайно выбранном) порядке на V. Вершина d рассматривается как *все еще опасная*, если существует некоторое ребро $e \in H$ (их может быть несколько), содержащее d и являвшееся монохроматическим при первой раскраске, для которого ни одна вершина еще не изменила цвет. Если d не является все еще опасной, то не делаем ничего. Но если она все еще опасна, то бросаем вторую монету. Если выпадает герб, то изменяем цвет вершины d, иначе ничего не меняем. Назовем раскраску *финальной*, если она стабилизировалась и далее не меняется.

Скажем, что алгоритм срабатывает неудачно, если в результате раскраски некоторое ребро оказывается монохроматическим. Оценим вероятность ошибки сверху величиной $k(1-p)^n + k^2 p$. Предположение теоремы 3.5.1 тогда гарантирует нам, что с положительной вероятностью алгоритм заканчивается успешно. Отсюда с помощью наших обычных рассуждений получаем, что существует способ применения алгоритма, приводящий к финальной раскраске без монохроматических ребер e, а тем самым, существует 2-раскраска множества V без монохроматических ребер. Для определенности оценим вероятность того, что некоторое ребро $e \in H$ в финальной раскраске является красным. Вероятность неудачного срабатывания алгоритма превышает ее не более чем в два раза.

Ребро $e \in E$ может оказаться красным в финальной раскраске в двух случаях. Либо e было красным после первой раскраски и остается таковым до финальной раскраски, либо e не было красным после первой раскраски, но стало красным при финальной раскраске. (Из определения алгоритма следует, что вершина не может изменить цвет более одного раза.) Пусть A_e будет первым событием, а C_e — вторым. Тогда

$$\Pr[A_e] = 2^{-n}(1-p)^n.$$

Первый множитель есть вероятность того, что e стало красным при первой раскраске, т. е. на первом шаге при всех соответствующих подбрасываниях монеты выпал герб. Второй множитель есть вероятность того, что при всех оставшихся подбрасываниях монеты выпала решетка. Если все подбрасывания сделаны, то ни одна из вершин $v \in e$ не перекрасится на шаге 2. Наоборот, если при каком-либо подбрасывании на втором шаге выпал герб для некоторой вершины $v \in e$, то существует *первая* (в данном упорядочении) вершина v, для которой выпал герб. Когда эта вершина v была еще опасной, ребро e было монохроматическим, и поэтому v ждала выпадения герба на втором шаге для изменения своего цвета. Имеем

$$2\sum_{e \in H} \Pr[A_e] = k(1-p)^n,$$

что является первым слагаемым вероятности неудачи.

В доказательстве Бека, приведенном в первом издании нашей книги, в котором понятие «все еще опасная» отсутствовало, каждая вершина меняла цвет тогда и только тогда, когда на втором шаге выпадал герб. Значения $\Pr[A_e] = 2^{-n}(1-p)^n$ одинаковые в обоих доказательствах. Бек получил оценку $\Pr[C_e] \leqslant k^2 p e^{pn}$. Новое доказательство позволяет избежать лишних перекрасок и приводит к лучшей оценке для $\Pr[C_e]$. Мы переходим к хитроумной оценке вероятности $\Pr[C_e]$.

Для различных ребер $e, f \in E$ скажем, что e *осуждает* f если:

- ребра e, f пересекаются в точности по одной вершине (обозначим ее через v);
- при первой раскраске ребро f было синим, а ребро e при финальной раскраске стало красным;
- на шаге 2 вершина v была *последней* из принадлежащих ребру e, которая изменила синий цвет на красный;
- когда v меняет свой цвет, ребро f все еще полностью синее.

Предположим, что имеет место событие C_e. Некоторые вершины ребра e меняют цвет с синего на красный так, что существует *последняя* вершина v, изменившая цвет. Но почему v требует бросания монеты? Она должна быть все еще опасной. То есть v должна лежать в некотором (возможно, в нескольких) множествах f, которые являются синими при первой раскраске и остаются все еще синими в момент, когда v рассматривается. Могут ли e и f пересекаться по некоторой другой вершине v'? Нет! Такой вершине v' было бы необходимо быть синей при первой раскраске (так как $v' \in f$) и красной при финальной раскраске (поскольку $v' \in e$). Но тогда v' меняет цвет раньше, чем v. Значит, f не является полностью синим, когда v рассматривается, и поэтому e не может осуждать f. Следовательно, когда C_e имеет место, e осуждает некоторое f. Обозначим через B_{ef} событие «e осуждает f». Тогда $\sum_e \Pr[C_e] \leqslant \sum_{e \neq f} \Pr[B_{ef}]$. Поскольку существует меньше чем $(5^{n-1}k)^2$ пар $e \neq f$, теперь достаточно показать, что $\Pr[B_{ef}] \leqslant 2^{8-2n}p$.

Пусть зафиксированы e, f, такие, что $e \cap f = \{v\}$ (иначе B_{ef} не может произойти). Случайное упорядочение множества V порождает случайный порядок σ на $e \cup f$. Пусть $i = i(\sigma)$ обозначает число вершин $v' \in e$, встречающихся перед v в этом порядке, и пусть $j = j(\sigma)$ есть число вершин $v' \in f$, встречающихся перед v. Фиксируя σ, мы имеем

$$\Pr[B_{ef}|\sigma] \leqslant \frac{p}{2} 2^{-n+1}(1-p)^j 2^{-n+1+i} \left(\frac{1+p}{2}\right)^i .$$

Будем рассматривать множители по одному. Во-первых, вершина v сама должна быть сначала синей и превратиться в красную. Во-вторых, все другие $v' \in f$ должны быть сначала синими. В-третьих, все $v' \in f$ встречаются раньше чем v при перекраске. В-четвертых, вершины $v' \in e$, встречающиеся после v, должны быть вначале красными (так как v — последняя вершина множества e, которая меняет цвет). Наконец, все $v' \in e$, встречающиеся перед v, должны либо сначала быть красными, либо сначала быть синими, но перекраситься в красный. Последний множитель можно оценить сверху достаточно точно. Те

вершины $v' \in e$, которые встречаются перед v и являлись вначале синими, не только требуют второго подбрасывания монеты и выпадения герба, но, кроме того, сами должны лежать в некотором $e' \in H$, ставшем монохроматическим после первой раскраски. Попытки дальнейшего улучшения оценок величины $m(n)$ часто концентрируются на этом, но (до сих пор!) они оказывались безуспешными.

Теперь мы имеем

$$\Pr[B_{ef}] \leqslant 2^{1-2n} p E\left[(1+p)^i (1-p)^j\right],$$

где математическое ожидание берется при равновероятном выборе упорядочения σ. Следующая жемчужина завершает доказательство.

Лемма 3.5.3. *Справедлива оценка:* $\mathbf{E}\left[(1+p)^i(1-p)^j\right] \leqslant 1$.

Доказательство. Зафиксируем некоторое соответствие между $e \setminus \{v\}$ и $f \setminus \{v\}$, держа в уме м-ра и миссис Джонс, м-ра и миссис Смит, и т. д. Никаких ограничений на количество лиц, которые приходят перед v, в паре нет (двое Джонсов, один Смит, никого из Тейлоров, и т. д.). Условное математическое ожидание величины $(1+p)^i(1-p)^j$ раскладываем на множители для каждой пары. В случае отсутствия Тейлоров соответствующий множитель отсутствует. Если Джонсов двое, присутствует множитель $(1+p)(1-p) \leqslant 1$. Если Смит один, то с равной вероятностью присутствует ровно один из множителей $1+p$ или $1-p$. Таким образом, условное математическое ожидание заменяется множителем 1. Каждый множитель не превышает 1, поэтому их произведение не больше 1. ■

Отсюда вытекает требуемый результат. ■

3.6. НЕПРЕРЫВНОЕ ВРЕМЯ

Дискретные случайные процессы могут иногда изучаться путем помещения их в рамки непрерывного времени. Это позволяет применить мощные методы математического анализа (например, интегрирование). Такой подход оказывается наиболее эффективным, когда речь идет о случайных упорядочениях. Ниже приводятся два примера.

Свойство B. Мы модифицируем доказательство оценки $m(n) = \Omega(2^n n^{1/2} \ln^{-1/2} n)$ из предыдущего раздела. Припишем каждой вершине $v \in V$ «момент рождения» x_v. Тогда x_v — это независимые действительные переменные, каждая из которых равномерно распределена на отрезке $[0,1]$. При этом порядок на множестве V задается упорядочением (относительно $\leqslant$) величин x_v. Мы утверждаем, что

$$\Pr[B_{ef}] \leqslant \sum_{l=0}^{n-1} \binom{n-1}{l} 2^{1-2n} \int_0^1 x^l p^{l+1} (1-xp)^{n-1}\, dx.$$

Для $T \subseteq e \setminus \{v\}$ обозначим через B_{efT} следующее событие: «выполнено B_{ef}, и при первой раскраске ребро e содержит в точности $T \cup \{v\}$ синих вершин». Имеется $\binom{n-1}{l}$ способов выбора l-множества T при l, меняющемся от 0 до $n-1$. Тогда первая раскраска множества $e \cup f$ определена и имеет вероятность появления, равную 2^{1-2n}. Предположим, что v имеет момент рождения $x_v = e$. Для всех $w \in T \cup \{v\}$ при втором подбрасывании монеты должен выпадать герб. Вероятность этого равна p^{l+1}. Все $w \in T$ должны родиться раньше v, т. е. $x_w < x$, что выполняется с вероятностью x^l. Ни для какой вершины $w \in f \setminus \{v\}$, рождающейся прежде v, при бросании монеты не может выпасть герб. Каждая такая w имеет вероятность рождения xp, а вероятность того, что ни одна из w не родится, равна $(1-xp)^{n-1}$. Поскольку величина $x_v = x$ равномерно распределена в $[0,1]$, мы интегрируем по x. Суммируя по l, имеем

$$\Pr[B_{ef}] \leqslant 2^{1-2n} p \int_0^1 (1+xp)^{n-1}(1-xp)^{n-1}\, dx.$$

Подынтегральное выражение всегда не превосходит единицы, поэтому $\Pr[B_{ef}] \leqslant 2^{1-2n}p$. Остающаяся часть доказательства не меняется.

Случайная жадная упаковка. Пусть H — $(k+1)$-однородный гиперграф на множестве вершин V размера N. Элементы $e \in H$, называемые ребрами, являются просто $(k+1)$-подмножествами множества V. Мы предполагаем, что выполнены следующие условия:

Условие на степени: каждая вершина $v \in V$ принадлежит в точности D ребрам.

Условие на ко-степени: любая пара $v, v' \in V$ имеет не более $o(D)$ общих ребер.

Мы предполагаем, что k фиксировано ($k = 2$ является иллюстративным примером), а асимптотика рассматривается при $N, D \to \infty$ без какой-либо связи между N и D.

Упаковка — это семейство P непересекающихся по вершинам ребер $e \in H$. Ясно, что $|P| \leqslant N/(k+1)$. Определим рандомизированный алгоритм получения (не обязательно оптимальной) упаковки. Припишем каждому ребру $e \in H$ равномерно и независимо момент рождения $x_e \in [0, D)$. (Выбор $[0, D)$ вместо $[0, 1]$ сделан по техническим соображениям. Заметим, что поскольку x_e — действительные величины, с вероятностью 1 они независимы.) В нулевой момент $P = \varnothing$. При изменении времени от 0 к D в момент рождения ребра e оно добавляется к P по возможности — т. е. всегда, когда не существует ребра $e' \in P$, пересекающегося с e. Пусть P_c обозначает величину P в точности перед моментом c, когда все e с временем рождения $t_e < c$ уже рассмотрены. Положим $P^{\text{FINAL}} = P_D$. Заметим, что к моменту D все ребра появились, а их появления совершались в случайном порядке. Таким образом, P^{FINAL} идентична дискретному процессу (часто называемому вероятностным жадным алгоритмом), при котором ребра гиперграфа H сначала случайно упорядочиваются, а затем рассматриваются последовательно.

Теорема 3.6.1 [Спенсер (1995)]. *Математическое ожидание величины* $|P^{\text{FINAL}}|$ *асимптотически равно* $N/(k+1)$.

Скажем, что вершина $v \in V$ *живет* в момент c, если ни одно из ребер $e \in P_c$ не содержит v. Обозначим через S_c множество таких $v \in V$. Вместо P^{FINAL} будем рассматривать P_c, где c — произвольное действительное число. Пусть

$$f(c) = \overset{*}{\lim} \Pr[v \in S_c],$$

где, формально, мы имеем ввиду, что для всякого $\varepsilon > 0$ существуют D_0, N_0 и $\delta > 0$, такие, что если H является $(k+1)$-однородным гиперграфом на $N > N_0$ вершинах, каждая вершина v содержится в $D > D_0$ ребрах, а каждая пара вершин $v, v' \in V$ имеет меньше чем δD общих ребер, то $|f(c) - \Pr[v \in S_c]| < \varepsilon$ для всех $v \in V$.

Главная часть доказательства состоит в том, чтобы показать, что $f(c)$ существует как величина, определяемая некоторым процессом, состоящим из актов рождения (и смерти) и происходящим в непрерывном времени. Теперь мы опишем такой процесс, опуская некоторые эпсилон-дельта манипуляции, необходимые для формального доказательства существования предела.

Наш процесс начинается в момент c, а время движется к 0. Он начинается с корня Ева (это наша «очеловеченная» вершина v). Ева производит потомство во временном интервале $[0, c)$. Число родов распределено по закону Пуассона со средним значением c, а моменты родов распределены равномерно. [Это стандартный пуассоновский процесс с интенсивностью 1. Его можно понимать также следующим образом: на произвольно малом интервале времени $[x, x + dx)$ Ева может с вероятностью dx произвести потомство, и эти события независимы для непересекающихся интервалов]. При каждых родах наша плодовитая Ева всегда производит на свет k потомков. Плодовитость каждого потомка подчиняется тем же правилам. Так, если длительность репродуктивного периода Алисы равна x, то число родов (в нашей однополой модели) имеет пуассоновское распределение со средним x, а моменты времени родов равномерно распределены в промежутке $[0, x)$.

Результирующее случайное дерево $T = T_c$, как можно показать, является конечным (ввиду конечности интервала) с вероятностью 1. В произвольно заданном конечном дереве T каждая вершина (назовем ее, скажем, «Алисой») остается в живых или умирает согласно следующей схеме.

Правило Менендеса: если Алиса дала жизнь множеству (или, быть может, нескольким множествам) из k потомков, каждый из которых остается живым, она умирает. В противном случае она остается живой.

В частности, если Алиса бездетна, она продолжает жить. Таким образом, по дереву мы можем определить для каждой вершины, жива она или нет.

Пример. Пусть $c = 10, k = 2$. Ева производит на свет Алису и Барбару в момент 8.3, а затем Рэчел и Сиенну в момент 4.3. Алиса рожает Нэнси и Оливию в момент 5.7, а Рэчел рожает Линду и Маявати в момент 0.4. Других родов нет. Остаются в живых Нэнси, Оливия, Линда, Маявати, Барбара и Сиенна. В соответствии с деревом, Алиса и Рэчел умирают. Ни при каком из родов оба потомка не остаются в живых, поэтому Ева продолжает жить.

Обозначим через $f(c)$ вероятность того, что корневая вершина Ева остается живой в случайном родословном дереве $T = T_c$.

Определим дерево $T = T_c(v)$ для произвольной вершины $v \in H$. Для каждого ребра e, содержащего вершину v с моментом рождения $t = t_e < c$, скажем, что $e \setminus \{v\}$ есть множество k потомков, порожденных вершиной v в момент t. Мы действуем рекурсивно. Если w родилась в момент t, то для всякого e', содержащего w с моментом рождения $t' = t_{e'} < t$, мы скажем, что $e' \setminus \{w\}$ есть множество k потомков, рожденных вершиной w в момент t'. Возможно, что этот процесс не дает дерево, так как одна и та же вершина w может достигаться более чем одним путем. Простейшим примером является ситуация, когда $v \in e, e'$, где оба имеют момент рождения, меньший чем c, и ребра e, e' имеют еще одну общую вершину w. Тогда процесс является вырожденным, а $T_c(v)$ не определено. Докажем, что для произвольно заданного дерева T

$$\lim^{*} \Pr[T_c(v) \cong T] = \Pr[T_c = T]. \tag{3.1}$$

Поскольку $\sum_T \Pr[T_c = T] = 1$, это является косвенным аргументом в пользу того, что процесс определения $T_c(v)$ почти никогда не является вырожденным.

Мы построим $T_c(v)$ поэтапно. Сначала рассмотрим D ребер e, содержащих вершину v. Число таких ребер с моментом рождения $t_e < c$ имеет биномиальное распределение $B[D, \frac{c}{D}]$, сходящееся к пуассоновскому со средним значением c. При условии, что существует l таких ребер e, их моменты рождения t_e распределены равномерно. По ко-степенному условию существует $o(D^2)$ пар e, e', содержащих v и еще некоторую другую вершину. Значит, вероятность того, что два ребра e, e' имеют момент рождения меньше чем c, есть $o(1)$. Теперь предположим, что дерево $T_c(v)$ построено до определенного уровня, и вершина w рождена в момент t. Существует только $o(D)$ общих ребер между w и любыми из конечного числа уже рожденных к этому моменту вершин w'. Поэтому существуют все еще $\sim D$ ребер e, содержащих w и не содержащих никакой другой вершины w'. Далее рассмотрим моменты их рождения. Число вершин с $t_e < x$ имеет биномиальное распределение $B[D - o(D), \frac{x}{D}]$, сходящееся к распределению Пуассона со средним значением x. Как и прежде, почти наверное не существует двух таких ребер e, e' с общей вершиной, отличной от w. Для любого фиксированного T величина $\Pr[T_c(v) \cong T]$ является суммой конечного числа таких пределов, что позволяет нам сделать заключение о справедливости равенства (3.1).

Для $c < d$ случайное дерево T_d включает T_c, состоящее из вершин, соответствующих тем родам Евы, которые происходят в интервале $[0, c)$. Если Ева жива в дереве T_d, она должна быть жива и в T_c. Следовательно, $f(d) \leqslant f(c)$. Мы утверждаем, что

$$\lim_{c \to \infty} f(c) = 0,$$

ибо в противном случае невозрастающая функция f имела бы предел $L > 0$ и $f(x) \geqslant L$. Предположим, что к моменту T_c Ева производит потомство i раз. Для каждого акта рождения существует вероятность по меньшей мере L^k, что все k потомков живы. Вероятность того, что Ева жива, не больше $(1 - L^k)^i$. Так как число родов Евы распределено по Пуассону со средним значением c,

имеем

$$f(c) \leqslant \sum_{i=0}^{\infty} e^{-c}\frac{c^i}{i!}(1-L^k)^i = e^{-L^k c}.$$

Но тогда $\lim_{c\to\infty} f(c) = 0$. Приходим к противоречию.

В силу линейности математического ожидания $\mathbf{E}[|S_c|] \to f(c)n$. Поскольку $(k+1)|P_c|+|S_c| = n$, имеем $\mathbf{E}[|P_c|] \to (1-f(c))n/(k+1)$. Но $\mathbf{E}[|P^{\text{FINAL}}|] \geqslant \mathbf{E}[|P_c|]$. Мы делаем $f(c)$ произвольно малой, выбирая константу c достаточно большой, так что $\mathbf{E}[|P^{\text{FINAL}}|] \geqslant (1-o(1))n/(k+1)$. Поскольку $|P^{\text{FINAL}}| \leqslant n/(k+1)$ всегда, теорема доказана. ■

Замечание. В действительности, мы можем сказать о $f(c)$ больше. Для малого Δc имеем $f(c+\Delta c) - f(c) \sim -(\Delta c)f(c)^{k+1}$. Это, грубо говоря, означает, что Ева, начиная с момента $c + \Delta c$, могла бы производить потомство во временном интервале $[c, c+\Delta c)$ (причем потомки, родившиеся в этом интервале живы), и в то же время не иметь родов в интервале $[0, c)$ с остающимися в живых потомками. При $\Delta c \to 0$ приходим к дифференциальному уравнению $f'(c) = -f(c)^{k+1}$. Начальное значение $f(0) = 1$ дает единственное решение $f(c) = (1+ck)^{-1/k}$. Интересно было бы положить $c = D$. Но это недопустимо, поскольку наши предельные соображения распространяются на фиксированные c при $N, D \to \infty$. Тем не менее это *бы* повлекло $E[|S_D|] = O(ND^{-1/k})$, а это означает, что вероятностный жадный алгоритм оставлял бы $O(ND^{-1/k})$ вершин непокрытыми. Предположим, мы заменили ко-степенное условие на более сильное, что каждая пара различных вершин $v, v' \in V$ имеет не более одной общей вершины. Компьютерное моделирование показывает, что в подобных случаях жадный алгоритм оставляет непокрытыми $O(ND^{-1/k})$ вершин. Это означает, что вопрос по-прежнему открыт, хотя в работе [Alon, Kim and Spencer (1997)] доказано, что в таких случаях неплохо работает модифицированная версия жадного алгоритма.

Следствие 3.6.2. *В условиях теоремы существует упаковка P размера $\sim N/(k+1)$.*

Доказательство. Мы определили случайный процесс, который дает упаковку с ожидаемым размером $\sim N/(k+1)$, и теперь с помощью наших обычных приемов убеждаемся в том, что такая упаковка P должна существовать. ■

В частности, это дает альтернативное доказательство гипотезы Эрдёша—Ханани, впервые доказанной Рёдлем (см. разд. 4.7). Мы используем определения этого раздела и определяем число $m(n,k,l)$ как максимальный размер семейства F из k-элементных подмножеств множества $[n] = \{1,\ldots,n\}$, такой, что не существует l-множества, содержащегося в более чем одном k-множестве. Определим гиперграф $H = H(n,k,l)$ следующим образом: вершины гиперграфа H суть l-элементные подмножества множества $[n]$. Для каждого k-элементного множества $A \subset [n]$ определим ребро e_A как множество l-элементных подмножеств множества A. Тогда семейство F, удовлетворяющее указанным

выше условиям, соответствует упаковке $P = \{e_A : A \in F\}$ в H. Гиперграф H имеет $N = \binom{n}{l}$ вершин. Каждое ребро e_A имеет размер $K+1 = \binom{k}{l}$. Каждая вершина содержится в $D = \binom{n-l}{k-l}$ ребрах. Число ребер, содержащих две вершины v, v', зависит от их пересечения. Оно максимально (при $v \neq v'$), когда v, v' (рассматриваемые как l-множества) пересекаются по $l-1$ точкам, т. е. $\binom{n-l-1}{k-l-1}$. Мы предполагаем (как в разд. 4.7), что k, l фиксированы и $n \to \infty$, а значит, число общих ребер есть $o(D)$. Предположение из разд. 4.7 дает при фиксированном $K+1$ и $N, D \to \infty$ существование упаковки P, такой, что

$$m(n,k,l) = |P| \sim N/(K+1) \sim \frac{\binom{n}{l}}{\binom{k}{l}}.$$

3.7. УПРАЖНЕНИЯ

1. Доказать, что число Рамсея $r(k,k)$ для каждого целого n удовлетворяет неравенству
$$r(k,k) > n - \binom{n}{k} 2^{1-\binom{k}{2}},$$
откуда следует, что
$$r(k,k) \geqslant (1-o(1))\frac{k}{e}2^{k/2}.$$

2. Доказать, что число Рамсея $r(4,k)$ удовлетворяет неравенству
$$r(4,k) \geqslant \Omega((k/\ln k)^2).$$

3. Доказать, что каждый 3-однородный гиперграф с n вершинами и $m \geqslant n/3$ ребрами содержит независимое множество размера по меньшей мере $\frac{2n^{3/2}}{3\sqrt{3}\sqrt{m}}$.

4*. Показать, что существует конечное n_0, такое, что любой ориентированный граф на $n > n_0$ вершинах, в котором степень исхода не меньше $\log_2 n - \frac{1}{10}\log_2\log_2 n$, содержит четный простой цикл.

ВЕРОЯТНОСТНЫЙ ВЗГЛЯД:

Большой обхват и большое хроматическое число

Многие рассматривают следующее ниже утверждение как одно из самых красивых применений вероятностного метода, поскольку сам результат удивительный и, на первый взгляд, не предполагает неконструктивного доказательства.

Обхватом $girth(G)$ графа G называется размер его кратчайшего цикла. Через $\alpha(G)$ обозначается максимальный размер независимого множества в G, а через $\chi(G)$ — хроматическое число G.

Теорема [Erdős (1959)]. *Для любых k, l существует граф G с обхватом $girth(G) > l$ и $\chi(G) > k$.*

Доказательство. Зафиксируем $\theta < 1/l$ и положим $G \sim G(n,p)$ с $p = n^{\theta-1}$. (То есть G — случайный граф на n вершинах, порождаемый выбором каждой пары вершин в качестве ребра случайно и независимо с вероятностью p). Пусть X — количество циклов размера, не превосходящего l. Тогда

$$\mathbf{E}[X] = \sum_{i=3}^{l} \frac{(n)_i}{2i} p^i \leqslant \sum_{i=3}^{l} \frac{n^{\theta i}}{2i} = o(n),$$

так как $\theta l < 1$. В частности,

$$\Pr[X \geqslant n/2] = o(1).$$

Положим $x = \lceil \frac{3}{p} \ln n \rceil$, тогда

$$\Pr[\alpha(G) \geqslant x] \leqslant \binom{n}{x} (1-p)^{\binom{x}{2}} < \left[ne^{-p(x-1)/2}\right]^x = o(1).$$

Пусть n настолько велико, что оба события имеют вероятность, меньшую чем 0.5. Тогда существует граф G с менее чем $n/2$ циклами длины, не большей l, и с $\alpha(G) < 3n^{1-\theta} \ln n$. Удалим из G по вершине из каждого цикла длины, не превышающей l. Получим граф G^* по крайней мере с $n/2$ вершинами. При этом G^* имеет обхват больше чем l и $\alpha(G^*) \leqslant \alpha(G)$. Таким образом,

$$\chi(G^*) \geqslant \frac{|G^*|}{\alpha(G^*)} \geqslant \frac{n/2}{3n^{1-\theta} \ln n} = \frac{n^\theta}{6 \ln n}.$$

Чтобы завершить доказательство, возьмем n достаточно большим с тем, чтобы последнее было больше k. ■

4

Второй момент

Вы можете не верить в Бога, но должны верить в Книгу.[1)]

Пол Эрдёш

4.1. ОСНОВЫ

После математического ожидания наиболее используемым статистическим понятием для случайной величины X является *дисперсия*. Мы обозначаем ее через $\operatorname{Var}[X]$. Она определяется равенством

$$\operatorname{Var}[X] = \mathbf{E}[(X - \mathbf{E}[X])^2]$$

и является мерой отклонения величины X от своего математического ожидания. Следуя обычной практике, будем обозначать через μ математическое ожидание, а через σ^2 — дисперсию. Положительный квадратный корень σ из дисперсии называется *стандартным отклонением*. При этих обозначениях нашим основным инструментом будет следующее утверждение.

Теорема 4.1.1 (неравенство Чебышёва). *Для любого положительного* λ

$$\Pr[|X - \mu| \geqslant \lambda\sigma] \leqslant \frac{1}{\lambda^2}.$$

Доказательство. Имеем

$$\sigma^2 = \operatorname{Var}[X] = \mathbf{E}[(X - \mu)^2] \geqslant \lambda^2\sigma^2 \Pr[|X - \mu| \geqslant \lambda\sigma]. \qquad \blacksquare$$

Использование неравенства Чебышёва принято называть *методом второго момента*.

Неравенство Чебышёва неулучшаемо, если нет никаких дополнительных ограничений на X. Например, если X принимает значения $\mu + \lambda\sigma$ и $\mu - \lambda\sigma$ с вероятностью $\lambda^2/2$, а значение μ — с вероятностью $1 - \lambda^2$, то неравенство превращается в равенство. Заметим, однако, что если X имеет нормальное

[1)] Смысл высказывания станет понятным после прочтения приложения В.4. — *Прим. ред.*

распределение со средним μ и стандартным отклонением σ, то

$$\Pr[|X-\mu| \geqslant \lambda\sigma] = 2\int_{\lambda}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt.$$

Для больших λ эта вероятность асимптотически равна $\sqrt{2/\pi}e^{-\lambda^2/2}/\lambda$, что существенно меньше $1/\lambda^2$. В гл. 7 и 8 будут рассмотрены примеры, когда X является суммой «почти независимых» случайных величин, и подобные улучшенные оценки оказываются справедливыми.

Предположим, что

$$X = X_1 + \ldots + X_m.$$

Тогда $\mathrm{Var}[X]$ может быть подсчитана по формуле

$$\mathrm{Var}[X] = \sum_{i=1}^{m} \mathrm{Var}[X_i] + \sum_{i \neq j} \mathrm{Cov}[X_i, X_j].$$

Здесь вторая сумма берется по упорядоченным парам $(i,j) \neq (j,i)$, а *ковариация* $\mathrm{Cov}[Y,Z]$ определяется следующим образом:

$$\mathrm{Cov}[Y,Z] = \mathbf{E}[YZ] - \mathbf{E}[Y]\mathbf{E}[Z].$$

Если случайные величины Y и Z независимы, то $\mathrm{Cov}[Y,Z] = 0$. Это часто существенно упрощает вычисление дисперсии. В дальнейшем мы предполагаем (это всегда будет выполняться в последующем), что X_i являются индикаторами некоторых событий, т. е. $X_i = 1$, если событие A_i происходит, и $X_i = 0$ иначе. Если X_i равна единице с вероятностью $p_i = \Pr[A_i]$, то

$$\mathrm{Var}[X_i] = p_i(1-p_i) \leqslant p_i = \mathbf{E}[X_i],$$

и поэтому

$$\mathrm{Var}[X] \leqslant \mathbf{E}[X] + \sum_{i \neq j} \mathrm{Cov}[X_i, X_j].$$

4.2. ТЕОРИЯ ЧИСЕЛ

Метод второго момента оказывается эффективным в теории чисел. Пусть $\nu(n)$ обозначает количество простых чисел p, делящих n. (Мы не учитываем кратность, хотя это приводит к некоторым отличиям.) Следующий результат утверждает, грубо говоря, что «почти все» n имеют число простых делителей «весьма близкое» к $\ln\ln n$. Это впервые было обнаружено Харди и Рамануджаном в 1920 г. с помощью весьма сложного доказательства. Мы представим здесь удивительно простое доказательство, принадлежащее Турану [Turán (1934)] и сыгравшее ключевую роль в развитии вероятностных методов в теории чисел.

Теорема 4.2.1. *Пусть $\omega(n) \to \infty$ произвольно медленно. Тогда число тех x из множества $\{1,\ldots,n\}$, для которых*

$$|\nu(x) - \ln\ln n| > \omega(n)\sqrt{\ln\ln n},$$

есть $o(n)$.

Доказательство. Пусть x случайно выбирается из множества $\{1,\dots,n\}$. Для простого p положим

$$X_p = \begin{cases} 1, & \text{если } p|x, \\ 0 & \text{иначе.} \end{cases}$$

Пусть $M = n^{1/10}$ и $X = \sum X_p$, где суммирование ведется по всем простым $p \leqslant M$. Поскольку никакое $x \leqslant n$ не может иметь более десяти простых делителей, больших M, мы имеем $\nu(x) - 10 \leqslant X(x) \leqslant \nu(x)$, и, тем самым, границы больших уклонений для X переходят в асимптотически равные им границы для ν. (Здесь число 10 может быть заменено любой другой достаточно большой константой.) Имеем

$$\mathbf{E}[X_p] = \frac{\lfloor n/p \rfloor}{n}.$$

Поскольку $y - 1 < \lfloor y \rfloor \leqslant y$, то

$$\mathbf{E}[X_p] = 1/p + O(1/n).$$

В силу линейности математического ожидания

$$\mathbf{E}[X] = \sum_{p \leqslant M} \left(\frac{1}{p} + O\left(\frac{1}{n}\right)\right) = \ln\ln n + O(1).$$

Здесь мы использовали хорошо известный факт, что $\sum_{p\leqslant x} \frac{1}{p} = \ln\ln x + O(1)$, который может быть доказан комбинированием формулы Стирлинга с суммированием по Абелю.

Теперь мы найдем асимптотическое выражение для

$$\operatorname{Var}[X] = \sum_{p\leqslant M} \operatorname{Var}[X_p] + \sum_{p\neq q} \operatorname{Cov}[X_p, X_q].$$

Поскольку $\operatorname{Var}[X_p] = \frac{1}{p}(1 - \frac{1}{p}) + O(\frac{1}{n})$, то

$$\sum_{p\leqslant M} \operatorname{Var}[X_p] = \left(\sum_{p\leqslant M} \frac{1}{p} \right) + O(1) = \ln\ln n + O(1).$$

Для различных простых p и q равенство $X_p X_q = 1$ выполняется тогда и только тогда, когда $p|x$ и $q|x$, что эквивалентно $pq|x$. Следовательно,

$$\operatorname{Cov}[X_p, X_q] = \mathbf{E}[X_p X_q] - \mathbf{E}[X_p]\mathbf{E}[X_q] = \frac{\lfloor n/pq \rfloor}{n} - \frac{\lfloor n/p \rfloor}{n}\frac{\lfloor n/q \rfloor}{n} \leqslant$$
$$\leqslant \frac{1}{pq} - (\frac{1}{p} - \frac{1}{n})(\frac{1}{q} - \frac{1}{n}) \leqslant \frac{1}{n}(\frac{1}{p} + \frac{1}{q}).$$

Таким образом,

$$\sum_{p\neq q} \operatorname{Cov}[X_p, X_q] \leqslant \frac{1}{n} \sum_{p\neq q} \left(\frac{1}{p} + \frac{1}{q} \right) \leqslant \frac{2M}{n} \sum \frac{1}{p}.$$

Отсюда

$$\sum_{p\neq q} \mathrm{Cov}[X_p, X_q] \leqslant O(n^{-9/10}\ln\ln n) = o(1),$$

и, аналогично,

$$\sum_{p\neq q} \mathrm{Cov}[X_p, X_q] \geqslant -o(1).$$

То есть ковариации не влияют на дисперсию, $\mathrm{Var}[X] = \ln\ln n + O(1)$ и неравенство Чебышёва приводит к соотношению

$$\Pr\left[|X - \ln\ln n| > \lambda\sqrt{\ln\ln n}\right] < \lambda^{-2} + o(1)$$

для любой константы $\lambda > 0$. Так как $|X - \nu| \leqslant 10$, то же самое справедливо для ν. ■

В классической работе [Erdős and Kac (1940)] показано, что распределение случайной величины ν приблизительно нормально с математическим ожиданием и дисперсией, равными $\ln\ln n$. Сформулируем точный результат.

Теорема 4.2.2. *Пусть λ — произвольное фиксированное число (положительное, отрицательное или ноль). Тогда*

$$\lim_{n\to\infty} \frac{1}{n}\left|\{x : 1 \leqslant x \leqslant n, \nu(x) \geqslant \ln\ln n + \lambda\sqrt{\ln\ln n}\}\right| = \int_\lambda^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt.$$

Доказательство. Мы представим набросок доказательства, подчеркивающий связь с доказательством Турана. Зафиксируем функцию $s(n)$ такую, что $s(n) \to \infty$ и $s(n) = o((\ln\ln n)^{1/2})$. Например, $s(n) = \ln\ln\ln n$. Пусть $M = n^{1/s(n)}$. Положим $X = \sum X_p$, где суммирование ведется по всем простым $p \leqslant M$. Так как никакое $x \leqslant n$ не может иметь больше $s(n)$ простых множителей, превосходящих M, мы имеем $\nu(x) - s(n) \leqslant X(x) \leqslant \nu(x)$. Поэтому достаточно доказать теорему 4.2.2 с заменой ν на X. Пусть Y_p — независимые случайные величины, для которых $\Pr[Y_p = 1] = p^{-1}$, $\Pr[Y_p = 0] = 1 - p^{-1}$, и пусть $Y = \sum Y_p$ — сумма по всем простым $p \leqslant M$. Такое Y представляет собой идеализированную версию X. Положим

$$\mu = \mathbf{E}[Y] = \sum_{p\leqslant M} p^{-1} = \ln\ln n + o((\ln\ln n)^{1/2}),$$

$$\sigma^2 = \mathrm{Var}[Y] = \sum_{p\leqslant M} p^{-1}(1 - p^{-1}) \sim \ln\ln n$$

и определим нормированную величину $\tilde{Y} = (Y - \mu)/\sigma$. По центральной предельной теореме, $\tilde{Y}$ сходится к случайной величине N со стандартным нормальным распределением, при этом $\mathbf{E}[\tilde{Y}^k] \to \mathbf{E}[N^k]$ для всякого положительного целого k. Положим $\tilde{X} = (X - \mu)/\sigma$. Сравним $\tilde{X}$ и $\tilde{Y}$. Для любых различных простых чисел $p_1, \ldots, p_s \leqslant M$ имеем

$$\mathbf{E}[X_{p_1} \times \ldots \times X_{p_s}] - \mathbf{E}[Y_{p_1} \times \ldots \times Y_{p_s}] = \frac{\left\lfloor \frac{n}{p_1\ldots p_s} \right\rfloor}{n} - \frac{1}{p_1 \cdots p_s} = O(n^{-1}).$$

Пусть k — произвольное фиксированное натуральное число; сравним $\mathbf{E}[\tilde{X}^k]$ и $\mathbf{E}[\tilde{Y}^k]$. Величина $\tilde{X}^k$ является полиномом от X с коэффициентами, не превосходящими $n^{o(1)}$. Дальнейшее разложение каждого $X^j = (\sum X_p)^j$ (всегда сводящее X_p^a к X_p, когда $a \geqslant 2$) дает сумму $O(M^k) = n^{o(1)}$ слагаемых вида $X_{p_1} \times \ldots \times X_{p_s}$. То же самое разложение применяется к $\tilde{Y}$. Поскольку математические ожидания соответствующих слагаемых имеют порядок $O(n^{-1})$, общая разность есть

$$\mathbf{E}[\tilde{X}^k] - \mathbf{E}[\tilde{Y}^k] = n^{-1+o(1)} = o(1).$$

Следовательно, каждый момент величины $\tilde{X}$ сходится к соответствующему моменту стандартно нормально распределенной случайной величины N. Согласно стандартной, хотя и нетривиальной теореме из теории вероятностей из этого следует, что распределение величины $\tilde{X}$ должно стремиться к нормальному. ■

Мы процитируем здесь известное изречение Харди:

«Число 317 является простым вовсе не потому, что мы так решили или наш мозг устроен так, а не иначе, но потому, что это **на самом деле так**, потому, что такова математическая реальность.»

Может показаться странным (хотя это ничему и не противоречит), что именно методы теории вероятностей позволяют прийти к лучшему пониманию свойств разложения чисел на простые множители..

4.3. ДОПОЛНИТЕЛЬНЫЕ ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ

Рассмотрим неотрицательную целочисленную случайную величину X. Предположим, нам требуется оценить вероятность $\Pr[X = 0]$ при заданном значении $\mu = \mathbf{E}[X]$. Если $\mu < 1$, то можно использовать неравенство

$$\Pr[X > 0] \leqslant \mathbf{E}[X]\,,$$

так что, если $\mathbf{E}[X] \to 0$, то $X = 0$ почти всегда. (При этом мы мысленно имеем дело с бесконечной последовательностью X-ов, зависящей от некоторого параметра n, стремящегося к бесконечности.) Предположим теперь, что $\mathbf{E}[X] \to \infty$. Вовсе *не* обязательно при этом окажется, что $X > 0$ почти всегда. Например, пусть X — число смертей, причиненных ядерной войной в течение двенадцати месяцев, начиная с данного момента. Вычисление математического ожидания $\mathbf{E}[X]$ может породить оживленную дискуссию, но мало кто может утверждать, что оно окажется очень большим. Более того, мы можем верить (или надеяться), что вероятность $\Pr[X \neq 0]$ очень близка к нулю. Можно даже иногда утверждать, что $X > 0$ почти всегда, если имеется некоторая информация о Var[X].

Теорема 4.3.1. *Имеет место неравенство*

$$\Pr[X = 0] \leqslant \frac{\operatorname{Var}[X]}{\mathbf{E}[X]^2}.$$

Доказательство. Пусть $\lambda = \mu/\sigma$ в неравенстве Чебышёва. Тогда

$$\Pr[X = 0] \leqslant \Pr\left[|X - \mu| \geqslant \lambda\sigma\right] \leqslant \frac{1}{\lambda^2} = \frac{\sigma^2}{\mu^2}.$$

■

По большей части мы будем применять этот результат в асимптотической форме.

Следствие 4.3.2. *Если* $\operatorname{Var}[X] = o(\mathbf{E}[X]^2)$, *то* $X > 0$ *почти всегда.*

Доказательство теоремы 4.3.1 в действительности позволяет утверждать, что для любого $\varepsilon > 0$

$$\Pr\left[|X - \mathbf{E}[X]| \geqslant \varepsilon\mathbf{E}[X]\right] \leqslant \frac{\operatorname{Var}[X]}{\varepsilon^2\mathbf{E}[X]^2}$$

и, таким образом, в асимптотической форме, мы и в самом деле имеем следующий, более сильный результат.

Следствие 4.3.3. *Если* $\operatorname{Var}[X] = o(\mathbf{E}[X]^2)$, *то* $X \sim \mathbf{E}[X]$ *почти всегда.*

Предположим, что $X = X_1 + \ldots + X_m$, где X_i является индикатором события A_i. Для индексов i, j мы пишем $i \sim j$, если $i \neq j$ и события A_i, A_j не являются независимыми. Положим (сумма ведется по упорядоченным парам)

$$\Delta = \sum_{i \sim j} \Pr[A_i \wedge A_j].$$

Заметим, что при $i \sim j$

$$\operatorname{Cov}[X_i, X_j] = \mathbf{E}[X_i X_j] - \mathbf{E}[X_i]\mathbf{E}[X_j] \leqslant \mathbf{E}[X_i X_j] = \Pr[A_i \wedge A_j],$$

а при $i \neq j$ и $i \sim j$ имеем $\operatorname{Cov}[X_i, X_j] = 0$. Итак,

$$\operatorname{Var}[X] \leqslant \mathbf{E}[X] + \Delta.$$

Следствие 4.3.4. *Если* $\mathbf{E}[X] \to \infty$ *и* $\Delta = o(\mathbf{E}[X]^2)$, *то* $X > 0$ *почти всегда. Более того,* $X \sim \mathbf{E}[X]$ *почти всегда.*

Будем говорить, что индикаторы $X_1, \ldots, X_m$ *симметричны*, если для любых $i \neq j$ существует автоморфизм рассматриваемого вероятностного пространства, отображающий событие A_i в событие A_j. Примеры будут приведены в следующем разделе[2)]. В нашем случае мы пишем

$$\Delta = \sum_{i \sim j} \Pr[A_i \wedge A_j] = \sum_i \Pr[A_i] \sum_{j \sim i} \Pr[A_j | A_i],$$

[2)] В приводимых примерах рассматриваются симметричные случаи. Несимметричная система случайных величин рассматривалась, например, в работе [Сапоженко (1968)]. — *Прим. ред.*

причем внутреннее суммирование не зависит от i. Мы полагаем

$$\Delta^* = \sum_{j \sim i} \Pr[A_j | A_i],$$

где i — фиксировано. Тогда

$$\Delta = \sum_i \Pr[A_i] \Delta^* = \Delta^* \sum_i \Pr[A_i] = \Delta^* \mathbf{E}[X].$$

Следствие 4.3.5. *Если* $\mathbf{E}[X] \to \infty$ *и* $\Delta^* = o(\mathbf{E}[X])$, *то* $X > 0$ *почти всегда. Кроме того,* $X \sim \mathbf{E}[X]$ *почти всегда.*

Условие следствия 4.3.5 имеет следующий интуитивный смысл. Учет того условия, что некоторое конкретное событие A_i происходит, несущественно увеличивает среднее число $\mathbf{E}[X]$ происходящих событий.

4.4. СЛУЧАЙНЫЕ ГРАФЫ

Определение случайного графа $G(n, p)$ и «пороговой функции» приведено в разд. 10.1, гл. 10. Результаты этого раздела существенно перекрываются результатами гл. 10. Однако они были исторически первыми и являются хорошей иллюстрацией метода второго момента. Мы начнем с частного примера. Начиная с этого момента через $\omega(G)$ будем обозначать максимальное число вершин клики в графе G.

Теорема 4.4.1. *Пороговая функция для свойства* $\omega(G) \geqslant 4$ *равна* $n^{-2/3}$.

Доказательство. Для всякого четырехэлементного множества S вершин графа $G(n, p)$ обозначим через A_S событие «S является кликой». Пусть X_S — индикатор события A_S. Тогда

$$\mathbf{E}[X_S] = \Pr[A_S] = p^6,$$

поскольку графу $G(n, p)$ должны принадлежать шесть различных ребер. Пусть

$$X = \sum_{|S|=4} X_S,$$

так что X — число 4-клик в графе G, а $\omega(G) \geqslant 4$ тогда и только тогда, когда $X > 0$. В силу линейности математического ожидания имеем

$$\mathbf{E}[X] = \sum_{|S|=4} \mathbf{E}[X_S] = \binom{n}{4} p^6 \sim \frac{n^4 p^6}{24}.$$

При $p(n) \ll n^{-2/3}$ выполнено $\mathbf{E}[X] = o(1)$, и, следовательно, $X = 0$ почти наверное.

Теперь предположим, что $p(n) \gg n^{-2/3}$, а значит, $\mathbf{E}[X] \to \infty$. Рассмотрим Δ^* из следствия 4.3.5. (Все четырехэлементные множества «неотличимы друг

от друга», поэтому множество всех X_S симметрично.) Здесь $S \sim T$ тогда и только тогда, когда $S \neq T$ и у S и T есть общие ребра, т. е. тогда и только тогда, когда $|S \cap T| = 2$ или 3. Зафиксируем S. Существует $O(n^2)$ множеств T, таких, что $|S \cap T| = 2$, причем для каждого из них $\Pr[A_T|A_S] = p^5$. Имеется $O(n)$ множеств T, таких, что $|S \cap T| = 3$, причем для каждого из них $\Pr[A_T|A_S] = p^3$. Таким образом,

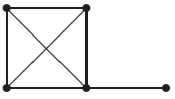
$$\Delta^* = O(n^2 p^5) + O(np^3) = o(n^4 p^6) = o(\mathbf{E}[X])$$

так как $p \gg n^{-2/3}$. Значит, следствие 4.3.5 применимо и $X > 0$, т.е. клика размера 4 *существует* почти всегда. ■

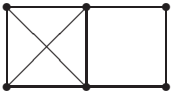
В доказательстве теоремы 4.4.1 мы уже столкнулись с вычислением Δ^*. Следующие понятия необходимы для доказательства более общей теоремы 4.4.2.

Определение 1. *Рассмотрим граф H с v вершинами и e ребрами. Назовем отношение $\rho(H) = e/v$* ***плотностью*** *графа H. Граф H называется* ***сбалансированным****, если для каждого его подграфа H' выполнено неравенство $\rho(H') \leqslant \rho(H)$. Назовем граф H* ***строго сбалансированным****, если для каждого собственного подграфа H' выполнено строгое неравенство $\rho(H') < \rho(H)$.*

Примеры. Граф K_4 и, вообще, графы K_k строго сбалансированы. Граф



не является сбалансированным, поскольку он имеет плотность $7/5$, в то время как его подграф K_4 имеет плотность $3/2$. Граф



сбалансирован, но не строго, ибо он и его подграф K_4 имеют плотность $3/2$.

Теорема 4.4.2. *Рассмотрим сбалансированный граф H с v вершинами и e ребрами. Пусть $A(G)$ — событие «H является подграфом (не обязательно порожденным) графа G». Тогда $p = n^{-v/e}$ является пороговой функцией для A.*

Доказательство. Мы следуем доказательству теоремы 4.4.1. Для каждого v-множества S пусть A_S — событие «$G|_S$ содержит H в качестве подграфа».

Тогда

$$p^e \leqslant \Pr[A_S] \leqslant v!p^e.$$

(Любое конкретное размещение подграфа H (на некотором фиксированном v-подмножестве вершин графа G) имеет вероятность p^e и при этом имеется не более $v!$ возможных размещений. Точный подсчет $\Pr[A_S]$ в общем случае затруднителен в силу взаимного наложения копий графа H.) Обозначим через X_S индикатор события A_S. Положим

$$X = \sum_{|S|=v} X_S.$$

Ясно, что событие $A(G)$ происходит тогда и только тогда, когда $X > 0$. Из линейности математического ожидания вытекает, что

$$\mathbf{E}[X] = \sum_{|S|=v} \mathbf{E}[X_S] = \binom{n}{v} \Pr[A_S] = \Theta(n^v p^e).$$

Если $p \ll n^{-v/e}$, то $\mathbf{E}[X] = o(1)$ и $X = 0$ почти наверное.

Пусть теперь $p \gg n^{-v/e}$ и, следовательно, $\mathbf{E}[X] \to \infty$. Рассмотрим Δ^* из следствия 4.3.5. (Все v-множества равноправны, поэтому X_S симметричны.) Здесь $S \sim T$ в том и только том случае, когда $S \neq T$ и у S и T есть общие ребра, т. е. тогда и только тогда, когда $|S \cap T| = i$ с $2 \leqslant i \leqslant v-1$. Зафиксируем S. Имеем

$$\Delta^* = \sum_{T \sim S} \Pr[A_T | A_S] = \sum_{i=2}^{v-1} \sum_{|T \cap S|=i} \Pr[A_T | A_S].$$

Для каждого i существует $O(n^{v-i})$ способов выбора T. Зафиксируем S, T и рассмотрим $\Pr[A_T | A_S]$. Имеется $O(1)$ возможных копий графа H на множестве T. Каждая из них имеет (поскольку граф H сбалансированный) не более $\frac{ie}{v}$ ребер с обоими концами в S, а, значит, не менее $e - \frac{ie}{v}$ ребер другого типа. Отсюда

$$\Pr[A_T | A_S] = O(p^{e - \frac{ie}{v}})$$

и

$$\Delta^* = \sum_{i=2}^{v-1} O(n^{v-i} p^{e-\frac{ie}{v}}) = \sum_{i=2}^{v-1} O((n^v p^e)^{1-\frac{i}{v}}) =$$
$$= \sum_{i=2}^{v-1} o(n^v p^e) = o(\mathbf{E}[X]),$$

так как $n^v p^e \to \infty$. Это означает, что применимо следствие 4.3.5. ■

Теорема 4.4.3. *В обозначениях теоремы* 4.4.2*, если H* ***не*** *является сбалансированным, то $p = n^{-v/e}$* ***не*** *является пороговой функцией для события A.*

Доказательство. Пусть H_1 — подграф графа H с v_1 вершинами и e_1 ребрами, $e_1/v_1 > e/v$. Пусть α удовлетворяет неравенствам $v/e < \alpha < v_1/e_1$.

Положим $p = n^{-\alpha}$. Математическое ожидание числа копий H_1 есть тогда $o(1)$. Поэтому почти всегда $G(n,p)$ не содержит копий графа H_1. Но если он не содержит копий графа H_1, то он, конечно, не содержит и копий графа H. ■

Пороговая функция для свойства содержания копий H для произвольного H изучалась в классической статье Эрдёша и Реньи [Erdős and Rényi (1960)]. (Она все еще служит великолепным введением в теорию случайных графов.) Пусть H_1 — подграф максимальной плотности, т. е. $\rho(H_1) = e_1/v_1$ (если H сбалансирован, можно положить $H_1 = H$). Эрдёш и Реньи показали, что функция $p = n^{-v_1/e_1}$ является пороговой. Мы не доказываем это здесь, хотя это непосредственно вытекает из обсуждаемых методов.

Мы заканчиваем этот раздел двумя усилениями теоремы 4.4.2.

Теорема 4.4.4. *Пусть H — строго сбалансированный граф с v вершинами, e ребрами и a автоморфизмами. Обозначим через X число копий графа H в $G(n,p)$. Предположим, что $p \gg n^{-v/e}$. Тогда для почти всех графов*

$$X \sim \frac{n^v p^e}{a}.$$

Доказательство. Пронумеруем вершины графа H числами $1, \dots, v$. Для каждого упорядоченного множества $x_1, \dots, x_v$ вершин графа G обозначим через $A_{x_1,\dots,x_v}$ событие «$\{x_1, \dots, x_v\}$ содержит копию графа H именно в этом порядке». А именно, мы полагаем

$$A_{x_1,\dots,x_v} : \{i,j\} \in E(H) \Rightarrow \{x_i, x_j\} \in E(G).$$

Пусть $I_{x_1,\dots,x_v}$ — соответствующий индикатор. Определим класс эквивалентности на v-кортежах следующим образом. Положим $(x_1, \dots, x_v) \equiv (y_1, \dots, y_v)$, если существует автоморфизм σ на множестве $V(H)$, такой, что $y_{\sigma(i)} = x_i$ для $1 \leqslant i \leqslant v$. Тогда величина

$$X = \sum I_{x_1,\dots,x_v},$$

где в сумме участвует по одному представителю из каждого класса эквивалентности, есть число копий графа H в G. Поскольку существует $(n)_v/a$ слагаемых, то

$$\mathbf{E}[X] = \frac{(n)_v}{a}\mathbf{E}[I_{x_1,\dots,x_v}] = \frac{(n)_v p^e}{a} \sim \frac{n^v p^e}{a}.$$

Так как по предположению $p \gg n^{-v/e}$, то $\mathbf{E}[X] \to \infty$. Теперь достаточно показать, что $\Delta^* = o(\mathbf{E}[X])$. Фиксируя $x_1, \dots, x_v$, имеем

$$\Delta^* = \sum_{(y_1,\dots,y_v)\sim(x_1,\dots,x_v)} \Pr[A_{(y_1,\dots,y_v)} | A_{(x_1,\dots,x_v)}].$$

Существует $v!/a = O(1)$ слагаемых с $\{y_1, \dots, y_v\} = \{x_1, \dots, x_v\}$, причем для каждого из них условная вероятность не превышает 1 (в действительности, не больше p), внося, тем самым вклад $O(1) = o(\mathbf{E}[X])$ в Δ^*. Когда $\{y_1, \dots, y_v\} \cap \{x_1, \dots, x_v\}$ имеет i элементов, $2 \leqslant i \leqslant v-1$, рассуждения теоремы 4.4.2 показывают, что вклад в Δ^* равен $o(\mathbf{E}[X])$. В итоге $\Delta^* = o(\mathbf{E}[X])$ и мы можем применить следствие 4.3.5. ■

Теорема 4.4.5. *Пусть H — **произвольно** фиксированный граф. Для каждого подграфа H' графа H (включая сам H) обозначим через $X_{H'}$ число копий графа H' в $G(n,p)$. Пусть p таково, что $E[X_{H'}] \to \infty$ для каждого H'. Тогда*

$$X_H \sim \mathbf{E}[X_H]$$

для почти всех графов.

Доказательство. Пусть H имеет v вершин и e ребер. Как в теореме 4.4.4, достаточно показать, что $\Delta^* = o(\mathbf{E}[X])$. Представим Δ^* суммой конечного числа слагаемых. Для каждого H' с w вершинами и f ребрами мы имеем такие $(y_1,\dots,y_v)$, которые пересекаются с фиксированными $(x_1,\dots,x_v)$ по некоторой копии графа H'. Эти слагаемые вносят вклад в Δ^*, с точностью до константы равный

$$n^{v-w}p^{e-f} = \Theta\left(\frac{\mathbf{E}[X_H]}{\mathbf{E}[X_{H'}]}\right) = o(\mathbf{E}[X_H]).$$

Поэтому можно применить следствие 4.3.5. ■

4.5. МАКСИМАЛЬНЫЙ РАЗМЕР КЛИКИ

Зафиксируем вероятность $p = \frac{1}{2}$ появления ребра и пусть $\omega(G)$ — максимальный размер клики. Обозначим через

$$f(k) = \binom{n}{k} 2^{-\binom{k}{2}}$$

математическое ожидание числа k-клик. Функция $f(k)$ становится меньше единицы при $k \sim 2\log_2 n$. (Грубо говоря, $f(k)$ ведет себя как $n^k 2^{-k^2/2}$.)

Теорема 4.5.1. *Пусть $k = k(n)$ таково, что $k \sim 2\log_2 n$, и $f(k) \to \infty$. Тогда для почти всех графов $\omega(G) \geqslant k$.*

Доказательство. Для каждого k-множества S пусть A_S — событие вида «S является кликой», а X_S — его индикатор. Положим

$$X = \sum_{|S|=k} X_S.$$

Заметим, что $\omega(G) \geqslant k$ тогда и только тогда, когда $X > 0$. Тогда $\mathbf{E}[X] = f(k) \to \infty$, и мы изучим Δ^* из следствия 4.3.5. Зафиксируем S и заметим, что $T \sim S$ в том и только том случае, когда $|T \cap S| = i$, где $2 \leqslant i \leqslant k-1$. Следовательно,

$$\Delta^* = \sum_{i=2}^{k-1} \binom{k}{i}\binom{n-k}{k-i} 2^{\binom{i}{2}-\binom{k}{2}},$$

и поэтому

$$\frac{\Delta^*}{\mathbf{E}[X]} = \sum_{i=2}^{k-1} g(i),$$

где мы полагаем

$$g(i) = \frac{\binom{k}{i}\binom{n-k}{k-i}}{\binom{n}{k}} 2^{\binom{i}{2}}.$$

Заметим, что $g(i)$ можно представлять себе как вероятность того, что случайно выбранное T пересекает фиксированное S в i точках, умноженную на возрастающую по $\Pr[A_T]$ функцию. Полагая $i = 2$, имеем

$$g(2) = 2\frac{\binom{k}{2}\binom{n-k}{k-2}}{\binom{n}{k}} \sim \frac{k^4}{n^2} \leqslant o(n^{-1}).$$

На другом конце, при $i = k - 1$, имеем

$$g(k-1) = \frac{k(n-k)2^{-(k-1)}}{\binom{n}{k}2^{-\binom{k}{2}}} \sim \frac{2kn2^{-k}}{\mathbf{E}[X]}.$$

Поскольку $k \sim 2\log_2 n$, числитель равен $n^{-1+o(1)}$. Знаменатель стремится к бесконечности, поэтому $g(k-1) \leqslant o(n^{-1})$. Более детальные вычисления (которые мы опускаем) показывают, что остальными $g(i)$ и их суммой можно пренебречь, так что следствие 4.3.5 применимо. ■

Теорема 4.5.1 приводит к сильной концентрации для $\omega(G)$. При $k \sim 2\log_2 n$

$$\frac{f(k+1)}{f(k)} = \frac{n-k}{k+1}2^{-k} = n^{-1+o(1)} = o(1).$$

Пусть $k_0 = k_0(n)$ — то значение, при котором $f(k_0) \geqslant 1 > f(k_0+1)$. Для «большинства» n функция $f(k)$ резко изменяет большое значение $f(k_0)$ на малое $f(k_0+1)$. Вероятность того, что G содержит клику размера k_0+1, не превосходит весьма малой величины $f(k_0+1)$. Когда $f(k_0)$ велико, из теоремы 4.5.1 следует, что G содержит клику размера k_0 с вероятностью, близкой к 1. Таким образом, с очень большой вероятностью $\omega(G) = k_0$. Для некоторых n одна из величин $f(k_0), f(k_0+1)$ может оказаться средних размеров. В этом случае эти соображения не применимы. Тем не менее можно доказать строгую концентрацию, полученную независимо в работах [Bollobás and Erdős (1976)] и [Matula (1976)].

Следствие 4.5.2. *Существует $k = k(n)$, такое, что*

$$\Pr[\omega(G) = k \text{ или } k+1] \to 1.$$

Мы еще получим более сильные результаты относительно распределения величины $\omega(G)$ в разделе 10.2.

4.6. РАЗЛИЧНЫЕ СУММЫ

Говорят, что множество $\{x_1, \ldots, x_k\}$ натуральных чисел имеет различные суммы, если все суммы

$$\sum_{i \in S} x_i, \quad \text{где} \quad S \subseteq \{1, \ldots, k\},$$

попарно различны. Через $f(n)$ обозначим максимальное k, для которого существует множество

$$\{x_1, \ldots, x_k\} \subset \{1, \ldots, n\}$$

с различными суммами. Простейший пример множества с различными суммами: $\{2^i : i \leqslant \log_2 n\}$. Этот пример показывает, что

$$f(n) \geqslant 1 + \lfloor \log_2 n \rfloor.$$

Пол Эрдёш обещал \$300 за доказательство или опровержение того, что

$$f(n) \leqslant \log_2 n + C$$

для некоторой константы C. Из сказанного следует, что все $2^{f(n)}$ сумм различны и меньше nk, а значит,

$$2^{f(n)} < nk = nf(n).$$

Отсюда

$$f(n) < \log_2 n + \log_2 \log_2 n + O(1).$$

Метод второго момента дает небольшое улучшение. Зафиксируем $\{x_1, \ldots, x_k\} \subset \{1, \ldots, n\}$ с различными суммами. Пусть $\varepsilon_1, \ldots, \varepsilon_k$ выбираются независимо и

$$\Pr[\varepsilon_i = 1] = \Pr[\varepsilon_i = 0] = \frac{1}{2}.$$

Положим

$$X = \varepsilon_1 x_1 + \ldots + \varepsilon_k x_k.$$

(Мы представляем X в виде случайной суммы.) Положим

$$\mu = E[X] = \frac{x_1 + \ldots + x_k}{2}$$

и $\sigma^2 = \operatorname{Var}[X]$. Заметим, что

$$\sigma^2 = \frac{x_1^2 + \ldots + x_k^2}{4} \leqslant \frac{n^2 k}{4},$$

так что $\sigma \leqslant n\sqrt{k}/2$. По неравенству Чебышёва для любого $\lambda > 1$ имеем

$$\Pr\left[|X - \mu| \geqslant \lambda n\sqrt{k}/2\right] \leqslant \lambda^{-2}.$$

Тогда для вероятности обратного события получаем неравенство

$$1 - \frac{1}{\lambda^2} \leqslant \Pr\left[|X - \mu| < \lambda n\sqrt{k}/2\right].$$

Но X принимает каждое конкретное значение с вероятностью либо ноль, либо 2^{-k}, поскольку каждое значение представляется в виде суммы не более чем одним способом. Отсюда

$$\Pr\left[|X - \mu| < \lambda n\sqrt{k}/2\right] \leqslant 2^{-k}(\lambda n\sqrt{k} + 1)$$

и

$$n \geqslant \frac{2^k(1 - \lambda^{-2}) - 1}{\sqrt{k}\lambda}.$$

Хотя $\lambda = \sqrt{3}$ максимизирует правую часть последнего неравенства, при любом выборе $\lambda > 1$ из этого неравенства вытекает следующий результат.

Теорема 4.6.1. *Справедливо соотношение*

$$f(n) \leqslant \log_2 n + \frac{1}{2} \log_2 \log_2 n + O(1).$$

4.7. ПОДХОД РЁДЛЯ

Для чисел $2 \leqslant l < k < n$ пусть $M(n,k,l)$, — минимальный размер семейства $\mathcal{K}$ из k-элементных подмножеств множества $\{1,\dots,n\}$, обладающего тем свойством, что каждое l-множество содержится по крайней мере в одном $A \in \mathcal{K}$. Ясно, что $M(n,k,l) \geqslant \binom{n}{l}/\binom{k}{l}$, так как каждое k-множество покрывает $\binom{k}{l}$ l-множеств и каждое l-множество должно быть покрыто. Равенство выполняется тогда и только тогда, когда семейство $\mathcal{K}$ обладает тем свойством, что каждое l-множество содержится ровно в одном $A \in \mathcal{K}$. Такое семейство называется (n,k,l)-тактической конфигурацией (или блок-схемой). Например, $(n,3,2)$-тактическая конфигурация больше известна как система троек Штейнера. Вопрос о существовании тактических конфигураций является центральным в комбинаторике, но одним из тех, в которых вероятностные методы играют (по крайней мере, в настоящий момент) незначительную роль. В 1963 г. Эрдёш и Ханани предположили, что для фиксированных $2 \leqslant l < k$

$$\lim_{n\to\infty} \frac{M(n,k,l)}{\binom{n}{l}/\binom{k}{l}} = 1.$$

Их предположение состояло, грубо говоря в том, что можно асимптотически приблизиться к тактической конфигурации. Хотя эта гипотеза представлялась идеальной для вероятностного анализа, потребовалось целое поколение, прежде чем Рёдль [Rödl (1985)] нашел доказательство, которое приводится в данном разделе. (Можно подобным образом определить число $m(n,k,l)$ как максимальный размер семейства $\mathcal{K}$ из k-элементных подмножеств множества $\{1,\dots,n\}$, обладающего тем свойством, что каждое l-множество содержится не более чем в одном $A \in \mathcal{K}$. С помощью элементарных соображений Эрдёш и Ханани получили, что

$$\lim_{n\to\infty} \frac{M(n,k,l)}{\binom{n}{l}/\binom{k}{l}} = 1 \iff \lim_{n\to\infty} \frac{m(n,k,l)}{\binom{n}{l}/\binom{k}{l}} = 1.$$

Хотя результат Рёдля можно формулировать как в терминах упаковки, так и в терминах покрытий, здесь мы имеем дело только с задачей о покрытии.)

Многие исследователи осознавали, что метод Рёдля применим в более общей ситуации, когда рассматриваются покрытия в однородных гиперграфах. Это было замечено Франклем и Рёдлем, и впоследствии было упрощено Пиппенджером и Спенсером [Pippenger and Spencer (1989)], а также Каном [Kahn (1996)]. В нашем изложении мы следуем Пиппенджеру и Спенсеру [Pippenger and Spencer (1989)] и основываемся на обзоре Фюреди [Fűredi (1988)], в котором основным инструментом является метод второго момента.

Для r-однородного гиперграфа $H = (V,E)$ и вершины $x \in V$ мы используем обозначение $d_H(x)$ (или просто $d(x)$, когда это не приводит к недоразумению) для степени вершины x в гиперграфе H, т. е. числа ребер, содержащих x. Аналогично, для $x,y \in V$ пусть $d(x,y) = d_H(x,y)$ — число ребер гиперграфа H, содержащих как x, так и y. Покрытие гиперграфа H — это множество ребер, объединение которых содержит все вершины. В дальнейшем, везде под $\pm\delta$

мы подразумеваем произвольное число между $-\delta$ и δ. Следующая теорема принадлежит Пиппенджеру. Изложим ее, следуя Франклю и Рёдлю.

Теорема 4.7.1. *Для каждого целого $r \geqslant 2$, действительного $k \geqslant 1$ и $a > 0$ существуют $\gamma = \gamma(r,k,a) > 0$ и $d_0 = d_0(r,k,a)$, такие, что для всех $n \geqslant D \geqslant d_0$ выполнено следующее.*

Пусть $H = (V,E)$ является r-однородным гиперграфом на множестве V из n вершин, в котором все вершины имеют положительные степени, и удовлетворяет следующим условиям:

(1) *для всех $x \in V$, за исключением не более чем γn, $d(x) = (1 \pm \gamma)D$,*

(2) *для всех $x \in V$ выполнено неравенство $d(x) < kD$,*

(3) *для любых двух различных $x, y \in V$ выполнено $d(x,y) < \gamma D$.*

Тогда H содержит покрытие из не более $(1+a)\frac{n}{r}$ ребер.

Основная идея доказательства проста. Зафиксируем малое $\varepsilon > 0$. Можно показать, что случайное множество приблизительно из $\varepsilon n/r$ ребер с высокой вероятностью содержит только $O(\varepsilon^2 n)$ вершин, покрываемых более одного раза, и, следовательно, покрывает не менее чем $\varepsilon n - O(\varepsilon^2 n)$ вершин. Кроме того, после удаления покрытых вершин гиперграф, порожденный оставшимися вершинами, все еще удовлетворяет свойствам (1), (2) и (3) из условия теоремы (для некоторых других значений n, γ, k и D). Следовательно, можно снова выбрать случайное множество ребер в этом гиперграфе, покрывающее приблизительно ε-долю его вершин почти без пересечений. Действуя таким образом достаточно большое число раз, мы получим εn оставшихся непокрытыми вершин. Они могут быть покрыты тривиально, путем выбора для каждой из них произвольного содержащего ее ребра. Поскольку ε достаточно мало, последний шаг является вполне приемлемым, несмотря на явную его неэффективность.

Технические детали требуют аккуратного применения метода второго момента, используемого несколько раз в доказательстве следующей леммы.

Лемма 4.7.2. *Для любого целого $r \geqslant 2$, действительных $K \geqslant 1$ и $\varepsilon > 0$, а также для любого действительного $\delta' > 0$ существуют $\delta = \delta(r,K,\varepsilon,\delta') > 0$ и $D_0 = D_0(r,K,\varepsilon,\delta')$, такие, что для всякого $n \geqslant D \geqslant D_0$ выполняется следующее.*

Пусть $H = (V,E)$ является r-однородным гиперграфом на множестве V из n вершин и удовлетворяет следующим условиям:

(*i*) *для всех $x \in V$, за исключением не более чем δn, $d(x) = (1 \pm \delta n)D$,*

(*ii*) *для всех $x \in V$ выполнено $d(x) < KD$,*

(*iii*) *для любых двух различных $x, y \in V$ выполнено $d(x,y) < \delta D$.*

Тогда H содержит подмножество E' ребер со следующими свойствами:

(*iv*) $|E'| = \frac{\varepsilon n}{r}(1 \pm \delta')$,

(*v*) *множество $V' = V - \bigcup_{e \in E'} e$ имеет мощность $|V'| = ne^{-\varepsilon}(1 \pm \delta')$,*

(*vi*) *для всех вершин $x \in V'$, за исключением не более чем $\delta'|V'|$, степень $d'(x)$ вершины x в индуцированном гиперграфе H на множестве V' удовлетворяет условию $d'(x) = De^{-\varepsilon(r-1)}(1 \pm \delta')$.*

Доказательство. На протяжении доказательства мы предполагаем, когда это необходимо, что D (а, значит, и n) достаточно велики. Мы обозначаем через $\delta_1, \delta_2, \dots$ положительные константы (которые могут быть точно оценены), которые стремятся к нулю, когда δ стремится к нулю, и D стремится к бесконечности (при фиксированных r, K, ε). Следовательно, выбирая δ и D_0 подходящим образом, мы можем быть уверены, что каждая из них будет меньше δ'.

Пусть E' — случайное подмножество множества E, полученное путем случайного и независимого включения в E' каждого ребра из E с вероятностью $p = \frac{\varepsilon}{D}$. Мы должны показать, что с положительной вероятностью свойства (iv), (v) и (vi) выполняются.

Доказать, что выполняется равенство (iv), легко. Заметим, что по предположению гиперграф H имеет не менее $(1-\delta)n$ вершин степени, не меньшей $(1-\delta)D$. Поэтому число его ребер не меньше $\frac{(1-\delta)^2 nD}{r}$. Аналогично, число ребер в H не превосходит $\frac{(1+\delta)Dn+\delta nKD}{r}$. Следовательно, $|E| = (1 \pm \delta_1)\frac{Dn}{r}$. Отсюда вытекает, что математическое ожидание величины $|E'|$ удовлетворяет условию $\mathbf{E}[|E'|] = |E|p = (1\pm\delta_1)\frac{\varepsilon n}{r}$, а дисперсия равна $\mathrm{Var}[|E'|] = |E|p(1-p) \leqslant (1\pm\delta_1)\frac{\varepsilon n}{r}$. По неравенству Чебышёва, для подходящего $\delta_2 > 0$

$$\Pr\left[|E'| = (1 \pm \delta_2)\frac{\varepsilon n}{r}\right] > 0.99.$$

Отсюда следует равенство (iv).

Чтобы доказать утверждение (v), определим для каждой вершины $x \in V$ индикатор I_x так, что $I_x = 1$, если $x \notin \cup_{e\in E'} e$, и $I_x = 0$ иначе. Заметим, что $|V'| = \sum_{x\in V} I_x$. Назовем вершину $x \in V$ *хорошей*, если $d(x) = (1 \pm \delta)D$; в противном случае назовем ее *плохой*. Если x — хорошая вершина, то

$$\mathbf{E}[I_x] = \Pr[I_x = 1] = (1-p)^{d(x)} = \left(1 - \frac{\varepsilon}{D}\right)^{(1\pm\delta)D} = e^{-\varepsilon}(1 \pm \delta_3).$$

Если x — плохая, то, ясно, что $0 \leqslant \mathbf{E}[I_x] \leqslant 1$. Поскольку имеется не более δn плохих вершин, отсюда с использованием линейности математического ожидания получаем, что математическое ожидание величины $|V'|$ равно $ne^{-\varepsilon}(1\pm\delta_4)$. Чтобы вычислить дисперсию величины $|V'| = \sum_{x\in V} I_x$, заметим, что

$$\begin{aligned}\mathrm{Var}[|V'|] &= \sum_{x\in V} \mathrm{Var}[I_x] + \sum_{x,y\in V, x\neq y} \mathrm{Cov}[I_x, I_y] \leqslant \\ &\leqslant \mathbf{E}[|V'|] + \sum_{x,y\in V, x\neq y} \mathrm{Cov}[I_x, I_y]. \qquad (4.1)\end{aligned}$$

Однако

$$\begin{aligned}\mathrm{Cov}[I_x, I_y] &= \mathbf{E}[I_x I_y] - \mathbf{E}[I_x]\mathbf{E}[I_y] = \\ &= (1-p)^{d(x)+d(y)-d(x,y)} - (1-p)^{d(x)+d(y)} \leqslant \\ &\leqslant (1-p)^{-d(x,y)} - 1 \leqslant (1 - \frac{\varepsilon}{D})^{-\delta D} \leqslant \delta_5.\end{aligned}$$

Отсюда следует, что

$$\mathrm{Var}[|V'|] \leqslant \mathbf{E}[|V'|] + \delta_5 n^2 \leqslant \delta_6(\mathbf{E}[|V'|])^2.$$

Теперь с использованием неравенства Чебышёва получаем, что с вероятностью, не меньшей чем 0.99

$$|V'| = (1 \pm \delta_7)\mathbf{E}[|V'|] = (1 \pm \delta_8)ne^{-\varepsilon},$$

как и требовалось в утверждении (v).

Остается доказать (vi). Заметим прежде всего, что все вершины x, за исключением не более чем $\delta_9 n$, удовлетворяют следующим двум условиям:
(A) $d(x) = (1 \pm \delta)D$ и
(B) все ребра $e \in E$ с $x \in e$, за исключением, быть может, $\delta_{10}D$ ребер, удовлетворяют условию

$$|\{f \in E \,:\, x \notin f, f \cap e \neq \varnothing\}| = (1 \pm \delta_{11})(r-1)D. \tag{4.2}$$

В самом деле, (A) выполняется по предположению для всех вершин, за исключением не более чем $\delta n < \delta_9 n/2$. Более того, общее число ребер, содержащих вершины, степени которых лежат вне интервала $(1\pm\delta)D$, не превышает δnKD а, значит, число вершин, содержащихся в более чем $\delta_{10}D$ таких ребрах, не больше $\delta nKDr/(\delta_{10}D) \leqslant \delta_9 n/2$ для подходящего выбора δ_9, δ_{10}. Заметим далее, что если $x \in e$ и e не содержит вершин, степени которых лежат вне интервала $(1 \pm \delta)D$, то, поскольку $d(y,z) < \delta D$ для всех y, z, число ребер f, не содержащих x и пересекающих e, не больше $(r-1)(1 \pm \delta)D$ и не меньше $(r-1)(1 \pm \delta)D - \binom{r-1}{2}\delta D$. Следовательно, e удовлетворяет (4.2).

Таким образом, достаточно показать, что для большинства вершин x, удовлетворяющих (A) и (B), $d'(x)$ удовлетворяет (vi). Зафиксируем такую вершину x. Назовем ребро e, такое, что $x \in e$, *хорошим*, если оно удовлетворяет (4.2). Обозначим через E_x подмножество хороших ребер e, для которых $x \in e$. При условии, что $x \in V'$, вероятность того, что ребро $e \in E_x$ остается в гиперграфе, порожденном множеством V', равна $(1-p)^{(1\pm\delta_{11})(r-1)D}$. Следовательно, математическое ожидание величины $d'(x)$ равно

$$\mathbf{E}[d'(x)] = (1 \pm \delta_{10} \pm \delta)D(1-p)^{(1\pm\delta_{11})(r-1)D} \pm \delta_{10}D = e^{-\varepsilon(r-1)}D(1 \pm \delta_{12}).$$

Для каждого ребра e, содержащего x, обозначим через I_e индикатор, принимающий значение 1, если и только если e содержится в V'. Тогда степень $d'(x)$ есть просто сумма этих индикаторов при условии, что $x \in V'$. Отсюда следует, что

$$\mathrm{Var}[d'(x)] \leqslant \mathbf{E}[d'(x)] + \sum_{x\in e, x\in f} \mathrm{Cov}[I_e, I_f] \leqslant$$
$$\leqslant \mathbf{E}[d'(x)] + 2\delta_{10}D^2(1 \pm \delta) + \sum_{e,f\in E_x} \mathrm{Cov}[I_e, I_f]. \tag{4.3}$$

Остается оценить сумму $\sum_{e,f\in E_x} \mathrm{Cov}[I_e, I_f]$. Для каждого фиксированного хорошего ребра e такая сумма имеет вид $\sum_{f\in E_x} \mathrm{Cov}[I_e, I_f]$. Имеется не более $(r-1)\delta D$ ребер f в последней сумме, для которых $|e \cap f| > 1$, и их вклад

в сумму не превышает $(r-1)\delta D$. Если $e \cap f = \{x\}$, то пусть $t(e,f)$ — число ребер гиперграфа H, которые пересекают как e, так и f и не содержат x. В этом случае, очевидно, $t(e,f) \leqslant (r-1)^2\delta D$. Поэтому для таких e и f выполнено $\mathrm{Cov}[I_e, I_f] \leqslant (1-p)^{-t(e,f)} - 1 \leqslant \delta_{13}$. Отсюда вытекает, что для каждого хорошего ребра e

$$\sum_{f \in E_x} \mathrm{Cov}[I_e, I_f] \leqslant (r-1)\delta D + D(1+\delta)\delta_{13} \leqslant \delta_{14} D.$$

Поскольку $\sum_{e,f \in E_x} \mathrm{Cov}[I_e, I_f]$ состоит из не более чем $D(1+\delta)$ таких величин (слагаемых), мы заключаем, что

$$\mathrm{Var}[d'(x)] \leqslant \mathbf{E}[d'(x)] + \delta_{15} D^2 \leqslant \delta_{16} (\mathbf{E}[d'(x)])^2.$$

Отсюда в силу неравенства Чебышёва следует, что с вероятностью, не превышающей δ_{17}, степень $d'(x)$ лежит вне интервала $(1 \pm \delta_{18}) D e^{-\varepsilon(r-1)}$, и, следовательно, по неравенству Маркова, с вероятностью, не меньшей чем, скажем, 0.99, для всех вершин, за исключением не более $\delta_{19} n$, $d'(x) = (1 \pm \delta_{18}) D e^{-\varepsilon(r-1)}$. Это завершает доказательство леммы. ■

Доказательство теоремы 4.7.1. Зафиксируем $\varepsilon > 0$, такое, что

$$\frac{\varepsilon}{1 - e^{-\varepsilon}} + r\varepsilon < 1 + a,$$

и $1/10 > \delta > 0$, такое, что

$$(1+4\delta)\frac{\varepsilon}{1 - e^{-\varepsilon}} + r\varepsilon < 1 + a.$$

Зафиксируем также целое t, такое, что $e^{-\varepsilon t} < \varepsilon$. Теорема доказывается t-кратным применением леммы. Положим $\delta = \delta_t$ и определим с помощью обратной индукции последовательность $\delta_t > \delta_{t-1} > \cdots > \delta_0$, такую, что $\delta_i \leqslant \delta_{i+1} e^{-\varepsilon(r-1)}$ и $\prod_{i=0}^{t}(1+\delta_i) < 1 + 2\delta$. Тогда для $n \geqslant D \geqslant R_i$ можно применить лемму с r, $K = ke^{\varepsilon i(r-1)}$, ε, $\delta' = \delta_{i+1}$ и $\delta = \delta_i$. Это доказывает утверждение теоремы с $\gamma = \delta_0$, $d_0 = \max\ R_i$. В самом деле, применяя повторно лемму, мы получаем убывающую последовательность множеств вершин $V = V_0, V_1, \ldots, V_t$, каждое из которых вложено в предыдущее, и последовательность множеств ребер $E_1, E_2, \ldots, E_t$, где E_i — множество ребер E', полученное применением леммы к гиперграфу, порожденному множеством V_{i-1}. Здесь

$$|V_i| = |V_{i-1}| e^{-\varepsilon}(1 \pm \delta_i)\ (\ = |V_0| e^{-i\varepsilon}(1 \pm 2\delta)\ ),$$

$$|E_i| = \frac{\varepsilon |V_{i-1}|}{r}(1 \pm \delta_i) \leqslant (1+4\delta)\frac{\varepsilon n}{r} e^{-(i-1)\varepsilon},$$

и

$$D_i = D_{i-1} e^{-\varepsilon(r-1)} = D e^{-\varepsilon i (r-1)}.$$

Покроем каждую вершину из множества V_t произвольным содержащим ее ребром. Мы получим покрытие, общее число ребер в котором не больше

$$(1+4\delta)\sum_{i=0}^{t-1} \frac{\varepsilon n}{r} e^{-i\varepsilon} + |V_t| \leqslant (1+4\delta)\frac{\varepsilon n}{r}\frac{1}{1 - e^{-\varepsilon}} + (1+2\delta) n e^{-\varepsilon t} \leqslant$$

$$\leqslant \frac{n}{r}[(1+4\delta)(\frac{\varepsilon}{1 - e^{-\varepsilon}} + r\varepsilon)] < (1+a)\frac{n}{r}.$$

Тем самым теорема доказана. ■

Мы завершим раздел демонстрацией того, как легко из доказанной теоремы вытекает предложенное Рёдлем решение проблемы Эрдёша—Ханани, упоминавшейся в начале раздела.

Теорема 4.7.3 (Рёдль). *Для фиксированных k, l выполнено неравенство*

$$M(n,k,l) \leqslant (1+o(1))\binom{n}{l} \Big/ \binom{k}{l},$$

где $o(1)$ стремится к нулю при n, стремящемся к бесконечности.

Доказательство. Положим $r = \binom{k}{l}$, и пусть H — r-однородный гиперграф, вершинами которого являются все l-подмножества множества $\{1, 2, \ldots, n\}$, а ребрами — семейства из $\binom{k}{l}$ l-множеств, содержащихся в k-множествах. Гиперграф H содержит $\binom{n}{l}$ вершин, каждая из которых имеет степень $D = \binom{n-l}{k-l}$, и любые две различные вершины лежат не более чем в $\binom{n-l-1}{k-l-1} = o(D)$ общих ребрах. Следовательно, по теореме 4.7.1 гиперграф H имеет покрытие размера, не превышающего $(1+o(1))\binom{n}{l} \Big/ \binom{k}{l}$, что и требовалось. ■

4.8. УПРАЖНЕНИЯ

1. Пусть X — случайная величина, принимающая целые неотрицательные значения, $\mathbf{E}[X^2]$ — математическое ожидание ее квадрата и $\mathrm{Var}[X]$ — ее дисперсия. Доказать, что

$$\Pr[X = 0] \leqslant \frac{\mathrm{Var}[X]}{\mathbf{E}[X^2]}.$$

2*. Доказать справедливость следующего утверждения. Пусть n действительных чисел $a_1, a_2, \ldots, a_n$ удовлетворяют условию $\sum_{i=1}^n a_i^2 = 1$, а $(\varepsilon_1, \ldots, \varepsilon_n)$ — случайный $\{-1, 1\}$-вектор, полученный путем случайного и независимого выбора ε_i, с равной вероятностью принимающих значения -1 или 1. Тогда существует положительная константа c такая, что

$$\Pr[|\sum_{i=1}^n \varepsilon_i a_i| \leqslant 1] \geqslant c.$$

3*. Доказать справедливость следующего утверждения. Пусть n векторов $a_1, a_2, \ldots, a_n \in \mathbb{R}^2$ удовлетворяют условию $\sum_{i=1}^n ||a_i||^2 = 1$ и $||a_i|| \leqslant 1/10$, где $||\cdot||$ обозначает обычную евклидову норму. Пусть $(\varepsilon_1, \ldots, \varepsilon_n)$ — случайный $\{-1, 1\}$-вектор, полученный путем случайного и независимого выбора ε_i, с равной вероятностью принимающих значения -1 или 1. Тогда существует положительная константа c, такая, что

$$\Pr\left[\left|\left|\sum_{i=1}^n \varepsilon_i a_i\right|\right| \leqslant 1/3\right] \geqslant c.$$

4. Пусть X — случайная величина с математическим ожиданием $\mathbf{E}[X] = 0$ и дисперсией σ^2. Доказать, что для всех $\lambda > 0$ справедливо неравенство

$$\Pr[X \geqslant \lambda] \leqslant \frac{\sigma^2}{\sigma^2 + \lambda^2}.$$

5. Рассмотрим n двумерных векторов $v_1 = (x_1, y_1), \ldots, v_n = (x_n, y_n)$, где каждое x_i и каждое y_i является целым числом с абсолютной величиной, не превышающей $\frac{2^{n/2}}{100\sqrt{n}}$. Показать, что существуют два непересекающихся множества $I, J \subset \{1, 2, \ldots, n\}$, такие, что

$$\sum_{i \in I} v_i = \sum_{j \in J} v_j.$$

6* Доказать, что для каждого множества X из не менее чем $4k^2$ различных классов вычетов по простому модулю p существует целое число a, такое, что множество $\{ax \pmod p : x \in X\}$ имеет непустое пересечение с каждым интервалом длины, не меньшей p/k, множества $\{0, 1, \ldots, p-1\}$.

ВЕРОЯТНОСТНЫЙ ВЗГЛЯД:

Гамильтоновы пути

Каково максимально возможное число ориентированных гамильтоновых путей в турнире на n вершинах? Обозначим это число через $P(n)$. Первым применением вероятностного метода в комбинаторике был результат Селе [Szele (1943)], представленный в гл. 2. Результат утверждает, что $P(n) \geqslant n!/2^{n-1}$. Эта оценка немедленно следует из того, что правая часть неравенства является средним числом путей в случайном турнире на n вершинах. В той же статье Селе показал, что существует предел величины $\left(\frac{P(n)}{n!}\right)^{1/n}$ и доказал, что

$$\frac{1}{2} \leqslant \lim_{n\to\infty} \left(\frac{P(n)}{n!}\right)^{1/n} \leqslant \frac{1}{2^{3/4}}.$$

Он высказал также гипотезу о том, что точное значение предела равно $1/2$.

Эта гипотеза доказана Алоном [Alon (1990a)]. Доказательство приводится ниже. Основной инструмент — это доказательство Брегмана гипотезы Минка о перманенте $(0,1)$-матрицы, представленное в разд. «Вероятностный взгляд» после гл. 2.

Теорема. *Существует положительная константа c, такая, что для всякого n выполнено неравенство*

$$P(n) \leqslant cn^{3/2}\frac{n!}{2^{n-1}}.$$

Доказательство. Для турнира T пусть $P(T)$ — число ориентированных гамильтоновых путей в T, $C(T)$ — число гамильтоновых циклов в T, а $F(T)$ — число остовных подграфов в T, у которых для каждой вершины существуют в точности одна заходящая дуга и в точности одна исходящая дуга. Очевидно,

$$C(T) \leqslant F(T). \tag{1}$$

Если $T = (V, E)$ является турниром на множестве $V = \{1, 2, \ldots, n\}$ из n вершин, то *матрица смежности* турнира T представляет собой $(0,1)$-матрицу размера $n \times n$, обозначаемую через $A_T = (a_{ij})$ и определяемую равенствами $a_{ij} = 1$ при $(i,j) \in E$ и $a_{ij} = 0$ в противном случае. Обозначим через r_i количество единиц в строке с номером i. Ясно, что

$$\sum_{i=1}^{n} r_i = \binom{n}{2}. \tag{2}$$

Комбинаторная интерпретация слагаемых из разложения перманента $\mathrm{per}(\mathrm{A_T})$ влечет равенство

$$\mathrm{per}(A_T) = F(T). \tag{3}$$

Нам понадобится следующая техническая лемма.

Лемма. *Для любых двух целых чисел a, b, таких, что $b - 2 \geqslant a \geqslant 1$ выполняется неравенство*

$$(a!)^{1/a} \cdot (b!)^{1/b} < ((a+1)!)^{1/(a+1)} \cdot ((b-1)!)^{1/(b-1)}.$$

Доказательство. Утверждение состоит просто в том, что $f(a) < f(b-1)$ для функции f, определенной равенством $f(a) = (a!)^{1/a}/((a+1)!)^{1/(a+1)}$. Таким образом, достаточно показать, что для всякого целого числа $x \geqslant 2$ выполнено неравенство $f(x-1) < f(x)$. Подставим сюда выражение для f и возведем обе части в степень $x(x-1)(x+1)$. Теперь достаточно убедиться в том, что для всех $x \geqslant 2$ выполнено соотношение

$$((x-1)!)^{x(x+1)} \cdot ((x+1)!)^{x(x-1)} < (x!)^{2(x^2-1)},$$

т. е.

$$\left(\frac{x^x}{x!}\right)^2 > \left(\frac{x+1}{x}\right)^{x(x-1)}.$$

Это, конечно, справедливо для $x = 2$. При $x \geqslant 3$ это следует из того, что $4^x > e^{x+1}$, $x! < (\frac{x+1}{2})^x$ и $e^{x-1} > (\frac{x+1}{x})^{x(x-1)}$. ■

Следствие. *Положим $g(x) = (x!)^{1/x}$. Для всякого целого $S \geqslant n$ и целочисленных $x_i \geqslant 1$ при ограничении $\sum_{i=1}^n x_i = S$ максимум функции $\prod_{i=1}^n g(x_i)$ достигается тогда и только тогда, когда x_i различаются не более чем на единицу (т. е. тогда и только тогда, когда x_i равны либо $\lfloor S/n \rfloor$, либо $\lceil S/n \rceil$.)*

Доказательство. Если существуют два индекса i и j, такие, что $x_i \geqslant x_j + 2$ то по лемме значение произведения возрастает при добавлении единицы к x_j и вычитании единицы из x_i. ■

Возвращаясь к нашему турниру T, мы заметим, что числа r_i, определенные выше, являются степенями исхода вершин турнира T. Если хотя бы одна из них равна нулю, то ясно, что $C(T) = F(T) = 0$. В противном случае, из теоремы Брегмана, только что доказанного следствия, а также из соотношений (2) и (3), вытекает, что $F(T)$ является минимальным значением функции $\prod_{i=1}^n (r_i!)^{1/r_i}$, где целочисленные переменные r_i удовлетворяют условию (2) и являются равными, сколь это возможно. Прямым, хотя и слегка нудным, дифференцированием асимптотики с использованием формулы Стирлинга получаем следующий результат.

Предложение. *Для всякого турнира T на n вершинах*

$$C(T) \leqslant F(T) \leqslant (1 + o(1)) \frac{\sqrt{\pi}}{\sqrt{2}e} n^{3/2} \frac{(n-1)!}{2^n}.$$

Чтобы закончить доказательство теоремы, мы должны получить оценку для числа гамильтоновых путей в турнире из полученного выше результата. Рассмотрим турнир S на n вершинах. Обозначим через T случайный турнир, полученный из S добавлением новой вершины y и присвоением ориентации ребрам, связывающим y с вершинами из S, случайно и независимо. Для каждого фиксированного гамильтонова пути в S вероятность того, что он расширится до гамильтонова цикла в T, равна $1/4$. Значит, математическое

ожидание числа гамильтоновых циклов в T равно $\frac{1}{4}P(S)$, и, следовательно, существует T, для которого $C(T) \geqslant \frac{1}{4}P(S)$. Но в силу предложения выполнено неравенство $C(T) \leqslant (1+o(1))\frac{\sqrt{\pi}}{\sqrt{2}e}(n+1)^{3/2}\frac{n!}{2^{n+1}}$, и, таким образом, $P(S) \leqslant O(n^{3/2}\frac{n!}{2^{n-1}})$. Тем самым теорема доказана. ■

5

Локальная лемма

Математика, на самом деле, — это почти полностью предмет эстетики. Эта мысль непостижима для нематематиков.

Джон Конвей

5.1. ЛЕММА

В типичном вероятностном доказательстве комбинаторного результата обычно требуется показать, что вероятность некоторого события положительна. Однако многие из таких доказательств в действительности дают больше и показывают, что вероятность рассматриваемого события не только положительная, но и достаточно большая. Фактически, большинство вероятностных доказательств имеют дело с событиями, которые происходят с большой вероятностью (т. е. вероятность стремится к 1), когда размерность задачи возрастает. Например, рассмотрим приведенное в гл. 1 доказательство того, что для каждого $k \geqslant 1$ существуют турниры, в которых для любого множества из k игроков существует игрок, побеждающий каждого из них. Доказательство действительно показывает, что для любого фиксированного k, при достаточно большом числе игроков n почти все турниры с n игроками обладают этим свойством; т. е. вероятность того, что случайный турнир с n игроками обладает требуемым свойством, стремится к 1 при n, стремящемся к бесконечности.

С другой стороны, существует тривиальный случай, в котором можно показать, что определенное событие происходит с положительной, хотя и очень малой, вероятностью. В самом деле, если мы имеем n взаимно независимых событий, и каждое из них происходит с вероятностью, не меньшей $p > 0$, то вероятность того, что все события происходят одновременно, равна по меньшей мере p^n, т. е. положительна, хотя может быть экспоненциально малой по n.

Естественно ожидать что случай взаимной независимости может быть обобщен на слабые зависимости и дать более общий способ доказательства того, что определенные события происходят с положительной, хотя и малой вероятностью. Такое обобщение возможно, и это утверждается в следующей лемме, известной как локальная лемма Ловаса. Эта простая лемма, впервые

доказанная в работе Эрдёша и Ловаса [Erdős and Lovász (1975)], является чрезвычайно мощным инструментом, поскольку дает возможность работать с редкими событиями.

Лемма 5.1.1 (локальная лемма: общий случай). *Пусть $A_1, A_2, \ldots, A_n$ — события в произвольном вероятностном пространстве. Ориентированный граф $D = (V, E)$ на множестве вершин $V = \{1, 2, \ldots, n\}$ называется* ***орграфом зависимости*** *для событий $A_1, \ldots, A_n$, если для всякого i, $1 \leqslant i \leqslant n$, каждое событие A_i взаимно независимо с событиями $\{A_j : (i,j) \notin E\}$. Предположим, что $D = (V, E)$ является орграфом зависимости для определенных выше событий и пусть существуют действительные числа $x_1, \ldots, x_n$, такие, что $0 \leqslant x_i < 1$ и $\Pr[A_i] \leqslant x_i \prod_{(i,j)\in E}(1 - x_j)$ для всех $1 \leqslant i \leqslant n$. Тогда $\Pr\left[\bigwedge_{i=1}^n \overline{A}_i\right] \geqslant \prod_{i=1}^n (1 - x_i)$. В частности, с положительной вероятностью ни одно из событий A_i не происходит.*

Доказательство. Сначала докажем индукцией по s, что для любого набора чисел $S \subset \{1, \ldots, n\}$, $|S| = s < n$ и любого $i \notin S$ справедливо неравенство

$$\Pr\left[A_i \middle| \bigwedge_{j\in S} \overline{A}_j\right] \leqslant x_i. \tag{5.1}$$

Это, конечно, справедливо для $s = 0$. Предполагая, что неравенство выполняется для всех чисел $s' < s$, докажем его для s. Положим $S_1 = \{j \in S; (i,j) \in E\}$, $S_2 = S \backslash S_1$. Тогда

$$\Pr\left[A_i \middle| \bigwedge_{j\in S} \overline{A}_j\right] = \frac{\Pr\left[A_i \wedge \left(\bigwedge_{j\in S_1} \overline{A}_j\right) \middle| \bigwedge_{\ell\in S_2} \overline{A}_\ell\right]}{\Pr\left[\bigwedge_{j\in S_1} \overline{A}_j \middle| \bigwedge_{\ell\in S_2} \overline{A}_\ell\right]}. \tag{5.2}$$

Чтобы оценить числитель, заметим, что так как каждое событие A_i взаимно независимо с событиями $\{A_\ell : \ell \in S_2\}$, то

$$\begin{aligned}\Pr\left[A_i \wedge \left(\bigwedge_{j\in S_1} \overline{A}_j\right) \mid \bigwedge_{\ell\in S_2} \overline{A}_\ell\right] &\leqslant \Pr\left[A_i \middle| \bigwedge_{\ell\in S_2} \overline{A}_\ell\right] = \\ &= \Pr[A_i] \leqslant x_i \prod_{(i,j)\in E} (1 - x_j).\end{aligned} \tag{5.3}$$

С другой стороны, знаменатель может быть оценен с использованием предположения индукции. В самом деле, предположим $S_1 = \{j_1, j_2, \ldots, j_r\}$. Если $r = 0$, то знаменатель равен 1, и неравенство (5.1) выполняется. В противном

случае

$$\Pr\left[\overline{A}_{j_1}\wedge\overline{A}_{j_2}\wedge\ldots\wedge\overline{A}_{j_r}\Big|\bigwedge_{\ell\in S_2}\overline{A}_\ell\right]=$$

$$=\left(1-\Pr[A_{j_1}|\bigwedge_{\ell\in S_2}\overline{A}_\ell]\right)\cdot\left(1-\Pr\left[A_{j_2}|\overline{A}_{j_1}\wedge\bigwedge_{\ell\in S_2}\overline{A}_\ell\right]\right)\times\ldots$$

$$\ldots\times\left(1-\Pr\left[A_{j_r}|\overline{A}_{j_1}\wedge\ldots\wedge\overline{A}_{j_{r-1}}\wedge\bigwedge_{\ell\in S_2}\overline{A}_\ell\right]\right)\geqslant$$

$$\geqslant(1-x_{j_1})(1-x_{j_2})\ldots(1-x_{j_r})\geqslant\prod_{(i,j)\in E}(1-x_j). \tag{5.4}$$

Подставляя соотношения (5.3) и (5.4) в равенство (5.2), заключаем, что $\Pr\left[A_i|\bigwedge_{j\in S}\overline{A}_j\right]\leqslant x_i$, что завершает индуктивное доказательство.

Теперь легко получить утверждение леммы 5.1.1, поскольку

$$\Pr\left[\bigwedge_{i=1}^n\overline{A}_i\right]=(1-\Pr[A_1])\cdot(1-\Pr[A_2|\overline{A}_1])\times\ldots$$

$$\ldots\times(1-\Pr[A_n|\bigwedge_{i=1}^{n-1}\overline{A}_i]\geqslant\prod_{i=1}^n(1-x_i). \tag{5.5}$$

Лемма доказана. ■

Следствие 5.1.2 (локальная лемма: симметричный случай). *Пусть $A_1,A_2,\ldots,A_n$ — события в произвольном вероятностном пространстве. Пусть каждое событие A_i взаимно независимо со всеми событиями, за исключением тех не более чем d событий A_j, для которых $\Pr[A_i]\leqslant p$ при всех $1\leqslant i\leqslant n$. Если*

$$ep(d+1)\leqslant 1, \tag{5.6}$$

то $\Pr\left[\bigwedge_{i=1}^n\overline{A}_i\right]>0$.

Доказательство. Если $d=0$, то утверждение тривиально. В противном случае рассмотрим орграф зависимости $D=(V,E)$ для событий $A_1,\ldots,A_n$, в котором по предположению для каждого i выполнено $\big|\{j:(i,j)\in E\}\big|\leqslant d$. Результат теперь вытекает из леммы 5.1.1, если положить $x_i=1/(d+1)(<1)$ для всех i и использовать тот факт, что $\left(1-\frac{1}{d+1}\right)^d>1/e$ для всякого $d\geqslant 1$. ■

Как показал Ширер в 1985 г., константа «e» является наилучшей из возможных в неравенстве (5.6). Заметим также, что доказательство леммы 5.1.1 показывает, что утверждение остается справедливым, даже если заменить два предположения о том, что каждое A_i является взаимно независимым с событиями $\{A_j:(i,j)\notin E\}$ и что $\Pr[A_i]\leqslant x_i\prod_{(i,j)\in E}(1-x_j)$, более слабым предположением о том, что для каждого i и любого $S_i\subset\{1,\ldots,n\}\backslash\{j:(i,j)\in E\}$

выполнено $\Pr\left[A_i \middle| \bigwedge_{j\in S_i} \overline{A}_j\right] \leqslant x_i \prod_{(i,j)\in E}(1-x_j)$. Это оказывается полезным в некоторых приложениях.

В нескольких следующих разделах мы рассмотрим различные применения локальной леммы для получения комбинаторных результатов. Доказательства этих результатов, не использующие локальную лемму, неизвестны.

5.2. СВОЙСТВО B И РАЗНОЦВЕТНЫЕ МНОЖЕСТВА ДЕЙСТВИТЕЛЬНЫХ ЧИСЕЛ

Напомним, что гиперграф $H = (V, E)$ обладает свойством B, (т. е. является 2-раскрашиваемым), если существует раскраска множества V в два цвета, при которой никакое ребро $f \in E$ не является монохроматическим.

Теорема 5.2.1. *Пусть $H = (V, E)$ — гиперграф, в котором каждое ребро содержит по меньшей мере k элементов и каждое ребро пересекается не более чем с d другими ребрами. Если $e(d+1) \leqslant 2^{k-1}$, то H обладает свойством B.*

Доказательство. Раскрасим каждую вершину v из H случайно и независимо либо синим, либо красным цветом с равной вероятностью. Для каждого ребра $f \in E$ пусть A_f — событие «f является монохроматическим». Ясно, что $\Pr[A_f] = 2/2^{|f|} \leqslant 1/2^{k-1}$. Кроме того, каждое событие A_f, очевидно, не зависит от событий $A_{f'}$, соответствующих тем ребрам f', которые не пересекаются с f. Теперь результат вытекает из следствия 5.1.2. ■

Специальный случай теоремы 5.2.1 состоит в том, что для всех $k \geqslant 9$ любой k-однородный гиперграф H обладает свойством B. В самом деле, любое ребро f такого гиперграфа H содержит k вершин, каждая из которых инцидентна k ребрам (включая f). Поэтому ребро f пересекает не более $d = k(k-1)$ других ребер. Требуемый результат вытекает отсюда, поскольку $e(k(k-1)+1) < 2^{k-1}$ для каждого $k \geqslant 9$.

Следующий результат, который мы рассмотрим, появился в работе Эрдёша и Ловаса, и касается k-раскрасок действительных чисел. Пусть даны k-раскраска $c : \mathbb{R} \to \{1, 2, \ldots, k\}$ действительных чисел в цвета $1, 2, \ldots, k$ и подмножество $T \subset \mathbb{R}$. Скажем, что T *разноцветно* (относительно c), если $c(T) = \{1, 2, \ldots, k\}$, т. е. T содержит элементы всех цветов.

Теорема 5.2.2. *Пусть m и k — два натуральных числа, удовлетворяющих условию*

$$e\,(m(m-1)+1)\,k\left(1-\frac{1}{k}\right)^m \leqslant 1. \tag{5.7}$$

Тогда для любого множества S из m действительных чисел существует k-раскраска, такая, что каждый сдвиг $x + S$ (для $x \in \mathbb{R}$) является разноцветным.

Заметим, что неравенство (5.7) справедливо для всех $m > (3 + o(1))k \log k$.

Доказательство. Сначала зафиксируем *конечное* подмножество $X \subseteq \mathbb{R}$ и докажем существование k-раскраски, для которой каждый сдвиг $x + S$ (для $x \in X$) является разноцветным. Это простое следствие локальной леммы. В самом деле, положим $Y = \bigcup_{x \in X}(x + S)$ и пусть $c : Y \to \{1, 2, \ldots, k\}$ — случайная k-раскраска множества Y, полученная выбором для каждого $y \in Y$, случайно и независимо $c(y) \in \{1, 2, \ldots, k\}$ в соответствии с равномерным распределением на множестве $\{1, 2, \ldots, k\}$. Для каждого $x \in X$ пусть A_x — событие «$x + S$ не является разноцветным (относительно c)». Ясно, что $\Pr[A_x] \leqslant k\left(1 - \frac{1}{k}\right)^m$. Кроме того, каждое событие A_x является взаимно независимым со всеми другими событиями $A_{x'}$ кроме тех, для которых $(x + S) \cap (x' + S) \neq \varnothing$. Поскольку существует не более $m(m-1)$ таких событий, требуемый результат вытекает из следствия 5.1.2.

Мы можем теперь доказать существование раскраски множества всех действительных чисел, удовлетворяющей требуемым свойствам, с помощью стандартного понятия компактности. Дискретное пространство с k точками (тривиально) является компактом. Теорема Тихонова (эквивалентная аксиоме выбора) утверждает, что произвольное произведение таких пространств также является компактом. В частности, пространство всех функций из $\mathbb{R}$ в $\{1, 2, \ldots, k\}$ с обычной топологией произведения, является компактом. В этом пространстве для каждого фиксированного $x \in \mathbb{R}$ множество C_x всех раскрасок c, таких, что $x + S$ является разноцветным, замкнуто. (В действительности, оно и открыто и замкнуто, так как базис для открытых множеств — это множество всех раскрасок, значения которых определены в конечном числе точек.) Как доказано выше, пересечение конечного числа множеств C_x не является пустым. Отсюда следует, ввиду компактности, что пересечение всех множеств C_x не пусто. Любая раскраска из этого пересечения обладает свойствами из утверждения теоремы 5.2.2. ■

Заметим, что, в общем, невозможно применить локальную лемму к бесконечному множеству событий и сделать заключение о том, что в некоторой точке вероятностного пространства ни одно из них не выполняется. В самом деле, существуют тривиальные примеры счетного множества взаимно независимых событий A_i, удовлетворяющих условиям $\Pr[A_i] = 1/2$ и $\bigwedge_{i \geqslant 1} \overline{A}_i = \varnothing$. Поэтому соображения компактности в предыдущем доказательстве являются существенными.

5.3. НИЖНИЕ ОЦЕНКИ ДЛЯ ЧИСЕЛ РАМСЕЯ

Получение Эрдёшем в 1947 г. нижних оценок для чисел Рамсея было одним из первых применений вероятностного метода. Локальная лемма дает простой способ улучшения этих оценок. Сначала получим нижнюю оценку для диагональных чисел Рамсея $R(k, k)$. Рассмотрим случайную 2-раскраску ребер полного графа K_n. Для каждого множества S из k вершин графа K_n обозначим через A_S событие «полный граф на множестве S является монохроматическим». Ясно, что $\Pr[A_S] = 2^{1-\binom{k}{2}}$. Очевидно, что каждое событие A_S является

взаимно независимым со всеми другими событиями A_T кроме тех, для которых $|S \cap T| \geqslant 2$, поскольку только в этом случае соответствующие полные подграфы имеют общее ребро. Значит, можно применить следствие 5.1.2 с $p = 2^{1-\binom{k}{2}}$ и $d = \binom{k}{2}\binom{n}{k-2}$ с тем, чтобы заключить:

Предложение 5.3.1. *Если* $e\left(\binom{k}{2}\binom{n}{k-2}+1\right)\cdot 2^{1-\binom{k}{2}} < 1$, *то* $R(k,k) > n$.

Отсюда легко следует, что $R(k,k) > \frac{\sqrt{2}}{e}(1+o(1))\,k2^{k/2}$. Это всего лишь в два раза лучше оценки, полученной прямым вероятностным методом. Хотя столь небольшое улучшение несколько разочаровывает, это не удивительно. Локальная лемма наиболее эффективна, когда зависимости между событиями редки, а в данном случае это не так. В самом деле, существует всего $K = \binom{n}{k}$ рассматриваемых событий, а максимальная степень исхода d в орграфе зависимости равна приблизительно $\binom{k}{2}\binom{n}{k-2}$. Для больших k и много бóльших n (что как раз и представляет для нас интерес) мы имеем $d > K^{1-O(1/k)}$, т. е. массу зависимостей. С другой стороны, если мы рассмотрим малые множества S, например, множества размера 3, то увидим, что всего их $K = \binom{n}{3}$, и каждое из них имеет общее ребро только с $3(n-3) \approx K^{1/3}$ множествами размера 3. Это подсказывает, что локальная лемма может дать больше для улучшения оценок внедиагональных чисел Рамсея $R(k,l)$, особенно если один из параметров, скажем l, мал. Рассмотрим например, следуя Спенсеру [Spencer (1977)], число Рамсея $R(k,3)$. Здесь, конечно, мы должны применить несимметричную форму локальной леммы. Раскрасим ребра графа K_n в два цвета случайно, независимо и так, что каждое ребро окрашивается в синий цвет с вероятностью p. Для каждого множества из трех вершин T обозначим через A_T событие «треугольник на множестве T синий». Аналогично, для каждого множества S из k вершин обозначим через B_S событие «полный граф на множестве S красный». Ясно, что $\Pr[A_T] = p^3$ и $\Pr[B_S] = (1-p)^{\binom{k}{2}}$. Построим орграф зависимости для событий A_T и B_S путем соединения двух вершин дугами (в обоих направлениях) тогда и только тогда, когда соответствующие графы имеют общее ребро. Ясно, что каждая вершина A_T орграфа зависимости S смежна с $3(n-3) < 3n$ вершинами $A_{T'}$ и не более чем с $\binom{n}{k}$ вершинами $B_{S'}$. Аналогично, каждая вершина B_S смежна с $\binom{k}{2}(n-k) < k^2n/2$ вершинами A_T и не более чем с $\binom{n}{k}$ вершинами $B_{S'}$. Как вытекает из общего случая локальной леммы 5.1.1, если мы сможем найти $0 < p < 1$ и два действительных числа $0 \leqslant x < 1$ и $0 \leqslant y < 1$, таких, что

$$p^3 \leqslant x(1-x)^{3n}(1-y)^{\binom{n}{k}}$$

и

$$(1-p)^{\binom{k}{2}} \leqslant y(1-x)^{k^2n/2}(1-y)^{\binom{n}{k}},$$

то $R(k,3) > n$.

Наша цель — найти наибольшее возможное $k = k(n)$, для которого существуют такие числа p, x и y. Элементарное (но утомительное) вычисление показывает, что наилучшим является выбор, при котором $p = c_1 n^{-1/2}$,

$k = c_2 n^{1/2} \log n$, $x = c_3/n^{3/2}$ и $y = \frac{c_4}{e^{n^{1/2}\log^2 n}}$. Это дает оценку $R(k,3) > c_5 k^2/\log^2 k$. Подобные же соображения дают $R(k,4) > k^{5/2+o(1)}$. В обоих случаях количество вычислений соизмеримо. Однако, тяжкий труд оплачивается. Оценка $R(k,3) > c_5 k^2/\log^2 k$ сравнима с нижней оценкой Эрдёша, доказанной в 1961 г. с помощью весьма сложных вероятностных соображений. Она была улучшена Кимом [Kim (1995)] до $R(k,3) > c_6 k^2/\log k$. Данная оценка для $R(k,4)$ лучше, чем любая известная из тех, что получены без применения локальной леммы.

5.4. ГЕОМЕТРИЧЕСКИЙ РЕЗУЛЬТАТ

Семейство $\mathcal{F}$ открытых шаров в трехмерном евклидовом пространстве $\mathbb{R}^3$ называется *k-кратным покрытием* пространства $\mathbb{R}^3$, если любая точка $x \in \mathbb{R}^3$ принадлежит не менее чем k шарам. В частности, однократное покрытие называется просто *покрытием*. Назовем k-кратное покрытие $\mathcal{F}$ *разложимым*, если существует разбиение семейства $\mathcal{F}$ на два попарно непересекающихся семейства $\mathcal{F}_1$ и $\mathcal{F}_2$, каждое из которых является покрытием пространства $\mathbb{R}^3$. Мани-Левицка и Пач [Mani-Levitska and Pach (1988)] построили неразложимые k-кратные покрытия пространства $\mathbb{R}^3$ открытыми единичными шарами для любых целых чисел $k \geqslant 1$. С другой стороны, они доказали, что любое k-кратное покрытие пространства $\mathbb{R}^3$, при котором никакая точка не покрывается более чем $c2^{k/3}$ шарами, является разложимым. Обнаруживается неожиданное явление: труднее разложить те покрытия, которые покрывают некоторые точки пространства $\mathbb{R}^3$ слишком часто, чем те покрытия, которые покрывают каждую точку равномерно. Точное утверждение теоремы Мани-Левицки и Пача состоит в следующем.

Теорема 5.4.1. *Пусть $\mathcal{F} = \{B_i\}_{i\in I}$ является k-кратным покрытием трехмерного евклидова пространства открытыми единичными шарами. Пусть каждая точка пространства $\mathbb{R}^3$ содержится не более чем в t шарах из $\mathcal{F}$. Если*

$$e \cdot t^3 2^{18}/2^{k-1} \leqslant 1,$$

то покрытие $\mathcal{F}$ разложимо.

Доказательство. Определим бесконечный гиперграф $H = (V(H), E(H))$ следующим образом. Множество $V(H)$ вершин гиперграфа H — это просто $\mathcal{F} = \{B_i\}_{i\in I}$. Для каждого $x \in \mathbb{R}^3$ положим $E_x = \{B_i \in \mathcal{F} : x \in B_i\}$. Множество $E(H)$ ребер гиперграфа H есть просто множество семейств E_x с учетом того, что при $E_x = E_y$ ребро берется один раз.

Мы утверждаем, что каждое ребро E_x пересекает менее $t^3 2^{18}$ других ребер из H. Если $x \in B_i$, то центр шара B_i находится на расстоянии не больше 1 от x. Если же $B_j \cap B_i \neq \varnothing$, то центр шара B_j находится внутри шара радиуса 3 с центром в точке x, а сам B_j полностью содержится в шаре радиуса 4 с центром в x. Шар B_j покрывает долю, в точности равную $4^{-3} = 2^{-6}$ от объема этого

шара. Поскольку ни одна из вершин не покрывается больше t раз, существует не более 2^6t таких шаров. Нетрудно проверить, что m шаров в $\mathbb{R}^3$ разделяют пространство $\mathbb{R}^3$ меньше чем на m^3 компонент связности. Поэтому существует не более $(2^6t)^3$ различных шаров E_y, пересекающих E_x.

Рассмотрим теперь произвольный конечный подгиперграф L гиперграфа H. Каждое ребро в L имеет по меньшей мере k вершин и пересекает не более $d < t^3 2^{18}$ других ребер из L. Так как по условию $e(d+1) \leqslant 2^{k-1}$, из теоремы 5.2.1 (которая является простым следствием локальной леммы) следует, что L является 2-раскрашиваемым. Это означает, что можно раскрасить вершины гиперграфа L в синий и красный цвета так, что ни одно из ребер в L не является монохроматическим. Это справедливо для любого конечного гиперграфа L. Соображения компактности, аналогично тому, как они использовались в доказательстве теоремы 5.2.2, показывают что H является 2-раскрашиваемым. При заданной 2-раскраске гиперграфа H без монохроматических ребер мы просто выбираем в качестве $\mathcal{F}_1$ множество всех синих шаров, а в качестве $\mathcal{F}_2$ — множество красных. Ясно, что каждое $\mathcal{F}_i$ является покрытием для $\mathbb{R}^3$, что и завершает доказательство теоремы. ■

Стоит заметить, что теорема 5.4.1 может быть легко обобщена на случай бóльших размерностей. Мы опускаем детальную формулировку этого обобщения.

5.5. ЛИНЕЙНАЯ ДРЕВЕСНОСТЬ ГРАФОВ

Линейным называется лес (т. е. ациклический простой граф), в котором каждая компонента связности является цепью. *Линейная древесность* $\mathrm{la}(G)$ графа G есть минимальное число линейных лесов, объединение которых содержит все ребра графа G. Это понятие было введено Харари как один из типов покрытий, являющийся инвариантом графа. Следующая гипотеза, известная как *гипотеза линейной древесности* возникла в статье [Akiyama, Exoo and Harary (1981)].

Гипотеза 5.5.1 (гипотеза линейной древесности). *Линейная древесность любого d-регулярного графа равна $\lceil (d+1)/2 \rceil$.*

Заметим, что поскольку каждый d-регулярный граф G с n вершинами имеет $nd/2$ ребер, каждый линейный лес в нем имеет не более $n-1$ ребер. Отсюда немедленно следует неравенство

$$\mathrm{la}(G) \geqslant \frac{nd}{2(n-1)} > \frac{d}{2}.$$

Так как $\mathrm{la}(G)$ — целое, $\mathrm{la}(G) \geqslant \lceil (d+1)/2 \rceil$. Трудность доказательства гипотезы 5.5.1 лежит в обосновании справедливости обратного неравенства: $\mathrm{la}(G) \leqslant \lceil (d+1)/2 \rceil$. Заметим, что поскольку каждый граф G с максимальной степенью вершины Δ является подграфом Δ-регулярного графа (который, быть может, имеет больше вершин, а также большее число ребер, чем G), то

гипотеза линейной древесности эквивалентна утверждению о том, что линейная древесность каждого графа G с максимальной степенью Δ не превышает $\lceil(\Delta+1)/2\rceil$.

Хотя эта гипотеза привлекла большое внимание, наилучший общий результат в этом направлении, полученный без применения вероятностных соображений, состоит в том, что $\mathrm{la}(G) \leqslant \lceil 3\Delta/5\rceil$ для четных Δ и $\mathrm{la}(G) \leqslant \lceil(3\Delta+2)/5\rceil$ для нечетных Δ. В этом разделе мы докажем, что для любого $\varepsilon > 0$ существует $\Delta_0 = \Delta_0(\varepsilon)$, такое, что для всякого $\Delta \geqslant \Delta_0$ и любого графа с максимальной степенью Δ его линейная древесность меньше $\left(\frac{1}{2}+\varepsilon\right)\Delta$. Этот результат (с несколько более сложным доказательством) был получен Алоном в [Alon (1988)]. Его доказательство существенно опирается на локальную лемму. Заметим, что его доказательство более сложное, чем другие, приводимые в этой главе, и требует определенных предварительных сведений. Некоторые из них представляют самостоятельный интерес.

Нам будет удобнее вывести результат для неориентированных графов из его аналога для ориентированных графов. Ориентированный граф называется *d-регулярным*, если степени захода и исхода каждой вершины равна d. Линейный ориентированный лес — это ориентированный граф, в котором каждая компонента является ориентированной цепью. *Ориентированная линейная древесность* $\mathrm{dla}(G)$ ориентированного графа G — это минимальное число линейных ориентированных лесов в G, объединение которых содержит все дуги графа G. «Ориентированная версия» гипотезы линейной древесности, впервые сформулированная в 1987 г. в статье [Nakayama and Peroche (1987)], гласит:

Гипотеза 5.5.2. *Для каждого d-регулярного орграфа D*

$$\mathrm{dla}(D) = d+1.$$

Заметим, что ребра любого (связного) $2d$-регулярного неориентированного графа G могут быть ориентированы вдоль эйлерова цикла так, что полученный в результате ориентированный граф является d-регулярным. Справедливость гипотезы ориентированной линейной древесности 5.5.2 для d влечет справедливость гипотезы линейной древесности 5.5.1 для $2d$.

Легко доказать, что произвольный граф с n вершинами и максимальной степенью d содержит независимое множество размера не меньше $n/(d+1)$. Следующее утверждение показывает, что ценой уменьшения размера такого множества в некоторое постоянное число раз можно гарантировать, что оно будет обладать некоторым дополнительным свойством.

Предложение 5.5.3. *Рассмотрим граф $H = (V, E)$ с максимальной степенью d и разбиение $V = V_1 \cup V_2 \cup \ldots \cup V_r$ множества V на r попарно непересекающихся множеств. Пусть каждое множество V_i имеет мощность $|V_i| \geqslant 2ed$, где e — основание натурального логарифма. Тогда существует независимое множество вершин $W \subseteq V$, такое, что оно содержит по одной вершине из каждого V_i.*

Доказательство. Мы можем предположить, что каждое множество V_i имеет мощность, в точности равную $g = \lceil 2ed \rceil$ (иначе просто заменим каждое V_i подмножеством мощности g, а H — графом, который индуцирован объединением этих r новых множеств). Выберем из каждого множества V_i случайно и независимо по одной вершине в соответствии с равномерным распределением. Пусть W — случайное множество выбранных таким образом вершин. Чтобы закончить доказательство, покажем, что с положительной вероятностью W является независимым множеством вершин в графе H.

Для каждого ребра графа H определим событие A_f, состоящее в том, что W содержит оба конца ребра f. Очевидно, $\Pr[A_f] \leqslant 1/g^2$. Кроме того, если концы ребра f лежат в V_i и V_j, то A_f взаимно независимо со всеми событиями, соответствующих ребрам, концы которых не лежат в $V_i \cup V_j$. Следовательно, максимальная степень вершины орграфа зависимости меньше $2gd$. Поскольку $e \cdot 2gd \cdot 1/g^2 = 2ed/g < 1$, то в силу следствия 5.1.2 мы заключаем, что с положительной вероятностью ни одно из событий A_f не происходит. Но это означает, что W является независимым множеством, содержащим по одной вершине из каждого V_i, что и требовалось. ■

Предложения 5.5.3 достаточно, чтобы доказать гипотезу линейной древесности 5.5.2 для орграфов без коротких ориентированных циклов. Напомним, что ориентированный обхват орграфа — это минимальная длина ориентированного цикла в нем.

Теорема 5.5.4. *Пусть $G = (U, F)$ — d-регулярный орграф с ориентированным обхватом $g \geqslant 8ed$. Тогда*

$$\mathrm{dla}(G) = d + 1.$$

Доказательство. Как хорошо известно, F может быть разбито на d попарно непересекающихся 1-регулярных остовных подграфов $F_1, \ldots, F_d$ графа G. (Это — следствие теоремы Кенига—Холла. Рассмотрим двудольный граф H, у которого долями вершин A и B являются копии множества U, и в котором вершина $u \in A$ соединена с $v \in B$ тогда и только тогда, когда $(u, v) \in F$. Поскольку H является d-регулярным, множество его ребер может быть разбито на d совершенных паросочетаний, соответствующих d 1-регулярным остовным подграфам графа G.) Каждое F_i является объединением непересекающихся по вершинам направленных циклов $C_{i1}, C_{i2}, \ldots, C_{ir_i}$. Пусть V_1, $V_2, \ldots, V_r$ — множества *ребер* всех циклов $\{C_{i,j} : 1 \leqslant i \leqslant d, 1 \leqslant j \leqslant r_i\}$. Очевидно, $V_1, V_2, \ldots, V_r$ является разбиением множества F всех ребер G, по условию на обхват имеем $|V_i| \geqslant g \geqslant 8ed$ для всех $1 \leqslant i \leqslant r$. Рассмотрим линейный граф H графа G, т. е. граф, множеством вершин которого является множество F ребер графа G, причем два ребра смежны тогда и только тогда, когда у них существует общая вершина в G. Очевидно, H является $(4d-2)$-регулярным. Поскольку мощность каждого V_i не меньше $8ed \geqslant 2e(4d-2)$, то в силу предложения 5.5.3 существует независимое множество графа H, содержащее по одному элементу из каждого V_i. Но это означает существование в G

паросочетания M, содержащего не менее одного ребра из каждого цикла C_{ij} из 1-факторов $F_1, \ldots, F_d$. Следовательно, $M, F_1 \backslash M, F_2 \backslash M, \ldots, F_d \backslash M$ — это $d+1$ ориентированных лесов в G (один из них является паросочетанием), которые покрывают все его ребра). Следовательно,

$$\mathrm{dla}(G) \leqslant d + 1.$$

Поскольку G имеет $|U| \cdot d$ ребер, а каждый ориентированный линейный лес может иметь не более $|U| - 1$ ребер, выполнено

$$\mathrm{dla}(G) \geqslant |U|d/(|U| - 1) > d.$$

Таким образом, $\mathrm{dla}(G) = d + 1$. Теорема доказана. ■

Последняя теорема показывает, что утверждение гипотезы 5.5.2 справедливо для орграфов с достаточно большим (ориентированным) обхватом. Для того, чтобы оперировать с графами малого обхвата, мы покажем, что большинство ребер каждого регулярного орграфа может быть разложено на относительно небольшое число регулярных орграфов с большим обхватом. Для этого нам потребуется следующее утверждение, доказываемое с помощью локальной леммы.

Лемма 5.5.5. *Пусть $G = (V, E)$ — d-регулярный граф с достаточно большим d, а p — целое число, удовлетворяющее неравенствам $10\sqrt{d} \leqslant p \leqslant 20\sqrt{d}$. Тогда существует p-раскраска вершин графа G цветами 0,1,2, …, $p-1$ со следующим свойством: для всякой вершины $v \in V$ и любого цвета i величины $N^+(v,i) = \big|\{u \in V : (v,u) \in E,$ и раскрашена в цвет $i\}\big|$ и $N^-(v,i) = \big|\{u \in V : (u,v) \in E,$ и раскрашена в цвет $i\}\big|$ удовлетворяют неравенствам*

$$\begin{aligned} \left|N^+(v,i) - \frac{d}{p}\right| &\leqslant 3\sqrt{d/p}\sqrt{\ln d}, \\ \left|N^-(v,i) - \frac{d}{p}\right| &\leqslant 3\sqrt{d/p}\sqrt{\ln d}. \end{aligned} \tag{5.8}$$

Доказательство. Пусть $f : V \to \{0, 1, \ldots, p-1\}$ — случайная раскраска множества V в p цветов, при которой $f(v) \in \{0, 1, \ldots, p-1\}$ выбирается для каждого $v \in V$ в соответствии с равномерным распределением. Для каждой вершины $v \in V$ и каждого цвета $i, 0 \leqslant i < p$, обозначим через $A^+_{v,i}$ событие «число $N^+(v,i)$ соседей вершины v в G, окрашенных в цвет i, не удовлетворяет неравенству (5.8)». Очевидно, $N^+(v,i)$ — биномиальная случайная величина с математическим ожиданием $\frac{d}{p}$ и стандартным отклонением $\sqrt{\frac{d}{p}(1-\frac{1}{p})} < \sqrt{\frac{d}{p}}$. Следовательно, в силу стандартных оценок для биномиального распределения (см. Приложение А), для всякого $v \in V$ и $0 \leqslant i < p$

$$\Pr[A^+_{v,i}] < 1/d^4.$$

Аналогично, если $A^-_{v,i}$ — событие «число $N^-(v,i)$ соседей вершины v в G, окрашенных в цвет i, не удовлетворяет неравенству (5.8)», то

$$\Pr[A^-_{v,i}] < 1/d^4.$$

Ясно, что каждое из событий $A^+_{v,i}$ или $A^-_{v,i}$ взаимно независимо со всеми событиями $A^+_{u,j}$ или $A^-_{u,j}$ для всех вершин $u \in V$, которые не имеют общих соседей с v в графе G. Следовательно, орграф зависимости для всех наших событий имеет максимальную степень, не превосходящую $(2d)^2 \cdot p$. Поскольку $e \cdot \frac{1}{d^4}((2d)^2 p + 1) < 1$, следствие 5.1.2 (т. е. симметричная форма локальной леммы) утверждает, что с положительной вероятностью ни одно из событий $A^+_{v,i}$ или $A^-_{v,i}$ не происходит. Следовательно, существует раскраска f, удовлетворяющая неравенствам (5.8) для всех $v \in V$ и $0 \leqslant i < p$. Что и требовалось. ■

Теперь мы можем работать с произвольными регулярными орграфами. Пусть $G = (V, E)$ — произвольный d-регулярный орграф. Везде далее в доказательстве мы предполагаем, где это необходимо, что d существенно велико. Пусть p — простое число, удовлетворяющее неравенствам $10d^{1/2} \leqslant p \leqslant 20d^{1/2}$ (хорошо известно, что для всякого n существует простое число между n и $2n$). По лемме 5.5.5 существует вершинная раскраска $f : V \to \{0, 1, \ldots, p-1\}$, удовлетворяющая неравенствам (5.8). Для всякого i, $0 \leqslant i < p$, пусть $G_i = (V, E_i)$ — остовный подграф орграфа G, определяемый равенством $E_i = \{(u, v) \in E : f(v) \equiv (f(u) + i) \bmod p\}$. По неравенству (5.8) максимальная степень захода Δ^-_i и максимальная степень исхода Δ^+_i в каждом G_i не превосходит $\frac{d}{p} + 3\sqrt{\frac{d}{p}}\sqrt{\ln d}$. Кроме того, для каждого $i > 0$ длина каждого ориентированного цикла в G_i делится на p. Поэтому ориентированный обхват g_i графа G_i не меньше p. Поскольку каждый G_i может быть дополнен путем добавления вершин и ребер до Δ_i-регулярного орграфа с тем же ориентированным обхватом g_i и с $\Delta_i = \max(\Delta^+_i, \Delta^-_i)$, и поскольку $g_i > 8e\Delta_i$ (для достаточно больших d), мы заключаем по теореме 5.5.4, что $\mathrm{dla}(G_i) \leqslant \Delta_i + 1 \leqslant \frac{d}{p} + 3\sqrt{\frac{d}{p}}\sqrt{\ln d} + 1$ для всех $1 \leqslant i < p$. Для G_0 мы применяем тривиальное неравенство

$$\mathrm{dla}(G_0) \leqslant 2\Delta_0 \leqslant 2\frac{d}{p} + 6\sqrt{\frac{d}{p}}\sqrt{\ln d},$$

получаемое, например, вложением G_0 как подграфа в Δ_0-регулярный граф, разделяющий ребра этого графа на Δ_0 1-регулярных остовных подграфов, и разбивая каждый из этих 1-регулярных остовных подграфов на два линейных ориентированных леса. Последние два неравенства вместе с $10\sqrt{d} \leqslant p \leqslant 20\sqrt{d}$ дают

$$\mathrm{dla}(G) \leqslant d + 2\frac{d}{p} + 3\sqrt{pd}\sqrt{\ln d} + 3\sqrt{\frac{d}{p}}\sqrt{\ln d} + p - 1 \leqslant d + c \cdot d^{3/4}(\ln d)^{1/2}.$$

Таким образом доказан следующий результат.

Теорема 5.5.6. *Существует абсолютная константа $c > 0$, такая, что для каждого d-регулярного орграфа G*

$$\mathrm{dla}(G) \leqslant d + cd^{3/4}(\ln d)^{1/2}.$$

Заметим, что, действуя более аккуратно, мы можем довести остаточный член до $c'd^{2/3}(\ln d)^{1/3}$. Поскольку ребра неориентированного d-регулярного графа с $d = 2f$ могут быть ориентированы так, что полученный орграф станет f-регулярным, и поскольку всякий $(2f-1)$-регулярный неориентированный граф является подграфом некоторого $2f$-регулярного графа, последняя теорема влечет следующий результат.

Теорема 5.5.7. *Существует абсолютная константа $c > 0$, такая, что для всякого неориентированного d-регулярного графа G*

$$\mathrm{la}(G) \leqslant \frac{d}{2} + cd^{3/4}(\ln d)^{1/2}.$$

5.6. ЛАТИНСКИЕ ТРАНСВЕРСАЛИ

Обсуждая доказательство локальной леммы, мы заметили, что предположение взаимной независимости может быть заменено более слабым предположением о достаточно малой условной вероятности каждого события, при условии взаимной «невстречаемости» произвольного множества событий, каждое из которых несмежно с ним в орграфе зависимости. В этом разделе мы покажем, как эта модифицированная версия леммы была применена в статье Эрдёша и Спенсера [Erdős and Spencer (1991)]. Пусть $A = (a_{ij})$ — матрица размера $n \times n$ скажем, из целых чисел. Перестановка π называется *латинской трансверсалью* (для A), если все элементы $a_{i\pi(i)}$ $(1 \leqslant i \leqslant n)$ различны между собой.

Теорема 5.6.1. *Пусть $k \leqslant (n-1)/(4e)$ и никакое целое число не встречается в качестве элемента матрицы A более k раз. Тогда A обладает латинской трансверсалью.*

Доказательство. Пусть π — случайная перестановка на множестве $\{1, 2, \ldots, n\}$, выбранная в соответствии с равномерным распределением из всех возможных $n!$ перестановок. Обозначим через T множество всех упорядоченных четверок (i, j, i', j'), таких, что $i < i', j \neq j'$ и $a_{ij} = a_{i'j'}$. Для каждого $(i, j, i', j') \in T$, пусть $A_{iji'j'}$ означает событие «$\pi(i) = j$ и $\pi(i') = j'$». Существование латинской трансверсали эквивалентно утверждению, что с положительной вероятностью ни одно из этих событий не происходит. Определим симметричный орграф (т. е. просто граф) G на вершинах множества T, полагая, что вершина (i, j, i', j') смежна с (p, q, p', q') тогда и только тогда, когда либо $\{i, i'\} \cap \{p, p'\} \neq \varnothing$, либо $\{j, j'\} \cap \{q, q'\} \neq \varnothing$. Таким образом, эти две четверки несмежны тогда и только тогда, когда четыре клетки $(i, j), (i', j'), (p, q)$ и (p', q') встречаются в четырех различных строках и в четырех различных столбцах матрицы A. Максимальная степень графа G меньше $4nk$. В самом деле, для данной четверки $(i, j, i', j') \in T$ существует не более $4n$ способов выбрать пару (s, t) такую, что или $s \in \{i, i'\}$ или $t \in \{j, j'\}$, и для каждой пары (s, t) существует не более k способов выбора $(s', t') \neq (s, t)$ с $a_{st} = a_{s't'}$. Всякая такая четверка (s, t, s', t') может

быть однозначно представлена в виде (p, q, p', q') при $p < p'$. Поскольку $e \cdot 4nk \cdot \frac{1}{n(n-1)} \leqslant 1$, требуемый результат вытекает из упомянутого выше усиления симметричной версии локальной леммы, если мы покажем, что

$$\Pr\left[A_{iji'j'} \Big| \bigwedge_S \overline{A}_{pqp'q'}\right] \leqslant \frac{1}{n(n-1)} \tag{5.9}$$

для любой четверки $(i, j, i', j') \in T$ и любого множества S элементов множества T, не смежных в G с (i, j, i', j'). В силу симметрии мы можем предположить, что $i = j = 1, i' = j' = 2$, и что, следовательно, ни одна из p или q не равна ни 1, ни 2. Назовем перестановку π *хорошей*, если $\bigwedge_S \overline{A}_{pqp'q'}$. Пусть S_{ij} обозначает множество всех хороших перестановок π, удовлетворяющих условию $\pi(1) = i$ и $\pi(2) = j$. Мы утверждаем, что $\big|S_{12}\big| \leqslant \big|S_{ij}\big|$ для всех $i \neq j$. В самом деле, предположим сначала, что $i, j > 2$. Для каждой хорошей $\pi \in S_{12}$ определим перестановку π^* следующим образом. Предположим, что $\pi(x) = i$, $\pi(y) = j$. Тогда определим $\pi^*(1) = i, \pi^*(2) = j, \pi^*(x) = 1, \pi^*(y) = 2$ и $\pi^*(t) = \pi(t)$ для всех $t \neq 1, 2, x, y$. Можно легко проверить, что π^* — хорошая, так как ни одна из пар $(1, i), (2, j), (x, 1), (y, 2)$ не является частью ни одной $(p, q, p', q') \in S$. Поэтому $\pi^* \in S_{ij}$, и так как отображение $\pi \to \pi^*$ инъективно, то $\big|S_{12}\big| \leqslant \big|S_{ij}\big|$, как и обещано. Аналогично можно определить инъективное отображение, показывающее, что $\big|S_{12}\big| \leqslant \big|S_{ij}\big|$, даже когда $\{i, j\} \cap \{1, 2\} \neq \varnothing$. Отсюда следует, что $\Pr[A_{1122} \wedge \bigwedge_S \overline{A}_{pqp'q'}] \leqslant \Pr[A_{1i2j} \wedge \bigwedge_S \overline{A}_{pqp'q'}]$ для всех $i \neq j$, а, значит, $\Pr\left[A_{1122} \big| \bigwedge_S \overline{A}_{pqp'q'}\right] \leqslant \frac{1}{n(n-1)}$. В силу симметрии отсюда следует неравенство (5.9). Тем самым доказательство завершено. ■

5.7. АЛГОРИТМИЧЕСКИЙ АСПЕКТ

Когда вероятностный метод применяется с целью доказать, что определенное событие имеет место с высокой вероятностью, это часто дает эффективный детерминированный алгоритм или, по крайней мере, рандомизированный алгоритм решения соответствующей проблемы.

Применение локальной леммы Ловаса часто дает нам возможность доказать, что данное событие происходит с положительной вероятностью, хотя эта вероятность может быть экспоненциально малой от размерности задачи. Поэтому неясно, можно ли извлечь из таких доказательств полиномиальные алгоритмы решения соответствующих алгоритмических проблем.

Довольно долго не существовало методов преобразования доказательств какой-либо из обсуждаемых в этой главе задач в эффективный алгоритм. В 1991 г. Д. Бек предложил метод, позволяющий решать эту проблему для некоторых таких задач с небольшими потерями в константах. Он продемонстрировал в работе [Beck (1991)] свой метод на примере задачи о 2-раскраске

гиперграфа. Для простоты мы опишем здесь только случай ребер фиксированного размера, когда каждое ребро пересекается с фиксированным числом других ребер.

Пусть n, d — фиксированные натуральные числа. Под (n,d)-задачей подразумевается следующее. Пусть заданы множества $A_1, \dots, A_N \subseteq \Omega$, такие, что $|A_i| = n$, и при этом каждое из A_i пересекается не больше чем с d другими A_j. Требуется найти 2-раскраску множества Ω, при которой не существует монохроматических A_i. Теорема 5.2.1 гарантирует разрешимость нашей задачи при $e(d+1) < 2^{n-1}$. Вопрос: можем ли мы найти раскраску за полиномиальное (от N при фиксированных n, d) время? Бек дал утвердительный ответ на этот вопрос при некоторых более сильных предположениях. Предположим, что $\Omega = \{1, \dots, m\}$, $m \leqslant Nn$, а начальная структура данных состоит из списка элементов множеств A_i и списка, указывающего для каждого элемента j те i, для которых $j \in A_i$. Обозначим через G граф зависимости с A_i в качестве вершин и парами пересекающихся A_i, A_j в качестве ребер.

Теорема 5.7.1. *Пусть n, d таковы, что для величины $D = d(d-1)^3$ существует разложение $n = n_1 + n_2 + n_3$, где*

$$16D(1+d) < 2^{n_1},$$
$$16D(1+d) < 2^{n_2},$$
$$2e(1+d) < 2^{n_3}.$$

Тогда существует рандомизированный алгоритм с математическим ожиданием времени работы $O(N(\ln N)^c)$ для (n,d)-задачи, где c — константа (зависящая только от n и d).

Для фиксированного $\varepsilon < 1/11$ заметим, что упомянутые выше условия выполняются для достаточно больших n, если $d < 2^{n\varepsilon}$ и $n_1 = n_2 \sim 5n/11$, $n_3 \sim n/11$. Подчеркнем еще раз, что анализ алгоритма проводится для *фиксированных* n, d и N, стремящегося к бесконечности, хотя рассуждения могут быть распространены на случай, когда указанные переменные не являются фиксированными.

Бек указал детерминированный алгоритм решения (n,d)-задачи. Рандомизированный алгоритм, который здесь предлагается, может быть дерандомизирован с использованием приемов из гл. 15. Время работы остается полиномиальным, и, по-видимому, не большим $N^{1+o(1)}$. Более того, алгоритм может быть распараллелен с использованием техники из гл. 15 и некоторой модификации алгоритма.

Доказательство. Первый проход. Во время этого прохода точки будут красными, синими, нераскрашенными или резервными. Мы движемся по вершинам $j \in \Omega$, раскрашивая их красным или синим случайно, выбирая цвет подбрасыванием правильной монеты. После того, как некоторая вершина j раскрашена, мы проверяем все A_i, содержащие j. Если A_i в этот момент имеет n_1 точек одного цвета и ни одной точки другого цвета, мы назовем A_i *опасной*.

Все нераскрашенные $k \in A_i$ объявляются *резервными*. Когда резервные точки k впоследствии встречаются при раскраске, они не окрашиваются, а просто пропускаются. В результате первого прохода точки становятся красными, синими или резервными. Мы скажем, что множество A_i *выживает*, если оно не состоит только из красных или только из синих точек. Обозначим через S (случайное) множество выживших точек из G.

Утверждение 5.7.2. *Почти наверное все компоненты C из $G|_S$ имеют размер $O(\ln N)$.*

Доказательство. Множество $A_i \in S$ может быть опасным или, возможно, многие из его точек были резервными ввиду того, что соседствующие с ними (в G) множества оказались опасными. Вероятность того, что некоторое A_i становится опасным, не превышает 2^{1-n_1}, так как должно произойти n_1 подбрасываний монеты (определяющих цвета точек $j \in A_i$) с одним и тем же исходом. (Мы имеем только неравенство, так как, к тому же, все эти n_1 точек из A_i должны быть окрашены, а не стать резервными.) Пусть V — независимое множество в G, т. е. $A_i \in V$ попарно не пересекаются. Тогда вероятность того, что все $A_i \in V$ становятся опасными, не превышает $(2^{1-n_1})^{|V|}$, поскольку бросания монеты проводятся для непересекающихся множеств. Пусть теперь $V \subseteq G$ таково, что все расстояния между $A_i \in V$ не меньше 4. (Расстояние — это длина кратчайшей цепи между вершинами в G). Мы утверждаем, что

$$\Pr[V \subseteq S] \leqslant (d+1)^{|V|}(2^{1-n_1})^{|V|}.$$

Это выполнено потому, что для каждого $A_i \in V$ существуют не более $d+1$ способов выбора опасного соседа $A_{i'}$, а следовательно, не более $(d+1)^{|V|}$ способов выбрать $A_{i'}$. Поскольку A_i, находящееся на расстоянии не меньшем четырех от $A_{i'}$, не может быть смежным с ним, вероятность того, что они все опасны, не больше $(2^{1-n_1})^{|V|}$, как и утверждалось.

Назовем $T \subseteq G$ 4-*деревом*, если $A_i \in T$ таковы, что расстояния между ними в графе G не меньше четырех, и при этом, соединив ребром каждые два множества $A_i, A_j \in T$, находящиеся на расстоянии четыре, мы получим связный граф. Мы оценим сначала число 4-деревьев размера u. Граф «дистанции четыре», определенный на T, должен содержать дерево. Существует не больше чем 4^j деревьев (с точностью до изоморфизма) на j вершинах. Зафиксируем одно из них. Мы можем пометить вершины дерева числами $1, \ldots, u$ так, что каждое $j > 1$ смежно с некоторой $i < j$. Теперь рассмотрим число семейств $(A^1, \ldots, A^u)$, для которых граф дистанции четыре соответствует этому дереву. Существует N способов выбора A^1. При выбранном A^i для всех $i < j$ множество A^j должно быть на расстоянии четыре от A^i в G. Имеется не более D таких точек. Следовательно, число 4-деревьев размера u не больше $4^u N D^{u-1} < N(4D)^u$. Для заданного 4-дерева T мы уже доказали, что $\Pr[T \subseteq S] \leqslant [(d+1)2^{1-n_1}]^u$. Значит, математическое ожидание числа 4-деревьев $T \subseteq S$ не больше

$$N\left[8D(d+1)2^{-n_1}\right]^u.$$

Поскольку содержимое скобок по предположению меньше $1/2$, при $u = c_1 \ln N$ все выражение есть $o(1)$. Итак, почти наверное $G|_S$ не содержит 4-деревьев размера, большего чем $c_1 \ln N$. В действительности, мы хотим ограничить размер компонент C графа $G|_S$. Максимальное 4-дерево T в компоненте C обладает тем свойством, что каждое $A_i \in C$ лежит в дереве $A_j \in T$. Имеется меньше чем d^3 (константа) множеств A_i с тремя A_j, следовательно, $c_1 \ln N \geqslant |T| \geqslant |C| d^{-3}$, и поэтому (поскольку d — константа)

$$|C| \leqslant c_2 \ln N,$$

что и требовалось. ■

Если после первого прохода остаются компоненты размера, большего чем $c_2 \ln N$, мы просто полностью повторяем всю процедуру. В ожидаемое *линейное* время первый проход оказывается успешным. Точки, окрашенные красным и синим, отныне зафиксированы (в своем статусе). Множества A_i, в которых встречаются как красные, так и синие точки теперь игнорируются. Для каждого выжившего A_i зафиксируем подмножество B_i его $n - n_1$ резервных точек. Теперь достаточно раскрасить резервные точки так, чтобы ни одно из B_i не было монохроматическим. Множество B_i разбито на компоненты размера $O(\ln N)$, поэтому достаточно раскрасить каждую компоненту в отдельности. При втором проходе мы применяем метод первого прохода к каждой компоненте множества B_i. Назовем множество B_i *опасным*, если оно содержит n_2 точек одного цвета и не содержит точек другого цвета. Математическое ожидание времени, требуемого при втором проходе для раскраски компоненты размера M, равно $O(M)$. Следовательно, математическое ожидание времени, требуемого для раскраски всех компонент, равно $O(N)$. (Для того чтобы такая раскраска была успешной, мы должны потребовать, чтобы компонента размера M разбивалась на компоненты размера не больше $c_2 \ln M$. Во избежание тривиальностей, при $M < \ln \ln N$ мы не проводим второй проход для соответствующей компоненты). После осуществления второго прохода (все еще за линейное время!) существует семейство дважды выживших множеств $C_i \subset B_i \subset A_i$ размера n_3, максимальный размер компоненты которых не больше $O(\ln \ln N)$.

Нам все еще нужно раскрасить эти $O(N)$ компонент из множеств размера n_3, каждая компонента которых имеет размер $O(\ln \ln N)$. В силу локальной леммы (или прямо по теореме 5.2.1), каждая из этих компонент может быть 2-раскрашиваемой. Мы найдем теперь 2-раскраску *с помощью грубой силы!* Перебирая все 2-раскраски компонент размера M, мы тратим время $O(M 2^M)$, которое в нашем случае равно $O((\ln N)^c)$. Проводя это для каждой компоненты, мы тратим время $O(N(\ln N)^c)$. Раскраска завершена. ■

Заметим, что при чуть больших ограничениях на n, d с помощью третьего прохода можно было бы добиться общей затраты времени порядка $O(N(\ln \ln N)^c)$. Заметим также, что подобный прием может быть применен для преобразования некоторых других приложений локальной леммы в эффективные алгоритмы.

5.8. УПРАЖНЕНИЯ

1*. Доказать, что для любого целого $d > 1$ существует конечное $c(d)$, такое, что ребра любого двудольного графа с максимальной степенью вершины d, в котором каждый цикл имеет не менее $c(d)$ ребер, могут быть раскрашены $d+1$ красками так, что не существует двух смежных ребер одного цвета, и не существует 2-раскрашенных циклов.

2*. Доказать, что для каждого $\varepsilon > 0$ существует $l_0 = l_0(\varepsilon)$ и бесконечная двоичная последовательность $a_1, a_2, a_3, \ldots$, $a_i \in \{0,1\}$, такая, что для всякого числа $l > l_0$ и любого номера $i \geqslant 1$ два бинарных вектора $u = (a_i, a_{i+1}, \ldots, a_{i+l-1})$ и $v = (a_{i+l}, a_{i+l+1}, \ldots, a_{i+2l-1})$ отличаются не менее чем в $(\frac{1}{2} - \varepsilon)l$ координатах.

3. Пусть $G = (V, E)$ — простой граф, и пусть каждой вершине $v \in V$ сопоставлено множество $S(v)$ цветов, такое, что $|S(v)| \leqslant 10d|$, где $d \geqslant 1$. Предположим вдобавок, что для любой вершины $v \in V$ и цвета $c \in S(v)$ существует не более d соседей u вершины v, таких, что $c \in S(u)$. Доказать, что существует правильная раскраска графа G, окрашивающая каждую вершину v в цвет из $S(v)$.

4. Пусть $G = (V, E)$ — цикл длины $4n$ и $V = V_1 \cup V_2 \ldots \cup V_n$ — разбиение его $4n$ вершин на n попарно непересекающихся подмножеств, каждое из которых имеет мощность 4. Верно ли, что обязательно существует независимое множество графа G, содержащее ровно одну вершину из каждого V_i ? (Доказать или привести контрпример.)

5*. Доказать, что существует абсолютная константа $c > 0$, такая, что для любого k существует множество S_k, состоящее по меньшей мере из $ck \ln k$ целых чисел, такое, что для любой раскраски целых чисел в k цветов существует целое x, для которого множество $x + S_k$ не пересекает все классы цветов.

ВЕРОЯТНОСТНЫЙ ВЗГЛЯД:

Ориентированные циклы

Пусть $D = (V, E)$ — простой ориентированный граф с минимальной степенью исхода δ и максимальной степенью захода Δ.

Теорема [Alon and Linial (1989)]. *Если* $e(\Delta\delta + 1)\left(1 - \frac{1}{k}\right)^{\delta} < 1$, *то граф* D *содержит ориентированный простой цикл длины* $0 (\mathrm{mod}\, k)$.

Доказательство. Очевидно, можно предполагать, что степень исхода каждой вершины в точности равна δ, так как иначе можно рассмотреть подграф графа D с таким свойством.

Пусть $f : V \to \{0, 1, \ldots, k-1\}$ — случайная раскраска множества V, полученная путем выбора для каждой вершины $v \in V$ независимо и равновероятно значения $f(v) \in \{0, \ldots, k-1\}$. Для всякой вершины $v \in V$ обозначим через A_v событие «не существует вершины $u \in V$, для которой $(v, u) \in E$ и $f(u) \equiv (f(v) + 1) (\mathrm{mod}\, k)$». Ясно, что $\Pr[A_v] = \left(1 - \frac{1}{k}\right)^{\delta}$. Нетрудно проверить, что каждое A_v взаимно независимо со всеми событиями A_u, кроме тех, которые удовлетворяют условию

$$N^+(v) \cap \left(u \bigcup N^+(u)\right) \neq \varnothing,$$

где $N^+(v) = \{w \in V : (v, w) \in E\}$. Число таких u не превосходит $\Delta\delta$ и, значит, по нашему предположению и локальной лемме (следствие 5.1.2) вероятность $\Pr\left[\bigwedge_{v \in V} \overline{A}_v\right] > 0$. Следовательно, существует $f : V \to \{0, 1, \ldots, k-1\}$, такое, что для каждого $v \in V$ существует вершина $u \in V$, для которой

$$(v, u) \in E \quad \text{и} \quad f(u) \equiv (f(v) + 1) \,(\mathrm{mod}\, k). \tag{1}$$

Начав с произвольной вершины $v = v_0 \in V$ и применяя раз за разом (1), мы получим последовательность $v_0, v_1, v_2, \ldots$ вершин из D, такую, что $(v_i, v_{i+1}) \in E$ и $f(v_{i+1}) \equiv (f(v_i) + 1) \,(\mathrm{mod}\, k)$ для всех $i \geqslant 0$. Пусть j — минимальное целое, для которого существует $\ell < j$, такое, что $v_\ell = v_j$. Последовательность $v_\ell v_{\ell+1} v_{\ell+2} \cdots v_j = v_\ell$ представляет собой простой ориентированный цикл в D, длина которого делится на k. ■

6

Корреляционные неравенства

Думай, Буч, думай, — это именно то, в чем ты силен.

Роберт Редфорд Полу Ньюмену,
Буч Кассиди и Санденс Кид

Пусть $G = (V, E)$ — случайный граф на множестве вершин $V = \{1, 2, \ldots, n\}$, в котором произвольная пара $i, j \in V, i \neq j$ является ребром независимо (от других) с вероятностью p, где $0 < p < 1$. Обозначим через H событие «граф G является гамильтоновым», а через P — событие «граф G является планарным». Предположим, мы хотим сравнить две величины $\Pr[P \wedge H]$ и $\Pr[P] \cdot \Pr[H]$. Интуитивно кажется, что знание того, что граф G — гамильтонов, предполагает, что в нем много ребер и, значит, G скорее всего не является планарным. Следовательно, представляется естественным ожидать, что выполнено неравенство $\Pr[P|H] \leqslant \Pr[P]$, влекущее

$$\Pr[P \wedge H] \leqslant \Pr[H] \cdot \Pr[P].$$

Это неравенство, которое и в самом деле верно, является специальным случаем FKG-неравенства, принадлежащего Фортуину, Кастелейну и Джинибру [Fortuin, Kasteleyn and Ginibre (1971)]. В данной главе мы приводим доказательство этого неравенства и некоторых родственных результатов, которые имеют отношение к корреляции между определенными событиями в вероятностных пространствах. Доказательства всех этих результатов достаточно простые, и, тем не менее, они порождают много интересных следствий. Первое неравенство такого типа принадлежит Харрису [Harris (1960)]. Результат, наиболее близкий к тем, что рассматриваются здесь, — это лемма Клейтмана [Kleitman (1966b)], утверждающая, что если $\mathcal{A}$ и $\mathcal{B}$ — два *монотонно убывающих* семейства подмножеств множества $\{1, 2, \ldots, n\}$ (т. е., $A \in \mathcal{A}$ и $A' \subseteq A \Rightarrow A' \in \mathcal{A}$ и, аналогично, $B \in \mathcal{B}$ и $B' \subseteq B \Rightarrow B' \in \mathcal{B}$), то

$$|\mathcal{A} \cap \mathcal{B}| \cdot 2^n \geqslant |\mathcal{A}| \cdot |\mathcal{B}|.$$

За этой леммой последовала масса усилений и обобщений, пока Альсведе и Дайкин [Ahlswede and Daykin (1978)] не получили весьма общий результат,

из которого все эти обобщения вытекают. В следующем разделе мы приведем этот результат и его доказательство. Некоторые из многочисленных его приложений обсуждаются в оставшейся части главы.

6.1. ТЕОРЕМА О ЧЕТЫРЕХ ФУНКЦИЯХ АЛЬСВЕДЕ И ДАЙКИНА

Для $n \geqslant 1$ положим $N = \{1, 2, \ldots, n\}$. Через $P(N)$ обозначим семейство всех подмножеств множества N, а через $\mathbb{R}^+$ — множество неотрицательных действительных чисел. Для функции $\varphi : P(N) \to \mathbb{R}^+$ и семейства $\mathcal{A}$ подмножеств множества N положим $\varphi(\mathcal{A}) = \sum_{A \in \mathcal{A}} \varphi(A)$. Для двух семейств $\mathcal{A}$ и $\mathcal{B}$ подмножеств множества N определим $\mathcal{A} \cup \mathcal{B} = \{A \cup B : A \in \mathcal{A}, B \in \mathcal{B}\}$ и $\mathcal{A} \cap \mathcal{B} = \{A \cap B : A \in \mathcal{A}, B \in \mathcal{B}\}$.

Теорема 6.1.1 (теорема о четырех функциях). *Пусть четыре функции $\alpha, \beta, \gamma, \delta$: $P(N) \to \mathbb{R}^+$ отображают семейство всех подмножеств множества N во множество неотрицательных действительных чисел. Если для любых двух подмножеств $A, B \subseteq N$ выполняется неравенство*

$$\alpha(A)\beta(B) \leqslant \gamma(A \cup B)\delta(A \cap B), \tag{6.1}$$

то для любых двух семейств подмножеств $\mathcal{A}, \mathcal{B} \subseteq P(N)$

$$\alpha(\mathcal{A})\beta(\mathcal{B}) \leqslant \gamma(\mathcal{A} \cup \mathcal{B})\delta(\mathcal{A} \cap \mathcal{B}). \tag{6.2}$$

Доказательство. Сначала модифицируем четыре функции $\alpha, \beta, \gamma, \delta$, положив $\alpha(A) = 0$ для всех $A \notin \mathcal{A}$, $\beta(B) = 0$ для всех $B \notin \mathcal{B}$, $\gamma(C) = 0$ для всех $C \notin \mathcal{A} \cup \mathcal{B}$ и $\delta(D) = 0$ для всех $D \notin \mathcal{A} \cap \mathcal{B}$. Ясно, что соотношение (6.1) остается справедливым для модифицированных функций, и в неравенстве (6.2) мы можем считать теперь, что $\mathcal{A} = \mathcal{B} = \mathcal{A} \cup \mathcal{B} = \mathcal{A} \cap \mathcal{B} = P(N)$.

Докажем неравенство (6.2) индукцией по n. Единственный случай, когда требуются некоторые вычисления, это $n = 1$. Тогда $P(N) = \{\varnothing, N\}$. Для всякой функции $\varphi \in \{\alpha, \beta, \gamma, \delta\}$ определим $\varphi_0 = \varphi(\varnothing)$ и $\varphi_1 = \varphi(N)$. В силу неравенства (6.1) имеем

$$\begin{aligned} \alpha_0\beta_0 &\leqslant \gamma_0\delta_0, \\ \alpha_0\beta_1 &\leqslant \gamma_1\delta_0, \\ \alpha_1\beta_0 &\leqslant \gamma_1\delta_0, \\ \alpha_1\beta_1 &\leqslant \gamma_1\delta_1. \end{aligned} \tag{6.3}$$

Согласно замечанию из предыдущего раздела мы должны только доказать неравенство (6.2), где $\mathcal{A} = \mathcal{B} = P(N)$, т. е. доказать, что

$$(\alpha_0 + \alpha_1)(\beta_0 + \beta_1) \leqslant (\gamma_0 + \gamma_1)(\delta_0 + \delta_1). \tag{6.4}$$

Если $\gamma_1 = 0$ или $\delta_0 = 0$, утверждение сразу следует из условий (6.3). В противном случае в силу соотношений (6.3) выполнены неравенства $\gamma_0 \geqslant \frac{\alpha_0\beta_0}{\delta_0}$ и $\delta_1 \geqslant \frac{\alpha_1\beta_1}{\gamma_1}$. Значит, достаточно показать, что $\left(\frac{\alpha_0\beta_0}{\delta_0} + \gamma_1\right)\left(\delta_0 + \frac{\alpha_1\beta_1}{\gamma_1}\right) \geqslant (\alpha_0 + \alpha_1)(\beta_0 + \beta_1)$, или, что то же, $(\alpha_0\beta_0 + \gamma_1\delta_0)(\delta_0\gamma_1 + \alpha_1\beta_1) \geqslant (\alpha_0 + \alpha_1)(\beta_0 + \beta_1)\delta_0\gamma_1$. Последнее эквивалентно

неравенству

$$(\gamma_1\delta_0 - \alpha_0\beta_1)(\gamma_1\delta_0 - \alpha_1\beta_0) \geqslant 0,$$

которое вытекает из соотношений (6.3), поскольку оба сомножителя из левой части неотрицательны. Это завершает доказательство для случая $n = 1$.

Пусть теперь утверждение справедливо при $n-1$. Докажем его для $n \geqslant 2$. Положим $N' = N\backslash\{n\}$ и определим для каждой функции $\varphi \in \{\alpha, \beta, \gamma, \delta\}$ и любых $A \subseteq N'$ функцию $\varphi'(A) = \varphi(A) + \varphi(A \cup \{n\})$. Ясно, что для каждой $\varphi \in \{\alpha, \beta, \gamma, \delta\}$ выполнено равенство $\varphi'(P(N')) = \varphi(P(N))$. Поэтому требуемое неравенство (6.3) будет следовать из предположения индукции, примененного к функциям $\alpha', \beta', \gamma', \delta' : P(N') \to \mathbb{R}^+$. Однако чтобы применить это предположение, необходимо убедиться, что эти новые функции удовлетворяют условиям теоремы 6.1.1 на множестве N', т. е. что для любых $A', B' \subseteq N'$ выполнено неравенство

$$\alpha'(A')\beta'(B') \leqslant \gamma'(A' \cup B')\delta'(A' \cap B'). \tag{6.5}$$

Оно легко вытекает из случая $n = 1$, который уже рассмотрен. В самом деле, пусть T является одноэлементным множеством. Определим $\overline{\alpha}(\varnothing) = \alpha(A')$, $\overline{\alpha}(T) = \alpha(A' \cup \{n\})$, $\overline{\beta}(\varnothing) = \beta(B')$, $\overline{\beta}(T) = \beta(B' \cup \{n\})$, $\overline{\gamma}(\varnothing) = \gamma(A' \cup B')$, $\overline{\gamma}(T) = \gamma(A' \cup B' \cup \{n\})$ и $\overline{\delta}(\varnothing) = \delta(A' \cap B')$, $\overline{\delta}(T) = \delta((A' \cap B') \cup \{n\})$. По предположению (6.1), для всех $S, R \subseteq T$ выполнено соотношение $\overline{\alpha}(S)\overline{\beta}(R) \leqslant \overline{\gamma}(S \cup R)\overline{\delta}(S \cap R)$, и, следовательно, в силу доказанного для $n = 1$, имеем

$$\alpha'(A')\beta'(B') = \overline{\alpha}(P(T))\overline{\beta}(P(T)) \leqslant \overline{\gamma}(P(T))\overline{\delta}(P(T)) = \gamma'(A' \cup B')\delta'(A' \cap B'),$$

которое является требуемым неравенством (6.5). Следовательно, неравенство (6.2) выполняется, что и завершает доказательство. ■

Теорема Альсведе—Дайкина может быть обобщена на произвольные конечные дистрибутивные решетки. *Решетка* — это частично упорядоченное множество, в котором любые два элемента x и y имеют единственную верхнюю грань, обозначаемую через $x \vee y$ и называемую *объединением* x и y, а также единственную нижнюю грань, обозначаемую через $x \wedge y$ и называемую *пересечением* x и y. Решетка L является дистрибутивной, если для всех $x, y, z \in L$ выполнено равенство

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z),$$

или, что эквивалентно, если для всех $x, y, z \in L$ верно соотношение

$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z).$$

Для двух множеств $X, Y \subseteq L$ определим

$$X \vee Y = \{x \vee y : x \in X, y \in Y\},$$

и

$$X \wedge Y = \{x \wedge y : x \in X, y \in Y\}.$$

Любое подмножество L множества $P(N)$, где $N = \{1, 2, \ldots, n\}$, упорядоченное по включению и замкнутое относительно операций объединения и пересечения, является *дистрибутивной* решеткой. Здесь объединение двух элементов $A, B \in L$ есть просто объединение множеств $A \cup B$, а их пересечение есть обычное пересечение множеств $A \cap B$. Нечто более удивительное (но легко проверяемое) состоит в том, что для каждой конечной дистрибутивной решетки L существует n, такое, что L изоморфна некоторой подрешетке решетки $P(\{1, 2, \ldots, n\})$. (Чтобы в этом убедиться, назовем элемент $x \in L$ *объединительно-неизбыточным*, если из $x = y \vee z$ вытекает, что $x = y$ или $x = z$. Пусть $x_1, x_2, \ldots, x_n$ — множество всех объединительно-неизбыточных элементов в L. Каждому элементу $x \in L$ поставим в соответствие множество $A = A(x) \subseteq N$, где $x = \bigvee\limits_{i \in A} x_i$ и $\{x_i : i \in A\}$ — все объединительно-неизбыточные y, такие,что $y \leqslant x$. Отображение $x \to A(x)$ является искомым вложением.) Этот факт позволяет нам обобщить теорему 6.1.1 на произвольные конечные дистрибутивные решетки следующим образом.

Следствие 6.1.2. *Пусть L — конечная дистрибутивная решетка, а α, β, γ и δ — четыре функции из L в $\mathbb{R}^+$. Если*

$$\alpha(x)\beta(y) \leqslant \gamma(x \vee y)\delta(x \wedge y)$$

для всех $x, y \in L$, то для всех $X, Y \subseteq L$ выполняется неравенство

$$\alpha(X)\beta(Y) \leqslant \gamma(X \vee Y)\delta(X \wedge Y).$$

Простейшим в последнем следствии является случай, когда все четыре функции α, β, γ и δ равны 1. Он формулируется следующим образом.

Следствие 6.1.3. *Пусть L — конечная дистрибутивная решетка и $X, Y \subseteq L$. Тогда*

$$|X| \cdot |Y| \leqslant |X \vee Y| \cdot |X \wedge Y|.$$

Завершая этот раздел, мы представим очень простое следствие из последнего утверждения, которое впервые было доказано в работе [Marica and Schonheim (1969)].

Следствие 6.1.4. *Пусть X — семейство всех подмножеств конечного множества. Положим*

$$X \backslash X = \{F \backslash F' : F, F' \in X\}.$$

Тогда $|X \backslash X| \geqslant |X|$.

Доказательство. Пусть L — дистрибутивная решетка всех подмножеств множества N. Применяя следствие 6.1.3 к X и $Y = \{N \backslash F : F \in X\}$, получаем

$$|X|^2 = |X| \cdot |Y| \leqslant |X \cup Y| \cdot |X \cap Y| = |X \backslash X|^2,$$

откуда и вытекает требуемое утверждение. ■

6.2. FKG-НЕРАВЕНСТВО

Функция $\mu : L \to \mathbb{R}^+$, где L — конечная дистрибутивная решетка, называется *log-супермодулярной*, если

$$\mu(x)\mu(y) \leqslant \mu(x \vee y)\mu(x \wedge y)$$

для всех $x, y \in L$. Функция $f : L \to \mathbb{R}^+$ называется возрастающей, если $f(x) \leqslant f(y)$ при $x \leqslant y$, и убывающей, если $f(x) \geqslant f(y)$ при $x \leqslant y$.

В работе [Fortuin et al. (1971)], мотивированной одной задачей из статистической механики, доказан следующий полезный результат, ставший известным как FKG-неравенство.

Теорема 6.2.1 (FKG-неравенство). *Пусть L — конечная дистрибутивная решетка, и пусть $\mu : L \to \mathbb{R}^+$ — log-супермодулярная функция. Тогда для любых двух возрастающих функций $f, g : L \to \mathbb{R}^+$ имеем*

$$\left(\sum_{x\in L} \mu(x)f(x)\right) \cdot \left(\sum_{x\in L} \mu(x)g(x)\right) \leqslant \left(\sum_{x\in L} \mu(x)f(x)g(x)\right) \cdot \left(\sum_{x\in L} \mu(x)\right). \quad (6.6)$$

Доказательство. Определим четыре функции $\alpha, \beta, \gamma, \delta : L \to \mathbb{R}^+$ следующим образом. Для всякого $x \in L$ положим

$$\alpha(x) = \mu(x)f(x), \qquad \beta(x) = \mu(x)g(x),$$
$$\gamma(x) = \mu(x)f(x)g(x), \qquad \delta(x) = \mu(x).$$

Мы утверждаем, что эти функции удовлетворяют условию теоремы Альсведе—Дайкина, сформулированной в следствии 6.1.2. В самом деле, если $x, y \in L$, то в силу супермодулярности функции μ и возрастания функций f и g имеем

$$\alpha(x)\beta(y) = \mu(x)f(x)\mu(y)g(y) \leqslant \mu(x \vee y)f(x)g(y)\mu(x \wedge y) \leqslant$$
$$\leqslant \mu(x \vee y)f(x \vee y)g(x \vee y)\mu(x \wedge y) = \gamma(x \vee y)\delta(x \wedge y).$$

Значит, в силу следствия 6.1.2 (с $X = Y = L$) выполнено неравенство

$$\alpha(L)\beta(L) \leqslant \gamma(L)\delta(L),$$

что и требовалось доказать. ■

Заметим, что утверждение теоремы 6.2.1 остается справедливым и в том случае, когда обе функции f и g являются убывающими (просто заменяем γ и δ друг на друга в доказательстве). В случае, когда f — возрастающая, а g — убывающая (или наоборот), имеет место противоположное неравенство:

$$\left(\sum_{x\in L} \mu(x)f(x)\right)\left(\sum_{x\in L} \mu(x)g(x)\right) \geqslant \left(\sum_{x\in L} \mu(x)f(x)g(x)\right)\left(\sum_{x\in L} \mu(x)\right).$$

Чтобы это доказать, просто применим теорему 6.2.1 к двум возрастающим функциям $f(x)$ и $k - g(x)$, где k есть константа $\max_{x\in L} g(x)$. (Эта константа необходима, чтобы гарантировать выполнение неравенства $k - g(x) \geqslant 0$ для всех $x \in L$.)

Полезно рассмотреть μ как меру на L. Предполагая, что μ не равна тождественно нулю, мы можем определить для каждой функции $f : L \to \mathbb{R}^+$

ее математическое ожидание

$$< f >= \frac{\sum_{x\in L} f(x)\mu(x)}{\sum_{x\in L} \mu(x)}.$$

В этих терминах FKG-неравенство утверждает, что, если μ является log-супермодулярной, а $f, g : L \to \mathbb{R}^+$ обе являются возрастающими или убывающими, то

$$< fg > \geqslant < f > \cdot < g > .$$

Аналогично, если f является возрастающей, а g — убывающей (или наоборот), то

$$< fg > \leqslant < f > \cdot < g > .$$

Эта формулировка ясно демонстрирует вероятностную природу неравенства. Некоторые из его интересных следствий представлены в оставшейся части главы.

6.3. МОНОТОННЫЕ СВОЙСТВА

Напомним, что семейство $\mathcal{A}$ подмножеств множества $N = \{1, 2, \ldots, n\}$ является *монотонно убывающим*, если $A \in \mathcal{A}$ и $A' \subseteq A \Rightarrow A' \in \mathcal{A}$. Аналогично, оно является *монотонно возрастающим*, если $A \in \mathcal{A}$ и $A \subseteq A' \Rightarrow A' \in \mathcal{A}$. Рассматривая показательное семейство $P(N)$ всех подмножеств множества N как симметричное вероятностное пространство, естественно определить вероятность множества $\mathcal{A}$ равенством

$$\Pr[\mathcal{A}] = \frac{|\mathcal{A}|}{2^n}.$$

Итак, $\Pr[\mathcal{A}]$ есть просто вероятность того, что случайно выбранное подмножество множества N лежит в $\mathcal{A}$.

Лемма Клейтмана, служившая стартовой точкой всех корреляционных неравенств, рассматриваемых в этой главе, формулируется следующим образом.

Предложение 6.3.1. *Пусть $\mathcal{A}$ и $\mathcal{B}$ — два монотонно возрастающих семейства подмножеств множества $N = \{1, 2, \ldots, n\}$, и пусть $\mathcal{C}$ и $\mathcal{D}$ — два монотонно убывающих семейства подмножеств множества N. Тогда*

$$\Pr[\mathcal{A} \cap \mathcal{B}] \geqslant \Pr[\mathcal{A}] \cdot \Pr[\mathcal{B}],$$
$$\Pr[\mathcal{C} \cap \mathcal{D}] \geqslant \Pr[\mathcal{C}] \cdot \Pr[\mathcal{D}],$$
$$\Pr[\mathcal{A} \cap \mathcal{C}] \leqslant \Pr[\mathcal{A}] \cdot \Pr[\mathcal{C}].$$

В терминах мощностей это можно записать так:

$$2^n|\mathcal{A} \cap \mathcal{B}| \geqslant |\mathcal{A}| \cdot |\mathcal{B}|,$$
$$2^n|\mathcal{C} \cap \mathcal{D}| \geqslant |\mathcal{C}| \cdot |\mathcal{D}|,$$
$$2^n|\mathcal{A} \cap \mathcal{C}| \leqslant |\mathcal{A}| \cdot |\mathcal{C}|\,.$$

Здесь и в дальнейшем через $\mathcal{A} \cap \mathcal{B}$, $\mathcal{C} \cap \mathcal{D}$ и $\mathcal{A} \cap \mathcal{C}$ обозначены обычные пересечения семейств.

Доказательство. Пусть $f : P(N) \to \mathbb{R}^+$ — характеристическая функция семейства $\mathcal{A}$, т. е. $f(A) = 0$, если $A \notin \mathcal{A}$, и $f(A) = 1$, если $A \in \mathcal{A}$. Аналогично, пусть g — характеристическая функция семейства $\mathcal{B}$. По предположению f и g — обе возрастающие. Применяя FKG-неравенство с тривиальной мерой $\mu \equiv 1$, получаем

$$\Pr[\mathcal{A} \cap \mathcal{B}] =< fg >\geqslant< f > \cdot < g >= \Pr[\mathcal{A}] \cdot \Pr[\mathcal{B}].$$

Два других неравенства подобным образом вытекают из теоремы 6.2.1 и следующего за ней абзаца.

Легко видеть, что предложение может быть выведено из теоремы Альсведе—Дайкина или из следствия 6.1.3. ■

Последнее предложение имеет некоторые интересные комбинаторные следствия, часть из которых появилась уже в оригинальной статье Клейтмана. Поскольку они являются прямыми следствиями и не содержат дополнительных вероятностных идей, мы опускаем их точные формулировки и возвращаемся к версии предложения 6.3.1 в более общем вероятностном пространстве.

Для действительного вектора $p = (p_1, \ldots, p_n)$ с координатами $0 \leqslant p_i \leqslant 1$ определим вероятностное пространство, элементами которого являются члены множества $P(N)$, где каждому $A \subseteq N$ приписана вероятность $\Pr[A] = \prod_{i \in A} p_i \prod_{j \notin A}(1 - p_j)$. Ясно, что это вероятностное распределение получается, если мы строим $A \subseteq N$ случайно, включая в A каждый элемент $i \in N$ независимо с вероятностью p_i. Обозначим для каждого семейства $\mathcal{A} \subseteq P(N)$ его вероятность в этом пространстве через $\Pr_p(\mathcal{A})$. В частности, если все вероятности p_i равны $1/2$, то $\Pr_p(\mathcal{A})$ — величина, обозначаемая через $\Pr[A]$ в предложении 6.3.1. Определим $\mu = \mu_p : P(N) \to \mathbb{R}^+$ равенством $\mu(A) = \prod_{i \in A} p_i \prod_{j \notin A}(1 - p_j)$.

Легко проверить, что функция μ является log-супермодулярной. В самом деле, $\mu(A)\mu(B) = \mu(A \cup B)\mu(A \cap B)$ для любых $A, B \subseteq N$. Это нетрудно проверить сравнением вклада, вносимого произвольным $i \in N$ в левую и правую части последнего равенства. Таким образом, можно применить FKG-неравенство и получить следующее обобщение предложения 6.3.1.

Теорема 6.3.2. *Пусть $\mathcal{A}$ и $\mathcal{B}$ — два монотонных возрастающих семейства подмножеств множества N и пусть $\mathcal{C}$ и $\mathcal{D}$ — два монотонных убывающих семейства подмножеств множества N. Тогда для действительного вектора $p = (p_1, , \ldots, , p_n)$, $0 \leqslant p_i \leqslant 1$, выполнены неравенства*

$$\Pr_p(\mathcal{A} \cap \mathcal{B}) \geqslant \Pr_p(\mathcal{A}) \cdot \Pr_p(\mathcal{B}),$$
$$\Pr_p(\mathcal{C} \cap \mathcal{D}) \geqslant \Pr_p(\mathcal{C}) \cdot \Pr_p(\mathcal{D}),$$
$$\Pr_p(\mathcal{A} \cap \mathcal{C}) \leqslant \Pr_p(\mathcal{A}) \cdot \Pr_p(\mathcal{C}).$$

Эта теорема может быть применена во многих случаях и будет использована нами в гл. 8 для вывода неравенства Янсона. В качестве иллюстрации

предположим, что $A_1, A_2, \ldots, A_k$ — произвольные подмножества множества N, причем каждое случайное подмножество A множества N строится путем выбора каждого $i \in N$ независимо с вероятностью p. Тогда из теоремы 6.3.2 легко следует, что

$$\Pr[A\text{пересекает каждое}A_i] \geqslant \prod_{i=1}^{k} \Pr[A\text{пересекает}A_i].$$

Заметим, что это, вообще говоря, неверно для других подобных вероятностных моделей. Например, если A — случайно выбранное ℓ-элементное подмножество множества N, то последнее неравенство неверно.

Рассматривая члены множества N как $n = \binom{m}{2}$ ребер полного графа на множестве вершин $V = \{1, 2, \ldots, m\}$, мы можем вывести корреляционное неравенство для случайных графов. Пусть $G = (V, E)$ — случайный граф на множестве вершин V, порожденный выбором для всех $i, j \in V, i \neq j$, независимо пары $\{i, j\}$ в качестве ребра с вероятностью p. (Эта модель случайного графа подробно обсуждается в гл. 10). *Свойство графов* — это подмножество множества всех графов на множестве V, замкнутое относительно изоморфизма. Так, например, связность — это свойство (соответствующее всем связным графам на V), планарность — также является свойством графов. Свойство Q называется *монотонно возрастающим*, если для графа G, обладающего свойством Q, граф H, полученный из G добавлением ребер, также обладает свойством Q. *Монотонно убывающее* свойство определяется аналогично. Интерпретируя члены множества N в теореме 6.3.2 как $\binom{m}{2}$ пар $\{i, j\}$ с $i, j \in V, i \neq j$, мы получаем следующий результат.

Теорема 6.3.3. *Пусть Q_1, Q_2, Q_3 и Q_4 — свойства графов, такие, что Q_1, Q_2 являются монотонно возрастающими, а Q_3, Q_4 — монотонно убывающими. Пусть $G = (V, E)$ — случайный граф на множестве вершин V, полученный путем случайного выбора каждого ребра с вероятностью p. Тогда*

$$\Pr[G \in Q_1 \cap Q_2] \geqslant \Pr[G \in Q_1] \cdot \Pr[G \in Q_2],$$
$$\Pr[G \in Q_3 \cap Q_4] \geqslant \Pr[G \in Q_3] \cdot \Pr[G \in Q_4],$$
$$\Pr[G \in Q_1 \cap Q_3] \leqslant \Pr[G \in Q_1] \cdot \Pr[G \in Q_3].$$

Так, например, вероятность того, что G является и гамильтоновым и планарным, не превосходит произведения вероятности того, что он гамильтонов, на вероятность того, что он планарный. Представляется безнадежным пытаться доказать такое утверждение непосредственно, без использования корреляционных неравенств.

6.4. ЛИНЕЙНЫЕ РАСШИРЕНИЯ ЧАСТИЧНО УПОРЯДОЧЕННЫХ МНОЖЕСТВ

Пусть $(P, \leqslant)$ — частично упорядоченное множество, состоящее из n элементов. Линейное расширение P — это взаимно-однозначное отображение

$\sigma : P \to \{1, 2, \ldots, n\}$, сохраняющее порядок, т. е. если $x, y \in P$ и $x \leqslant y$, то $\sigma(x) \leqslant \sigma(y)$. Интуитивно, σ — это ранжирование элементов множества P, которое сохраняет частичный порядок P. Рассмотрим вероятностное пространство всех линейных расширений множества P, в котором каждое возможное расширение равновероятно. В этом пространстве можно рассматривать, например, события вида $x \leqslant y$ или $(x \leqslant y) \wedge (x \leqslant z)$ (для $x, y, z \in P$) и вычислять их вероятности. Оказывается, что FKG-неравенство является очень полезным инструментом для изучения корреляций между такими событиями. Наилучший известный результат такого типа был предугадан Райвалом и Сэндсом и доказан Шепом [Shepp (1982)]. (См. также работу [Fishburn (1992)], где доказан более сильный результат.) Он утверждает, что для всякого частично упорядоченного множества P и любых трех элементов $x, y, z \in P$ выполнено неравенство $\Pr[x \leqslant y \wedge x \leqslant z] \geqslant \Pr[x \leqslant y] \Pr[x \leqslant z]$.

Этот результат стал известен как XYZ-теорема. Хотя она выглядит интуитивно очевидной, ее доказательство нетривиально и основано на остроумном применении FKG-неравенства. В этом разделе мы представим этот результат и его элегантное доказательство.

Теорема 6.4.1. *Пусть P — частично упорядоченное множество, состоящее из n элементов $a_1, a_2, \ldots, a_n$. Тогда*

$$\Pr[a_1 \leqslant a_2 \wedge a_1 \leqslant a_3] \geqslant \Pr[a_1 \leqslant a_2] \Pr[a_1 \leqslant a_3].$$

Доказательство. Пусть m — большое целое число (которое в дальнейшем будет стремиться к бесконечности), и пусть L — множество всех упорядоченных наборов длины n вида $\mathbf{x} = (x_1, \ldots, x_n)$, где $x_i \in M = \{1, 2, \ldots, m\}$. (Заметим, что числа x_i не обязательно различны между собой.) Определим отношение порядка $\leqslant$ на L следующим образом. Для $\mathbf{y} = (y_1, \ldots, y_n) \in L$ и $\mathbf{x}$, указанных выше, $\mathbf{x} \leqslant \mathbf{y}$ тогда и только тогда, когда $x_1 \geqslant y_1$ и $x_i - x_1 \leqslant y_i - y_1$ для всех i, $2 \leqslant i \leqslant n$. Нетрудно проверить, что $(L, \leqslant)$ является решеткой, в которой i-я компонента пересечения $\mathbf{x} \wedge \mathbf{y}$ равна $(\mathbf{x} \wedge \mathbf{y})_i = \min(x_i - x_1, y_i - y_1) + \max(x_1, y_1)$, а i-я компонента объединения $\mathbf{x} \vee \mathbf{y}$ равна $(\mathbf{x} \vee \mathbf{y})_i = \max(x_i - x_1, y_i - y_1) + \min(x_1, y_1)$.

Кроме того, решетка L — дистрибутивная. Это является простым следствием того факта, что тривиальная решетка целых чисел (относительно обычного порядка) является дистрибутивной, а, значит, для любых трех чисел a, b и c

$$\min(a, \max(b, c)) = \max(\min(a, b), \min(a, c)) \tag{6.7}$$

и

$$\max(a, \min(b, c)) = \min(\max(a, b), \max(a, c)). \tag{6.8}$$

Покажем, как отсюда следует, что L дистрибутивна. Пусть $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n)$ и $\mathbf{z} = (z_1, \ldots, z_n)$ — три элемента решетки L. Мы должны показать, что

$$\mathbf{x} \wedge (\mathbf{y} \vee \mathbf{z}) = (\mathbf{x} \wedge \mathbf{y}) \vee (\mathbf{x} \wedge \mathbf{z}).$$

Заметим, что i-я компонента в $\mathbf{x} \wedge (\mathbf{y} \vee \mathbf{z})$ равна

$$\begin{aligned}(\mathbf{x} \wedge (\mathbf{y} \vee \mathbf{z}))_i &= \min(x_i - x_1, (\mathbf{y} \vee \mathbf{z})_i - (\mathbf{y} \vee \mathbf{z})_1) + \\ &\quad + \max(x_1, (\mathbf{y} \vee \mathbf{z})_1) = \\ &= \min(x_i - x_1, \max(y_i - y_1, z_i - z_1)) + \\ &\quad + \max(x_1, \min(y_1, z_1)).\end{aligned}$$

Аналогично, i-я компонента в $(\mathbf{x} \wedge \mathbf{y}) \vee (\mathbf{x} \wedge \mathbf{z})$ равна

$$\begin{aligned}&((\mathbf{x} \wedge \mathbf{y}) \vee (\mathbf{x} \wedge \mathbf{z}))_i = \\ &\quad = \max((\mathbf{x} \wedge \mathbf{y})_i - (\mathbf{x} \wedge \mathbf{y})_1, (\mathbf{x} \wedge \mathbf{z})_i - (\mathbf{x} \wedge \mathbf{z})_1) + \\ &\qquad + \min((\mathbf{x} \wedge \mathbf{y})_1, (\mathbf{x} \wedge \mathbf{z})_1) = \\ &\quad = \max(\min(x_i - x_1, y_i - y_1), \min(x_i - x_1, z_i - z_1)) + \\ &\qquad + \min(\max(x_1, y_1), \max(x_1, z_1)).\end{aligned}$$

Указанные две величины равны, что легко показать применением равенства (6.7) с $a = x_i - x_1$, $b = y_i - y_1, c = z_i - z_1$ и равенства (6.8) с $a = x_1, b = y_1, c = z_1$.

Таким образом, решетка L дистрибутивна. Чтобы применить FKG-неравенство, нам понадобится функция меры μ и две функции f и g. Определим μ как характеристическую функцию множества P, т. е. для $\mathbf{x} = (x_1, \ldots, x_n) \in L$ выполнено равенство $\mu(\mathbf{x}) = 1$, если $x_i \leqslant x_j$ как только $a_i \leqslant a_j$ в P, и $\mu(\mathbf{x}) = 0$ иначе. Чтобы показать, что μ является log-супермодулярной, достаточно убедиться в том, что, если $\mu(\mathbf{x}) = \mu(\mathbf{y}) = 1$, то $\mu(\mathbf{x} \vee \mathbf{y}) = \mu(\mathbf{x} \wedge \mathbf{y}) = 1$. В самом деле, если $\mu(\mathbf{x}) = \mu(\mathbf{y}) = 1$ и $a_i \leqslant a_j$ в P, то $x_i \leqslant x_j$ и $y_i \leqslant y_j$ и поэтому

$$\begin{aligned}(x \vee y)_i &= \max(x_i - x_1, y_i - y_1) + \min(x_1, y_1) \leqslant \\ &\leqslant \max(x_j - x_1, y_j - y_1) + \min(x_1, y_1) = (x \vee y)_j,\end{aligned}$$

т. е. $\mu(\mathbf{x} \vee \mathbf{y}) = 1$. Аналогично, из $\mu(\mathbf{x}) = \mu(\mathbf{y}) = 1$ следует $\mu(\mathbf{x} \wedge \mathbf{y}) = 1$.

Вполне естественным будет определить f и g как характеристические функции двух событий $x_1 \leqslant x_2$ и $x_1 \leqslant x_3$, соответственно, т. е. $f(\mathbf{x}) = 1$, если $x_1 \leqslant x_2$, и $f(\mathbf{x}) = 0$ иначе, и $g(\mathbf{x}) = 1$, если $x_1 \leqslant x_3$, и $g(\mathbf{x}) = 0$ иначе. Обе функции f и g, очевидно, являются возрастающими. В самом деле, если $\mathbf{x} \leqslant \mathbf{y}$ и $f(\mathbf{x}) = 1$, то $0 \leqslant x_2 - x_1 \leqslant y_2 - y_1$, а, значит, $f(\mathbf{y}) = 1$. То же справедливо для g.

Таким образом, мы имеем все необходимое для применения FKG-неравенства (теорема 6.2.1). Отсюда следует, что вероятность того, что n-ка $(x_1, \ldots, x_n)$ из L, которая удовлетворяет в P обоим неравенствам $x_1 \leqslant x_2$ и $x_1 \leqslant x_3$, по крайней мере столь же велика, как произведение вероятностей того, что $x_1 \leqslant x_2$ и того, что $x_1 \leqslant x_3$. Заметим, что это еще не то, что мы хотели доказать. Ведь n-ки в L не являются n-ками *различных* целых чисел, и поэтому не соответствуют линейным расширениям множества P. Однако, при $m \to \infty$ вероятность того, что $x_i = x_j$ для некоторых $i \neq j$ в элементе $\mathbf{x} = (x_1, \ldots, x_n)$ решетки L, стремится к нулю. Отсюда следует утверждение теоремы. ■

6.5. УПРАЖНЕНИЯ

1. Обозначим через P вероятность того, что случайный подграф графа G, полученный независимым выбором каждого ребра графа G с вероятностью $1/2$, является связным (и остовным). Рассмотрим случайную раскраску ребер графа G в синий и красный цвета, при которой выбор цвета производится случайно и равновероятно. Обозначим через Q вероятность того, что при такой раскраске графа G оба графа, синий и красный, являются связными (и остовными). Верно ли, что $Q \leqslant P^2$?

2. Семейство подмножеств $\mathcal{G}$ называется *пересекающимся*, если $G_1 \cap G_2 \neq \varnothing$ для всех $G_1, G_2 \in \mathcal{G}$. Пусть $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_k$ — k пересекающихся семейств подмножеств множества $\{1, 2, \ldots, n\}$. Доказать, что

$$\left| \bigcup_{i=1}^{k} \mathcal{F}_i \right| \leqslant 2^n - 2^{n-k}.$$

3. Показать, что вероятность того, что в случайном графе $G(2k, 1/2)$ максимальная степень вершины не превышает $k-1$, не меньше $1/4^k$.

ВЕРОЯТНОСТНЫЙ ВЗГЛЯД:

Теорема Турана

Обозначим через d_v степень вершины v в графе $G = (V, E)$. Пусть $\alpha(G)$ — максимальный размер независимого множества вершин графа G. Каро и Уэй доказали следующий результат.

Теорема. *Справедливо неравенство*

$$\alpha(G) \geqslant \sum_{v \in V} \frac{1}{d_v + 1}.$$

Доказательство. Пусть «$<$» — линейный порядок на множестве V, выбранный случайно и равновероятно. Положим

$$I = \{v \in V : \{v, w\} \in E \Rightarrow v < w\}.$$

Пусть X_v — индикатор события $v \in I$, а $X = \sum_{v \in V} X_v = |I|$. Для всякого v

$$\mathbf{E}[X_v] = \Pr[v \in I] = \frac{1}{d_v + 1},$$

так как $v \in I$ тогда и только тогда, когда v является наименьшим среди v и вершин из ее окрестности. Следовательно,

$$\mathbf{E}[X] = \sum_{v \in V} \frac{1}{d_v + 1}$$

и поэтому существует порядок «$<$», для которого

$$|I| \geqslant \sum_{v \in V} \frac{1}{d_v + 1}.$$

Но если $x, y \in I$ и $\{x, y\} \in E$, то $x < y$ *и* $y < x$. Приходим к противоречию. Таким образом, I является независимым и $\alpha(G) \geqslant |I|$. ■

Для всякого $m \leqslant n$ пусть q, r удовлетворяют условиям $n = mq + r$, $0 \leqslant r < m$, и пусть $e = r\binom{q+1}{2} + (m - r)\binom{q}{2}$. Определим граф $G = G_{n,e}$ на n вершинах и e ребрах, разделив множество вершин на m максимально близких по мощности классов и соединив две вершины в том и только том случае, когда они принадлежат одному классу. Ясно, что $\alpha(G_{n,e}) = m$.

Теорема [Turán (1941)]. *Пусть граф H имеет n вершин и e ребер. Тогда $\alpha(H) \geqslant m$ и $\alpha(H) = m \Leftrightarrow H \cong G_{n,e}$.*

Доказательство. Граф $G_{n,e}$ удовлетворяет условию $\sum_{v \in V}(d_v + 1)^{-1} = m$, поскольку каждая клика вносит в эту сумму 1. Положив $e = \sum_{v \in V} d_v/2$, мы минимизируем $\sum_{v \in V}(d_v + 1)^{-1}$, выбирая d_v близкими друг к другу, насколько

это возможно. Таким образом, для всякого H

$$\alpha(H) \geqslant \sum_{v \in V} \frac{1}{d_v + 1} \geqslant m.$$

При $\alpha(H) = m$ мы должны иметь равенство в обоих случаях. Из второго равенства следует, что d_v должны быть близки друг к другу насколько это возможно. Положив $X = |I|$, как в предыдущей теореме, получим $\alpha(H) = \mathbf{E}[X]$. Но $\alpha(H) \geqslant X$ для любого знака неравенства $<$, поэтому X должно быть константой. Предположим, H не является объединением клик. Тогда существуют $x, y, z \in V$ с $\{x, y\}, \{x, z\} \in E, \{y, z\} \notin E$. Пусть $<$ — порядок, в котором наименьшими являются элементы x, y, z и $<'$ — тот же самый порядок за исключением минимальных элементов, порядок которых — y, z, x. Пусть I, I' — соответствующие множества вершин, все соседи которых «больше» их самих. Тогда I, I' идентичны за исключением того, что $x \in I, y, z \notin I$, в то время как $x \notin I', y, z \in I'$. Таким образом, X не является константой. То есть из $\alpha(H) = \mathbf{E}[X]$ следует, что H — это объединение клик, и поэтому $H \cong G_{n,e}$. ■

7

Мартингалы и плотная концентрация

Математика представляется мне гораздо более реальной, чем бизнес. К примеру: ну что за реальность в рекламном стенде МакДональдса? Сегодня он здесь, а завтра его нет. А целые числа — это реальность. Доказав теорему, вы действительно сделали что-то, что имеет субстанцию, и с чем никакая деловая афера не может сравниться по реальности.

Джим Саймонс

7.1. ОПРЕДЕЛЕНИЯ

Мартингал — это последовательность $X_0, \ldots, X_m$ случайных величин, такая, что для $0 \leqslant i < m$ выполняется равенство

$$\mathbf{E}[X_{i+1}|X_i, X_{i-1}, \ldots, X_0] = X_i.$$

Представим себе игрока, пришедшего в казино с X_0 долларами. Казино предоставляет ассортимент азартных игр. Все игры «справедливые» в том смысле, что математические ожидания выигрыша в них равны нулю. Игрок может учесть предысторию при планировании своей игры и расчете ставки. Он может использовать игровое определение мартингала — удваивай ставку, пока не выиграешь. Он может играть в рулетку до трех выигрышей, а затем переключиться на кено. Пусть X_i — выигрыш игрока в момент i. При $X_i = a$ условное математическое ожидание величины X_{i+1} должно быть равно a, следовательно, это — мартингал.

Простой, но поучительный пример мартингала — это игра с бросанием монеты с постоянной ставкой в один доллар. Пусть $Y_1, \ldots, Y_m$ — независимые бросания монеты, при которых выигрыш $+1$ и проигрыш -1 имеют одинаковую вероятность $\frac{1}{2}$. Пусть начальная сумма $X_0 = 0$, и при этом игрок имеет неограниченный кредит. Тогда $X_i = Y_1 + \ldots + Y_i$ имеет распределение $\mathbf{S}_i$.

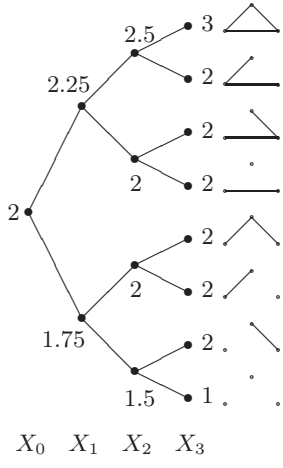
Наши мартингалы будут выглядеть совершенно различными, по крайней мере, внешне.

Мартингал проявления ребер. Рассмотрим случайный граф $G(n,p)$ в качестве вероятностного пространства. Пометим потенциальные ребра $\{i,j\} \subseteq [n]$ произвольно символами $e_1, \ldots, e_m$, положив для удобства $m = \binom{n}{2}$. Пусть f — произвольная теоретико-графовая функция. Определим мартингал $X_0, \ldots, X_m$ заданием значений $X_i(H)$. Величина $X_m(H)$ есть просто $f(H)$, а $X_0(H)$ — математическое ожидание величины $f(G)$ с $G \sim G(n,p)$. Заметим, что X_0

является константой. В общем (включая случаи $i = 0$ и $i = m$),

$$X_i(H) = \mathbf{E}[f(G)|e_j \in G \longleftrightarrow e_j \in H, 1 \leqslant j \leqslant i].$$

Словом, чтобы найти $X_i(H)$, мы сначала проявим первые i пар $e_1, \ldots, e_i$ и посмотрим, принадлежат ли они графу H. Остальные ребра невидимы и рассматриваются как случайные. Тогда $X_i(H)$ является условным математическим ожиданием величины $f(G)$ с учетом этой частичной информации. При $i = 0$ ни одно ребро не проявляется, и X_0 является константой. При $i = m$ все ребра проявились, и X_m есть функция f. Наш мартингал продвигается малыми шагами от полного отсутствия информации к полной информации.



На рисунке представлен *мартингал проявления ребер* с $n = m = 3$ и хроматическим числом в качестве f. При этом ребра проявляются в следующем порядке: «нижнее, левое, правое». Величины $X_i(H)$ определяются по ходу движения от центральной вершины к листу, помеченному символом H.

Рисунок показывает, почему это мартингал. Математическое ожидание $f(H)$ при условии, что первые $i-1$ ребер известны, есть взвешенное среднее математических ожиданий величины $f(H)$ при условии, что i-е ребро проявилось. В более общем случае (на который иногда ссылаются как на мартингальный процесс Дуба) величина X_i может рассматриваться как математическое ожидание величины $f(H)$ при условии, что определенная информация остается открытой до тех пор, пока информация, известная к моменту i, включает информацию, известную к моменту $i - 1$.

Мартингал проявления вершин. Вновь пусть $G(n, p)$ — вероятностное пространство и f — произвольная теоретико-графовая функция. Определим величины $X_1, \ldots, X_n$ равенством

$$X_i(H) = \mathbf{E}\Big[f(G)|\text{для } x, y \leqslant i, \{x, y\} \in G \longleftrightarrow \{x, y\} \in H\Big].$$

Словом, чтобы найти $X_i(H)$, мы проявляем первые i вершин, а также все их внутренние ребра и вычисляем условное математическое ожидание значения $f(G)$ с этой частичной информацией. При подходящем упорядочении ребер мартингал проявления вершин может рассматриваться как подпоследовательность мартингала проявления ребер. Заметим, что $X_1(H) = \mathbf{E}[f(G)]$ является константой, если ни одно из ребер не проявилось, и $X_n(H) = f(H)$ — если проявились все ребра.

7.2. БОЛЬШИЕ УКЛОНЕНИЯ

Мори [Maurey (1979)] применил неравенство больших уклонений для мартингалов при доказательстве изопериметрического неравенства для симметрической группы. Это неравенство оказалось полезным при изучении нормированных пространств (см. также [Milman and Schechtman (1986)]). Все применения мартингалов в теории графов также опираются на тот же самый основной мартингальный результат, использованный Мори и заключающийся в следующем.

Теорема 7.2.1 (неравенство Ацумы). *Пусть $0 = X_0, \ldots, X_m$ — мартингал, удовлетворяющий условию*

$$|X_{i+1} - X_i| \leqslant 1$$

для всех $0 \leqslant i < m$. Пусть $\lambda > 0$. Тогда

$$\Pr[X_m > \lambda\sqrt{m}] < e^{-\lambda^2/2}.$$

В мартингале «подбрасывания монеты» X_m имеет распределение $\mathbf{S}_m$, и этот результат есть теорема A.1.1. На самом деле доказательство общего случая совершенно аналогично.

Доказательство. Предусмотрительно положим $\alpha = \lambda/\sqrt{m}$. Пусть $Y_i = X_i - X_{i-1}$, так что $|Y_i| \leqslant 1$ и $\mathbf{E}[Y_i | X_{i-1}, X_{i-2}, \ldots, X_0] = 0$. Тогда, подобно тому как это сделано в теореме A.1.16, имеем

$$\mathbf{E}[e^{\alpha Y_i} | X_{i-1}, X_{i-2}, \ldots, X_0] \leqslant \operatorname{ch}(\alpha) \leqslant e^{\alpha^2/2}.$$

Следовательно,

$$\mathbf{E}[e^{\alpha X_m}] = \mathbf{E}\left[\prod_{i=1}^{m} e^{\alpha Y_i}\right] = \mathbf{E}\left[\left(\prod_{i=1}^{m-1} e^{\alpha Y_i}\right)\mathbf{E}[e^{\alpha Y_m} | X_{m-1}, X_{m-2}, \ldots, X_0]\right] \leqslant$$
$$\leqslant \mathbf{E}\left[\prod_{i=1}^{m-1} e^{\alpha Y_i}\right] e^{\alpha^2/2} \leqslant e^{\alpha^2 m/2}.$$

Значит,

$$\Pr[X_m > \lambda\sqrt{m}] = \Pr[e^{\alpha X_m} > e^{\alpha\lambda\sqrt{m}}] < \mathbf{E}[e^{\alpha X_m}] e^{-\alpha\lambda\sqrt{m}} \leqslant$$
$$\leqslant e^{\alpha^2 m/2 - \alpha\lambda\sqrt{m}} = e^{-\lambda^2/2},$$

что и требовалось. ■

Следствие 7.2.2. *Пусть $c = X_0, \dots, X_m$ — мартингал, удовлетворяющий условию*

$$|X_{i+1} - X_i| \leqslant 1$$

для всех $0 \leqslant i < m$. Тогда

$$\Pr[|X_m - c| > \lambda\sqrt{m}] < 2e^{-\lambda^2/2}.$$

Скажем, что теоретико-графовая функция f *удовлетворяет реберному условию Липшица,* если для любых графов H и H', отличающихся в точности одним ребром, выполняется неравенство $|f(H) - f(H')| \leqslant 1$. Функция f *удовлетворяет вершинному условию Липшица,* если для любых графов H и H', отличающихся в точности одной вершиной, выполняется неравенство $|f(H) - f(H')| \leqslant 1$.

Теорема 7.2.3. *Если f удовлетворяет реберному условию Липшица, то соответствующий мартингал проявления ребер удовлетворяет неравенству $|X_{i+1} - X_i| \leqslant 1$. Если f удовлетворяет вершинному условию Липшица, то соответствующий мартингал проявления вершин удовлетворяет неравенству $|X_{i+1} - X_i| \leqslant 1$.*

Мы докажем эти результаты в более общей форме позднее. Они имеют тот интуитивный смысл, что если проявление некоторой частной вершины или ребра не могут изменить f больше, чем на единицу, то проявление вершины или ребра не может изменить математическое ожидание больше, чем на единицу. Теперь мы приведем простой пример применения этих результатов.

Теорема 7.2.4 [Shamir and Spencer (1987)]. *Пусть n, p — произвольные числа, и пусть $c = \mathbf{E}[\chi(G)]$, где $G \sim G(n,p)$. Тогда*

$$\Pr\left[|\chi(G) - c| > \lambda\sqrt{n-1}\right] < 2e^{-\lambda^2/2}.$$

Доказательство. Рассмотрим мартингал проявления вершин $X_1, \dots, X_n$ на графе $G(n,p)$ с $f(G) = \chi(G)$. Одной вершине всегда можно присвоить новый цвет. Поэтому вершинное условие Липшица выполняется. Теперь применим неравенство Ацумы в форме следствия 7.2.2. ■

Пусть $\lambda \to \infty$ сколь угодно медленно. Этот результат показывает, что распределение величины $\chi(G)$ является «плотно сконцентрированным» около ее среднего. Однако доказательство не дает рецепта для вычисления самого́ среднего.

7.3. ХРОМАТИЧЕСКОЕ ЧИСЛО

В теореме 10.3.1 мы докажем, что $\chi(G) \sim n/2\log_2 n$ почти всегда при условии, что $G \sim G(n, 1/2)$. Здесь мы представим оригинальное доказательство Белы Боллобаша, использующее мартингалы. Мы используем определения гл. 10, разд. 10.3. Положим $f(k) = \binom{n}{k} 2^{-\binom{k}{2}}$. Пусть k_0 таково, что $f(k_0 - 1) > 1 > f(k_0)$. Пусть далее $k = k_0 - 4$, так что $k \sim 2\log_2 n$ и $f(k) > n^{3+o(1)}$. Наша цель —

показать, что

$$\Pr[\omega(G) < k] = e^{-n^{2+o(1)}},$$

где $\omega(G)$ — это размер максимальной клики графа G. В действительности, мы докажем в теореме 7.3.2 более точную оценку. Остальные соображения приводятся в разд. 10.3.

Пусть $Y = Y(H)$ — максимальный размер семейства непересекающихся по ребрам k-клик в H. Этот остроумный и необычный выбор функции является ключом к применению мартингалов в доказательстве.

Лемма 7.3.1. *Справедливо неравенство* $\mathbf{E}[Y] \geqslant \frac{n^2}{2k^4}(9 + o(1))$.

Доказательство. Обозначим через $\mathcal{K}$ семейство k-клик графа G так, что $f(k) = \mu = \mathbf{E}[|\mathcal{K}|]$. Через W обозначим количество неупорядоченных пар $\{A, B\}$, составленных из k-клик графа G с $2 \leqslant |A \cap B| < k$. Тогда $\mathbf{E}[W] = \Delta/2$ при Δ, определенном как в гл. 10, разд. 10.3 (см. также гл. 4, разд. 4.5), имеем $\Delta \sim \mu^2 k^4 n^{-2}$. Пусть $\mathcal{C}$ — случайное подсемейство семейства $\mathcal{K}$, определяемое тем, что для всякого $A \in \mathcal{K}$

$$\Pr[A \in \mathcal{C}] = q,$$

где q будет определено позже. Пусть W' — количество неупорядоченных пар $\{A, B\}$, $A, B \in \mathcal{C}$ с $2 \leqslant |A \cap B| < k$. Тогда

$$\mathbf{E}[W'] = \mathbf{E}[W]q^2 = \Delta q^2/2.$$

Удалим из $\mathcal{C}$ по одному множеству из каждой такой пары $\{A, B\}$. В результате получим семейство $\mathcal{C}^*$ непересекающихся по ребрам k-клик графа G и

$$\mathbf{E}[Y] \geqslant \mathbf{E}[|\mathcal{C}^*|] \geqslant \mathbf{E}[|\mathcal{C}|] - \mathbf{E}[W'] = \mu q - \Delta q^2/2 = \mu^2/2\Delta \sim n^2/2k^4,$$

где мы выбираем $q = \mu/\Delta$ (неважно, что это меньше единицы!), чтобы минимизировать квадратичную функцию. ■

Мы предполагаем, что неравенство в лемме 7.3.1 можно улучшить, заменив правую его часть на $\mathbf{E}[Y] > cn^2/k^2$. То есть с положительной вероятностью существует семейство непересекающихся по ребрам k-клик, которое покрывает положительную долю ребер.

Теорема 7.3.2. *Справедливо неравенство*

$$\Pr[\omega(G) < k] < e^{-(c+o(1))\frac{n^2}{\ln^8 n}},$$

где c — некоторая положительная константа.

Доказательство. Пусть $Y_0, \ldots, Y_m$, $m = \binom{n}{2}$, — мартингал проявления ребер на графе $G(n, 1/2)$ с функцией Y, определенной перед леммой. Функция Y удовлетворяет реберному условию Липшица, так как добавление одного ребра может добавить не более одной клики к семейству клик, непересекающихся по ребрам. (Заметим, что условие Липшица могло бы не выполняться для обычного числа k-клик, поскольку добавление одного ребра могло бы повлечь появление нескольких клик.) Граф G не имеет k-клик тогда и только

тогда, когда $Y = 0$. Применим неравенство Ацумы с $m = \binom{n}{2} \sim n^2/2$ и $\mathbf{E}[Y] \geqslant \frac{n^2}{2k^4}(1 + o(1))$. Тогда

$$\Pr[\omega(G) < k] = \Pr[Y = 0] \leqslant \Pr\left[Y - \mathbf{E}[Y] \leqslant -\mathbf{E}[Y]\right] \leqslant$$
$$\leqslant e^{-\mathbf{E}[Y]^2/2\binom{n}{2}} \leqslant e^{-(c'+o(1))n^2/k^8} = e^{-(c+o(1))n^2/\ln^8 n},$$

что и требовалось. ■

Ниже — еще один пример того, какой изобретательности требует мартингальный подход при выборе теоретико-графовой функции.

Теорема 7.3.3. *Пусть $p = n^{-\alpha}$, где α — фиксированное, $\alpha > \frac{5}{6}$. Пусть $G = G(n,p)$. Тогда существует $u = u(n,p)$, такое, что почти всегда*

$$u \leqslant \chi(G) \leqslant u + 3.$$

То есть $\chi(G)$ сконцентрировано в четырех значениях.

Для начала нам потребуется одна хорошо известная техническая лемма.

Лемма 7.3.4. *Пусть α, c — фиксированные числа, $\alpha > \frac{5}{6}$. Пусть $p = n^{-\alpha}$. Тогда почти наверное каждые $c\sqrt{n}$ вершин графа $G = G(n,p)$ могут быть правильно раскрашены в три цвета.*

Доказательство. Пусть это не так. Обозначим через T множество минимального размера, которое не является 3-раскрашиваемым. Поскольку для всякого $x \in T$ подграф, порожденный множеством $T \setminus \{x\}$, является 3-раскрашиваемым, а для множества T это не так, x должен иметь по меньшей мере трех соседей в T. Отсюда следует, что если T имеет t вершин, то граф, порожденный множеством T, должен иметь по меньшей мере $\frac{3t}{2}$ ребер. Вероятность того, что это происходит для некоторого T с $t \leqslant c\sqrt{n}$, ограничено сверху числом

$$\sum_{t=4}^{c\sqrt{n}} \binom{n}{t} \binom{\binom{t}{2}}{\frac{3t}{2}} p^{3t/2}.$$

Поскольку

$$\binom{n}{t} \leqslant \left(\frac{ne}{t}\right)^t \text{ и } \binom{\binom{t}{2}}{\frac{3t}{2}} \leqslant \left(\frac{te}{3}\right)^{3t/2},$$

то каждое слагаемое не превышает

$$\left[\frac{ne}{t}\frac{t^{3/2}e^{3/2}}{3^{3/2}}n^{-3\alpha/2}\right]^t \leqslant \left[c_1 n^{2-\frac{3\alpha}{2}} t^{1/2}\right]^t \leqslant \left[c_2 n^{1-\frac{3\alpha}{2}} n^{1/4}\right]^t = \left[c_2 n^{-\varepsilon}\right]^t,$$

где $\varepsilon = \frac{3\alpha}{2} - \frac{5}{4} > 0$. Следовательно, их сумма равна $o(1)$. ■

Доказательство теоремы 7.3.3. Пусть $\varepsilon > 0$ произвольно мало и пусть $u = u(n,p,\varepsilon)$ — наименьшее число, такое, что

$$\Pr[\chi(G) \leqslant u] > \varepsilon.$$

Пусть теперь $Y(G)$ — минимальный размер множества вершин S, для которого $G - S$ может быть u-раскрашено. Такое Y удовлетворяет вершинному условию Липшица, так как в худшем случае можно добавить одну вершину к S. Применим мартингал проявления вершин на $G(n,p)$ к функции Y. Положив

$\mu = \mathbf{E}[Y]$, имеем

$$\Pr[Y \leqslant \mu - \lambda\sqrt{n-1}] < e^{-\lambda^2/2},$$
$$\Pr[Y \geqslant \mu + \lambda\sqrt{n-1}] < e^{-\lambda^2/2}.$$

Пусть λ удовлетворяет соотношению $e^{-\lambda^2/2} = \varepsilon$. Тогда вероятность каждого из этих «хвостовых» событий меньше чем ε. Мы определили u так, что с вероятностью по меньшей мере ε, граф G мог бы быть u-раскрашиваемым и, значит, $Y = 0$. То есть, $\Pr[Y = 0] > \varepsilon$. Из первого неравенства следует, что $\mu \leqslant \lambda\sqrt{e-1}$. Теперь с использованием второго неравенства имеем

$$\Pr[Y \geqslant 2\lambda\sqrt{n-1}] \leqslant \Pr[Y \geqslant \mu + \lambda\sqrt{n-1}] \leqslant \varepsilon.$$

С вероятностью не меньшей $1-\varepsilon$ существует u-раскраска всех, кроме не более чем $c'\sqrt{n}$ вершин. По лемме 7.3.4 почти всегда, а значит, с вероятностью, не меньшей чем $1-\varepsilon$, эти вершины могут быть окрашены тремя дополнительными цветами, что дает в результате $(u+3)$-раскраску графа G. Минимальность u гарантирует, что с вероятностью, не меньшей $1-\varepsilon$, требуется как минимум u цветов для раскраски графа G. В итоге

$$\Pr[u \leqslant \chi(G) \leqslant u+3] \geqslant 1 - 3\varepsilon.$$

Осталось вспомнить, что ε произвольно мало. ■

С использованием этой же техники аналогичные результаты могут быть получены и для других значений α. Вместе с некоторыми похожими идеями это приводит к доказательству того, что для любых фиксированных $\alpha > \frac{1}{2}$, значение $\chi(G)$ сконцентрировано не более, чем в двух значениях (см. детали доказательств в работах [Luczak (1991)] и [Alon and Krivelevich (1997)]).

7.4. ДВА ОБОБЩЕНИЯ

Мартингалы, полезные при изучении случайных графов, могут быть помещены в следующую общую схему, которая подробно исследовалась в работах [Maurey (1979)] и [Milman and Schechtman (1986)]. Обозначим через $\Omega = A^B$ множество функций $g : B \to A$. (Если, например, B — множество пар элементов некоторого n-множества и $A = \{0, 1\}$, мы можем сопоставить $g \in A^B$ граф на n вершинах.) Мы введем меру, определяя значения w_{ab} равенствами вида

$$\Pr[g(b) = a] = w_{ab},$$

считая при этом величины $g(b)$ взаимно независимыми. (В графе $G(n,p)$ для всех b полагаем $w_{1b} = p, w_{0b} = 1-p$.) Теперь зафиксируем последовательность, называемую далее *шкалой*

$$\varnothing = B_0 \subset B_1 \subset \ldots \subset B_m = B.$$

Пусть $L : A^B \to \mathbb{R}$ — функционал (например, размер максимальной клики). Определим мартингал X_0, X_1, $\ldots$ X_m, полагая

$$X_i(h) = \mathbf{E}[L(g) | g(b) = h(b) \text{ для всех } b \in B_i].$$

Здесь X_0 — константа, равная математическому ожиданию величины L для случайного графа g, а X_m есть само L. Величины $X_i(g)$ приближаются к $L(g)$ по мере того как величины $g(b)$ «проявляются». Скажем, что функционал L *удовлетворяет условию Липшица относительно введенной шкалы*, если для всех $0 \leqslant i < m$

$$h, h' \text{ отличаются только на } B_{i+1} - B_i \Rightarrow |L(h') - L(h)| \leqslant 1.$$

Теорема 7.4.1. *Пусть L удовлетворяет условию Липшица. Тогда соответствующий мартингал удовлетворяет неравенству*

$$|X_{i+1}(h) - X_i(h)| \leqslant 1$$

для всех $0 \leqslant i < m$, $h \in A^B$.

Доказательство. Пусть H — семейство тех h', которые совпадают с h на B_{i+1}. Тогда

$$X_{i+1}(h) = \sum_{h' \in H} L(h') w_{h'},$$

где $w_{h'}$ — условная вероятность того, что $g = h'$ при условии, что $g = h$ на B_{i+1}. Для каждого $h' \in H$ обозначим через $H[h']$ семейство тех h^*, которые совпадают с h' везде, исключая, возможно, $B_{i+1} - B_i$. Семейство $H[h']$ разбивает семейство функций h^*, совпадающих с h на B_i. Таким образом, справедливо представление

$$X_i(h) = \sum_{h' \in H} \sum_{h^* \in H[h']} [L(h^*) q_{h^*}] w_{h'},$$

где q_{h^*} — условная вероятность того, что g совпадает с h^* на B_{i+1} при том, что g совпадает с h на B_i. (Это потому, что для $h^* \in H[h']$ величина $w_{h'}$ также является вероятностью того, что $g = h^*$ при условии, что $g = h^*$ на B_{i+1}.) Таким образом,

$$\begin{aligned} |X_{i+1}(h) - X_i(h)| &= \left| \sum_{h' \in H} w_{h'} \left[L(h') - \sum_{h^* \in H[h']} L(h^*) q_{h^*} \right] \right| \leqslant \\ &\leqslant \sum_{h' \in H} w_{h'} \sum_{h^* \in H[h']} \left| q_{h^*} [L(h') - L(h^*)] \right|. \end{aligned}$$

По условию Липшица $|L(h') - L(h^*)| \leqslant 1$, поэтому

$$|X_{i+1}(h) - X_i(h)| \leqslant \sum_{h' \in H} w_{h'} \sum_{h^* \in H[h']} q_{h^*} = \sum_{h' \in H} w_{h'} = 1. \qquad \blacksquare$$

Теперь мы можем сформулировать неравенство Ацумы в общем виде.

Теорема 7.4.2. *Пусть L удовлетворяет условию Липшица относительно шкалы длины m, и пусть $\mu = \mathbf{E}[L(g)]$. Тогда для всех $\lambda > 0$*

$$\Pr[L(g) \geqslant \mu + \lambda\sqrt{m}] < e^{-\lambda^2/2},$$
$$\Pr[L(g) \leqslant \mu - \lambda\sqrt{m}] < e^{-\lambda^2/2}.$$

Второе обобщение взято из работы [Alon et al. (1997)]. Мы предполагаем, что наше вероятностное пространство порождено конечным множеством взаимно независимых выборов типа ДА/НЕТ, помеченных индексами $i \in I$. Рассмотрим случайную величину Y на этом пространстве. Обозначим через p_i вероятность того, что выбор с индексом i — ДА. Пусть c_i таково, что при изменении выбора i (и сохранении всех остальных индексов прежними) величина Y изменяется не более чем на c_i. Назовем c_i *результатом* выбора i. Пусть C — максимальное число среди c_i. Назовем $p_i(1-p_i)c_i^2$ *дисперсией* выбора i.

Теперь рассмотрим игру «солитер», в которой Пол находит величину Y, задавая вопросы всегда правдивому оракулу Кэрол. Вопросы всегда касаются выбора $i \in I$. Выбор вопроса Полом может зависеть от предыдущего ответа Кэрол. Стратегия Пола может быть представлена в форме дерева решений. «Линия опроса» — это путь от корня к листу этого дерева, т. е. последовательность вопросов и ответов, которая определяет значение величины Y. Полная дисперсия линии опроса есть сумма дисперсий составляющих ее вопросов.

Теорема 7.4.3. *Для всякого $\varepsilon > 0$ существует $\delta > 0$, для которого выполнено следующее. Предположим, Пол использует стратегию для нахождения Y, при которой каждая линия опроса имеет полную дисперсию, не превышающую σ^2. Тогда*

$$\Pr[|Y - \mathbf{E}[Y]| > \alpha\sigma] \leqslant 2e^{-\frac{\alpha^2}{2(1+\varepsilon)}} \tag{7.1}$$

для всех положительных α, таких, что $\alpha C < \sigma(1+\varepsilon)\delta$.

Приложения. Для получения конкретной субоптимальной оценки можно положить $\varepsilon = \delta = 1$. Если $C = O(1)$, $\alpha \to \infty$ и $\alpha = o(\sigma)$, то верхняя граница (7.1) равна $\exp[-\Omega(\alpha^2)]$. Во многих случаях *все* вопросы Пола лежат в I. Тогда мы можем взять σ, такое, что $\sigma^2 = \sum_{i\in I} p_i(1-p_i)c_i^2$. Например, рассмотрим реберную Липшицеву величину Y на $G(n,p)$ с $p = p(n) \to 0$. Множество I состоит из $m = \binom{n}{2}$ потенциальных ребер, все $p_i = p$, $C = 1$, так что $\sigma = \Theta(\sqrt{n^2 p})$. Если $\alpha \to \infty$ с $\alpha = o(\sqrt{n^2 p})$, то верхняя оценка (7.1) снова равна $\exp[-\Omega(\alpha^2)]$.

Доказательство. Для простоты заменим Y на $Y - \mathbf{E}[Y]$, с тем, чтобы впредь предполагать $\mathbf{E}[Y] = 0$. В силу симметрии будем оценивать только верхний хвост величины Y. Предусмотрительно положим $\lambda = \alpha/[\sigma(1+\varepsilon)]$. В условиях теоремы имеем $C\lambda < \delta$. Покажем, что

$$\mathbf{E}[e^{\lambda Y}] \leqslant e^{(1+\varepsilon)\lambda^2\sigma^2/2}. \tag{7.2}$$

Тогда мартингальное неравенство будет следовать из неравенства Маркова

$$\Pr[Y > \alpha\sigma] < e^{-\lambda\alpha\sigma}\mathbf{E}[e^{\lambda Y}] \leqslant e^{-\alpha^2/2(1+\varepsilon)}.$$

Прежде всего, мы утверждаем, что для всякого $\varepsilon > 0$ существует $\delta > 0$, такое, что при $0 \leqslant p \leqslant 1$ и $|a| \leqslant \delta$

$$pe^{(1-p)a} + (1-p)e^{-pa} \leqslant e^{(1+\varepsilon)p(1-p)a^2/2}. \tag{7.3}$$

Разложим левую часть неравенства в ряд Тейлора по a. Константа в этом разложении равна 1, линейный член равен 0, коэффициент при a^2 равен $\frac{1}{2}p(1-p)$, а при $j \geqslant 3$ коэффициент при a^j не превосходит $\frac{1}{j!}p(1-p)[p^{j-1}+(1-p)^{j-1}] \leqslant \frac{1}{j!}p(1-p)$. Выберем δ так, чтобы при $|a| \leqslant \delta$ выполнялось соотношение

$$\sum_{j=3}^{\infty} \frac{a^j}{j!} < \varepsilon a^2/2.$$

(В частности, это справедливо для $\varepsilon = \delta = 1$.) Тогда

$$pe^{(1-p)a} + (1-p)e^{-pa} \leqslant 1 + p(1-p)\frac{a^2}{2}(1+\varepsilon),$$

и соотношение (7.3) вытекает из неравенства $1 + x \leqslant e^x$.

Используя такое δ, мы докажем неравенство (7.2) индукцией по глубине M дерева решений. Для $M = 0$ значение Y есть константа, и (7.2) следует немедленно. В противном случае обозначим соответственно через $p, c, v = p(1-p)c^2$ вероятность, результат и дисперсию первого вопроса Пола. Через μ_y, μ_n обозначим условное математическое ожидание величины Y при ответах Кэрол ДА и НЕТ, соответственно. Тогда $0 = \mathbf{E}[Y]$ может быть представлено в виде

$$0 = p\mu_y + (1-p)\mu_n.$$

Разность $\mu_y - \mu_n$ есть ожидаемое *изменение* величины Y при условии, что все другие выборы были сделаны независимо с соответствующими вероятностями, а корневой выбор изменился с ДА на НЕТ. Поскольку это всегда изменяет Y не более чем на c, имеем

$$|\mu_y - \mu_n| \leqslant c.$$

Поэтому мы можем ввести параметр b, $|b| \leqslant c$, для которого

$$\mu_y = (1-p)b \quad \text{и} \quad \mu_n = -pb.$$

В силу соотношения (7.3) имеем

$$pe^{\lambda\mu_y} + (1-p)e^{\lambda\mu_n} \leqslant e^{(1+\varepsilon)p(1-p)b^2\lambda^2/2} \leqslant e^{(1+\varepsilon)v\lambda^2/2}.$$

Обозначим через A_y математическое ожидание величины $e^{\lambda(Y-\mu_y)}$ при условии, что первый ответ Кэрол был ДА, а через A_n — аналогичную величину при ее ответе НЕТ. Если ответ Кэрол на первый вопрос известен, Пол имеет дерево решений (одно из двух основных поддеревьев), которое определяет Y с полной дисперсией, не превосходящей $\sigma^2 - v$, а глубина дерева не больше $M-1$. Отсюда по предположению индукции $A_y, A_n \leqslant A^-$, где

$$A^- = e^{(1+\varepsilon)\lambda^2(\sigma^2-v)/2}.$$

Теперь мы разлагаем

$$\begin{aligned}\mathbf{E}[e^{\lambda Y}] &= pe^{\lambda\mu_y}A_y + (1-p)e^{\lambda\mu_n}A_n \leqslant \\ &\leqslant [pe^{\lambda\mu_y} + (1-p)e^{\lambda\mu_n}]A^- \leqslant e^{(1+\varepsilon)\lambda^2(v+(\sigma^2-v))/2}, \qquad (7.4)\end{aligned}$$

что завершает доказательство неравенства (7.2), а, значит, и теоремы 7.4.3. ■

Заметим, что это формальное индуктивное доказательство слегка затеняет роль мартингала. Мартингал $\mathbf{E}[Y] = Y_0, \ldots, Y_M = Y$ можно определить как условное математическое ожидание Y_t величины Y после первых t вопросов и ответов. Теорему 7.4.3 можно понимать как оценку хвоста величины Y хвостом нормального распределения с большей или равной дисперсией. Для очень больших уклонений от среднего и больших α эта оценка оказывается неверной.

7.5. ЧЕТЫРЕ ПРИМЕРА

Пусть g — случайная функция, отображающая множество $\{1, \ldots, n\}$ само на себя, причем все n^n возможных функций равновероятны. Пусть $L(g) = n - val(g)$, где $val(g)$ — размер области значений функции g, или что то же, число тех y, для которых уравнение $g(x) = y$ имеет решение. Из линейности математического ожидания следует, что

$$\mathbf{E}[L(g)] = n\left(1 - \frac{1}{n}\right)^n \sim \frac{n}{e}.$$

Положим $B_i = \{1, \ldots, i\}$. Функционал L удовлетворяет условию Липшица относительно этой шкалы, так как изменение значения $g(i)$ может изменить $L(g)$ не более, чем на 1. Следовательно, имеем следующий результат.

Теорема 7.5.1. *Справедливо неравенство*

$$\Pr\left[|L(g) - \frac{n}{e}| > \lambda\sqrt{n} + 1\right] < 2e^{-\lambda^2/2}.$$

Выводить эти асимптотические оценки без использования мартингалов довольно обременительно.

Второй пример: пусть B — произвольное нормированное пространство, и пусть $v_1, \ldots, v_n \in B$, причем все $|v_i| \leqslant 1$. Пусть случайные величины $\varepsilon_1, \ldots, \varepsilon_n$ независимы, и пусть

$$\Pr[\varepsilon_i = +1] = \Pr[\varepsilon_i = -1] = \frac{1}{2}.$$

Положим

$$X = |\varepsilon_1 v_1 + \ldots + \varepsilon_n v_n|.$$

Теорема 7.5.2. *Имеют место соотношения*

$$\Pr[X - \mathbf{E}[X] > \lambda\sqrt{n}] < e^{-\lambda^2/2},$$
$$\Pr[X - \mathbf{E}[X] < -\lambda\sqrt{n}] < e^{-\lambda^2/2}.$$

Доказательство. В качестве вероятностного пространства рассмотрим множество $\{-1, +1\}^n$, в котором все элементы $(\varepsilon_1, \ldots, \varepsilon_n)$ равновероятны. Тогда X является случайной величиной на этом пространстве. Определим мартингал $X_0, \ldots, X_n = X$, проявляя по одному ε_i для очередного $i > 0$. При этом $X_0 = \mathbf{E}[X]$. Значение ε_i может изменить X на 2, так что прямое применение теоремы 7.4.1 дает $|X_{i+1} - X_i| \leqslant 2$. Пусть $\varepsilon, \varepsilon'$ — два вектора длины n,

различающиеся только в i-й координате:

$$X_i(\varepsilon) = \frac{1}{2}\left[X_{i+1}(\varepsilon) + X_{i+1}(\varepsilon')\right],$$

так что

$$|X_i(\varepsilon) - X_{i+1}(\varepsilon)| = \frac{1}{2}\,|X_{i+1}(\varepsilon') - X_{i+1}(\varepsilon)| \leqslant 1.$$

Теперь применяем неравенство Ацумы. ■

Третий пример: пусть ρ — метрика Хэмминга на множестве $\{0,1\}^n$. Для $A \subseteq \{0,1\}^n$ обозначим через $B(A,s)$ множество таких $y \in \{0,1\}^n$, что $\rho(x,y) \leqslant s$ для некоторого $x \in A$. (Ясно, что $A \subseteq B(A,s)$, так как мы можем взять $x = y$.)

Теорема 7.5.3. *Пусть величины $\varepsilon, \lambda > 0$ связаны соотношением $e^{-\lambda^2/2} = \varepsilon$. Тогда*

$$|A| \geqslant \varepsilon 2^n \Rightarrow |B(A, 2\lambda\sqrt{n})| \geqslant (1-\varepsilon)2^n.$$

Доказательство. Рассмотрим в качестве вероятностного пространства множество $\{0,1\}^n$, в котором все точки равновероятны. Для $y \in \{0,1\}^n$ положим

$$X(y) = \min_{x \in A} \rho(x,y).$$

Пусть $X_0, X_1, \ldots, X_n = X$ — мартингал, полученный проявлением одной координаты из $\{0,1\}^n$ на каждом шаге. Условие Липшица для X выполняется, ибо, если y, y' различаются только в одной координате, то $|X(y) - X(y')| \leqslant 1$. Следовательно, если $\mu = \mathbf{E}[X]$, то

$$\Pr[X < \mu - \lambda\sqrt{n}] < e^{-\lambda^2/2} = \varepsilon,$$
$$\Pr[X > \mu + \lambda\sqrt{n}] < e^{-\lambda^2/2} = \varepsilon.$$

Но

$$\Pr[X = 0] = |A|2^{-n} \geqslant \varepsilon,$$

так что $\mu \leqslant \lambda\sqrt{n}$. Значит,

$$\Pr[X > 2\lambda\sqrt{n}] < \varepsilon$$

и

$$|B(A, 2\lambda\sqrt{n})| = 2^n \Pr[X \leqslant 2\lambda\sqrt{n}] \geqslant 2^n(1-\varepsilon).$$ ■

В действительности, известен гораздо более сильный результат. Через $B(s)$ обозначим шар радиуса s с центром в точке $(0,\ldots,0)$. Изопериметрическое неравенство, доказанное Харпером [Harper (1966)], заключается в том, что

$$|A| \geqslant |B(r)| \Rightarrow |B(A,s)| \geqslant |B(r+s)|.$$

Фактически, это неравенство можно использовать как начало альтернативного доказательства того, что $\chi(G) \sim n/2\log_2 n$, а также для вывода еще нескольких результатов, которые мы доказали с помощью мартингалов.

Мы иллюстрируем теорему 7.4.3 ключевой технической леммой (в упрощенной форме) из [Alon et al. (1997)]. Пусть $G = (V, E)$ — граф на N вершинах, степень каждой вершины равна D. Асимптотика рассматривается при $N, D \to \infty$.

Положим $p = 1/D$. Определим случайный подграф $H \subseteq G$, включая в него каждое ребро $e \in E$ независимо с вероятностью p. Обозначим через M (от слова matching — паросочетание) множество изолированных ребер графа H. Пусть множество V^* состоит из вершин $v \in V$, *не* входящих ни в одну пару $\{v, w\} \in M$. Для $v \in V$ положим $\deg^*(v)$ равным числу таких $w \in V^*$, что $\{v, w\} \in E$. Так как

$$\Pr[v \notin V^*] = \sum_{\{v,w\} \in E} p(1-p)^{2D-1} = e^{-2} + O(D^{-1}),$$

из линейности математического ожидания следует, что

$$\mathbf{E}[\deg^*(v)] = D(1 - e^{-2}) + O(1).$$

Нам нужно, чтобы $\deg^*(v)$ было плотно сосредоточено вокруг своего среднего значения.

В обозначениях теоремы 7.4.3 вероятностное пространство определяется выбором для каждого $e \in E$, верно ли $e \in H$. Все $p_i = p$. Изменение $e \in H$ на $e \notin H$ может изменить $\deg^*(v)$ не более, чем на $C = 4$.

Полу нужно найти $\deg^*(v)$ с помощью запросов вида «$e \in H$?» Для всех таких w, что $\{v, w\} \in E$, он определяет, верно ли, что $w \in V^*$, с помощью следующей линии опроса. Прежде всего, для всех таких u, что $\{w, u\} \in E$, он запрашивает, верно ли, что $\{w, u\} \in H$. Если это неверно для всех u, то $w \in V^*$. Если два (или более) $\{w, u_1\}, \{w, u_2\} \in H$, то w не может быть в *изолированном* ребре H, так что $w \in V^*$. Теперь предположим, что $\{w, u\} \in H$ в точности для одного u. Затем Пол спрашивает (используя полученную информацию!) для каждого $z \neq w$ с $\{u, z\} \in E$, верно ли, что $\{u, z\} \in H$. Ответы определяют, является ли $\{w, u\}$ изолированным ребром в графе H, а значит, верно ли, что $w \in V^*$. Пол сделал не более $D + (D-1)$ запросов для каждого w, всего не более $D(2D-1) = O(D^2)$ запросов. Мы делаем вывод, что

$$\Pr[|\deg^*(v) - D(1 - e^{-2})| > \lambda D^{1/2}] = \exp[-\Omega(\lambda^2)]$$

при $\lambda \to \infty$ и $\lambda = o(D^{1/2})$.

В приложениях бывает желательно итерировать эту процедуру (применяемую здесь при ограничении графа G на множестве V^*) для получения большого паросочетания. Это отчасти напоминает метод Рёдля из разд. 4.7. Есть много других трудностей, но плотная концентрация $\deg^*(v)$ вокруг своего среднего значения играет особую роль.

7.6. НЕРАВЕНСТВО ТАЛАГРАНА

Пусть $\Omega = \prod_{i=1}^n \Omega_i$, где каждое Ω_i — вероятностное пространство, и мера на Ω задана как произведение мер. Пусть $A \subseteq \Omega$ и $\vec{x} = (x_1, \ldots, x_n) \in \Omega$. В статье [Talagrand (1996)] дано необычное, изящное и чрезвычайно мощное понятие расстояния $\rho(A, \vec{x})$ от $\vec{x}$ до A. Представим себе движение из $\vec{x}$ к некоторому $\vec{y} = (y_1, \ldots, y_n) \in A$ посредством изменения координат. Обозначим

через $\rho(A, \vec{x})$ минимальную цену такого движения, когда определенным образом ограниченный в возможностях соперник задает цену каждого изменения.

Определение 2. *Обозначим через $\rho(A, \vec{x})$ наименьшую величину, для которой при любом $\vec{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$ с $|\vec{\alpha}| = 1$ существует такой вектор $\vec{y} = (y_1, \ldots, y_n) \in A$, что*

$$\sum_{x_i \neq y_i} \alpha_i \leqslant \rho(A, \vec{x}).$$

Заметим, что вектор $\vec{y}$ может зависеть, и, как правило, зависит от $\vec{\alpha}$.

Для любого действительного $t \geqslant 0$ положим

$$A_t = \{\vec{x} \in \Omega : \rho(A, \vec{x}) \leqslant t\}.$$

Заметим, что $A_0 = A$, так как при $\vec{x} \in A$ можно выбрать $\vec{y} = \vec{x}$.

Неравенство Талаграна: $\Pr[A](1 - \Pr[A_t]) \leqslant e^{-t^2/4}$.

В частности, если $\Pr[A] \geqslant \frac{1}{2}$ (или больше любой фиксированной константы) и t «очень велико», то все элементы пространства Ω, за исключением очень малой их доли, находятся от A на «расстоянии», не превышающем t.

Пример. Рассмотрим пространство $\Omega = \{0,1\}^n$ с равномерным распределением, и обозначим через τ метрику Хэмминга (L^1). Тогда $\rho(A, \vec{x}) \geqslant \min\limits_{\vec{y} \in A} \tau(\vec{x}, \vec{y}) n^{-1/2}$, так как соперник может выбрать все $\alpha_i = n^{-1/2}$. Предположим, что для того, чтобы переместиться из $\vec{x}$ в A, должны быть изменены значения $x_1, \ldots, x_l$ (или любых других l координат). Тогда $\rho(A, \vec{x}) \geqslant l^{1/2}$, так как соперник может выбрать $\alpha_i = l^{-1/2}$ при $1 \leqslant i \leqslant l$, и ноль в противном случае.

Обозначим через $U(A, \vec{x})$ множество наборов $\vec{s} = (s_1, \ldots, s_n) \in \{0,1\}^n$, обладающих свойством, что существует $\vec{y} \in A$, для которого

$$x_i \neq y_i \Rightarrow s_i = 1.$$

Можно считать, что $U(A, \vec{x})$ — множество возможных путей из $\vec{x}$ в A. Заметим, что когда $s_i = 1$, мы, по некоторым техническим соображениям, не требуем условия $x_i \neq y_i$. В этих обозначениях $\rho(A, \vec{x})$ — это наименьшее из таких действительных чисел, что при всех $\vec{\alpha}$ с $|\vec{\alpha}| = 1$ существует $\vec{s} \in U(A, \vec{x})$, для которого выполнено неравенство $\vec{\alpha} \cdot \vec{s} \leqslant \rho(A, \vec{x})$.

Пусть $V(A, \vec{x})$ — это выпуклая оболочка множества $U(A, \vec{x})$. Следующий результат придает другой смысл величине ρ, открывающий бóльшую глубину этого понятия.

Теорема 7.6.1. *Справедливо равенство*

$$\rho(A, \vec{x}) = \min_{\vec{v} \in V(A, \vec{x})} |\vec{v}|.$$

Доказательство. Пусть этот минимум достигается в $\vec{v} \in V(A, \vec{x})$. Тогда гиперплоскость, проходящая через $\vec{v}$ и перпендикулярная прямой, соединяющей

начало координат с $\vec{v}$, отделяет $V(A,\vec{x})$ от начала координат. Следовательно, для всех $\vec{s}\in V(A,\vec{x})$ выполнено неравенство $\vec{s}\cdot\vec{v}\geqslant\vec{v}\cdot\vec{v}$. Положим $\vec{\alpha}=\vec{v}/|\vec{v}|$. Тогда для всех $\vec{s}\in U(A,\vec{x})\subseteq V(A,\vec{x})$ верно соотношение $\vec{s}\cdot\vec{\alpha}\geqslant\vec{v}\cdot\vec{v}/|\vec{v}|=|\vec{v}|$. С другой стороны, возьмем любое $\vec{\alpha}$ с $|\vec{\alpha}|=1$. Тогда $\vec{\alpha}\cdot\vec{v}\leqslant|\vec{v}|$. Так как $\vec{v}\in V(A,\vec{x})$, мы можем записать, что $\vec{v}=\sum\lambda_i\vec{s}_i$ для некоторых $\vec{s}_i\in U(A,\vec{x})$, где $\lambda_i\geqslant 0$ и $\sum\lambda_i=1$. Тогда

$$|\vec{v}|\geqslant\sum\lambda_i(\vec{\alpha}\cdot\vec{s}_i),$$

а, значит, некоторое произведение $\vec{\alpha}\cdot\vec{s}_i\leqslant|\vec{v}|$. ■

Случай $\Omega=\{0,1\}^n$ особенно важен и поучителен. Здесь $\rho(A,\vec{x})$ — просто евклидово расстояние от $\vec{x}$ до выпуклой оболочки A.

Теорема 7.6.2. *Справедливо неравенство*

$$\int_\Omega\exp\left[\frac{1}{4}\rho^2(A,\vec{x})\right]d\vec{x}\leqslant\frac{1}{\Pr[A]}.$$

Теорема Талаграна — прямое следствие этого результата. Действительно, зафиксируем A и рассмотрим случайную величину $X=\rho(A,\vec{x})$. Тогда

$$\Pr[\overline{A_t}]=\Pr[X\geqslant t]=\Pr[e^{X^2/4}\geqslant e^{t^2/4}]\leqslant\mathbf{E}[e^{X^2/4}]e^{-t^2/4},$$

и по утверждению теоремы $\mathbf{E}[e^{X^2/4}]\leqslant\frac{1}{\Pr[A]}$.

Доказательство теоремы 7.6.2. Будем вести индукцию по размерности n. Пусть $n=1$. Тогда $\rho(A,\vec{x})=1$, если $\vec{x}\notin A$, и $\rho(A,\vec{x})=0$ в противном случае, так что

$$\int\exp\left[\frac{1}{4}\rho^2(A,\vec{x})\right]d\vec{x}=\Pr[A]+(1-\Pr[A])e^{1/4}\leqslant\frac{1}{\Pr[A]},$$

ибо, как нетрудно видеть, $u+(1-u)e^{1/4}\leqslant u^{-1}$ при $0<u\leqslant 1$.

Предположим, что утверждение верно для n. Положим $\mathrm{OLD}=\prod_{i=1}^n\Omega_i$, $\mathrm{NEW}=\Omega_{n+1}$, так что $\Omega=\mathrm{OLD}\times\mathrm{NEW}$ и любое $\vec{z}\in\Omega$ может быть единственным образом представлено в виде $\vec{z}=(\vec{x},\omega)$ при $\vec{x}\in\mathrm{OLD}$, $\omega\in\mathrm{NEW}$. Положим

$$B=\{\vec{x}\in\mathrm{OLD}:(\vec{x},\omega)\in A\text{ для некоторого }\omega\in\mathrm{NEW}\},$$

и для любого $\omega\in\mathrm{NEW}$ пусть

$$A_\omega=\{\vec{x}\in\mathrm{OLD}:(\vec{x},\omega)\in A\}.$$

При $\vec{z}=(\vec{x},\omega)\in\Omega$ мы можем двигаться к A двумя основными путями — изменяя ω, что сводит задачу к перемещению из $\vec{x}$ в B, или, не изменяя ω, что сводит задачу к перемещению из $\vec{x}$ в A_ω. Таким образом,

$$\vec{s}\in U(B,\vec{x})\Rightarrow(\vec{s},1)\in U(A,(\vec{x},\omega))$$

и

$$\vec{t}\in U(A_\omega,\vec{x})\Rightarrow(\vec{t},0)\in U(A,(\vec{x},\omega)).$$

Рассмотрим выпуклые оболочки. Если $\vec{s} \in V(B, \vec{x})$ и $\vec{t} \in V(A_\omega, \vec{x})$, то $(\vec{s}, 1)$ и $(\vec{t}, 0)$ принадлежат $V(A, (\vec{x}, \omega))$, а, значит, для любого $\lambda \in [0, 1]$

$$((1-\lambda)\vec{s} + \lambda\vec{t}, 1-\lambda) \in V(A, (\vec{x}, \omega)).$$

Тогда, в силу выпуклости, имеем

$$\rho^2(A, (\vec{x}, \omega)) \leqslant (1-\lambda)^2 + |(1-\lambda)\vec{s} + \lambda\vec{t}|^2 \leqslant (1-\lambda)^2 + (1-\lambda)|\vec{s}|^2 + \lambda|\vec{t}|^2.$$

Выбор $\vec{s}$ и $\vec{t}$ с минимальными нормами дает нам важное неравенство

$$\rho^2(A, (\vec{x}, \omega)) \leqslant (1-\lambda)^2 + \lambda\rho^2(A_\omega, \vec{x}) + (1-\lambda)\rho^2(B, \vec{x}).$$

Процитируем Талаграна: «Главный прием доказательства — избежать сейчас соблазна оптимизировать по λ.» Вместо этого сначала зафиксируем ω и оценим интеграл сверху:

$$\int \exp\left[\frac{1}{4}\rho^2(A, (\vec{x}, \omega))\right] d\vec{x} \leqslant$$
$$\leqslant e^{(1-\lambda)^2/4} \int \left(\exp\left[\frac{1}{4}\rho^2(A_\omega, \vec{x})\right]\right)^\lambda \left(\exp\left[\frac{1}{4}\rho^2(B, \vec{x})\right]\right)^{1-\lambda} d\vec{x}.$$

По неравенству Гёльдера это не превосходит значения

$$e^{(1-\lambda)^2/4} \left[\int \exp\left[\frac{1}{4}\rho^2(A_\omega, \vec{x})\right] d\vec{x}\right]^\lambda \left[\int \exp\left[\frac{1}{4}\rho^2(B, \vec{x})\right] d\vec{x}\right]^{1-\lambda},$$

что по предположению индукции не превышает

$$e^{(1-\lambda)^2/4} \left(\frac{1}{\Pr[A_\omega]}\right)^\lambda \left(\frac{1}{\Pr[B]}\right)^{1-\lambda} = \frac{1}{\Pr[B]} e^{(1-\lambda)^2/4} r^{-\lambda},$$

где $r = \Pr[A_\omega]/\Pr[B] \leqslant 1$. *Теперь* мы минимизируем функцию $e^{(1-\lambda)^2/4} r^{-\lambda}$, выбирая $\lambda = 1 + 2\ln r$ при $e^{-1/2} \leqslant r \leqslant 1$ и $\lambda = 0$ — в противном случае. Дальнейшие (несколько громоздкие, но простые) вычисления показывают, что $e^{(1-\lambda)^2/4} r^{-\lambda} \leqslant 2 - r$ при таком $\lambda = \lambda(r)$. Следовательно,

$$\int \exp\left[\frac{1}{4}\rho^2(A, (\vec{x}, \omega))\right] d\vec{x} \leqslant \frac{1}{\Pr[B]} \left(2 - \frac{\Pr[A_\omega]}{\Pr[B]}\right).$$

Интегрирование по ω дает

$$\iint \exp\left[\frac{1}{4}\rho^2(A, (\vec{x}, \omega))\right] d\vec{x} d\omega \leqslant \frac{1}{\Pr[B]} \left(2 - \frac{\Pr[A]}{\Pr[B]}\right) = \frac{1}{\Pr[A]} y(2-y),$$

где $y = \Pr[A]/\Pr[B] \in [0, 1]$. Но $y(2-y) \leqslant 1$, что завершает индукцию, а вместе с ней и доказательство теоремы. ■

7.7. ПРИЛОЖЕНИЯ НЕРАВЕНСТВА ТАЛАГРАНА

Пусть $\Omega = \prod_{i=1}^n \Omega_i$, где каждое Ω_i — вероятностное пространство, и мера на Ω задана как произведение мер. Рассмотрим функцию $h : \Omega \to \mathbb{R}$. Неравенство

Талаграна дает нам возможность при определенных условиях показать, что случайная величина $X = h(\cdot)$ плотно сконцентрирована. При этом неравенство Талаграна может играть ту же роль, что играет неравенство Ацумы для мартингалов, и во многих случаях оно дает гораздо более сильные результаты.

Говорят, что функция $h : \Omega \to \mathbb{R}$ *удовлетворяет условию Липшица*, если $|h(x) - h(y)| \leqslant 1$ для любых x и y, отличающихся не более чем в одной координате. Неравенство Талаграна наиболее эффективно на тех липшицевых функциях, для которых из выполнения неравенства $h(x) \geqslant s$ следует, что существует сравнительно небольшое число координат, обеспечивающих выполнение неравенства $h(x) \geqslant s$. Мы формализуем эту идею следующим образом.

Определение 3. *Пусть* $f : \mathbb{N} \to \mathbb{N}$. *Функция* h *называется* проверяемой со сложностью f, *если при* $h(x) \geqslant s$ *существует такое множество* $I \subseteq \{1, \ldots, n\}$ *с* $|I| \leqslant f(s)$, *что для всех* $y \in \Omega$, *совпадающих с* x *в координатах из* I, *выполняется неравенство* $h(y) \geqslant s$.

Пример. Рассмотрим граф $G(n, p)$ как результат $\binom{n}{2}$ подбрасываний монеты, и пусть $h(G)$ — число треугольников в G. Тогда h — функция, проверяемая со сложностью $f(s) = 3s$. Это объясняется тем, что если $h(G) \geqslant s$, то существует s треугольников, у которых в сумме не более $3s$ ребер, а у любого другого графа G' с этими $3s$ ребрами $h(G') \geqslant s$. Заметим, что множество I (здесь — множество индексов этих $3s$ ребер) очень сильно зависит от G. Отметим еще, что нам необходимы только нижние оценки для h.

Теорема 7.7.1. *При сделанных предположениях для всех* b *и* t *справедливо неравенство*

$$\Pr[X \leqslant b - t\sqrt{f(b)}] \Pr[X \geqslant b] \leqslant e^{-t^2/4}.$$

Доказательство. Положим $A = \{x : h(x) < b - t\sqrt{f(b)}\}$. Пусть теперь $h(y) \geqslant b$. Мы утверждаем, что $y \notin A_t$. Пусть I — множество индексов размера не больше $f(b)$, гарантирующее условие $h(y) \geqslant b$, как показано выше. Положим $\alpha_i = 0$ при $i \notin I$, $\alpha_i = |I|^{-1/2}$ при $i \in I$. Если $y \in A_t$, то существует $z \in A$, отличающееся от y не более чем в $t|I|^{1/2} \leqslant t\sqrt{f(b)}$ координатах из I (и в произвольном числе координат вне I). Пусть y' совпадает с y на I и совпадает с z вне I. Гарантируется, что $h(y') \geqslant b$. Далее, y' и z отличаются не более чем в $t\sqrt{f(b)}$ координатах, а значит из условия Липшица следует, что

$$h(z) \geqslant h(y') - t\sqrt{f(b)} \geqslant b - t\sqrt{f(b)},$$

но тогда $z \notin A$, противоречие. Так что $\Pr[X \geqslant b] \leqslant \Pr[\overline{A_t}]$, и по теореме Талаграна

$$\Pr[X < b - t\sqrt{f(b)}] \Pr[X \geqslant b] \leqslant e^{-t^2/4}.$$

Так как правая часть непрерывна по t, мы можем заменить $<$ на $\leqslant$, получив тем самым утверждение теоремы. ■

Иногда оказывается полезным небольшое обобщение. Функцию $h : \Omega \to \mathbb{R}$ назовем K-*липшицевой*, если $|h(x) - h(y)| \leqslant K$, когда x и y различаются только в одной координате. Применяя предыдущую теорему к функции h/K,

удовлетворяющей обычному условию Липшица, мы получаем

$$\Pr[X \leqslant b - tK\sqrt{f(b)}]\Pr[X \geqslant b] \leqslant e^{-t^2/4}.$$

На практике часто в качестве b выбирают медиану, так что для больших t вероятность оказаться на $t\sqrt{f(b)}$ ниже медианы резко стремится к нулю. С другой стороны, выбирая параметр так, что $m = b - t\sqrt{f(b)}$ — медиана, как правило получаем $b \sim m + t\sqrt{f(m)}$, и вероятность оказаться на $t\sqrt{f(b)}$ выше медианы резко стремится к нулю. Мартингалы с использованием неравенства Ацумы обычно дают результат, сконцентрированный вокруг μ, среднего значения X, в то время как неравенство Талаграна дает результат, сконцентрированный вокруг медианы m. Обычно средние значения легко вычислимы. Медианы, как известно, сложно вычислимы, но при плотной концентрации результата, как правило, можно показать, что среднее значение и медиана близки друг к другу.

Пусть $x = (x_1, \dots, x_n)$, где величины x_i выбраны независимо и равномерно из отрезка $[0, 1]$. Пусть $X = h(x)$ — длина самой длинной возрастающей подпоследовательности x. С помощью простейших методов можно выяснить, что почти наверное $c_1 n^{1/2} < X < c_2 n^{1/2}$ для некоторых положительных констант c_1, c_2, и что μ (среднее значение величины X), и m (медиана X) входят в этот диапазон. Кроме того, X удовлетворяет условию Липшица, так как изменение одного x_i может изменить X не более чем на единицу. Насколько сконцентрированы значения X? С помощью неравенства Ацумы можно доказать, что если $s \gg n^{1/2}$, то почти наверное $|X - \mu| \leqslant s$. Но этот результат не очень хорош, так как порядок самого X — лишь $n^{1/2}$. Теперь рассмотрим неравенство Талаграна. Случайная величина X проверяема со сложностью $f(s) = s$, так как, если в x есть возрастающая подпоследовательность длины s, то эти s координат гарантируют, что $X \geqslant s$. Тогда $\Pr[X < m - tm^{1/2}] \leqslant e^{-t^2/4}/\Pr[X \geqslant m] \leqslant 2e^{-t^2/4}$, так как m — медиана. Но $m = \Theta(n^{1/2})$. Значит, при $s \gg n^{1/4}$ почти наверное $X > m - s$. С другой стороны, пусть t медленно стремится к бесконечности, и пусть b удовлетворяет равенству $b - tb^{1/2} = m$. Тогда $\Pr[X \geqslant b] \leqslant e^{-t^2/4}/\Pr[X \leqslant m] \leqslant 2e^{-t^2/4}$. Следовательно, $X \leqslant b$ почти наверное. Но $b = m + (1 + o(1))tm^{1/2}$, так что $X \leqslant m + tm^{1/2}$ почти наверное. Объединяя, получим, что если $s \gg n^{1/4}$, то $|X - m| < s$ почти наверное. Гораздо более сильный результат, определяющий точное асимптотическое распределение X, был недавно получен в работе [Baik, Deift and Johansson (1999)] с помощью глубоких аналитических методов.

Рассмотрим еще раз теорему 7.3.2. Оценивается вероятность того, что граф $G(n, \frac{1}{2})$ не имеет клик определенного размера k. Пусть, как и в теореме 7.3.2, Y — максимальное число непересекающихся по ребрам k-клик. Было доказано, что $\mathbf{E}[Y] = \Omega(n^2 k^{-4})$ и Y плотно сконцентрировано вокруг $\mathbf{E}[Y]$, так что m, являющееся медианой Y, должно также удовлетворять равенству $m = \Omega(n^2 k^{-4})$. Как и раньше, Y удовлетворяет условию Липшица. Более того, Y проверяемо со сложностью $f(s) = \binom{k}{2}s$, так как наличие в графе множества

ребер некоторых s клик размера k гарантирует, что $Y \geqslant s$. Следовательно,

$$\Pr\left[Y \leqslant m - tm^{1/2}\binom{k}{2}^{1/2}\right]\Pr[Y \geqslant m] < e^{-t^2/4}.$$

Положим $t = \Theta(m^{1/2}/k)$, так что $m = tm^{1/2}\binom{k}{2}^{1/2}$. Тогда

$$\Pr[\omega(G) < k] = \Pr[Y \leqslant 0] < 2e^{-t^2/4} < \exp\left[-\Omega\left(\frac{n^2}{\ln^6 n}\right)\right],$$

что улучшает оценку теоремы 7.3.2. Тем не менее, надо отметить, что с помощью обобщенного неравенства Янсона в разд. 10.3, оценку удается усилить еще больше.

7.8. ПОЛИНОМИАЛЬНАЯ КОНЦЕНТРАЦИЯ КИМА—ВУ

Многообещающе выглядит подход, использованный в недавней работе [Kim and Vu (2000)]. Пусть $H = (V(H), E(H))$ — гиперграф, и пусть у каждого ребра $e \in E(H)$ есть неотрицательный вес $w(e)$. Пусть t_i, $i \in V(H)$, — попарно независимые индикаторы с $\mathbf{E}[t_i] = p_i$. Рассмотрим многочлен от случайных величин

$$Y = \sum_{e \in E(H)} w_e \prod_{i \in e} t_i.$$

Мы допускаем $e = \varnothing$, в этом случае $\prod_{i\in e} t_i$ полагается равным 1. Нам нужно показать, что Y сосредоточено вокруг его среднего значения.

Пусть $S \subseteq V(H)$ — случайное множество, $\Pr[i \in S] = p_i$, эти события попарно независимы для $i \in V(H)$. Тогда Y — взвешенное число гиперребер e в подгиперграфе H, порожденном множеством S. На практике все веса, как правило, равны единице, так что Y просто равно числу гиперребер в случайном S. Но мы также можем рассматривать Y как произвольный абстрактный многочлен от индикаторов t_i, у которого все коэффициенты не отрицательны.

Пусть $n = |V(H)|$ — число вершин в гиперграфе H (число переменных t_i). Пусть k — наибольшее возможное значение для размера гиперребер (наибольшее возможное значение степени полинома Y).

Пусть $A \subseteq V(H)$ и $|A| \leqslant k$. Мы сокращаем Y до Y_A следующим образом. Для тех слагаемых $\prod_{i\in e} t_i$, у которых $A \subseteq e$, мы полагаем $t_i = 1$ для всех $i \in A$, заменяя слагаемое на $\prod_{i\in e-A} t_i$. Все остальные слагаемые (где e не содержит A) удаляются. Например, при $A = \{1\}$, многочлен $2t_1t_2 + 5t_1t_3t_4 + 7t_2t_4$ переходит в $2t_2 + 5t_3t_4$. Интересно, что, как полином от t_i, случайная величина Y_A представляет собой частную производную Y по переменным t_i, $i \in A$. Положим $E_A = \mathbf{E}[Y_A]$. То есть E_A — ожидаемое число содержащих A гиперребер в S, при условии, что все вершины A принадлежат S. Положим E_i равным максимальному E_A из всех $A \subseteq V(H)$ размера i. Для удобства положим $\mu = \mathbf{E}[Y]$. Пусть

$$E'' = \max_{1\leqslant i\leqslant k} E_i \text{ и } E' = \max[\mu, E''].$$

Теорема 7.8.1 (полиномиальная концентрация Кима—Ву). *В сделанных выше предположениях*

$$\Pr[|Y-\mu|>a_k(E'E'')^{1/2}\lambda^k]<d_ke^{-\lambda}n^{k-1}$$

для любого $\lambda>1$.

Для определенности мы можем здесь положить $a_k=8^kk!^{1/2}$ и $d_k=2e^2$.

Мы опускаем доказательство, в котором применяются неравенства для мартингалов, похожие на неравенства из теоремы 7.4.3, а также изящная индукция по степени k. Остается простор для улучшения оценок сомножителей a_k, d_k и n^{k-1}. На практике обычно k зафиксировано, и $\lambda\gg\ln n$, так что сомножитель $e^{-\lambda}$ является главным в оценке искомой вероятности.

Применение полиномиальной концентрации Кима—Ву, как правило, не вызывает затруднений. Пусть G — некоторая реализация случайного графа $G(n,p)$ с $p=n^{-\alpha}$, и пусть $0<\alpha<2/3$. Выберем вершину x в G и обозначим через $Y=Y(x)$ число треугольников, содержащих x. Положим $\mu=\mathbf{E}[Y]=\binom{n-1}{2}p^3\sim\frac12n^{2-3\alpha}$. Пусть $\delta>0$ зафиксировано. Нам нужно оценить сверху вероятность $\Pr[|Y-\mu|<\delta\mu]$.

Случайный граф G определен случайными величинами t_{ij}, индикаторами смежности пар вершин, по одной для каждой неупорядоченной пары вершин. В этом контексте

$$Y=\sum_{i,j\ne x}t_{xi}t_{xj}t_{ij}.$$

Это — полином степени $k=3$. Если A состоит из единственного ребра xi, получим $E_A=(n-2)p^2$; если оно состоит из трех ребер, составляющих треугольник, содержащий вершину x, получим $E_A=1$. Если $A=\varnothing$, $E_A=\mu$. Другие случаи дают меньшие E_A. Ясно, что $E''\sim\max[np^2,1]$. Вычисления показывают, что $E''\sim c\mu n^{-\varepsilon}$ для некоторого положительного ε (зависящего от α) на всем рассматриваемом диапазоне. Мы применяем полиномиальную концентрацию Кима—Ву с $\lambda=c'n^{\varepsilon/6}$, где c' — малая положительная константа, чтобы ограничить $\Pr[|Y-\mu|<\delta\mu]$ сверху величиной $\exp[-\Omega(n^{\varepsilon/6})]$. Заметим, что множитель n^{k-1} поглощается экспонентой.

В частности, так как эта вероятность равна $o(n^{-1})$, почти наверное каждая вершина x входит в $\sim\mu$ треугольников. Этот результат можно обобщить[1)]. Зафиксируем $\alpha\in(0,1)$, и пусть (R,H) — граф с корнем, безопасный в смысле разд. 10.7, относительно α. Пусть G — некоторая реализация случайного графа $G(n,p)$ с $p=n^{-\alpha}$. Для различных вершин $x_1,\dots,x_r$ пусть $Y=Y(x_1,\dots,x_r)$ обозначает число расширений G относительно H. Положим $\mu=\mathbf{E}[Y]$. Полиномиальная концентрация Кима—Ву дает экспоненциально малую верхнюю оценку вероятности того, что Y не близко к μ. В частности, эта вероятность равна $o(n^{-r})$. Следовательно, почти наверное у любых r вершин имеется $\sim\mu$ расширений относительно H.

[1)] Дальнейшее может быть понято после ознакомления с разд. 10.7. — *Прим. ред.*

7.9. УПРАЖНЕНИЯ

1. Пусть $G = (V, E)$ — граф, вершины которого — все 7^n векторов длины n над $\mathbb{Z}_7$, в котором две вершины смежны тогда и только тогда, когда они различаются ровно в одной координате. Пусть $U \subset V$ — множество из 7^{n-1} вершин графа G, и пусть W — множество всех вершин G, расстояние от которых до U превышает $(c+2)\sqrt{n}$, где $c > 0$ — константа. Доказать, что $|W| \leqslant 7^n \cdot e^{-c^2/2}$.

2*. Пусть $G = (V, E)$ — граф с хроматическим числом $\chi(G) = 1000$. Пусть $U \subset V$ — случайное подмножество множества вершин V, выбранное равномерно из всех $2^{|V|}$ подмножеств V. Пусть $H = G[U]$ — индуцированный подграф G на U. Доказать, что

$$\Pr[\chi(H) \leqslant 400] < 1/100.$$

3. Доказать, что существует такая абсолютная константа c, что для каждого $n > 1$ найдется интервал I_n не более чем из $c\sqrt{n}/\log n$ последовательных целых чисел, такой, что вероятность того, что хроматическое число графа $G(n, 0.5)$ принадлежит I_n, не меньше 0.99.

ВЕРОЯТНОСТНЫЙ ВЗГЛЯД:

Теорема Вейерштрасса о приближении

Хорошо известная теорема Вейерштрасса утверждает, что множество действительных полиномов на $[0,1]$ плотно в пространстве всех непрерывных действительных функций на $[0,1]$.

Теорема (теорема Вейерштрасса о приближении). *Для любой непрерывной действительной функции $f:[0,1]\mapsto\mathbb{R}$ и любого $\varepsilon>0$ существует полином $p(x)$, такой что $|p(x)-f(x)|\leqslant\varepsilon$ для всех $x\in[0,1]$.*

Работа [Bernstein (1912)] дает замечательное вероятностное доказательство этой теоремы, основанное на свойствах биномиального распределения. Оно заключается в следующем.

Доказательство. Так как непрерывная функция $f:[0,1]\mapsto\mathbb{R}$ равномерно непрерывна, найдется такое $\delta>0$, что если $x,x'\in[0,1]$ и $|x-x'|\leqslant\delta$, то $|f(x)-f(x')|\leqslant\varepsilon/2$. Кроме того, так как f должна быть ограничена, найдется такое $M>0$, что $|f(x)|\leqslant M$ на $[0,1]$.

Пусть $B(n,x)$ обозначает биномиальную случайную величину с n независимыми испытаниями и вероятностью успеха x для каждого из них. Тогда вероятность того, что $B(n,x)=j$, равна $\binom{n}{j}x^j(1-x)^{n-j}$. Математическое ожидание величины $B(n,x)$ равно nx, а стандартное отклонение составляет $\sqrt{nx(1-x)}\leqslant\sqrt{n}$. Следовательно, по неравенству Чебышёва, рассмотренному в гл. 4, для каждого целого n выполнено неравенство $\Pr[|B(n,x)-nx|>n^{2/3}]\leqslant\frac{1}{n^{1/3}}$. Значит, найдется такое целое n, что

$$\Pr[|B(n,x)-nx|>n^{2/3}]<\frac{\varepsilon}{4M}$$

и

$$\frac{1}{n^{1/3}}<\delta.$$

Положим

$$P_n(x)=\sum_{i=0}^{n}\binom{n}{i}x^i(1-x)^{n-i}f\left(\frac{i}{n}\right).$$

Мы утверждаем, что для каждого $x\in[0,1]$ выполнено требуемое неравенство $|P_n(x)-f(x)|\leqslant\varepsilon$. Действительно, так как $\sum_{i=0}^{n}\binom{n}{i}x^i(1-x)^{n-i}=1$, то

$$\begin{aligned}|P_n(x)-f(x)|\leqslant&\sum_{i;|i-nx|\leqslant n^{2/3}}\binom{n}{i}x^i(1-x)^{n-i}|f(\frac{i}{n})-f(x)|+\\&+\sum_{i;|i-nx|>n^{2/3}}\binom{n}{i}x^i(1-x)^{n-i}(|f(\frac{i}{n})|+|f(x)|)\leqslant\\&\leqslant\sum_{i;|i/n-x|\leqslant n^{-1/3}<\delta}\binom{n}{i}x^i(1-x)^{n-i}|f(\frac{i}{n})-f(x)|+\\&+\Pr\Big[|B(n,x)-nx|>n^{2/3}\Big]2M\leqslant\frac{\varepsilon}{2}+\frac{\varepsilon}{4M}2M=\varepsilon,\end{aligned}$$

что завершает доказательство. ■

8

Парадигма Пуассона

К решению математической задачи более всего побуждает призыв, всегда звучащий внутри нас: вот задача, ищи решение, ты можешь найти его одной только силой мысли, — в математике нет непознаваемого.

Давид Гильберт

Если случайная величина X является суммой большого числа индикаторов «почти независимых» случайных величин и $\mu = \mathbf{E}[X]$, мы будем говорить, что X имеет почти пуассоновское распределение со средним μ и, в частности, что $\Pr[X = 0]$ близко к $e^{-\mu}$. Мы называем это нестрогое утверждение *парадигмой Пуассона*. В настоящей главе рассматривается несколько случаев, когда эта парадигма может быть строго доказана.

8.1. НЕРАВЕНСТВА ЯНСОНА

Часто возникает необходимость в оценке вероятности того, что ни одно из «плохих» событий $B_i, i \in I$, не выполнено. Если события попарно независимы, то

$$\Pr\left[\bigwedge_{i\in I} \overline{B_i}\right] = \prod_{i\in I} \Pr\left[\,\overline{B_i}\,\right].$$

Когда B_i «почти» независимы, неравенства Янсона позволяют нам иногда утверждать, что эти две величины «почти» равны.

Пусть Ω является конечным универсальным множеством. Пусть R является случайным подмножеством множества Ω и задается равенствами

$$\Pr[r \in R] = p_r, r \in \Omega,$$

где все события с такими вероятностями попарно независимы. Пусть $A_i, i \in I$, являются подмножествами пространства Ω, I — конечное множество индексов. Обозначим через B_i событие «$A_i \subseteq R$». (То есть каждая точка $r \in \Omega$ «подбрасывает монету», чтобы определить принадлежит ли она R. Через B_i обозначим событие, состоящее в том, что при всех подбрасываниях монеты, соответствующих $r \in A_i$, выпал герб.) Пусть X_i — индикатор события B_i, а $X = \sum_{i\in I} X_i$ — количество $A_i \subseteq R$. Тогда события $\bigwedge_{i\in I} \overline{B_i}$ и $X = 0$ будут

идентичны. Мы пишем $i \sim j$, $i, j \in I$, если $i \neq j$ и $A_i \cap A_j \neq \varnothing$. Заметим, что если $i \neq j$, и не верно $i \sim j$, то события B_i и B_j являются независимыми, поскольку они зависят от подбрасывания различных монет. Более того (и это играет решающую роль в доказательствах), если $i \notin J \subset I$ и не верно $i \sim j$ для всех $j \in J$, то события B_i и $\{B_j | j \in J\}$ взаимонезависимы, т. е. B_i независимо с любой булевой функцией от этих B_j. В самом деле, «подбрасывания монет» элементами множества A_i и элементами множества $\cup_{j\in J} A_j$ независимы. Введем обозначение

$$\Delta = \sum_{i\sim j} \Pr[B_i \wedge B_j].$$

Здесь сумма берется по всем упорядоченным парам, поэтому, очевидно, $\Delta/2$ равно той же сумме, только взятой по неупорядоченным парам. Обозначим

$$M = \prod_{i\in I} \Pr\left[\ \overline{B_i}\ \right],$$

что было бы равно $\Pr[\bigwedge_{i\in I} \overline{B_i}]$, если бы B_i были независимы. И, наконец, пусть

$$\mu = \mathbf{E}[X] = \sum_{i\in I} \Pr[B_i].$$

Теорема 8.1.1 (неравенство Янсона). *Пусть события $B_i, i \in I$, и величины Δ, M, μ определены выше. Предположим, что все $\Pr[B_i] \leqslant \varepsilon$. Тогда*

$$M \leqslant \Pr\left[\bigwedge_{i\in I} \overline{B_i}\right] \leqslant M e^{\frac{1}{1-\varepsilon}\frac{\Delta}{2}}$$

и, более того,

$$\Pr\left[\bigwedge_{i\in I} \overline{B_i}\right] \leqslant e^{-\mu+\frac{\Delta}{2}}.$$

Для любого $i \in I$

$$\Pr\left[\ \overline{B_i}\ \right] = 1 - \Pr[B_i] \leqslant e^{-\Pr[B_i]},$$

поэтому произведение по всем $i \in I$ дает

$$M \leqslant e^{-\mu}.$$

Две верхние оценки в теореме 8.1.1 довольно близки. Однако для удобства мы будем в основном использовать вторую оценку. Часто для получения асимптотических оценок путем несложных вычислений можно показать, что $M \sim e^{-\mu}$. В частности, такую оценку можно получить, когда $\varepsilon = o(1)$ и $\varepsilon\mu = o(1)$.

Может быть, самым простым примером применения теоремы 8.1.1 служит асимптотика вероятности того, что граф $G(n, c/n)$ является свободным от треугольников. Соответствующий результат приведен в разд. 10.1. В этой

теореме, также как и во многих других случаях, $\varepsilon = o(1)$, $\Delta = o(1)$ и μ стремится к константе k. В этих случаях $\Pr[\bigwedge_{i\in I} \overline{B_i}] \to e^{-k}$. Далее мы не будем рассматривать случай, когда Δ становится большим. В самом деле, если $\Delta \geqslant 2\mu$, то верхняя оценка теоремы 8.1.1 становится бесполезной. Даже когда Δ немного меньше, эту оценку улучшает следующий результат.

Теорема 8.1.2 (обобщенное неравенство Янсона). *Пусть выполнены условия теоремы* 8.1.1 *и при этом* $\Delta \geqslant \mu$. *Тогда*

$$\Pr\left[\bigwedge_{i\in I} \overline{B_i}\right] \leqslant e^{-\frac{\mu^2}{2\Delta}}.$$

Теорема 8.1.2 (когда ее можно применить) как правило дает гораздо более сильные оценки, чем неравенство Чебышёва, которое применялось в гл. 4. В разд. 4.3 было показано, что $\mathrm{Var}[X] \leqslant \mu + \Delta$, откуда вытекает неравенство

$$\Pr\left[\bigwedge_{i\in I} \overline{B_i}\right] = \Pr[X = 0] \leqslant \frac{\mathrm{Var}[X]}{\mathbf{E}[X]^2} \leqslant \frac{\mu + \Delta}{\mu^2}.$$

Предположим, что $\mu \to \infty$, $\mu \ll \Delta$, и $\gamma = \frac{\mu^2}{\Delta} \to \infty$. Тогда неравенство Чебышёва дает оценку для $\Pr[X = 0]$, грубо говоря, порядка γ^{-1}, в то время как неравенство Янсона дает, грубо говоря, $e^{-\gamma}$.

8.2. ДОКАЗАТЕЛЬСТВА

Оригинальные доказательства Янсона базируются на оценках преобразования Лапласа соответствующей случайной величины. Мы приводим доказательство, полученное в работе [Boppana and Spencer (1989)]. Будем пользоваться неравенством

$$\Pr\left[B_i \Big| \bigwedge_{j\in J} \overline{B_j}\right] \leqslant \Pr[B_i],$$

которое верно для любого множества индексов $J \subset I, i \notin J$, а также неравенством

$$\Pr\left[B_i \big| B_k \wedge \bigwedge_{j\in J} \overline{B_j}\right] \leqslant \Pr[B_i | B_k],$$

справедливым для любого множества индексов $J \subset I, i, k \notin J$. Первое неравенство следует из теоремы 6.3.2. Второе же неравенство эквивалентно первому, поскольку условие, накладываемое на B_k, эквивалентно предположению, что $p_r = \Pr[r \in R] = 1$ для всех $r \in A_k$.

Доказательство теоремы 8.1.1. Нижняя оценка доказывается тривиально. Для удобства упорядочим множество индексов $I = \{1, \ldots, m\}$.

При $1 \leqslant i \leqslant m$

$$\Pr\left[B_i \,\middle|\, \bigwedge_{1\leqslant j<i} \overline{B_j}\right] \leqslant \Pr[B_i],$$

откуда вытекает соотношение

$$\Pr\left[\overline{B_i} \,\middle|\, \bigwedge_{1\leqslant j<i} \overline{B_j}\right] \geqslant \Pr[\overline{B_i}]$$

и, следовательно, выполняется неравенство

$$\Pr\left[\bigwedge_{i\in I} \overline{B_i}\right] = \prod_{i=1}^{m} \Pr\left[\overline{B_i} \,\middle|\, \bigwedge_{1\leqslant j<i} \overline{B_j}\right] \geqslant \prod_{i=1}^{m} \Pr\left[\,\overline{B_i}\,\right].$$

Докажем теперь первую верхнюю оценку. При заданном i для удобства перенумеруем индексы так, что $i \sim j$ верно при $1 \leqslant j \leqslant d$ и не верно при $d+1 \leqslant j < i$. Мы пользуемся неравенством $\Pr[A|B\wedge C] \geqslant \Pr[A\wedge B|C]$, справедливым для любых A, B, C. Пусть $A = B_i$, $B = \overline{B_1}\wedge\ldots\wedge\overline{B_d}$, $C = \overline{B_{d+1}}\wedge\ldots\wedge\overline{B_{i-1}}$, тогда

$$\Pr\left[B_i \middle|\, \bigwedge_{1\leqslant j<i} \overline{B_j}\right] = \Pr[A|B\wedge C] \geqslant \Pr[A\wedge B|C] = \Pr[A|C]\Pr[B|A\wedge C].$$

Из взаимной независимости событий следует, что $\Pr[A|C] = \Pr[A]$. Из корреляционного неравенства вытекает, что

$$\Pr[B|A\wedge C] \geqslant 1 - \sum_{j=1}^{d} \Pr[B_j|B_i\wedge C] \geqslant 1 - \sum_{j=1}^{d} \Pr[B_j|B_i].$$

Таким образом,

$$\Pr\left[B_i \middle|\, \bigwedge_{1\leqslant j<i} \overline{B_j}\right] \geqslant \Pr[B_i] - \sum_{j=1}^{d} \Pr[B_j\wedge B_i].$$

Откуда следует

$$\Pr\left[\overline{B_i} \middle|\, \bigwedge_{1\leqslant j<i} \overline{B_j}\right] \leqslant \Pr\left[\,\overline{B_i}\,\right] + \sum_{j=1}^{d} \Pr[B_j\wedge B_i] \leqslant$$

$$\leqslant \Pr\left[\,\overline{B_i}\,\right]\left(1 + \frac{1}{1-\varepsilon}\sum_{j=1}^{d} \Pr[B_j\wedge B_i]\right),$$

поскольку $\Pr\left[\,\overline{B_i}\,\right] \geqslant 1-\varepsilon$. Воспользовавшись неравенством $1+x \leqslant e^x$, получаем

$$\Pr\left[\overline{B_i} \middle|\, \bigwedge_{1\leqslant j<i} \overline{B_j}\right] \leqslant \Pr\left[\,\overline{B_i}\,\right] e^{\frac{1}{1-\varepsilon}\sum_{j=1}^{d} \Pr[B_j\wedge B_i]}.$$

Для каждого $1 \leqslant i \leqslant m$ мы подставляем последнее неравенство в соотношение

$$\Pr\left[\bigwedge_{i\in I}\overline{B_i}\right] = \prod_{i=1}^{m}\Pr\left[\overline{B_i}\Big|\bigwedge_{1\leqslant j<i}\overline{B_j}\right].$$

Члены $\Pr\left[\,\overline{B_i}\,\right]$ в произведении дают M. Показатели степеней складываются: для каждого $i, j \in I$ с $j < i$ и $j \sim i$ член $\Pr[B_j \wedge B_i]$ появляется однажды, поэтому в сумме они дают $\Delta/2$.

При доказательстве второй верхней оценки мы используем ограничения

$$\Pr\left[\overline{B_i}\Big|\bigwedge_{1\leqslant j<i}\overline{B_j}\right] \leqslant 1 - \Pr[B_i] + \sum_{j=1}^{d}\Pr[B_j \wedge B_i] \leqslant$$

$$\leqslant \exp\left(-\Pr[B_i] + \sum_{j=1}^{d}\Pr[B_j \wedge B_i]\right).$$

Теперь сумма членов $-\Pr[B_i]$ дает $-\mu$, в то время как сумма членов $\Pr[B_j \wedge B_i]$ снова дает $\Delta/2$. ■

Доказательство теоремы 8.1.2. Вторая верхняя оценка теоремы 8.1.1 может быть переписана в виде

$$-\ln\left(\Pr\left[\bigwedge_{i\in I}\overline{B_i}\right]\right) \geqslant \sum_{i\in I}\Pr[B_i] - \frac{1}{2}\sum_{i\sim j}\Pr[B_i \wedge B_j].$$

Для любого подмножества индексов $S \subset I$ то же неравенство, только примененное к $B_i, i \in S$, дает

$$-\ln\left(\Pr\left[\bigwedge_{i\in S}\overline{B_i}\right]\right) \geqslant \sum_{i\in S}\Pr[B_i] - \frac{1}{2}\sum_{i,j\in S, i\sim j}\Pr[B_i \wedge B_j].$$

Пусть теперь S является случайным подмножеством I и задается равенством

$$\Pr[i \in S] = p, \tag{8.1}$$

где p — константа, которая будет определена позже, и события с вероятностями (8.1) взаимно независимы. (Здесь мы будем использовать вероятностные методы для доказательства вероятностной теоремы!) Каждый член $\Pr[B_i]$ появляется с вероятностью p, а каждый член $\Pr[B_i \wedge B_j]$ — с вероятностью p^2. Откуда вытекает

$$\mathbf{E}\left[-\ln\left(\Pr\left[\bigwedge_{i\in S}\overline{B_i}\right]\right)\right] \geqslant \mathbf{E}\left[\sum_{i\in S}\Pr[B_i]\right] - \frac{1}{2}\mathbf{E}\left[\sum_{i,j\in S, i\sim j}\Pr[B_i \wedge B_j]\right] =$$

$$= p\mu - p^2\frac{\Delta}{2}.$$

Положим

$$p = \frac{\mu}{\Delta},$$

чтобы максимизировать оценку снизу. Из дополнительного условия теоремы 8.1.2 следует, что вероятность p не превосходит 1. Тогда

$$\mathbf{E}\left[-\ln\left(\Pr\left[\bigwedge_{i\in S}\overline{B_i}\right]\right)\right] \geqslant \frac{\mu^2}{2\Delta}.$$

Поэтому существует такое подмножество $S \subset I$, что

$$-\ln\left(\Pr\left[\bigwedge_{i\in S}\overline{B_i}\right]\right) \geqslant \frac{\mu^2}{2\Delta}.$$

То есть

$$\Pr\left[\bigwedge_{i\in S}\overline{B_i}\right] \leqslant e^{-\frac{\mu^2}{2\Delta}}.$$

Но

$$\Pr\left[\bigwedge_{i\in I}\overline{B_i}\right] \leqslant \Pr\left[\bigwedge_{i\in S}\overline{B_i}\right],$$

что и завершает доказательство. ■

8.3. РЕШЕТО БРУНА

Более традиционным подходом к парадигме Пуассона является метод *решета Бруна*, названный так в честь специалиста по теории чисел Т. Бруна, который пользовался этим методом. Пусть $B_1, \ldots, B_m$ — события, X_i — индикатор события B_i, и $X = X_1 + \ldots + X_m$ (количество наступивших событий B_i). Пусть есть скрытый параметр n, т. е. на самом деле $m = m(n)$, $B_i = B_i(n)$, $X = X(n)$. Это объясняет использование символов o и O. Введем обозначение

$$S^{(r)} = \sum \Pr[B_{i_1} \wedge \ldots \wedge B_{i_r}],$$

где сумма берется по всем множествам $\{i_1, \ldots, i_r\} \subseteq \{1, \ldots, m\}$. Из формулы включений-исключений следует, что

$$\Pr[X = 0] = \Pr\left[\,\overline{B_1} \wedge \ldots \wedge \overline{B_m}\,\right] = 1 - S^{(1)} + S^{(2)} - \ldots + (-1)^r S^{(r)} + \ldots\ .$$

Теорема 8.3.1. *Пусть существует константа μ, такая, что*

$$\mathbf{E}[X] = S^{(1)} \to \mu,$$

и для любого фиксированного r имеет место сходимость

$$\mathbf{E}[X^{(r)}/r!] = S^{(r)} \to \mu^r/r!.$$

Тогда

$$\Pr[X = 0] \to e^{-\mu}$$

и, более того, для каждого t

$$\Pr[X = t] \to \frac{\mu^t}{t!} e^{-\mu}.$$

Доказательство. Мы ограничимся рассмотрением только случая $t = 0$. Зафиксируем некоторое $\varepsilon > 0$. Выберем s так, что

$$\left| \sum_{r=0}^{2s} (-1)^r \frac{\mu^r}{r!} - e^{-\mu} \right| \leqslant \frac{\varepsilon}{2}.$$

Неравенства Бонферрони утверждают, что в общем случае формула включений-исключений попеременно оценивает то сверху, то снизу вероятность $\Pr[X = 0]$. В частности,

$$\Pr[X = 0] \leqslant \sum_{r=0}^{2s} (-1)^r S^{(r)}.$$

Выберем n_0 (скрытую переменную) так, чтобы при $n \geqslant n_0$ выполнялось неравенство

$$\left| S^{(r)} - \frac{\mu^r}{r!} \right| \leqslant \frac{\varepsilon}{2(2s+1)},$$

где $0 \leqslant r \leqslant 2s$. При таких n

$$\Pr[X = 0] \leqslant e^{-\mu} + \varepsilon.$$

Аналогично, суммируя до $2s + 1$, мы найдем n_0, такое, что при $n \geqslant n_0$ выполняется

$$\Pr[X = 0] \geqslant e^{-\mu} - \varepsilon.$$

Поскольку ε было выбрано произвольно, то $\Pr[X = 0] \to e^{-\mu}$. ■

Пороговые функции свойства содержания графом $G \sim G(n, p)$ копии заданного графа H, выводимые в разд. 10.1 с помощью неравенства Янсона, были в оригинале получены с помощью решета Бруна. В следующем примере задействованы оба метода. Пусть $G \sim G(n, p)$ — случайный граф, определенный в гл. 10. Введем обозначение $EPIT$ (Every Point In Triangle) для утверждения о том, что любая вершина лежит в некотором треугольнике.

Теорема 8.3.2. *Пусть зафиксирована константа* $c > 0$*, а величины* $p = p(n)$ *и* $\mu = \mu(n)$ *удовлетворяют условиям*

$$\binom{n-1}{2} p^3 = \mu, \qquad e^{-\mu} = \frac{c}{n}.$$

Тогда

$$\lim_{n\to\infty} \Pr\Big[G(n,p) \models EPIT\Big] = e^{-c}.$$

В статье [Spencer (1990a)] найдены пороговые функции для очень широкого класса «утверждений о расширениях»: все r вершин лежат в копии некоторого фиксированного графа H.

Доказательство. Сначала зафиксируем $x \in V(G)$. Для каждой неупорядоченной пары $y, z \in V(G) \setminus \{x\}$ пусть B_{xyz} обозначает событие, что тройка $\{x, y, z\}$ образует треугольник в графе G. Обозначим через C_x событие $\wedge\overline{B_{xyz}}$, а через X_x — соответствующий индикатор. Мы используем неравенство Янсона, чтобы оценить $\mathbf{E}[X_x] = \Pr[C_x]$. Поскольку $p = o(1)$, то и $\varepsilon = o(1)$. Согласно определению, данному выше, $\sum \Pr[B_{xyz}] = \mu$. Зависимость $xyz \sim xuv$ возникает тогда и только тогда, когда множества пересекаются (в точке, отличной от x). Следовательно,

$$\Delta = \sum_{y,z,z'} \Pr[B_{xyz} \wedge B_{xyz'}] = O(n^3)p^5 = o(1),$$

поскольку $p = n^{-2/3+o(1)}$. Итак,

$$\mathbf{E}[X_x] \sim e^{-\mu} = \frac{c}{n}.$$

Теперь определим

$$X = \sum_{x \in V(G)} X_x,$$

что равно числу вершин x, не лежащих в треугольниках. Тогда из линейности математического ожидания имеем

$$\mathbf{E}[X] = \sum_{x \in V(G)} \mathbf{E}[X_x] \to c.$$

Нам нужно показать, что парадигму Пуассона можно применить к X. Зафиксируем r. Тогда

$$\mathbf{E}[X^{(r)}/r!] = S^{(r)} = \sum \Pr[C_{x_1} \wedge \ldots \wedge C_{x_r}],$$

где сумма берется по всем множествам вершин $\{x_1, \ldots, x_r\}$. Нет разницы между r-множествами, поэтому

$$\mathbf{E}[X^{(r)}/r!] = \binom{n}{r} \Pr[C_{x_1} \wedge \ldots \wedge C_{x_r}] \sim \frac{n^r}{r!} \Pr[C_{x_1} \wedge \ldots \wedge C_{x_r}],$$

где $x_1, \ldots, x_r$ — некоторые конкретные вершины. Но

$$C_{x_1} \wedge \ldots \wedge C_{x_r} = \wedge\overline{B_{x_i yz}},$$

где пересечение берется по всем $1 \leqslant i \leqslant r$ и всем y, z. Мы применяем неравенство Янсона к этому пересечению. Снова $\varepsilon = p^3 = o(1)$. Количество троек $\{x_i, y, z\}$ равно $r\binom{n-1}{2} - O(n)$, где лишние тройки появляются из-за треугольников, содержащих две (или три) вершины x_i. (Здесь важную роль играет то, что r фиксировано.) Таким образом,

$$\sum \Pr[B_{x_i yz}] = p^3 \left(r\binom{n-1}{2} - O(n) \right) = r\mu + O(n^{-1+o(1)}).$$

Как и ранее, Δ равно произведению p^5 на количество пар $x_i yz \sim x_j y' z'$. Есть $O(rn^3) = O(n^3)$ членов с $i = j$ и $O(r^2 n^2) = O(n^2)$ членов с $i \neq j$. Таким

образом, снова получаем, что $\Delta = o(1)$. Поэтому

$$\Pr[C_{x_1} \wedge \ldots \wedge C_{x_r}] \sim e^{-r\mu}$$

и

$$\mathbf{E}[X^{(r)}/r!] \sim \frac{(ne^{-\mu})^r}{r!} = \frac{c^r}{r!}.$$

Следовательно, условия теоремы 8.3.1 выполнены для X. ■

8.4. БОЛЬШИЕ УКЛОНЕНИЯ

Вернемся к терминологии разд. 8.1. Наша цель состоит в получении для случайной величины X неравенств больший уклонений, аналогичных неравенствам, доказанным в приложении А. Пусть R — случайное подмножество множества Ω. Мы называем подмножество J множества индексов *непересекающимся множеством* (НМ), если

- Событие B_j выполняется для любого $j \in J$.
- Ни для каких $j, j' \in J$ не выполняется $j \sim j'$.

Если к тому же

- При выполнении события $B_{j'}$ из $j' \notin J$ следует, что $j \sim j'$ для некоторого $j \in J$,

то J называется *максимальным непересекающимся множеством* (МНМ). Приведем некоторые общие результаты о размерах максимальных непересекающихся множеств. Связь с X будет затем установлена с помощью специально разработанных для этого средств.

Лемма 8.4.1. *При введенных выше обозначениях для любого s*

$$\Pr[\exists \text{ НМ } J, |J| = s] \leqslant \frac{\mu^s}{s!}.$$

Доказательство. Пусть $\sum^*$ обозначает сумму по всем s-множествам $J \subseteq I$, не содержащим пар j, j', таких, что $j \sim j'$. Обозначим через $\sum^o$ сумму по всем упорядоченным наборам $(j_1, \ldots, j_s)$ из s элементов, образующим множество J. И, наконец, пусть $\sum^a$ обозначает сумму по всем упорядоченным наборам $(j_1, \ldots, j_s)$ из s элементов. Тогда

$$\begin{aligned}\Pr[\exists \text{ НМ } J,\ |J| = s] &\leqslant \sum{}^* \Pr[\bigwedge_{j\in J} B_j] = \\ &= \sum{}^* \prod_{j\in J} \Pr[B_j] = \frac{1}{s!}\sum{}^o \Pr[B_{j_1}]\ldots\Pr[B_{j_s}] \leqslant \\ &\leqslant \frac{1}{s!}\sum{}^a \Pr[B_{j_1}]\ldots\Pr[B_{j_s}] \leqslant \frac{1}{s!}\left[\sum_{i\in I}\Pr[B_i]\right]^s = \mu^s/s!.\end{aligned}$$

■

Лемма 8.4.1 дает эффективную верхнюю оценку при $\mu^s \ll s!$, в основном, когда $s > \mu\alpha$, где $\alpha > e$. Для небольших s мы учитываем условие максимальности НМ J. В этом случае обозначим через μ_s минимум, взятый по всем $j_1, \dots, j_s \in I$, от $\sum \Pr[B_i]$, где сумма берется по всем $i \in I$, кроме тех i, для которых $i \sim j_l$ при некотором $1 \leqslant l \leqslant s$. В приложениях s обычно мало́ (в противном случае мы используем лемму 8.4.1), и μ_s близко к μ. В некоторых случаях удобно ввести обозначение

$$\nu = \max_{j \in I} \sum_{i \sim j} \Pr[B_i]$$

и использовать то, что $\mu_s \geqslant \mu - s\nu$.

Лемма 8.4.2. *В предыдущих обозначениях для любого целого s справедливы неравенства*

$$\Pr[\exists \text{ МНМ } J,\ |J| = s] \leqslant \frac{\mu^s}{s!} e^{-\mu_s} e^{\frac{\Delta}{2}} \leqslant \frac{\mu^s}{s!} e^{-\mu} e^{s\nu} e^{\frac{\Delta}{2}}.$$

Доказательство. Также, как и в лемме 8.4.1 мы оцениваем эту вероятность сверху величиной $\sum^*$, где $J = \{j_1, \dots, j_s\}$ является МНМ. Для выполнения данного условия J должно быть, во-первых, непересекающимся семейством, а во-вторых, должно выполняться событие $\wedge^* \overline{B_i}$, где $\wedge^*$ обозначает пересечение по всем $i \in I$, кроме $i \sim j_l$ для некоторого $1 \leqslant l \leqslant s$. Применим неравенство Янсона, чтобы получить верхнюю оценку для $\Pr[\wedge^* \overline{B_i}]$. Соответствующие значения μ^*, Δ^* удовлетворяют неравенствам

$$\mu^* \geqslant \mu_s, \qquad \Delta^* \leqslant \Delta.$$

Последнее неравенство следует из того, что Δ^* имеет меньше слагаемых. Таким образом,

$$\Pr[\wedge^* \overline{B_i}] \leqslant e^{-\mu_s} e^{\frac{\Delta}{2}}$$

и

$$\sum{}^* \Pr[J \text{ является МНМ}] \leqslant e^{-\mu_s} e^{\frac{\Delta}{2}} \sum{}^* \Pr[\bigwedge_{j \in J} B_j] \leqslant$$

$$\leqslant e^{-\mu_s} e^{\frac{\Delta}{2}} \mu^s / s!. \qquad \blacksquare$$

Когда $\Delta = o(1)$ и $\nu\mu = o(1)$ или, более общо, $\mu_{3\mu} = \mu + o(1)$, лемма 8.4.2 дает хорошую аппроксимацию распределению Пуассона, поскольку

$$\Pr[\exists \text{ МНМ } J,\ |J| = s] \leqslant (1 + o(1)) \frac{\mu^s}{s!} e^{-\mu}$$

при $s \leqslant 3\mu$, и по лемме 8.4.1 эта вероятность гораздо меньше при больших s.

8.5. ОЦЕНКА ЧИСЛА РАСШИРЕНИЙ

Начнем со случая, где используется основной результат, касающийся больших уклонений из приложения А.

Теорема 8.5.1. *Пусть* $p = \frac{\ln n}{n}\omega(n)$, *где* $\omega(n) \to \infty$ *как угодно медленно. Тогда в графе* $G(n,p)$ *почти всегда*

$$\deg(x) \sim (n-1)p$$

для всех вершин x.

Это, на самом деле, результат по большим уклонениям. Достаточно доказать следующее утверждение.

Теорема 8.5.2. *Пусть* $p = \frac{\ln n}{n}\omega(n)$, *где* $\omega(n) \to \infty$ *сколь угодно медленно. Пусть точка* $x \in G$ *фиксирована. Зафиксируем некоторое* $\varepsilon > 0$. *Тогда*

$$\Pr\left[|\deg(x) - (n-1)p| > \varepsilon(n-1)p\right] = o(n^{-1}).$$

Доказательство. Поскольку $\deg(x) \sim B(n-1,p)$, т. е. это биномиальная случайная величина с указанными выше параметрами, то согласно следствию А.1.14 получаем, что

$$\Pr\left[|\deg(x) - (n-1)p| > \varepsilon(n-1)p\right] < 2e^{-c_\varepsilon(n-1)p} = o(n^{-1}),$$

так как c_ε фиксировано и $(n-1)p \gg \ln n$. ■

Этот результат показывает, почему логарифмические члены появляются так часто при изучении случайных графов. Мы хотим, чтобы каждое x обладало некоторым свойством, поэтому мы стараемся, чтобы вероятность того, что это свойство не выполняется, была $o(n^{-1})$. Когда можно применить парадигму Пуассона, вероятность этого нежелательного события, грубо говоря, экспоненциальна, поэтому мы хотим, чтобы показатель был логарифмическим. Это часто ведет к тому, что в выражении для вероятности появления ребра p возникает логарифм.

В гл. 3 мы нашли пороговую функцию условия, что любая вершина лежит в треугольнике. Это событие случалось, когда ожидаемое число расширений заданной вершины до треугольника достигало $\ln n$. Обозначим через $N(x)$ количество треугольников, содержащих x. Пусть $\mu = \binom{n-1}{2}p^3 = \mathbf{E}[N(x)]$.

Теорема 8.5.3. *Пусть* p *таково, что* $\mu \gg \ln n$. *Тогда почти всегда*

$$N(x) \sim \mu$$

для всех $x \in G(n,p)$.

Как и выше, данный результат относится, фактически, к большим уклонениям. Докажем следующее утверждение.

Теорема 8.5.4. *Пусть* p *таково, что* $\mu \gg \ln n$. *Зафиксируем* $x \in G$ *и* $\varepsilon > 0$. *Тогда*

$$\Pr[|N(x) - \mu| > \varepsilon\mu] = o(n^{-1}).$$

Доказательство. Мы будем доказывать это утверждение, предполагая, что $p = n^{-2/3+o(1)}$ (или, что эквивалентно, $\mu = n^{o(1)}$). Данного ограничения общности можно избежать техническими методами. Теперь в условиях лемм 8.4.1, 8.4.2 имеем $\nu, \mu, \Delta = o(1)$. Пусть P обозначает случайную величину,

распределенную по закону Пуассона с математическим ожиданием μ. Тогда

$$\Pr\Big[\exists \text{ МНМ } J,\ |J| \leqslant \mu(1-\varepsilon)\Big] \leqslant (1+o(1))\Pr[P \leqslant \mu(1-\varepsilon)],$$

$$\Pr\Big[\exists \text{ МНМ } J, \mu(1+\varepsilon) \leqslant |J| \leqslant 3\mu\Big] \leqslant (1+o(1))\Pr\Big[\mu(1+\varepsilon) \leqslant P \leqslant 3\mu\Big],$$

$$\Pr\Big[\exists \text{ МНМ } J, |J| \geqslant 3\mu\Big] \leqslant \Pr\Big[\exists \text{ НМ } J, |J| \geqslant 3\mu\Big] \leqslant \sum_{s=3\mu}^{\infty} \frac{\mu^s}{s!} = O((1-c)^{\mu}),$$

где $c > 0$ — некоторая постоянная. Поскольку $\mu \gg \ln n$, третий член равен $o(n^{-1})$. По теореме A.1.15 первый и второй члены равны $o(n^{-1})$. Каждое МНМ J с вероятностью $1 - o(n^{-1})$ лежит в интервале от $(1-\varepsilon)\mu$ до $(1+\varepsilon)\mu$.

Зафиксируем некоторое такое МНМ J. Заметим, что всегда существует некоторое МНМ, даже в случае, когда не выполняется ни одно из B_i, мы можем взять $J = \varnothing$. Элементами J являются тройки xyz, которые образуют треугольники. Следовательно, $N(x) \geqslant |J| \geqslant (1-\varepsilon)\mu$. *Верхняя оценка.* Вероятность того, что существует пять треугольников $xyz_1, xyz_2, xyz_3, xyz_4, xyz_5$, не превосходит $n^6 p^{11} = o(n^{-1})$. Вероятность того, что существуют треугольники xy_iz_i, xy_iz_i', $1 \leqslant i \leqslant 4$, где все вершины различны, не превосходит $n^{12}p^{20} = o(n^{-1})$. Рассмотрим граф, вершинами которого являются треугольники xyz. Пусть $\sim$ обозначает отношение смежности. Существует $N(x)$ вершин, МНМ J является максимальным независимым множеством. В этом графе с вероятностью $1 - o(n^{-1})$ каждая вершина xyz имеет степень не более 9, и не существует множества, состоящего из четырех несмежных ребер. Из этого вытекает, что для любого J выполнено неравенство $|J| \geqslant N(x) - 27$ и

$$N(x) \leqslant (1+\varepsilon)\mu + 27 \leqslant (1+\varepsilon')\mu. \qquad \blacksquare$$

Для любого графа H с «корнями» $x_1, \ldots, x_r$ мы можем подсчитать в графе $G(n,p)$ количество расширений $N(x_1, \ldots, x_r)$ заданного множества из r вершин до копии графа H. В работе [Spencer (1990b)] даны некоторые общие результаты, которые являются обобщениями теорем 8.5.2 и 8.5.4. При весьма широких предположениях (см. упр. 5, гл. 10), когда математическое ожидание μ количества расширений удовлетворяет условию $\mu \gg \ln n$, почти всегда выполняется отношение $N(x_1, \ldots, x_r) \sim \mu$.

8.6. ЧИСЛО ПРЕДСТАВЛЕНИЙ

Теоремы этого раздела основываются на следующем очень простом и очень полезном результате.

Лемма 8.6.1 (лемма Бореля—Кантелли). *Пусть события $A_n, n \in \mathbb{N}$, удовлетворяют неравенству*

$$\sum_{n=1}^{\infty} \Pr[A_n] < \infty.$$

Тогда

$$\Pr\left[\bigwedge_{i=1}^{\infty}\bigvee_{j=i}^{\infty} A_j\right] = 0.$$

То есть при достаточно больших n событие A_n почти всегда не выполнено. Для того чтобы применить эту лемму, мы будем стремиться получить неравенство $\Pr[A_n] < n^{-c}$, где $c > 1$.

В очередной раз мы начинаем со случая, который использует только результаты о больших уклонениях, полученные в приложении А. Для заданного множества S натуральных чисел обозначим (для любого $n \in \mathbb{N}$) через $f(n) = f_S(n)$ количество представлений $n = x + y$, $x, y \in S, x < y$.

Теорема 8.6.2 [Erdős (1956)]. *Существует множество S, для которого $f(n) = \Theta(\ln n)$. То есть существуют множество S и константы c_1, c_2 такие, что для всех достаточно больших n выполняются неравенства*

$$c_1 \ln n \leqslant f(n) \leqslant c_2 \ln n.$$

Доказательство. Определим случайное множество S равенством

$$\Pr[x \in S] = p_x = \min\left[10\sqrt{\frac{\ln x}{x}}, 1\right].$$

Зафиксируем n. Теперь $f(n)$ является случайной величиной с математическим ожиданием

$$\mu = \mathbf{E}[f(n)] = \frac{1}{2}\sum_{x+y=n, x\neq y} p_x p_y.$$

Примерно для n слагаемых выполняется неравенство $p_x p_y > p_n^2 = 100\frac{\ln n}{n}$. Мы получаем, что $p_x p_x = \Theta(\frac{\ln n}{n})$, за исключением случаев, когда $x = o(n), y = o(n)$. Нужно проследить, чтобы эти члены не давали существенного вклада в значение μ. Аккуратные асимптотические вычисления дают соотношение

$$\mu \sim (50 \ln n)\int_0^1 \frac{dx}{\sqrt{x(1-x)}} = 50\pi \ln n.$$

Незначительность вклада членов $x = o(n), y = o(n)$ отражает конечность неопределенного интеграла в точках $x = 0$ и $x = 1$. Возможные представления $x + y = n$ являются взаимно независимыми событиями. Поэтому из следствия А.1.14 получаем, что

$$\Pr[|f(n) - \mu| > \varepsilon\mu] < 2e^{-\delta\mu}$$

для некоторых констант ε и $\delta = \delta(\varepsilon)$. Для определенности мы можем взять $\varepsilon = 0.9, \delta = 0.1$; тогда

$$\Pr[|f(n) - \mu| > 0.9\mu] < 2e^{-5\pi \ln n} < n^{-1.1}$$

при достаточно больших n. Возьмем $c_1 < 0.1(50\pi)$ и $c_2 > 1.9(50\pi)$.

Пусть A_n — событие, заключающееся в том, что неравенства $c_1 \ln n \leqslant f(n) \leqslant c_2 \ln n$ не выполняются. При достаточно больших n справедливо неравенство $\Pr[A_n] < n^{-1.1}$. Из леммы Бореля—Кантелли следует, что почти всегда все события A_n не выполнены при достаточно больших n. Поэтому существует некоторая точка в вероятностном пространстве, т. е. некоторое множество S, для которого $c_1 \ln n \leqslant f(n) \leqslant c_2 \ln n$ при всех достаточно больших n. ■

Исследование бесконечного вероятностного пространства, использованного здесь и ниже, было аккуратно проведено в книге «Последовательности» Халберстама и Рота.

Использование бесконечного вероятностного пространства ставит много вопросов, связанных с экзистенциальной природой доказательств. Например, существует ли рекурсивное множество S, удовлетворяющее условию теоремы 8.6.2? Положительный ответ на этот вопрос был получен в статье [Kolountzakis (1999)].

Для заданного множества S натуральных чисел обозначим через $g(n) = g_S(n)$ количество представлений $n = x + y + z$, $x, y, z \in S$, $x < y < z$. Нижеследующий результат на самом деле был получен для представлений числа n в виде суммы k слагаемых при любом фиксированном k. Для простоты изложения мы приводим доказательство только для случая $k = 3$.

Теорема 8.6.3 [Erdős and Tetali (1990)]. *Существует множество S, для которого $g(n) = \Theta(\ln n)$. То есть существуют множество S и константы c_1, c_2 такие, что для всех достаточно больших n справедливы неравенства*

$$c_1 \ln n \leqslant g(n) \leqslant c_2 \ln n.$$

Доказательство. Зададим случайное множество S равенствами

$$\Pr[x \in S] = p_x = \min\left[10\left(\frac{\ln x}{x^2}\right)^{1/3}, \frac{1}{2}\right].$$

Зафиксируем n. Теперь $g(n)$ является случайной величиной с математическим ожиданием

$$\mu = \mathbf{E}[g(n)] = \sum_{x+y+z=n} p_x p_y p_z.$$

Аккуратные асимптотические вычисления дают

$$\mu \sim \frac{10^3}{6} \ln n \int_{x=0}^{1} \int_{y=0}^{1-x} \frac{dxdy}{[xy(1-x-y)]^{2/3}} = K \ln n,$$

где K велико. (Можно сделать K сколь угодно большим, увеличивая «10».) Применим лемму 8.4.2. Здесь

$$\Delta = \sum p_x p_y p_z p_{y'} p_{z'},$$

где сумма берется по всем пятеркам с $x + y + z = x + y' + z' = n$. Число слагаемых приблизительно равно n^3, и каждое из них асимптотически составляет $p_n^5 = n^{-10/3+o(1)}$, так что сумма есть $o(1)$. И снова необходимо проследить, чтобы слагаемые, в которые входят малые переменные, не

привносили большой вклад в сумму. В предположении $s \leqslant 3\mu = \Theta(\ln n)$ получим асимптотику для μ_s. Это минимально возможное значение величины $\sum p_x p_y p_z$, где сумма берется по всем x, y, z, таким, что $x + y + z = n$ и не пересекающимися с заданными s представлениями; ослабим данное условие, рассмотрим сумму по заданному множеству с $3s$ элементами. Вновь требуется показать, что основная составляющая суммы $\sum_{x+y+z=n} p_x p_y p_z$ приходится не на края, а «сосредоточена» вблизи центра, и что $\mu_s \sim \mu$. Теперь, как и в разд. 8.5, обозначим через P случайную величину, распределенную по закону Пуассона с параметром μ. Вероятность того, что существует МНМ J, размер которого меньше $\mu(1-\varepsilon)$, или заключен между $\mu(1+\varepsilon)$ и 3μ, асимптотически равна вероятности того, что P лежит в этих промежутках. Так как K велико, при умеренных ε эти вероятности, как и вероятность того, что существует непересекающееся семейство размера, большего чем 3μ, равны $o(n^{-c})$, где $c > 1$. По лемме Бореля—Кантелли почти всегда для всех достаточно больших n размер всех МНМ J находится в промежутке от $c_1 \ln n$ до $c_2 \ln n$. Отсюда вытекает оценка $g(n) \geqslant c_1 \ln n$.

Верхняя оценка. При заданном p пусть $f(n)$ является, как и раньше, количеством представлений n в виде суммы двух элементов из S. Мы используем только то, что $p_x = x^{-2/3+o(1)}$. Имеем

$$\mathbf{E}[f(n)] = \sum_{x+y=n} (xy)^{-2/3+o(1)} = n^{-1/3+o(1)},$$

учитывая также «полюс» в нуле. Здесь возможные представления взаимно независимы, так что

$$\Pr[f(n) \geqslant 4] \leqslant \mathbf{E}[f(n)]^4/4! = n^{-4/3+o(1)},$$

и по лемме Бореля—Кантелли почти всегда $f(n) \leqslant 3$ для всех достаточно больших n. Но тогда почти всегда существует такое C, что $f(n) \leqslant C$ для всех n. Для всех достаточно больших n существует МНМ (с представлениями в виде суммы трех слагаемых) размера, меньшего чем $c_2 \ln n$. Каждая тройка $x, y, z \in S$, такая, что $x+y+z = n$, должна содержать по крайней мере одну из этих не более чем $3c_2 \ln n$ точек. Количество троек $x, y, z \in S$ с $x+y+z = n$ при фиксированном x равно $f(n-x)$, т. е. числу представлений $n - x = y + z$ (быть может, на одно меньше, поскольку $y, z \neq x$), и поэтому не превосходит C. Но тогда всего существует не более $C(3c_2 \ln n)$ представлений $n = x + y + z$. ■

8.7. ДАЛЬНЕЙШИЕ ОБОБЩЕНИЯ

Здесь мы обсуждаем некоторые ситуации, когда применима парадигма Пуассона. Пусть $B_i, i \in I$, являются событиями в произвольном вероятностном пространстве. Так же как и в локальной лемме Ловаса из гл. 5, мы говорим, что симметрическое бинарное отношение $\sim$ на I называется *орграфом зависимостей*, если для всякого $i \in I$ событие B_i взаимно независимо с событием $\{B_j | \text{не выполняется } i \sim j\}$. (Для орграфа из гл. 5, разд. 5.1 имеем $E = \{(i,j) | i \sim j\}$.) Предположим, что события B_i удовлетворяют неравенствам

из разд. 8.2:

$$\Pr\left[B_i \Big| \bigwedge_{j\in J} \overline{B_j}\right] \leqslant \Pr[B_i],$$

верным для любых множеств индексов $J \subset I, i \notin J$, и неравенствам

$$\Pr\left[B_i \Big| B_k \wedge \bigwedge_{j\in J} \overline{B_j}\right] \leqslant \Pr[B_i|B_k],$$

верным для множеств индексов $J \subset I, i, k \notin J$. Тогда утверждения теорем 8.1.1 и 8.1.2 о неравенствах Янсона, а также лемм 8.4.1 и 8.4.2 выполняются. Доказательства идентичны; были использованы только перечисленные выше свойства событий B_i.

В работе [Suen (1990)] (см. также [Janson (1998)] для существенных вариаций) получен весьма общий результат, позволяющий аппроксимировать вероятность $\Pr[\bigwedge_{i\in I} \overline{B_i}]$ посредством произведения $M = \prod_{i\in I} \Pr[\overline{B_i}]$. Снова пусть $B_i, i \in I$, являются событиями в произвольном вероятностном пространстве. Будем говорить, что бинарное отношение $\sim$ на I является *орграфом суперзависимостей*, если выполняется следующее условие. Пусть $J_1, J_2 \subset I$ — такие непересекающиеся подмножества, что $j_1 \sim j_2$ не выполняется ни для каких $j_1 \in J_1, j_2 \in J_2$. Пусть B^1 является произвольной булевой комбинацией событий $B_j, j \in J_1$, и пусть B^2 является булевой комбинацией событий $B_j, j \in J_2$. Тогда события B^1 и B^2 независимы. Заметим, что отношение $\sim$, определенное в разд. 8.1, является на самом деле орграфом суперзависимостей.

Теорема 8.7.1 (Сен). *При указанных выше условиях*

$$\left|\Pr\left[\bigwedge_{i\in I} \overline{B_i}\right] - M\right| \leqslant M\left[e^{\sum_{i\sim j} y(i,j)} - 1\right],$$

где

$$y(i,j) = \Big(\Pr[B_i \wedge B_j] + \Pr[B_i]\Pr[B_j]\Big) \prod_{l\sim i \text{ или } l\sim j} (1 - \Pr[B_l])^{-1}.$$

Мы не будем доказывать теорему 8.7.1. Во многих случаях последнее произведение невелико. Предположим, что оно меньше двух для всех $i \sim j$. В этом случае

$$\sum_{i\sim j} y(i,j) \leqslant 2\Big[\Delta + \sum_{i\sim j} \Pr[B_i]\Pr[B_j]\Big].$$

Во многих случаях величина $\sum_{i\sim j} \Pr[B_i]\Pr[B_j]$ мала по сравнению с Δ (как и во многих случаях, когда $i \sim j$, события B_i, B_j положительно коррелированны). Если кроме того $\Delta = o(1)$, то теорема Сена дает аппроксимацию вероятности $\Pr[\bigwedge_{i\in I} \overline{B_i}]$ через M. Сен применил свой результат для оценки числа индуцированных копий фиксированного графа H в случайном графе $G(n,p)$.

В работе [Janson (1990)] получено одностороннее неравенство для больших уклонений величины X из разд. 8.1, которое несколько проще применить к леммам 8.4.1 и 8.4.2.

Теорема 8.7.2 (Янсон). *Для математического ожидания* $\mu = \mathbf{E}[X]$ *и произвольного* $\gamma > 0$ *справедливо соотношение*

$$\Pr[X \leqslant (1-\gamma)\mu] < e^{-\gamma^2 \mu/(2+\frac{\Delta}{\mu})}.$$

Когда $\Delta = o(\mu)$, эта оценка остаточного члена становится близка к оценке, задаваемой функцией нормального распределения, у которого среднее значение и стандартное отклонение равны μ. Мы не будем доказывать теорему 8.7.2. Доказательства теорем 8.7.1 и 8.7.2, как и оригинальные доказательства Янсона теорем 8.1.1 и 8.1.2, основаны на оценках преобразования Лапласа величины X, оценивающих $\mathbf{E}[e^{-tX}]$.

8.8. УПРАЖНЕНИЯ

1. Доказать, что для любого $\varepsilon > 0$ существует некоторое $n_0 = n_0(\varepsilon)$ такое, что для всех $n > n_0$ существует граф на n вершинах, содержащий в качестве индуцированного подграфа всякий граф на $k \leqslant (2-\varepsilon)\log_2 n$ вершинах.

2. Найти пороговую функцию для следующего свойства: граф $G(n,p)$ содержит по крайней мере $n/6$ попарно вершинно-независимых треугольников.

ВЕРОЯТНОСТНЫЙ ВЗГЛЯД:

Локальная раскраска

Этот результат Эрдёша [Erdős (1962)] является еще одним вероятностным свидетельством того, что хроматическое число не может быть найдено путем локального анализа.

Теорема. *Для всех k существует такое $\varepsilon > 0$, что при всех достаточно больших n существуют графы G на n вершинах с хроматическим числом $\chi(G) > k$, но, тем не менее, $\chi(G|_S) \leqslant 3$ для любого множества вершин S размера не большего εn.*

Доказательство. Для заданного k пусть $c, \varepsilon > 0$ удовлетворяют неравенствам

$$c > 2k^2 H(1/k) \ln 2,$$
$$\varepsilon < e^{-5} 3^3 c^{-3},$$

где $H(x) = -x \log_2 x - (1-x)\log_2(1-x)$ — функция энтропии. Пусть $p = c/n$ и пусть $G \sim G(n,p)$. Мы покажем, что граф G почти наверное удовлетворяет обоим условиям теоремы.

Если $\chi(G) \leqslant k$, то существует независимое множество размера n/k. Ожидаемое число таких множеств равно

$$\binom{n}{n/k}(1-p)^{\binom{n/k}{2}} < 2^{n(H(1/k)+o(1))} e^{-cn/2k^2(1+o(1))},$$

что есть $o(1)$, исходя из нашего выбора величины c. Следовательно, почти наверное $\chi(G) > k$.

Предположим, что для раскраски некоторого множества S, содержащего $t \leqslant \varepsilon n$ вершин, необходимо не менее четырех цветов. Тогда, как и в доказательстве леммы 7.3.4, существует минимальное такое множество S. Для любого $v \in S$ существует 3-раскраска множества $S \setminus \{v\}$. Если v имеет двух или менее соседей в S, то эта раскраска может быть продолжена до 3-раскраски S. Следовательно, любая $v \in S$ имеет степень не меньше трех в $G|_S$, и поэтому граф $G|_S$ содержит не менее $3t/2$ ребер. Вероятность того, что некоторые $t \leqslant \varepsilon n$ вершин имеют не меньше $3t/2$ ребер, меньше чем

$$\sum_{t \leqslant \varepsilon n} \binom{n}{t} \binom{\binom{t}{2}}{3t/2} \left(\frac{c}{n}\right)^{3t/2}.$$

Мы наметим путь дальнейших рассуждений. Когда $t = O(1)$, множителями можно пренебречь. В противном случае мы оцениваем каждый множитель сверху:

$$\left[\frac{ne}{t}\left(\frac{te}{3}\right)^{3/2}\left(\frac{c}{n}\right)^{3/2}\right]^t \leqslant \left(e^{5/2} 3^{-3/2} c^{3/2} \sqrt{t/n}\right)^t.$$

Теперь, поскольку $t \leqslant \varepsilon n$, последняя величина в скобках не превосходит значения $e^{5/2}3^{-3/2}c^{3/2}\varepsilon^{1/2}$. Из условия на ε следует, что последнее выражение меньше единицы. Вся сумма равна $o(1)$, т. е. почти наверное такого S не существует. ■

Многие заманчивые гипотезы легко опровергаются посредством вероятностного метода. Если всякие $n/(\ln n)$ вершин могут быть окрашены в три цвета, может ли тогда граф G на n вершинах быть раскрашен в четыре цвета? Данный результат показывает, что ответ является отрицательным.

9

Псевдослучайность

«Узел!» — воскликнула Алиса, всегда готовая помочь ближнему, и, с тревогой оглядываясь вокруг, продолжила: «Ах, позвольте мне помочь вам развязать его!»

Льюис Кэрролл,
Алиса в Стране Чудес

Как уже не раз говорилось в различных главах этой книги, вероятностный метод представляет собой мощный инструмент для установления факта существования комбинаторных объектов с определенными свойствами. Часто таких доказательств существования недостаточно, хотелось бы *построить объект явным образом*. Это желание вызвано не только тем, что явное построение может пролить больше света на соответствующую проблему. Часто случайная на вид конструкция может требоваться для некоторого алгоритмического метода. При этом хотелось бы не просто доказать существование алгоритма, а получить сам алгоритм.

Проблема поиска явных построений может показаться тривиальной. В конце концов, мы в основном имеем дело с конечными случаями. Получив вероятностное доказательство существования, можно найти явный пример полным перебором. Более того, многие вероятностные доказательства фактически показывают, что большинство элементов выбранного должным образом вероятностного пространства обладают желаемыми свойствами. Таким образом, можно ожидать, что найти один такой элемент не будет слишком сложно. Но несмотря на то, что в принципе это верно, ясно, что нецелесообразно перебирать все возможности. Поэтому обычно под явным построением комбинаторного объекта понимается построение, которое может быть выполнено эффективно. Например, за время, полиномиальное относительно некоторых параметров объекта.

Поясним эти идеи на примере одной из наиболее известных открытых проблем в области явных построений — проблемы построения явных *графов Рамсея*. Первый пример, приведенный в гл. 1 — доказательство Эрдёша того, что для каждого n существуют графы на n вершинах, не содержащие ни клики, ни независимого множества на $2\log_2 n$ вершинах. Это — доказательство существования. Можем ли мы явным образом указать такие графы? Эрдёш предложил награду в \$500 за явное построение бесконечного семейства графов,

в которых нет ни клики, ни независимого множества размера более чем некоторая абсолютная константа, умноженная на логарифм числа вершин. В принципе, мы, конечно, можем для любого фиксированного n перебирать все графы на n вершинах, пока не найдем подходящий, но это не эффективный способ построения желаемых графов, а, значит, и не явный. Несмотря на то, что этой проблеме было уделено значительное внимание, она до сих пор открыта. Наилучшее явное построение дано в статье [Frankl and Wilson (1981)], где явно описываются графы на n вершинах, не содержащие ни клику, ни независимое множество на более чем $2^{c\sqrt{\log n \log\log n}}$ вершинах, где c — некоторая абсолютная положительная константа.

Хотя проблема построения явных графов Рамсея до сих пор открыта, есть несколько других проблем, для которых явные построения известны. В этой главе мы приведем несколько примеров и кратко обсудим некоторые их алгоритмические приложения. Также мы опишем несколько, казалось бы, не связанных между собой свойств графа, которые окажутся эквивалентными. Все это — свойства случайных графов, поэтому граф, удовлетворяющий этим свойствам, принято называть *квазислучайным*. Эквивалентность всех этих свойств позволяет в некоторых случаях показать, что определенные явные графы обладают рядом псевдослучайных свойств, просто показав, что они обладают одним из свойств.

9.1. ТУРНИРЫ КВАДРАТИЧНЫХ ВЫЧЕТОВ

Напомним, что *турниром* на множестве V из n игроков называется ориентация $T = (V, E)$ множества ребер полного графа на множестве вершин V. Если (x, y) — ориентированное ребро, то мы говорим, что x *побеждает* y. Пусть дана перестановка π множества игроков. Ориентированное ребро (x, y) турнира *совместимо* с π, если x предшествует y в π. Если рассматривать π как турнирное ранжирование, то разумно пытаться найти турнирные ранжирования с максимально возможным числом совместимых дуг. Пусть $c(\pi, T)$ обозначает число дуг T, совместимых с π. Введем обозначение $c(T) = \max(c(\pi, T))$, где максимум берется по всем перестановкам π на множестве вершин T. Для каждого турнира T с n игроками если $\pi = 1, 2, \ldots, n$ и $\pi' = n, n-1, \ldots, 1$, то $c(\pi, T) + c(\pi', T) = \binom{n}{2}$. Следовательно, $c(T) \geqslant \frac{1}{2}\binom{n}{2}$. На самом деле, можно показать, что для каждого такого T выполнено неравенство $c(T) \geqslant \frac{1}{2}\binom{n}{2} + \Omega(n^{3/2})$. С другой стороны, с помощью вероятностных методов легко доказать, что существуют турниры T с n игроками, для которых $c(T) \leqslant (1 + o(1))\frac{1}{2}\binom{n}{2}$. (Наилучшая из известных оценок, которая дает точный порядок максимально возможного значения разности $c(T) - \frac{1}{2}\binom{n}{2}$, довольно сложна. Показано (см. [de la Vega (1983)]), что существуют турниры T с n игроками, для которых $c(T) \leqslant \frac{1}{2}\binom{n}{2} + O(n^{3/2})$.)

Можем ли мы явно указать те турниры T на n вершинах, для которых $c(T) \leqslant (1 + o(1))\frac{1}{2}\binom{n}{2}$? Эта задача упоминается в работах [Erdős and Moon (1965)] и [Spencer (1985b)]. Оказывается, можно получить несколько таких конструкций. Мы опишем одну из них.

Пусть $p \equiv 3 \pmod 4$ — простое число, и пусть $T = T_p$ — турнир, вершинами которого являются все элементы конечного поля $GF(p)$, а (i,j) является ориентированным ребром тогда и только тогда, когда $(i-j)$ — квадратичный вычет по модулю p. (Так как $p \equiv 3 \pmod 4$, -1 — квадратичный невычет по модулю p, а, значит, любые два различных элемента соединены единственным ребром, и турнир T_p определен корректно.[1])

Теорема 9.1.1. *Для описанных выше турниров T_p справедливо соотношение*

$$c(T_p) \leqslant \frac{1}{2}\binom{p}{2} + O(p^{3/2}\log p).$$

Для доказательства этой теоремы нам потребуется предварительная подготовка. Пусть χ — характер квадратичного вычета, определенный на элементах конечного поля $GF(p)$ как $\chi(y) = y^{(p-1)/2}$. Это равносильно тому, что $\chi(y)$ равно 1, если y — ненулевой квадрат, 0, если $y = 0$ и -1 иначе. Пусть $D = (d_{ij})_{i,j=0}^{p-1}$ — матрица размера $p \times p$ с элементами $d_{ij} = \chi(i-j)$.

Утверждение. *Для любых двух различных j и l выполнено соотношение*

$$\sum_{i \in GF(p)} d_{ij}d_{il} = -1.$$

Доказательство. Выпишем цепочку равенств:

$$\sum_i d_{ij}d_{il} = \sum_i \chi(i-j)\chi(i-l) = \sum_{i \neq j,l} \chi(i-j)\chi(i-l) =$$
$$= \sum_{i \neq j,l} \chi((i-j)/(i-l)) = \sum_{i \neq j,l} \chi(1 + (l-j)/(i-l)).$$

По мере того как i принимает все значения из поля $GF(p)$, кроме j и l, величина $(1 + (l-j)/(i-l))$ принимает все значения из $GF(p)$, кроме 0 и 1. Так как сумма $\chi(r)$ по всем r из $GF(p)$ равна 0, отсюда следует, что правая часть последнего соотношения равна $0 - \chi(0) - \chi(1) = -1$. Это завершает доказательство утверждения. ■

Пусть A и B — подмножества поля $GF(p)$. Обозначим через $e(A,B)$ количество ориентированных ребер T_p, начинающихся в вершине из A и заканчивающихся в вершине из B. Из определения матрицы D следует, что

$$\sum_{i \in A}\sum_{j \in B} d_{ij} = e(A,B) - e(B,A).$$

Следующая лемма была доказана в статье [Alon (1986b)].

Лемма 9.1.2. *Для любых двух подмножеств A и B поля $GF(p)$ справедливо неравенство*

$$\left|\sum_{i \in A}\sum_{j \in B} d_{ij}\right| \leqslant |A|^{1/2}|B|^{1/2}p^{1/2}.$$

[1] Число a называется квадратичным вычетом по модулю m, если сравнение $x^2 \equiv a \pmod m$ имеет решение. В противном случае a называется квадратичным невычетом. — *Прим. ред.*

Доказательство. Из неравенства Коши—Буняковского и из предыдущего утверждения следует, что

$$\left(\sum_{i\in A}\sum_{j\in B} d_{ij}\right)^2 \leqslant |A|\left(\sum_{i\in A}\left(\sum_{j\in B} d_{ij}\right)^2\right) \leqslant |A|\left(\sum_{i\in GF(p)}\left(\sum_{j\in B} d_{ij}\right)^2\right) =$$

$$= |A|\left(\sum_{i\in GF(p)}\left(|B| + 2\sum_{j<l\in B} d_{ij}d_{il}\right)\right) =$$

$$= |A||B|p + 2|A|\sum_{j<l\in B}\sum_{i\in GF(P)} d_{ij}d_{il} =$$

$$= |A||B|p - |A||B|\big(|B|-1\big) = |A||B|(p-|B|+1) \leqslant |A||B|p.$$

Этим завершается доказательство леммы. ■

Доказательство теоремы 9.1.1. Пусть r — наименьшее целое число, удовлетворяющее неравенству $2^r \geqslant p$. Пусть $\pi = \pi_1, \ldots, \pi_p$ — произвольная перестановка вершин турнира T_p. Введем обозначение $\pi' = \pi_p, \ldots, \pi_1$. Мы должны доказать, что $c(\pi, T_p) \leqslant \frac{1}{2}\binom{p}{2} + O(p^{3/2}\log p)$, или эквивалентное утверждение о том, что $c(\pi, T_p) - c(\pi', T_p) \leqslant O(p^{3/2}\log p)$. Пусть a_1 и a_2 — два целых числа, удовлетворяющих условиям $p = a_1 + a_2$ и $a_1 \leqslant 2^{r-1}, a_2 \leqslant 2^{r-1}$. Обозначим через A_1 множество первых a_1 вершин в перестановке π, а через A_2 — множество последних a_2 вершин в π. По лемме 9.1.2

$$e(A_1, A_2) - e(A_2, A_1) \leqslant (a_1 a_2 p)^{1/2} \leqslant 2^{r-1}p^{1/2}.$$

Далее, пусть $a_{11}, a_{12}, a_{21}, a_{22}$ — целые числа, каждое из которых не превосходит 2^{r-2}, такие, что $a_1 = a_{11} + a_{12}$ и $a_2 = a_{21} + a_{22}$. Пусть A_{11} — подмножество A_1, состоящее из элементов A_1, занимающих первые a_{11} позиций в π, и пусть A_{12} — множество из a_{12} оставшихся элементов A_1. Разбиение A_2 на два множества A_{21} и A_{22} определяется аналогично. Применяя лемму 9.1.2, получаем:

$$e(A_{11}, A_{12}) - e(A_{12}, A_{11}) + e(A_{21}, A_{22}) - e(A_{22}, A_{21}) \leqslant$$
$$\leqslant (a_{11}a_{12}p)^{1/2} + (a_{21}a_{22}p)^{1/2} \leqslant 2\cdot 2^{r-2}p^{1/2}.$$

Продолжая действовать таким же образом, мы получим на i-м шаге разбиение множества вершин на 2^i блоков, каждый из которых состоит не более чем из 2^{r-i} последовательных элементов перестановки π. Это разбиение получено делением каждого блока разбиения, соответствующего предыдущему шагу, на две части. Применяя лемму 9.1.2 для каждой такой пары $A_{\varepsilon 1}, A_{\varepsilon 2}$ (здесь ε — вектор длины $i-1$, состоящий из единиц и двоек), и суммируя, мы делаем вывод, что сумма разностей $e(A_{\varepsilon 1}, A_{\varepsilon 2}) - e(A_{\varepsilon 2}, A_{\varepsilon 1})$ по всем 2^{i-1} векторам ε не превышает

$$2^{i-1}2^{r-i}p^{1/2} = 2^{r-1}p^{1/2}.$$

Заметим, что сумма левых частей этих равенств по i от 1 до r в точности равна разности $c(\pi, T_p) - c(\pi', T_p)$. Следовательно, суммируя, мы получаем

$$c(\pi, T_p) - c(\pi', T_p) = 2^{r-1}p^{1/2}r = O(p^{3/2}\log p).$$

Доказательство завершено. ■

Заметим, что любая кососимметрическая матрица из элементов $+1$ и -1, в которой каждые две строки почти ортогональны, может быть использована для построения турнира, как это сделано выше. Некоторые родственные результаты появлялись ранее в статье [Frankl, Rödl and Wilson (1988)]. Однако, турниры T_p обладают более сильными псевдослучайными свойствами, чем другие турниры. Например, для каждого $k \leqslant \frac{1}{4}\log p$, и для каждого множества S из k вершин турнира T_p число вершин T_p, побеждающих все элементы S, равно $(1+o(1))p/2^k$. Это было доказано Грэхемом и Спенсером [Graham and Spencer (1971)] с помощью знаменитой теоремы Вейля, известной как гипотеза Римана о кривых над конечными полями [Weil (1948)]. При достаточно больших p приведенные рассуждения дают явное построение для задачи Шютте, упомянутой в гл. 1.

9.2. СОБСТВЕННЫЕ ЗНАЧЕНИЯ И РАСШИРИТЕЛИ

Граф $G = (V, E)$ называется (n, d, c)-*расширителем*, если в нем n вершин, максимальная степень вершины равна d, и для каждого множества вершин $W \subset V$ такого, что $|W| \leqslant n/2$, выполнено неравенство $|N(W)| \geqslant c|W|$, где $N(W)$ обозначает множество всех вершин в $V \setminus W$, смежных с какой-либо вершиной из W. Заметим, что иногда используется несколько иное определение, но разница несущественна. Расширители обладают многими свойствами разреженных случайных графов, и по ним имеется обширная литература. Семейство *линейных расширителей плотности* d *и расширения* c — это последовательность $\{G_i\}_{i=1}^{\infty}$, где G_i — (n_i, d, c)-расширитель, и $n_i \to \infty$ при $i \to \infty$.

Такое семейство является главным компонентом параллельной схемы сортировки (см. [Ajtai, Komlós and Szemerédi (1983)]). Оно может быть использовано для создания определенных помехоустойчивых линейных таблиц. Кроме того, оно образует основной строительный блок, используемый при создании графов с особыми свойствами связности и небольшим числом ребер. Некоторые другие примеры многочисленных приложений этих графов в различных задачах теоретической информатики можно найти, например, в работе [Alon (1986b)] и списке литературы к ней. Не слишком сложно доказать существование семейства линейных расширителей с помощью вероятностных методов. Впервые это сделал Пинскер [Pinsker (1973)]. Найти явное построение гораздо сложнее. Первым это сделал Маргулис (1973). Позже это построение было улучшено различными авторами; наиболее известными построениями являются графы Кэли определенных групп матриц. Их расширительные свойства доказываются с помощью оценки собственных значений матриц смежности графов, опираясь на тесную связь между расширительными свойствами графа и его спектральными свойствами. Впервые эта связь была независимо изучена в работах [Tanner (1984)] и [Alon and Milman (1984)]. Так как это несколько проще в случае регулярных графов, на этом случае мы и сосредоточим внимание.

Пусть $G = (V, E)$ — d-регулярный граф, и пусть $A = A_G = (a_{uv})_{u,v \in V}$ — его матрица смежности, в которой $a_{uv} = 1$, если $uv \in E$, и $a_{uv} = 0$ в противном

случае. Так как G — d-регулярный граф, то наибольшее собственное значение матрицы A равно d. Оно соответствует собственному вектору, целиком состоящему из единиц. Пусть $\lambda = \lambda(G)$ обозначает второе по величине собственное значение матрицы A. Для двух (не обязательно непересекающихся) подмножеств B и C множества V пусть $e(B,C)$ обозначает число упорядоченных пар (u,v), где $u \in B$, $v \in C$ и uv — ребро G. (Заметим, что если B и C не пересекаются, то это просто количество ребер графа G, соединяющих вершины из B с вершинами из C.)

Теорема 9.2.1. *Для любого разбиения множества вершин V на два непересекающихся подмножества B и C выполнено неравенство*

$$e(B,C) \geqslant \frac{(d-\lambda)|B||C|}{n}.$$

Доказательство. Положим $|V| = n, b = |B|, c = |C| = n - b$. Пусть $D = dI$ — скалярная матрица размера n на n со степенью вершин графа G на диагонали. Заметим, что для любого действительного вектора x длины n (рассматриваемого как функция $x : V \mapsto \mathbb{R}^n$)

$$((D-A)x,x) = \sum_{u\in V}(d(x(u))^2 - \sum_{v;uv\in E} x(v)x(u)) =$$
$$= d\sum_{u\in V}(x(u))^2 - 2\sum_{uv\in E} x(v)x(u) = \sum_{uv\in E}(x(v)-x(u))^2.$$

Теперь определим вектор x, полагая $x(v) = -c$ при $v \in B$ и $x(v) = b$ при $v \in C$. Заметим, что у матриц A и $D - A$ одинаковые собственные векторы, и что собственные значения матрицы $D - A$ в точности равны $d - \mu$, где μ пробегает все собственные значения матрицы A. Также заметим, что $\sum_{v\in V} x(v) = 0$, т. е. вектор x ортогонален собственному вектору матрицы $D - A$, отвечающему наименьшему собственному значению. Так как $D - A$ — симметричная матрица, ее собственные векторы попарно ортогональны и составляют базис n-мерного пространства. Следовательно, x — это линейная комбинация других собственных векторов матрицы $D - A$, а значит, из определения λ и из того факта, что $d - \lambda$ — второе снизу по величине собственное значение матрицы $D - A$, мы делаем вывод, что

$$((D-A)x,x) \geqslant (d-\lambda)(x,x) = (d-\lambda)(bc^2 + cb^2) = (d-\lambda)bcn.$$

Из второго абзаца доказательства следует, что левая часть последнего неравенства равна $\sum_{uv\in E}(x(u) - x(v))^2 = e(B,C)\cdot(b+c)^2 = e(B,C)\cdot n^2$. Следовательно,

$$e(B,C) \geqslant \frac{(d-\lambda)bc}{n}.$$ ■

Следствие 9.2.2. *Если λ — второе по величине собственное значение d-регулярного графа G с n вершинами, то G — (n,d,c)-расширитель при $c = \frac{d-\lambda}{2d}$.*

Доказательство. Пусть W — множество $w \leqslant n/2$ вершин графа G. По теореме 9.2.1 существует не менее $\frac{(d-\lambda)w(n-w)}{n} \geqslant \frac{(d-\lambda)w}{2}$ ребер из W в его дополнении. Так как ни одна из вершин дополнения не смежна с более чем d из этих ребер, получаем, что $|N(W)| \geqslant \frac{(d-\lambda)w}{2d}$. ■

На самом деле, оценка для c в последнем следствии может быть улучшена до $\frac{2(d-\lambda)}{3d-2\lambda}$, как показано в статье [Alon and Milman (1984)]. Каждая из этих оценок показывает, что если второе по величине собственное значение графа G далеко от первого, то G — хороший расширитель. Обратное утверждение также верно, хотя и более сложно доказывается. Доказательство можно найти в статье [Alon (1986a)]. Здесь мы приводим это утверждение без доказательства.

Теорема 9.2.3. *Если G — d-регулярный граф, являющийся (n, d, c)-расширителем, то $\lambda(G) \leqslant d - \frac{c^2}{4+2c^2}$.*

Два последних результата дают эффективный алгоритм для приближения расширительных свойств d-регулярного графа; мы просто вычисляем (или оцениваем) его второе по величине собственное значение. Чем больше разница между этим собственным значением и d, тем лучше следующие из этого расширительные свойства графа G. Поэтому естественно интересоваться, насколько d может быть удалено от этого второго собственного значения. Известно (см. [Nilli (1991)]), что второе по величине собственное значение любого d-регулярного графа с диаметром k не меньше $2\sqrt{d-1}(1 - O(1/k))$. Следовательно, в любом бесконечном семействе d-регулярных графов верхний предел второго по величине собственного значения не меньше $2\sqrt{d-1}$. Явные построения бесконечных семейств d-регулярных графов G_i со вторыми по величине собственными значениями $\lambda(G_i) \leqslant 2\sqrt{d-1}$ независимо были представлены в работах [Lubotzky, Phillips and Sarnak (1986)] и [Маргулис (1988)] (при этом $d = p + 1$, где p — простое число, сравнимое с 1 по модулю 4). Построенные графы являются графами Кэли фактор-групп группы всех обратимых матриц размера 2×2 над конечным полем. Их собственные значения оцениваются с помощью результатов Эйхлера и Игуса по гипотезе Рамануджана. Доказательство Эйхлера опирается на теорему Вейля, упомянутую в предыдущем разделе. Не двудольные графы G, построенные таким образом, удовлетворяют немного более сильному условию $\lambda(G) \leqslant 2\sqrt{d-1}$. В действительности, за исключением их наибольших собственных значений d, у них нет собственных значений, превосходящих по модулю $2\sqrt{d-1}$. Как показывают следующие результаты, из этого факта следуют некоторые сильные псевдослучайные свойства.

Теорема 9.2.4. *Пусть $G = (V, E)$ — d-регулярный граф с n вершинами. Предположим, что модули всех его собственных значений, кроме первого, не превосходят λ. Пусть $v \in V$, а B — подмножество V. Обозначим через $N(v)$ множество всех соседей вершины v в G, и пусть $N_B(v) = N(v) \cap B$ обозначает множество всех соседей вершины v в B. Тогда для любого подмножества*

$B \subseteq V$ мощности bn справедливо неравенство

$$\sum_{v\in V}(|N_B(v)| - bd)^2 \leqslant \lambda^2 b(1-b)n.$$

Заметим, что в случайном d-регулярном графе у каждой вершины v, как правило, примерно bd соседей в каждом множестве размера bn. Приведенная выше теорема показывает, что если λ гораздо меньше d, то $|N_B(v)|$ не очень далеко от bd для большинства вершин v.

Доказательство. Пусть A — матрица смежности графа G. Определим вектор $f : V \mapsto \mathbb{R}$, полагая $f(v) = 1-b$ при $v \in B$ и $f(v) = -b$ при $v \notin B$. Ясно, что $\sum_{v\in V} f(v) = 0$; т. е. вектор f ортогонален собственному вектору, отвечающему наибольшему собственному значению матрицы A. Следовательно,

$$(Af, Af) \leqslant \lambda^2 (f, f).$$

Правая часть последнего неравенства равна $\lambda^2(bn(1-b)^2+(1-b)nb^2) = \lambda^2 b(1-b)n$. Левая часть равна

$$\sum_{v\in V}((1-b)|N_B(v)| - b(d - |N_B(v)|))^2 = \sum_{v\in V}(|N_B(v)| - bd)^2.$$

Отсюда следует желаемый результат. ■

Следствие 9.2.5. *Пусть $G = (V, E), d, n$ и λ — те же, что и в теореме 9.2.4. Тогда для любых двух множеств B и C вершин графа G при $|B| = bn$ и $|C| = cn$ справедливо неравенство*

$$|e(B, C) - cbdn| \leqslant \lambda\sqrt{bc}\, n.$$

Доказательство. По теореме 9.2.4 имеем

$$\sum_{v\in C}(|N_B(v)| - bd)^2 \leqslant \sum_{v\in V}(|N_B(v)| - bd)^2 \leqslant \lambda^2 b(1-b)n.$$

Тогда согласно неравенству Коши—Буняковского

$$|e(B, C) - cbdn| \leqslant \sum_{v\in C}|N_B(v) - bd| \leqslant$$

$$\leqslant \sqrt{cn}\left(\sum_{v\in C}(|N_B(v)| - bd)^2\right)^{1/2} \leqslant \sqrt{cn}\lambda\sqrt{b(1-b)n} \leqslant \lambda\sqrt{bc}\, n.$$ ■

В частном случае $B = C$ получаем следующий результат. Немного более сильная оценка аналогичным образом доказана в статье [Alon and Chung (1988)].

Следствие 9.2.6. *Пусть $G = (V, E), d, n$ и λ — те же, что и в теореме 9.2.4. Обозначим через B произвольное множество из bn вершин графа G, и пусть $e(B) = \frac{1}{2}e(B, B)$ — число ребер в индуцированном подграфе G на B.*

Тогда

$$\left| e(B) - \frac{1}{2}b^2 dn \right| \leqslant \frac{1}{2}\lambda bn.$$

Путь длины l в графе G — это последовательность $v_0, \ldots, v_l$ вершин G, где для каждого $i, 1 \leqslant i \leqslant l$, $v_{i-1}v_i$ является ребром графа G. Ясно, что общее число путей длины l в d-регулярном графе с n вершинами в точности равно $n \cdot d^l$. Предположим теперь, что C — подмножество из, например, $n/2$ вершин графа G. Сколько этих путей не содержат ни одной вершины из C? Если граф G — несвязный, то может случиться, что половина этих путей избегает C. Тем не менее, как показано в [Ajtai, Komlós and Szemerédi (1987)], в случае, когда все собственные значения G, кроме наибольшего, малы, существует гораздо меньше таких путей. У этого результата и некоторых его обобщений есть несколько приложений в теоретической информатике, как показано в упомянутой выше работе (см. также [Cohen and Wigderson (1989))]. В завершение раздела мы приведем и докажем этот результат и одно из его приложений.

Теорема 9.2.7. *Пусть* $G = (V, E)$ — d*-регулярный граф с* n *вершинами. Предположим, что каждое из его собственных значений, кроме первого, не превосходит* λ*. Пусть* C — *множество из* cn *вершин графа* G*. Тогда для любого* l *число путей длины* l *в* G*, избегающих* C*, не превосходит* $(1-c)n((1-c)d + c\lambda))^l$*.*

Доказательство. Пусть A — матрица смежности графа G, и пусть A' — матрица смежности его индуцированного подграфа на дополнении C. Мы утверждаем, что максимальное собственное значение матрицы A' не превосходит $(1-c)d + c\lambda$. Для того, чтобы это доказать, мы должны показать, что для любого вектора $f : V \mapsto \mathbb{R}^n$ такого, что $f(v) = 0$ для каждого $v \in C$ и $\sum_{v \in V} f(v)^2 = 1$, верно неравенство $(Af, f) \leqslant (1-c)d + c\lambda$. Пусть $f_1, f_2, \ldots, f_n$ — ортонормальный базис из собственных векторов матрицы A, где f_1 — собственный вектор λ_i, $\lambda_1 = d$, и каждый элемент вектора f_1 равен $1/\sqrt{n}$. Тогда $f = \sum_{i=1}^n c_i f_i$, где $\sum_{i=1}^n c_i^2 = 1$ и

$$c_1 = \sum_{v \in V} f(v)/\sqrt{n} = \sum_{v \in V \setminus C} f(v)/\sqrt{n} \leqslant$$

$$\leqslant \left(\sum_{v \in V \setminus C} f(v)^2 \right)^{1/2} ((1-c)n/n)^{1/2} = \sqrt{1-c}.$$

Здесь мы использовали неравенство Коши—Буняковского. Следовательно, $\sum_{i=2}^n c_i^2 = c$, и

$$(Af, f) = \sum_{i=1}^n c_i^2 \lambda_i \leqslant (1-c)d + c\lambda,$$

что дает нам желаемую оценку для максимального собственного значения матрицы A'.

Пусть $\gamma_1 \geqslant \gamma_2 \geqslant \ldots \geqslant \gamma_m$ — собственные значения матрицы A', где $m = (1-c)n$. По теореме Фробениуса—Перрона отсюда следует, что модули

каждого из них не превосходят $\gamma_1 \leqslant (1-c)d + c\lambda$. Общее число путей длины l, избегающих C, в точности равно $(A'^l g, g)$, где g — вектор, целиком состоящий из индикаторов вершин множества $V \setminus C$. Выражая g как линейную комбинацию собственных векторов матрицы A', т. е. $g = \sum_{i=1}^m b_i g_i$, где g_i — собственный вектор с собственным значением γ_i, мы делаем вывод, что это число в точности равно

$$\sum_{i=1}^m b_i^2 \gamma_i^l \leqslant \gamma_1^l \sum_{i=1}^m b_i^2 = m\gamma_1^l \leqslant m((1-c)d + c\lambda)^l.$$

Подставляя $m = (1-c)n$, получаем желаемый результат. ■

Случайно выбранный путь длины l в графе G — это путь длины l в G, выбранный согласно равномерному распределению среди всех путей этой длины. Заметим, что если G — d-регулярный граф, то такой путь может быть выбран с помощью случайного выбора начальной точки v_0, и случайного выбора v_i для каждого $1 \leqslant i \leqslant l$ среди d соседей вершины v_{i-1}.

Следствие 9.2.8. *Пусть $G = (V, E), d, n, \lambda, C$ и c — те же, что и в теореме 9.2.7. Предположим, что*

$$(1-c)d + c\lambda \leqslant \frac{d}{\sqrt{2}}.$$

Тогда для любого l вероятность того, что случайно выбранный путь длины l в графе G, избегает множество C, не превосходит $2^{-l/2}$.

Доказательство. Число путей длины l в графе G, избегающих множество C, не больше $(1-c)n((1-c)d + c\lambda)^l \leqslant nd^l 2^{-l/2}$ по теореме 9.2.7. Так как общее число путей равно nd^l, получаем желаемый результат. ■

Приведенные выше результаты полезны для увеличения вероятностей в рандомизированных алгоритмах. Хотя такое увеличение может быть достигнуто для любого алгоритма Монте-Карло, мы для простоты рассмотрим один характерный пример — алгоритм проверки простоты, описанный в работе [Rabin (1980)].

Для нечетного q обозначим через a и b два целых числа, однозначно определяемых из соотношения $q-1 = 2^a b$, где b нечетно. Целое x, $1 \leqslant x \leqslant q-1$, называется *свидетелем* (непростоты q), если для последовательности $x_0, \ldots, x_a$, определенной равенствами $x_0 = x^b \pmod q$ и $x_i = x_{i-1}^2 \pmod q$ для $1 \leqslant i \leqslant a$, либо $x_a \neq 1$, либо существует такое i, что $x_i \neq -1, 1$ и $x_{i+1} = 1$. Можно показать, что если q — простое, то таких свидетелей для q не существует; если же q — нечетное составное число, то по крайней мере половина чисел между 1 и $q-1$ — свидетели для q. (На самом деле, не менее $3/4$ из них — свидетели, как показал Рабин.) Это наводит нас на мысль о следующем случайном алгоритме проверки того, что нечетное число q является простым (для четных целых чисел есть более простой алгоритм!).

Выберем случайно целое x между 1 и $q-1$ и проверим, является ли оно свидетелем. Если является, то сообщаем, что q — не простое. В противном случае сообщаем, что q — простое.

Заметим, что в случае, когда q — простое, алгоритм гарантированно сообщает, что оно простое, в то время как в случае, когда q не простое, вероятность того, что алгоритм делает ошибку и сообщает, что оно простое, не превосходит $1/2$. Что если нам хочется уменьшить вероятность такой ошибки? Ясно, что мы можем просто повторить алгоритм. Если мы независимо повторим его l раз, то вероятность сделать ошибку (т. е. сообщить, что составное число — простое) уменьшается до $1/2^l$. Тем не менее, количество случайных битов, необходимых для этой процедуры, равно $l \cdot \log_2(q-1)$.

Предположим, нам надо использовать меньше случайных битов. Применяя свойства случайно выбранного пути в подходящем графе, доказанные в двух предыдущих результатах, мы можем получить ту же оценку для вероятности ошибки, используя лишь $\log_2(q-1) + O(l)$ случайных битов. Это делается следующим образом.

Пусть G — d-регулярный граф с $q-1$ вершинами, помеченными всеми целыми числами от 1 до $q-1$. Предположим, что у G нет собственных значений, модуль которых больше λ, кроме первого, и предположим, что

$$\frac{d+\lambda}{2} \leqslant \frac{d}{\sqrt{2}}. \tag{9.1}$$

Теперь выберем случайно путь длины $2l$ в графе G и проверим для каждого числа, отмечающего его вершину, является ли оно свидетелем. Если q составное, то по крайней мере половина вершин графа G помечены его свидетелями. Значит, согласно следствию 9.2.8 и неравенству (9.1) вероятность того, что на пути нет свидетелей, не превосходит $2^{-2l/2} = 2^{-l}$. Таким образом, мы получили такое же уменьшение вероятности ошибки, что и при выборе l независимых свидетелей. Оценим количество случайных битов, необходимых для выбора такого случайного пути.

Известные построения расширителей, полученные в работах [Lubotzky et al. (1986)] и [Маргулис (1988)], дают явные семейства графов со степенью d и с $\lambda \leqslant 2\sqrt{d-1}$ для каждого $d = p+1$, где p — простое число, сравнимое с 1 по модулю 4. (Заметим, что у этих графов не ровно $q-1$ вершин, но это не вызывает никаких реальных трудностей, так как мы можем взять граф с n вершинами, где $q-1 \leqslant n \leqslant (1+o(1))(q-1)$, и пометить его i-ю вершину числом $i \pmod{(q-1)}$. В этом случае число вершин, помеченных свидетелями, все еще будет не меньше $(\frac{1}{2} + o(1))n$.) Нетрудно проверить, что, например, $d = 30$ и $\lambda = 2\sqrt{29}$ удовлетворяют неравенству (9.1), а, значит, можно использовать 30-регулярный граф. Число случайных битов, необходимых для выбора в нем случайного пути длины $2l$, меньше, чем $\log_2(q-1) + 10l + 1$. Это намного меньше, чем $l\log_2(q-1)$ битов, необходимых в процедуре с повторением.

9.3. КВАЗИСЛУЧАЙНЫЕ ГРАФЫ

В этом разделе мы опишем некоторые псевдослучайные свойства графов, которые довольно неожиданно оказываются эквивалентными. Всем этим свойствам удовлетворяет почти наверное случайный граф, в котором каждое ребро выбирается независимо с вероятностью $1/2$. Эквивалентность некоторых из этих свойств впервые была доказана несколькими авторами (см. [Thomason (1987)], [Frankl et al. (1988)] и [Alon and Chung (1988)]), но [Chung, Graham and Wilson (1989)] — первая работа, в которой появляются все указанные свойства (и некоторые другие). Мы следуем этой работе, хотя для упрощения изложения рассматриваем только регулярные графы.

Прежде всего, нам понадобятся некоторые обозначения. Для двух графов G и H пусть $N^*_G(H)$ — число помеченных вхождений H как индуцированного подграфа G (т. е. число инъекций $f : V(H) \mapsto V(G)$, сохраняющих смежность, образом которых является множество вершин индуцированной копии H в G). Аналогично, $N_G(H)$ обозначает число помеченных копий H как (не обязательно индуцированного) подграфа G. Заметим, что $N_G(H) = \sum_L N^*_G(L)$, где сумма берется по всем графам на множестве вершин H, полученным из H добавлением (возможно, пустого) множества ребер.

В этом разделе G всегда обозначает граф с n вершинами. Обозначим собственные значения его матрицы смежности (взятой с учетом кратности ребер) через $\lambda_1, \dots, \lambda_n$, где $|\lambda_1| \geqslant \dots \geqslant |\lambda_n|$. (Так как в этом разделе мы рассматриваем только собственные значения G, мы пишем просто λ_1, а не $\lambda_1(G)$.) Также напомним следующее обозначение, использованное в предыдущем разделе: для вершины v графа G множество ее соседей в G обозначается через $N(v)$. Если S — подмножество вершин графа G, то $e(S)$ обозначает число ребер в индуцированном подграфе G на S. Если B и C — два (необязательно непересекающихся) подмножества вершин графа G, то $e(B, C)$ обозначает количество упорядоченных пар (b, c), где $b \in B$, $c \in C$ и bc — ребро графа G. Тогда $e(S) = \frac{1}{2}e(S, S)$.

Теперь мы можем перечислить упомянутые псевдослучайные свойства. Все они относятся к графу $G = (V, E)$ с n вершинами. В этом разделе мы используем обозначения $o(\cdot)$ без упоминания точного поведения каждой такой величины. Таким образом, если в выражении встречаются два $o(1)$, то они не обязаны быть идентичными. Но если мы рассмотрим семейство графов G и устремим число вершин n к бесконечности, то каждое $o(1)$ стремится к нулю.

Свойство $P_1(s)$**.** Для каждого графа $H(s)$ на s вершинах

$$N^*_G(H(s)) = (1 + o(1))n^s 2^{-\binom{s}{2}}.$$

Свойство P_2**.** Для цикла $C(4)$ с 4 вершинами $N_G(C(4)) \leqslant (1 + o(1))(n/2)^4$.

Свойство P_3**.** $|\lambda_2| = o(n)$.

Свойство P_4**.** Для каждого подмножества S вершин графа G выполнено соотношение $e(S) = \frac{1}{4}|S|^2 + o(n^2)$.

Свойство P_5**.** Для любых двух множеств вершин B и C выполнено соотношение $e(B, C) = \frac{1}{2}|B||C| + o(n^2)$.

Свойство P_6**.** $\sum_{u,v \in V} \left| \, |N(u) \cap N(v)| - \frac{n}{4} \right| = o(n^3)$.

Легко проверить, что все эти свойства почти наверное выполнены для случайного графа на n вершинах. В этом разделе мы покажем, что все свойства эквивалентны для регулярного графа с n вершинами и степенью вершин около $n/2$. Один из интересных особых случаев этого результата заключается в том, что простое с виду свойство P_2 оказывается достаточным для выполнения свойства $P_1(s)$ при всех $s \geqslant 1$.

Графы, удовлетворяющие любому из этих свойств (а значит, и всем), называются *квазислучайными*. Как было сказано выше, предположение о том, что граф G регулярный, может быть опущено (при этом немного изменится свойство P_2, и немного усложнятся доказательства).

Теорема 9.3.1. *Положим $d = (\frac{1}{2} + o(1))n$. Если d-регулярный граф G на n вершинах удовлетворяет любому из семи свойств $P_1(4), P_1(s)$ для всех $s \geqslant 1$, P_2, P_3, P_4, P_5, P_6, то он удовлетворяет всем семи свойствам.*

Доказательство. Мы покажем, что

$$P_1(4) \Longrightarrow P_2 \Longrightarrow P_3 \Longrightarrow P_4 \Longrightarrow P_5 \Longrightarrow$$
$$\Longrightarrow P_6 \Longrightarrow P_1(s) \text{ для всех } s \geqslant 1 \ \ (\Longrightarrow P_1(4)).$$

1. $P_1(4) \Longrightarrow P_2$.
Предположим, что граф G удовлетворяет свойству $P_1(4)$. Тогда $N_G(C(4)) = \sum_L N^*_G(L)$, где сумма берется по четырем помеченным графам, полученным из помеченного $C(4)$ добавлением к нему (возможно, пустого) множества ребер. Так как G удовлетворяет свойству $P_1(4)$, то $N^*_G(L) = (1 + o(1))n^4 2^{-16}$ для каждого из этих графов L, а, значит, $N_G(C(4)) = (1+o(1))n^4 2^{-4}$. Следовательно, граф G удовлетворяет свойству P_2.
2. $P_2 \Longrightarrow P_3$.
Предположим, что граф G удовлетворяет свойству P_2, и пусть A — его матрица смежности. След матрицы A^4 в точности равен $\sum_{i=1}^n \lambda_i^4$. С другой стороны, нетрудно видеть, что этот след равен числу (помеченных) замкнутых путей длины 4 в графе G, т. е. числу последовательностей $v_0, v_1, v_2, v_3, v_4 = v_0$ вершин графа G, таких что $v_i v_{i+1}$ — ребро для каждого $0 \leqslant i \leqslant 3$. Это число равно $N_G((C(4))$ плюс число таких последовательностей, в которых $v_2 = v_0$, равное nd^2, плюс число таких последовательностей, в которых $v_2 \neq v_0$ и $v_3 = v_1$, равное $nd(d-1)$. Тогда

$$\sum_{i=1}^n \lambda_i^4 = d^4 + \sum_{i=2}^n \lambda_i^4 = (1 + o(1))(n/2)^4 + \sum_{i=2}^n \lambda_i^4 = N_G(C(4)) + O(n^3) =$$
$$= (1 + o(1))(n/2)^4.$$

Следовательно, $\sum_{i=2}^n \lambda_i^4 = o(n^4)$, а, значит, $|\lambda_2| = o(n)$, что и требовалось.
3. $P_3 \Longrightarrow P_4$.
Это напрямую вытекает из следствия 9.2.6.
4. $P_4 \Longrightarrow P_5$.
Предположим, что граф G удовлетворяет свойству P_4. Прежде всего, мы утверждаем, что для него выполнено свойство P_5 для непересекающихся множеств вершин B и C. Действительно, если множества B и C не пересекаются,

то

$$e(B,C) = e(B \cup C) - e(B) - e(C) = \\ = \frac{1}{4}(|B| + |C|)^2 - \frac{1}{4}|B|^2 - \frac{1}{4}|C|^2 + o(n^2) = \frac{1}{2}|B||C| + o(n^2).$$

Утверждение доказано.

В случае, когда множества B и C пересекаются, получаем

$$e(B,C) = e(B \setminus C, C \setminus B) + e(B \cap C, C \setminus B) + e(B \cap C, B \setminus C) + 2e(B \cap C).$$

Положим $|B| = b, |C| = c, |B \cap C| = x$. По предыдущему выражению для $e(B,C)$ и из того факта, что граф G удовлетворяет свойствам P_4 и P_5 для непересекающихся множеств B и C, получаем

$$e(B,C) = \frac{1}{2}(b-x)(c-x) + \frac{1}{2}x(c-x) + \frac{1}{2}x(b-x) + \frac{2}{4}x^2 + o(n^2) = \\ = \frac{1}{2}bc + o(n^2) = \frac{1}{2}|B||C| + o(n^2).$$

Следовательно, для графа G выполнено свойство P_5.

5. $P_5 \Longrightarrow P_6$.

Предположим, что граф G удовлетворяет свойству P_5 и напомним, что G — d-регулярный, где $d = (\frac{1}{2} + o(1))n$. Пусть v — фиксированная вершина G. Оценим сумму

$$\sum_{u \in V, u \neq v} \Big| \, |N(u) \cap N(v)| - \frac{n}{4} \Big| \, .$$

Положим

$$B_1 = \Big\{ u \in V, u \neq v : \ |N(u) \cap N(v)| \geqslant \frac{n}{4} \Big\}, \\ B_2 = \Big\{ u \in V, u \neq v : \ |N(u) \cap N(v)| < \frac{n}{4} \Big\}.$$

Пусть C — множество всех соседей вершины v в G. Заметим, что

$$\sum_{u \in B_1} \Big| \, |N(u) \cap N(v)| - \frac{n}{4} \Big| = \sum_{u \in B_1} |N(u) \cap N(v)| - |B_1|\frac{n}{4} = e(B_1, C) - |B_1|\frac{n}{4}.$$

Так как граф G удовлетворяет свойству P_5, а значит, $d = (\frac{1}{2} + o(1))n$, последняя разность равна $\frac{1}{2}|B_1|d + o(n^2) - |B_1|\frac{n}{4} = o(n^2)$.

Рассуждая аналогично, получим, что

$$\sum_{u \in B_2} \Big| \, |N(u) \cap N(v)| - \frac{n}{4} \Big| = o(n^2).$$

Следовательно, для каждой вершины v графа G

$$\sum_{u \in V, u \neq v} \Big| \, |N(u) \cap N(v)| - \frac{n}{4} \Big| = o(n^2),$$

и, суммируя по всем вершинам v, мы делаем вывод, что граф G удовлетворяет свойству P_6.

6. $P_6 \Longrightarrow P_1(s)$**для всех** $s \geqslant 1$.

Предположим, что граф $G = (V, E)$ удовлетворяет свойству P_6. Для любых двух различных вершин u и v графа G пусть $a(u, v)$ равно 1, если $uv \in E$, и 0 в противном случае. Положим $s(u, v) = |\{w \in V : a(u, w) = a(v, w)\}|$. Так как G является $d = (\frac{1}{2} + o(1))n$-регулярным, то $s(u, v) = 2|N(u) \cap N(v)| + n - 2d = 2|N(u) \cap N(v)| + o(n)$. Следовательно, из того, что G удовлетворяет свойству P_6, следует, что

$$\sum_{u,v \in V} \left| s(u, v) - \frac{n}{2} \right| = o(n^3). \tag{9.2}$$

Пусть $H = H(s)$ — произвольный фиксированный граф на s вершинах; положим $N_s = N^*_G(H(s))$. Мы должны показать, что

$$N_s = (1 + o(1))n^s 2^{-\binom{s}{2}}.$$

Обозначим множество вершин графа $H(s)$ через $\{v_1, \ldots, v_s\}$. Для каждого $1 \leqslant r \leqslant s$ положим $V_r = \{v_1, \ldots, v_r\}$, и пусть $H(r)$ — индуцированный подграф графа H на множестве вершин V_r. Мы докажем индукцией по r, что для $N_r = N^*_G(H(r))$ справедливо равенство

$$N_r = (1 + o(1))n_{(r)} 2^{-\binom{r}{2}}, \tag{9.3}$$

где $n_{(r)} = n(n-1)\cdots(n-r+1)$.

Это очевидно для $r = 1$. Предположим, что равенство справедливо при всех r, $1 \leqslant r < s$, и докажем его для $r + 1$. Для вектора $\alpha = (\alpha_1, \ldots, \alpha_r)$ из различных вершин графа G и для вектора $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_r)$ из нулей и единиц положим

$$f_r(\alpha, \varepsilon) = |\{v \in V : v \neq \alpha_1, \ldots \alpha_r \text{ и } a(v, \alpha_j) = \varepsilon_j \text{ для всех } 1 \leqslant j \leqslant r\}|.$$

Ясно, что N_{r+1} — это сумма величин $f_r(\alpha, \varepsilon)$ в количестве N_r, где $\varepsilon_j = a(v_{r+1}, v_j)$, и α пробегает все N_r индуцированных копий $H(r)$ в G.

Заметим, что всего существует в точности $n_{(r)}2^r$ величин $f_r(\alpha, \varepsilon)$. Удобно рассматривать $f_r(\alpha, \varepsilon)$ как случайную величину, определенную на том же пространстве из $n_{(r)}2^r$ точек, у каждой из которых равная вероятность. Чтобы завершить доказательство, вычислим математическое ожидание и дисперсию этой случайной величины. Покажем, что дисперсия настолько мала, что большинство величин $f_r(\alpha, \varepsilon)$ очень близки к математическому ожиданию, и получим, таким образом, достаточно точную оценку для величины N_{r+1}, равной сумме N_r таких величин.

Начнем с простого вычисления математического ожидания $\mathbf{E}[f_r]$ величины $f_r(\alpha, \varepsilon)$. Имеем:

$$\mathbf{E}[f_r] = \frac{1}{n_{(r)}2^r} \sum_{\alpha,\varepsilon} f_r(\alpha, \varepsilon) = \frac{1}{n_{(r)}2^r} \sum_{\alpha} \sum_{\varepsilon} f_r(\alpha, \varepsilon) =$$
$$= \frac{1}{n_{(r)}2^r} \sum_{\alpha} (n - r) = \frac{n - r}{2^r}.$$

Здесь мы использовали то, что вершина $v \neq \alpha_1, \dots, \alpha_r$ однозначно определяет ε.

Далее, оценим величину S_r, определенную равенством

$$S_r = \sum_{\alpha,\varepsilon} f_r(\alpha,\varepsilon)(f_r(\alpha,\varepsilon)-1).$$

Мы утверждаем, что

$$S_r = \sum_{u\neq v} s(u,v)_{(r)}. \tag{9.4}$$

Для доказательства этого утверждения заметим, что S_r можно рассматривать как количество упорядоченных троек $(\alpha,\varepsilon,(u,v))$, где $\alpha=(\alpha_1,\dots,\alpha_r)$ — упорядоченное множество из r различных вершин графа G, $\varepsilon=(\varepsilon_1,\dots,\varepsilon_r)$ — двоичный вектор длины r, и u,v — упорядоченная пара дополнительных вершин G, таких, что

$$a(u,\alpha_k)=a(v,\alpha_k)=\varepsilon_k \quad \text{для всех } 1\leqslant k\leqslant r.$$

Для любых фиксированных α и ε существует ровно $f_r(\alpha,\varepsilon)(f_r(\alpha,\varepsilon)-1)$ способов выбрать пару (u,v), а значит, S_r равно количеству таких троек.

Теперь вычислим это значение, сначала выбрав u и v. Когда u,v выбраны, все дополнительные вершины $\alpha_1,\dots,\alpha_r$ должны принадлежать множеству $\{w\in V : a(u,w)=a(v,w)\}$. Так как мощность этого множества равна $s(u,v)$, отсюда следует, что существует $s(u,v)_{(r)}$ способов выбрать $\alpha_1,\dots\alpha_r$. Когда они выбраны, вектор ε определяется однозначно, и отсюда следует формула (9.4).

Далее мы утверждаем, что из соотношения (9.2) вытекает равенство

$$\sum_{u\neq v} s(u,v)_{(r)} = (1+o(1))n^{r+2}2^{-r}. \tag{9.5}$$

Для доказательства этого утверждения введем величины $\varepsilon_{uv}=s(u,v)-\frac{n}{2}$. Заметим, что по формуле (9.2) выполнено равенство $\sum_{u\neq v}|\varepsilon_{uv}|=o(n^3)$, и что $|\varepsilon_{uv}|\leqslant n/2\leqslant n$ для любых u,v. Значит, для каждого фиксированного $a\geqslant 1$

$$\sum_{u\neq v}|\varepsilon_{uv}|^a \leqslant n^{a-1}\sum_{u\neq v}|\varepsilon_{uv}| = o(n^{a+2}).$$

Тогда

$$\sum_{u\neq v} s(u,v)_{(r)} = \sum_{u\neq v}\left(\frac{n}{2}+\varepsilon_{uv}\right)_{(r)} =$$

$$= \sum_{k=0}^{r}\sum_{u\neq v} c_k\left(\frac{n}{2}\right)^k \varepsilon_{uv}^{r-k} = \qquad \text{(для подходящих констант } c_k\text{)}$$

$$= \left(\frac{n}{2}\right)^r n_{(2)} + \sum_{k=0}^{r-1}\sum_{u\neq v} c_k\left(\frac{n}{2}\right)^k \varepsilon_{uv}^{r-k} \leqslant \left(\frac{n}{2}\right)^r n_{(2)} + \sum_{k=0}^{r-1}\sum_{u\neq v}|c_k|n^k|\varepsilon_{uv}|^{r-k} \leqslant$$

$$\leqslant n^{r+2}2^{-r} + c\sum_{k=0}^{r-1} n^k \sum_{u\neq v}|\varepsilon_{uv}|^{r-k} \leqslant \qquad \text{(для подходящей константы } c\text{)}$$

$$\leqslant n^{r+2}2^{-r} + c\sum_{k=0}^{r-1} n^k\cdot o(n^{r-k+2}) = n^{r+2}2^{-r}(1+o(1)).$$

Отсюда следует соотношение (9.5).

По формулам (9.4) и (9.5)

$$S_r = (1 + o(1))n^{r+2}2^{-r}.$$

Следовательно,

$$\begin{aligned}\sum_{\alpha,\varepsilon}(f_r(\alpha,\varepsilon) - \mathbf{E}[f_r])^2 &= \sum_{\alpha,\varepsilon} f_r^2(\alpha,\varepsilon) - \sum_{\alpha,\varepsilon}\mathbf{E}[f_r]^2 = \\ &= \sum_{\alpha,\varepsilon}(f_r^2(\alpha,\varepsilon) - f_r(\alpha,\varepsilon)) + \sum_{\alpha,\varepsilon} f_r(\alpha,\varepsilon) - n_{(r)}2^r(n-r)^2 2^{-2r} = \\ &= S_r + n_{(r)}2^r\mathbf{E}[f_r] - n_{(r)}2^r(n-r)^2 2^{-2r} = \\ &= S_r + n_{(r+1)} - n_{(r)}2^r(n-r)^2 2^{-2r} = o(n^{r+2}).\end{aligned}$$

Напомним, что N_{r+1} — сумма величин вида $f_r(\alpha,\varepsilon)$ в количестве N_r. Тогда

$$|N_{r+1} - N_r\mathbf{E}[f_r]|^2 = \left|\sum_{N_r \text{ членов}} (f_r(\alpha,\varepsilon) - \mathbf{E}[f_r])\right|^2.$$

По неравенству Коши—Буняковского последнее выражение не больше

$$\begin{aligned}N_r \sum_{\text{члены } N_r} (f_r(\alpha,\varepsilon) - \mathbf{E}[f_r])^2 &\leqslant N_r \sum_{\alpha,\varepsilon}(f_r(\alpha,\varepsilon) - \mathbf{E}[f_r])^2 = \\ &= N_r \cdot o(n^{r+2}) = o(n^{2r+2}).\end{aligned}$$

Следовательно

$$\left|N_{r+1} - N_r\mathbf{E}[f_r]\right| = o(n^{r+1}),$$

а значит, по предположению индукции

$$\begin{aligned}N_{r+1} &= N_r\mathbf{E}[f_r] + o(n^{r+1}) = \\ &= (1 + o(1))n_{(r)}2^{-\binom{r}{2}} \cdot (n-r)2^{-r} + o(n^{r+1}) = \\ &= (1 + o(1))n_{(r+1)}2^{-\binom{r+1}{2}}.\end{aligned}$$

Это завершает доказательство шага индукции, а значит, и теоремы 9.3.1. ■

Есть много примеров семейств квазислучайных графов. Возможно, наиболее широко используется семейство G_p графов Пэли, определенное следующим образом. Для простого числа p, сравнимого с 1 по модулю 4, обозначим через G_p граф, вершины которого суть целые числа $0, 1, 2, \ldots, p-1$, и в котором i и j соединены ребром тогда и только тогда, когда $(i-j)$ — квадратичный вычет по модулю p. Графы G_p, которые являются неориентированными аналогами турниров квадратичных вычетов, рассмотренных в разд. 9.1, $(p-1)/2$-регулярны. Для любых двух различных вершин i и j графа G_p количество вершин k, которые либо смежны и с i, и j, либо не смежны с ними, в точности равно числу раз, когда частное $\frac{k-i}{k-j}$ является квадратичным вычетом по модулю p. По мере того как k принимает все значения от 0 до $p-1$, кроме i и j, это частное принимает все значения, кроме 1 и 0, а значит, оно является квадратичным вычетом ровно $\frac{1}{2}(p-1) - 1$ раз. Таким образом, мы показали, что для любых

двух вершин i и j графа G_p выполнено равенство $s(i,j) = (p-3)/2$, а отсюда и из того факта, что G_p — $(p-1)/2$-регулярный, легко следует, что для него справедливо свойство P_6. Следовательно, он квазислучайный. Как и в случае с турнирами квадратичных вычетов, в действительности для G_p выполнены более сильные псевдослучайные свойства, которые справедливы не для любого квазислучайного графа, и которые могут быть доказаны с помощью теоремы Вейля.

9.4. УПРАЖНЕНИЯ

1. Пусть случайный двудольный 3-регулярный граф на $2n$ вершинах получен выбором трех случайных перестановок между двумя хроматическими классами. Доказать, что существует такое $c > 0$, что для любого n существует $(2n, 3, c)$-расширитель.

2. Пусть $G = (V, E)$ — (n, d, λ)-граф. Предположим, что n делится на k. Пусть $C : V \mapsto \{1, 2, \ldots, k\}$ — такая раскраска множества V в k цветов, что каждый цвет появляется ровно n/k раз. Доказать, что в графе G существует такая вершина, что у нее есть соседи всех k цветов, если $k\lambda \leqslant d$.

3. Пусть $G = (V, E)$ — граф, в котором есть как минимум одно ребро между любыми двумя непересекающимися множествами размера $a+1$. Доказать, что для каждого множества Y из $5a$ вершин существует такое множество X не более чем из a вершин, что для любого множества Z, удовлетворяющего условиям $Z \cap (X \cup Y) = \varnothing$ и $|Z| \leqslant a$, выполнено неравенство $|N(Z) \cap Y| \geqslant 2|Z|$.

4. Доказать, что для любого $\varepsilon > 0$ существует такое $n_0 = n_0(\varepsilon)$, что для любого $(n, n/2, 2\sqrt{n})$-графа $G = (V, E)$ с $n > n_0$ число M треугольников в графе G удовлетворяет неравенству $|M - n^3/48| \leqslant \varepsilon n^3$.

ВЕРОЯТНОСТНЫЙ ВЗГЛЯД:

Случайные блуждания

Вершинно-транзитивным графом называется такой граф $G = (V, E)$, что для любых двух вершин u и v из множества V существует автоморфизм графа G, отображающий u в v. *Случайное блуждание* длины l в графе G, начинающееся в вершине v, — это случайная последовательность $v = v_0, v_1, \ldots, v_l$, где каждая вершина v_{i+1} выбирается случайно и независимо среди соседей вершины v_i $(0 \leqslant i < l)$.

Следующая теорема утверждает, что для каждого вершинно-транзитивного графа G вероятность того, что случайное блуждание четной длины в G заканчивается в его начальной точке, не меньше вероятности того, что оно заканчивается в любой другой вершине. Заметим, что доказательство почти не требует вычислений. Заметим также, что это, вообще говоря, не верно для регулярных графов, и требование вершинной транзитивности необходимо.

Теорема. *Пусть $G = (V, E)$ — вершинно-транзитивный граф. Для целого k и двух (не обязательно различных) вершин u и v из G обозначим через $P^k(u, v)$ вероятность того, что случайное блуждание длины k, начинающееся в u, заканчивается в v. Тогда для каждого целого k и любых двух вершин $u, v \in V$ справедливо неравенство*

$$P^{2k}(u, u) \geqslant P^{2k}(u, v).$$

Доказательство. Нам потребуется следующее простое неравенство, иногда приписываемое Чебышёву.

Утверждение. *Для любой последовательности $(a_1, \ldots, a_n)$, состоящей из n действительных чисел, и для любой перестановки π на множестве $\{1, \ldots, n\}$ выполнено неравенство*

$$\sum_{i=1}^{n} a_i a_{\pi(i)} \leqslant \sum_{i=1}^{n} a_i^2.$$

Доказательство. Неравенство напрямую следует из того, что

$$\sum_{i=1}^{n} a_i^2 - \sum_{i=1}^{n} a_i a_{\pi(i)} = \frac{1}{2} \sum_{i=1}^{n} (a_i - a_{\pi(i)})^2 \geqslant 0. \qquad \blacksquare$$

Теперь рассмотрим случайное блуждание длины $2k$, начинающееся в u. Суммируя по всем возможным вершинам, которые блуждание достигает после k шагов, мы делаем вывод, что для любой вершины v выполнены соотношения

$$P^{2k}(u, v) = \sum_{w \in V} P^k(u, w) P^k(w, v) = \sum_{w \in V} P^k(u, w) P^k(v, w), \qquad (1)$$

где последнее равенство следует из того факта, что G — неориентированный регулярный граф.

Так как граф G — вершинно-транзитивный, два вектора $(P^k(u,w))_{w\in V}$ и $(P^k(v,w))_{w\in V}$ могут быть получены один из другого перестановкой координат. Следовательно, по предыдущему утверждению, максимально возможное значение суммы в правой части соотношения (1) достигается при $u = v$, чем завершается доказательство теоремы. ■

Часть II

Приложения

10

Случайные графы

Шесть часов утра. Дом спит. Играет прекрасная музыка. Я доказываю и делаю предположения.

из письма Пола Эрдёша Вере Шош

Пусть n является натуральным числом, $0 \leqslant p \leqslant 1$. Рассмотрим множество графов с множеством вершин $\{1, \ldots, n\}$. Вероятностное пространство $G(n, p)$, определенное в соответствии с распределением

$$\Pr[\{i, j\} \in G] = p,$$

где все события попарно независимы, называется *случайным графом*. Данная модель часто применяется в вероятностном методе для доказательства существования некоторых графов. В этой главе мы изучаем свойства графов как самостоятельные объекты.

Случайные графы представляют собой область активных исследований, которая находится на стыке теории вероятностей и теории графов. Предмет изучения возник в 1960 г. в основополагающей статье *Эволюция случайных графов* Пола Эрдёша и Альфреда Реньи. Книга Боллобаша *Случайные графы* [Bollobás (1985)] является стандартным пособием по этой тематике. Новая книга [Janson, Luczak and Rucinski (2000)] Сванте Янсона, Томаша Лучака и Анжея Ручиньски, тоже названная *Случайные графы*, также великолепна. В этой главе мы изучаем только некоторые из большого числа тем в этой привлекательной области.

Существует интересная динамическая модель для случайных графов. Для всех пар i, j пусть $x_{i,j}$ выбираются независимо и равномерно из отрезка $[0, 1]$. Представьте, что p возрастает от 0 до 1. В начале, все потенциальные ребра «выключены». Ребро от i до j (мы можем представлять это ребро в виде света в неоновой трубке) «включается», когда p достигает величины $x_{i,j}$, и остается включенным. При $p = 1$ все ребра «включены». В момент времени p граф со всеми «включенными» ребрами имеет распределение $G(n, p)$. С ростом p граф $G(n, p)$ расширяется от пустого до полного.

В своей оригинальной работе Эрдёш и Реньи рассматривают $G(n, e)$, случайный граф с n вершинами и ровно e ребрами. И снова есть соответствующая

динамическая модель. Начните с пустого графа и продолжайте добавлять ребра случайным образом, пока граф не станет полным. Как правило, при $p \sim \frac{e}{\binom{n}{2}}$ графы $G(n,e)$ и $G(n,p)$ имеют очень схожие свойства. Мы будем рассматривать только вероятностную модель.

10.1. ПОДГРАФЫ

Термин «случайный граф», строго говоря, некорректен: $G(n,p)$ является вероятностным пространством над графами. Пусть задано некоторое теоретико-графовое свойство A. Обозначим через $\Pr[G(n,p) \models A]$ вероятность того, что $G(n,p)$ обладает свойством A. Если A является монотонным, то и вероятность $\Pr[G(n,p) \models A]$ является монотонной функцией от p. Поучителен следующий пример. Обозначим A через событие «G свободен от треугольников». Пусть X — число треугольников в $G(n,p)$. Из линейности математического ожидания получаем, что

$$\mathbf{E}[X] = \binom{n}{3} p^3.$$

Из этого следует, что удобно ввести обозначение $p = c/n$. Тогда

$$\lim_{n\to\infty} \mathbf{E}[X] = \lim_{n\to\infty} \binom{n}{3} p^3 = c^3/6.$$

Получается, что распределение величины X асимптотически является пуассоновским. В частности,

$$\lim_{n\to\infty} \Pr[G(n,p) \models A] = \lim_{n\to\infty} \Pr[X = 0] = e^{-c^3/6}.$$

Заметим, что

$$\lim_{c\to 0} e^{-c^3/6} = 1, \qquad \lim_{c\to\infty} e^{-c^3/6} = 0.$$

Когда $p = 10^{-6}/n$, очень маловероятно, что граф $G(n,p)$ содержит треугольники. А когда $p = 10^6/n$, $G(n,p)$ содержит треугольники с большой вероятностью. С динамической точки зрения, треугольники начинают появляться, когда $p = \Theta(1/n)$. Если мы возьмем функцию $p(n) = n^{-0.9}$ (при этом $p(n) \gg n^{-1}$), то граф $G(n,p)$ почти наверное будет содержать треугольники. Иногда мы будем позволять себе не совсем корректные обозначения. Например, говоря, что $G(n, n^{-0.9})$ содержит треугольник, будем иметь в виду, что вероятность того, что этот граф содержит треугольник, стремится к 1 при $n \to \infty$. Аналогично, когда $p(n) \ll n^{-1}$, например, $p(n) = 1/(n \ln n)$, и $G(n,p)$ почти наверное не содержит треугольника, мы будем позволять себе вольность речи и говорить, что $G(n, 1/(n \ln n))$ является свободным от треугольников. Основным наблюдением Эрдёша и Реньи было то, что многие естественные теоретико-графовые свойства становятся верными в очень узком промежутке значений p. Они ввели следующее ключевое определение.

Определение 4. *Назовем $r(n)$ **пороговой функцией** теоретико-графового свойства A, если выполнены следующие два свойства:*

1) *если $p(n) \ll r(n)$, то $\lim_{n\to\infty} \Pr[G(n,p) \models A] = 0$.*

2) *если $p(n) \gg r(n)$, то $\lim_{n\to\infty} \Pr[G(n,p) \models A] = 1$.*

(Вероятности 0 и 1 можно поменять местами.)

В нашем примере функция $1/n$ является пороговой функцией для A. Заметим, что пороговая функция, когда она существует, не единственна. Мы могли бы сказать, что и $10/n$ является пороговой функцией для A.

Рассмотрим еще раз задачу о свойстве графа $G(n, c/n)$ быть свободным от треугольников. Для любого множества S, состоящего из трех вершин, обозначим через B_S событие «S является треугольником». Тогда $\Pr[B_S] = p^3$. Откуда следует, что «отсутствие треугольников» является в точности пересечением $\bigwedge \overline{B}_S$ по всем S. Если бы события B_S были взаимно независимы, то выполнялись бы равенства

$$\Pr\left[\bigwedge \overline{B}_S\right] = \prod [\overline{B}_S] = (1-p^3)^{\binom{n}{3}} \sim e^{-\binom{n}{3}p^3} \to e^{-c^3/6}.$$

Но на самом деле события B_S не являются взаимно независимыми. В то же время, если $|S \cap T| \leqslant 1$, то B_S и B_T взаимно независимы.

Применим неравенство Янсона (теорема 8.1.1). В обозначениях разд. 8.1 имеем $I = \{S \subset V(G) : |S| = 3\}$ и $S \sim T$ тогда и только тогда, когда $|S \cap T| = 2$. Здесь $\varepsilon = p^3 = o(1)$, $\mu = \binom{n}{3}p^3 \sim c^3/6$ и $M = e^{-\mu(1+o(1))} = e^{-c^3/6+o(1)}$. Существует $6\binom{n}{4} = O(n^4)$ пар S, T троек с условием, что $S \sim T$. Для каждой из них справедливо равенство $\Pr[B_S \wedge B_T] = p^5$. Таким образом,

$$\Delta = O(n^4)p^5 = n^{-1+o(1)} = o(1).$$

Когда $\Delta = o(1)$, неравенство Янсона дает следующую асимптотическую оценку:

$$\lim_{n\to\infty} \Pr[\wedge \overline{B}_S] = \lim_{n\to\infty} M = e^{-c^3/6}.$$

Можно ли повторить этот успех, если взять в качестве A свойство «несодержания графом G (не обязательно индуцированной) копии некоторого графа H»? Мы используем определения сбалансированного и строго сбалансированного графов из разд. 4.4.

Теорема 10.1.1. *Пусть H является строго сбалансированным графом с v вершинами, e ребрами и a автоморфизмами. Обозначим через A свойство «G не содержит копию H». Пусть $c > 0$ — произвольно выбранное число. Тогда при $p = cn^{-v/e}$ справедливо соотношение*

$$\lim_{n\to\infty} \Pr[G(n,p) \models A] = \exp[-c^e/a].$$

Доказательство. Пусть $A_\alpha, 1 \leqslant \alpha \leqslant \binom{n}{v}v!/a$, пробегает множества ребер возможных копий H, и пусть B_α является событием «$G(n,p) \supseteq A_\alpha$». Воспользуемся неравенством Янсона. Поскольку

$$\lim_{n\to\infty} \mu = \lim_{n\to\infty} \binom{n}{v} v!p^e/a = c^e/a,$$

то

$$\lim_{n\to\infty} M = \exp[-c^e/a].$$

Теперь рассмотрим (как и в теореме 4.4.2) величину

$$\Delta = \sum_{\alpha\sim\beta} \Pr\left[B_\alpha \wedge B_\beta\right].$$

Мы разобьем сумму по числу вершин в пересечении копий α и β. Предположим, что они пересекаются в j вершинах. Если $j = 0$ или $j = 1$, то $A_\alpha \cap A_\beta = \varnothing$. Таким образом, соотношение $\alpha \sim \beta$ не верно. При $2 \leqslant j \leqslant v$ обозначим через f_j максимальную мощность $|A_\alpha \cap A_\beta|$, где $\alpha \sim \beta$ и α и β пересекаются в j вершинах. Поскольку $\alpha \neq \beta$, то справедливо неравенство $f_v < e$. Когда $2 \leqslant j \leqslant v-1$, ключевым наблюдением является то, что $A_\alpha \cap A_\beta$ является подграфом графа H и, поскольку H строго сбалансирован, то

$$\frac{f_j}{j} < \frac{e}{v}.$$

Существует $O(n^{2v-j})$ способов выбора α и β, пересекающихся в j точках, так как α и β определены (с точностью до порядка) $2v-j$ точками. Для каждой такой пары α, β

$$\Pr[B_\alpha \wedge B_\beta] = p^{|A_\alpha \cup A_\beta|} = p^{2e-|A_\alpha\cap A_\beta|} \leqslant p^{2e-f_j}.$$

Таким образом,

$$\Delta = \sum_{j=2}^{v} O(n^{2v-j}) O(n^{-\frac{v}{e}(2e-f_j)}).$$

Но

$$2v - j - \frac{v}{e}(2e - f_j) = \frac{vf_j}{e} - j < 0,$$

поэтому каждое слагаемое есть $o(1)$ и, следовательно, $\Delta = o(1)$. Используя неравенство Янсона, получаем

$$\lim_{n\to\infty} \Pr[\wedge \overline{B}_\alpha] = \lim_{n\to\infty} M = \exp[-c^e/a],$$

что завершает доказательство. ■

10.2. РАЗМЕР МАКСИМАЛЬНОЙ КЛИКИ

В этом разделе мы фиксируем $p = 1/2$ (при других значения получаются аналогичные результаты) и рассматриваем размер максимальной клики $\omega(G(n,p))$. Зафиксируем некоторое $c > 0$, и пусть $n, k \to \infty$, так чтобы

$$\binom{n}{k} 2^{-\binom{k}{2}} \to c.$$

Первым приближением является

$$n \sim \frac{k}{e\sqrt{2}} \sqrt{2}^k$$

и

$$k \sim \frac{2\ln n}{\ln 2}.$$

Здесь $\mu \to c$, поэтому $M \to e^{-c}$. Величина Δ была оценена в разд. 4.5. При таком выборе k выполнено равенство $\Delta = o(\mathbf{E}[X]^2)$, и поэтому $\Delta = o(1)$. Отсюда вытекает, что

$$\lim_{n,k\to\infty} \Pr[\omega(G(n,p)) < k] = \exp[-c].$$

Пусть $n_0(k)$ обозначает минимальное значение n, при котором

$$\binom{n}{k} 2^{-\binom{k}{2}} \geqslant 1.$$

При более точном анализе можно заметить, что при таком значении n левая часть равна $1+o(1)$. Учтем, что $\binom{n}{k}$ растет относительно n, как n^k. При любом $\lambda \in (-\infty, +\infty)$, если

$$n = n_0(k)\left[1 + \frac{\lambda + o(1)}{k}\right],$$

то

$$\binom{n}{k} 2^{-\binom{k}{2}} = \left[1 + \frac{\lambda + o(1)}{k}\right]^k = e^{\lambda} + o(1).$$

Отсюда

$$\Pr[\omega(G(n,p)) < k] = e^{-e^{\lambda}} + o(1).$$

Когда λ возрастает от $-\infty$ к $+\infty$, величина $e^{-e^{\lambda}}$ убывает от 1 до 0. При $n_0(k+1) \sim \sqrt{2}n_0(k)$ промежутки не будут пересекаться при различных k. Пусть K произвольно велико. Положим

$$I_k = \left[n_0(k)\left[1 - \frac{K}{k}\right], n_0(k)\left[1 + \frac{K}{k}\right]\right].$$

Тогда $I_{k-1} \cap I_k = \varnothing$ при $k \geqslant k_0(K)$. Предположим, что $n \geqslant n_0(k_0(K))$. Если n лежит между интервалами (это верно для «большинства» n), что мы будем обозначать как $I_k < n < I_{k+1}$, то вероятность

$$\Pr[\omega(G(n,p)) < k] \leqslant e^{-e^{K}} + o(1)$$

почти равна нулю, а вероятность

$$\Pr[\omega(G(n,p)) < k+1] \geqslant e^{-e^{-K}} + o(1)$$

почти равна единице. Отсюда следует, что вероятность

$$\Pr[\omega(G(n,p)) = k] \geqslant e^{-e^{-K}} - e^{-e^{K}} + o(1)$$

почти равна единице. Когда $n \in I_k$, то все еще выполняется $I_{k-1} < n < I_{k+1}$. Таким образом, вероятность

$$\Pr[\omega(G(n,p)) = k \text{ или } k-1] \geqslant e^{-e^{-K}} - e^{-e^{K}} + o(1)$$

почти равна единице. Поскольку K можно выбрать произвольно большим, из этого следует знаменитая теорема о двухточечной концентрации размера максимальной клики (см. следствие 4.5.2 в разд. 4.5). Заметим, однако, что для большинства n концентрация $\omega(G(n,1/2))$ приходится на самом деле на одно значение!

10.3. ХРОМАТИЧЕСКОЕ ЧИСЛО

В этом разделе снова зафиксируем $p = 1/2$ (имеются аналогичные результаты для других значений p), и пусть G является случайным графом $G(n, 1/2)$. Мы оценим хроматическое число $\chi(G)$. Другой подход к получению основного результата этого раздела представлен в гл. 7, разд. 7.3. Введем функцию

$$f(k) = \binom{n}{k} 2^{-\binom{k}{2}}.$$

Пусть при $k_0 = k_0(n)$ выполнено двойное неравенство

$$f(k_0 - 1) > 1 > f(k_0).$$

Тогда $n = \sqrt{2}^{k(1+o(1))}$. Поэтому при $k \sim k_0$ имеем

$$f(k+1)/f(k) = \frac{n}{k} 2^{-k}(1 + o(1)) = n^{-1+o(1)}.$$

Пусть

$$k = k(n) = k_0(n) - 4$$

таково, что

$$f(k) > n^{3+o(1)}.$$

Для оценки вероятности $\Pr[\omega(G) < k]$ воспользуемся обобщенным неравенством Янсона (теорема 8.1.2). Здесь $\mu = f(k)$. (Заметим, что неравенство Янсона дает нижнюю оценку $2^{-f(k)} = 2^{-n^{3+o(1)}}$ для этой вероятности, но это неверно, поскольку с вероятностью $2^{-\binom{n}{2}}$ случайный граф G является пустым!) Значение Δ было оценено в разд. 4.5, где

$$\frac{\Delta}{\mu^2} = \frac{\Delta^*}{\mu} = \sum_{i=2}^{k-1} g(i).$$

Доминирующими слагаемыми являются $g(2) \sim k^4/n^2$ и $g(k-1) \sim 2kn2^{-k}/\mu$. В данном случае $\mu > n^{3+o(1)}$ и $2^{-k} = n^{-2+o(1)}$, поэтому $g(2)$ является доминирующим, и

$$\Delta \sim \frac{\mu^2 k^4}{n^2}.$$

Отсюда вытекает оценка размера максимальной клики

$$\Pr[\omega(G) < k] < e^{-\mu^2/2\Delta} = e^{-\Theta(n^2/(\ln n)^4)}$$

при $k = \Theta(\ln n)$. Вероятность того, что G пусто, дает нижнюю оценку. Поэтому можно сказать, что вероятность равна $e^{-n^{2+o(1)}}$, хотя $o(1)$ в показателе оставляет простор для размышлений.

Теорема 10.3.1 [Bollobás (1988)]. *Почти всегда*

$$\chi(G) \sim \frac{n}{2\log_2 n}.$$

Доказательство. Пусть $\alpha(G) = \omega(\overline{G})$ обозначает, как обычно, число независимости графа G[1]. Дополнение G имеет то же распределение $G(n, 1/2)$. Следовательно, почти всегда $\alpha(G) \leqslant (2 + o(1)) \log_2 n$. Поэтому почти всегда

$$\chi(G) \geqslant \frac{n}{\alpha(G)} \geqslant \frac{n}{2\log_2 n}(1 + o(1)).$$

Обратное неравенство было открытой проблемой целую четверть века! Введем обозначение $m = \lfloor n/\ln^2 n \rfloor$. Для любого множества S с m вершинами ограничение $G|_S$ имеет распределение $G(m, 1/2)$. Пусть, как и выше, $k = k(m) = k_0(m) - 4$. Заметим, что

$$k \sim 2\log_2 m \sim 2\log_2 n.$$

Тогда

$$\Pr\left[\alpha(G|_S) < k\right] < e^{-m^{2+o(1)}}.$$

Существует $\binom{n}{m} < 2^n = 2^{m^{1+o(1)}}$ таких множеств S. Следовательно,

$$\Pr\left[\alpha(G|_S) < k \text{ для некоторого } m\text{-множества } S\right] < 2^{m^{1+o(1)}} e^{-m^{2+o(1)}} = o(1).$$

То есть почти всегда каждые m вершин содержат k-элементное независимое множество.

Теперь предположим, что граф G обладает этим свойством. Мы берем k-элементные независимые множества и раскрашиваем их различными цветами пока не останется меньше m вершин. Тогда мы раскрашиваем оставшиеся вершины в различные цвета. Действуя таким образом, получаем

$$\chi(G) \leqslant \left\lceil \frac{n-m}{k} \right\rceil + m \leqslant \frac{n}{k} + m =$$
$$= \frac{n}{2\log_2 n}(1 + o(1)) + o\left(\frac{n}{\log_2 n}\right) = \frac{n}{2\log_2 n}(1 + o(1)),$$

что выполнено для почти всех G. ■

10.4. ВЕТВЯЩИЕСЯ ПРОЦЕССЫ

Пол Эрдёш и Альфред Реньи в своей основополагающей работе [Erdős and Rényi (1960)] обнаружили, что случайный граф $G(n, p)$ подвергается удивительному изменению при $p = 1/n$. Рассмотрим сначала случай $p = c/n$ при $c < 1$. Тогда $G(n, p)$ будет состоять из маленьких компонент, самая большая из которых будет иметь размер $\Theta(\ln n)$. Теперь предположим, что $p = c/n$, где $c > 1$. В этом небольшом промежутке «времени» многие компоненты объединятся, образовав таким образом «гигантскую компоненту» размера $\Theta(n)$. Остальные вершины все еще будут содержаться в небольших компонентах, самая большая из которых имеет размер $\Theta(\ln n)$. Этот феномен называют *большим прыжком*. Мы предпочитаем описательный термин *фазовый переход*

[1] Число независимости графа G — это максимальное число элементов в независимом множестве графа G. — *Прим. ред.*

из-за его связи с перколяцией в математической физике (например, превращением воды в лед).

Чтобы лучше понять фазовый переход, мы проделаем некоторый экскурс в теорию ветвящихся процессов. Рассмотрим однополую модель размножения. Вначале имеется некоторое существо, которое дает определенное число потомков, задаваемое некоторой случайной величиной Z. (В качестве Z будет взята пуассоновская случайная величина со средним c.) Затем эти потомки производят внуков, число которых вновь определяется случайной величиной Z. Эти внуки тоже становятся родителями, и так далее. Поскольку событие $Z = 0$ имеет ненулевую вероятность, то возможно, что весь род исчезнет полностью. Мы хотим изучить общее число организмов в этом процессе, обращая особое внимание на вопрос конечности данного процесса. (Первоначальным применением этой модели было, как это ни удивительно, изучение родословной мужской линии британского сословия пэров.)

Введем точные определения. Пусть случайные величины $Z_1, Z_2, \ldots$ независимы, и распределение каждой из них совпадает с распределением случайной величины Z. Определим $Y_0, Y_1, \ldots$ рекурсивно:

$$Y_0 = 1,$$
$$Y_i = Y_{i-1} + Z_i - 1,$$

и пусть T является наименьшим t, для которого $Y_t = 0$. Если такого t не существует (род продолжается бесконечно), будем говорить, что $T = +\infty$. Величины Y_i и Z_i отражают ветвящийся процесс следующим образом. Мы рассматриваем все организмы, как живые, так и мертвые (в том числе *еще* не живые). Изначально существует только один живой организм, а все остальные мертвы. В каждый момент времени мы выбираем какой-то из живых организмов, он имеет Z_i детей, а потом умирает. Тогда число Y_i живых организмов в момент времени i задается рекурсией. Процесс обрывается, когда $Y_t = 0$ (вырождение), но удобно определить рекурсию при всех t. Заметим, что данное определение не влияет на T, поскольку как только $Y_t = 0$, T становиться определенным. Значение случайной величины T (либо конечное, либо бесконечное) обозначает общее число организмов, включая предков, в этом процессе. (Во многих учебниках по теории вероятностей рассматривается естественный подход, заключающийся в предположении, что у всех живых организмов некоторого поколения дети рождаются одновременно, и изучается число детей каждого поколения. Мы можем считать, что организмы дают потомство поколениями, хоть это и не повлияет на нашу модель.)

Основным результатом ветвящихся процессов является следующий. Когда $\mathbf{E}[Z] = c < 1$, с вероятностью 1 род вырождается ($T < \infty$), но когда $\mathbf{E}[Z] = c > 1$, существует ненулевая вероятность того, что процесс будет продолжаться бесконечно ($T = \infty$). Интуитивно понятно, что это верно. Значения Y_i образуют цепь Маркова. Когда $c < 1$, вследствие сдвига «влево» произойдет событие $Y_i = 0$. Тогда как при $c > 1$, вследствие сдвига «вправо» $Y_i \to +\infty$, и с положительной вероятностью численность популяции никогда не достигнет нуля. Если процесс не закончится на ранней стадии, то популяция,

скорее всего, будет расти и расти. Приведем доказательство, основанное на результатах теоремы А.1.15 относительно больших уклонений. Заметим, что сумма $Z_1 + \ldots + Z_t$ имеет пуассоновское распределение со средним ct. Сначала предположим, что $c < 1$. Для любого t

$$\Pr[T > t] \leqslant \Pr[Y_t > 0] = \Pr[Z_1 + \ldots + Z_t \geqslant t] \leqslant (1-\delta)^t,$$

где $\delta > 0$. Вероятность $\Pr[T = \infty] = 0$, поскольку $\lim_{t\to\infty} \Pr[T > t] = 0$. Теперь предположим, что $c > 1$. Снова используя теорему А.1.15, получаем

$$\Pr[Y_t \leqslant 0] = \Pr[Z_1 + \ldots + Z_t \leqslant t] \leqslant (1-\delta)^t$$

с другим $\delta > 0$. Поскольку ряд $\sum_{t=1}^{\infty}(1-\delta)^t$ сходится, существует такое t_0, что

$$\sum_{t=t_0}^{\infty} \Pr[Y_t \leqslant 0] < 1.$$

Тогда

$$\sum_{t=0}^{\infty} \Pr[Y_t + t_0 - 1 \leqslant 0] < 1,$$

поскольку при $t < t_0$ справедливы неравенства $Y_t + t_0 - 1 \geqslant 1 - t + t_0 - 1 > 0$. Теперь мы наложим условие, что первый организм имел ровно t_0 детей. Условное распределение случайной величины

$$Y_t = t_0 + (Z_2 - 1) + \ldots + (Z_t - 1)$$

совпадает с безусловным распределением величины

$$Y_{t-1} + (t_0 - 1) = t_0 + (Z_1 - 1) + \ldots + (Z_{t-1} - 1),$$

и поэтому

$$\sum_{t=0}^{\infty} \Pr[Y_t \leqslant 0 | Z_1 = t_0] < 1.$$

С положительной вероятностью выполнено равенство $Z_1 = t_0$, и поэтому для всех t имеем $Y_t > 0$. Отсюда вытекает, что $\Pr[T = \infty] > 0$.

Метод производящих функций дает более точный результат. Пусть

$$p_i = \Pr[Z_j = i].$$

Определим производящую функцию

$$p(x) = \sum_{i=0}^{\infty} p_i x^i.$$

В нашем случае

$$p_i = e^{-c} c^i / i!,$$

так что

$$p(x) = \sum_{i=0}^{\infty} e^{-c} c^i x^i / i! = e^{c(x-1)}.$$

Пусть

$$q_i = \Pr[T = i]$$

и

$$q(x) = \sum_{i=0}^{\infty} q_i x^i$$

(в сумму не входит случай $i = \infty$). При условии, что первый организм имеет s детей, производящая функция общего числа потомков равна $x(q(x))^s$. Поэтому

$$q(x) = \sum_{s=0}^{\infty} p_s x q(x)^s = xp[q(x)].$$

Значит, производящая функция $y = q(x)/x$ удовлетворяет функциональному уравнению $y = p(xy)$, т.е.

$$y = e^{c(xy-1)}.$$

Вероятность вырождения

$$y = \Pr[T < \infty] = \sum_{i=0}^{\infty} q_i = q(1) = q(1)/1$$

должна удовлетворять уравнению

$$y = e^{c(y-1)}. \qquad (**)$$

При $c < 1$ это уравнение имеет единственное решение $y = 1$, соответствующее обязательному вырождению. (На самом деле, для $c = 1$ единственным решением также является $y = 1$, что доказывает вырождение и в граничном случае.) При $c > 1$ существует два решения: $y = 1$ и некоторое $y \in (0, 1)$. При $c > 1$ через $f(c)$ обозначим y, удовлетворяющее уравнению $(**)$, $0 < y < 1$. Когда $\Pr[T < \infty] < 1$, имеем

$$\Pr[T < \infty] = f(c).$$

Когда ветвящийся процесс является конечным, мы называем $H = (Z_1, \ldots, Z_T)$ *историей* этого процесса. Последовательность $(z_1, \ldots, z_t)$ является возможной историей тогда и только тогда, когда для последовательности y_i, заданной равенствами $y_0 = 1, y_i = y_{i-1} + z_i - 1$, выполнены неравенства $y_i > 0$ при $0 \leqslant i < t$ и $y_t = 0$. Когда Z является пуассоновской случайной величиной со средним λ,

$$\Pr[H = (z_1, \ldots, z_t)] = \prod_{i=1}^{t} \frac{e^{-\lambda}\lambda^{z_i}}{z_i!} = \frac{e^{-\lambda}(\lambda e^{-\lambda})^{t-1}}{\prod_{i=1}^{t} z_i!},$$

поскольку $z_1 + \ldots + z_t = t - 1$.

Назовем числа d и c, $d < 1 < c$, сопряженными, если

$$de^{-d} = ce^{-c}.$$

Функция $f(x) = xe^{-x}$ возрастает от 0 до e^{-1} в промежутке $[0, 1)$ и убывает обратно к нулю на интервале $(1, \infty)$, поэтому каждое $c \neq 1$ имеет единственный сопряженный элемент. Пусть $c > 1$ и $y = \Pr[T < \infty]$ таковы, что $y = e^{c(y-1)}$.

Тогда $(cy)e^{-cy} = ce^{-c}$, поэтому

$$d = cy.$$

Принцип двойственности. Пусть d и c, $d < 1 < c$, сопряжены. Ветвящийся процесс со средним c, обусловленный обязательным вырождением, имеет то же распределение, что и ветвящийся процесс со средним d.

Доказательство. Достаточно показать, что для любой истории $H = (z_1, \ldots, z_t)$

$$\frac{e^{-c}(ce^{-c})^{t-1}}{y\prod_{i=1}^{t} z_i!} = \frac{e^{-d}(de^{-d})^{t-1}}{\prod_{i=1}^{t} z_i!}.$$

Это легко следует из того, что $ce^{-c} = de^{-d}$ и $ye^{-d} = ye^{-cy} = e^{-c}$. ■

10.5. ГИГАНТСКАЯ КОМПОНЕНТА

Вернемся к случайным графам. Мы определим процедуру нахождения компоненты $C(v)$, содержащей заданную вершину v в заданном графе G. Здесь мы используем идеи из работы [Karp (1990)], в которой данный подход применен к случайным орграфам. В этой процедуре вершины будут трех видов: живые, мертвые и нейтральные. Первоначально вершина v является живой, а все остальные — нейтральными, время $t = 0$ и $Y_0 = 1$. В каждую единицу времени t мы берем живую вершину w и проверяем, принадлежат ли G все пары $\{w, w'\}$, где w' нейтральна. Если $\{w, w'\} \in G$, то отмечаем w' как живую, в противном случае она остается нейтральной. После отыскания всех нейтральных w' мы отмечаем w как мертвую, а Y_t принимает значение, равное новому числу живых вершин. Когда не остается ни одной живой вершины, процедура заканчивается, а $C(v)$ является множеством мертвых вершин. Пусть Z_t — количество тех w', для которых $\{w, w'\} \in G$. Таким образом,

$$Y_0 = 1,$$
$$Y_t = Y_{t-1} + Z_t - 1.$$

При $G = G(n, p)$ каждая нейтральная вершина w' независимо и с вероятностью p становится живой. Здесь никакая пара $\{w, w'\}$ не проверяется дважды, поэтому условная вероятность для $\{w, w'\} \in G$ всегда равняется p. Когда $t-1$ вершин мертвы, а Y_{t-1} — живы,

$$Z_t \sim B[n - (t-1) - Y_{t-1}, p].$$

Пусть T обозначает наименьшее t, при котором $Y_t = 0$. Тогда $T = |C(v)|$. Как и в разд. 10.4 мы вводим рекурсивное определение Y_t, на этот раз при $0 \leqslant t \leqslant n$.

Утверждение 10.5.1. *Для всех t*

$$Y_t \sim B[n-1, 1-(1-p)^t] + 1 - t.$$

Доказательство. Более удобно будет использовать величину

$$N_t = n - t - Y_t,$$

отвечающую количеству нейтральных вершин в момент времени t, и доказывать эквивалентное утверждение:

$$N_t \sim B[n-1,(1-p)^t].$$

Это целесообразно, поскольку каждое $w \neq v$ имеет независимую вероятность $(1-p)^t$ остаться нейтральной t раз. Так как $N_0 = n-1$ и

$$\begin{aligned} N_t = n - t - Y_t = n - t - B[n-(t-1)-Y_{t-1},p] - Y_{t-1} + 1 = \\ = N_{t-1} - B[N_{t-1},p] = B[N_{t-1},1-p], \end{aligned}$$

индукция завершает доказательство. ■

Пусть $p = c/n$. Когда t и Y_{t-1} малы, распределение случайной величины Z_t можно аппроксимировать биномиальным распределением $B[n,c/n]$, которое близко к пуассоновскому со средним c. В основном, размеры маленьких компонент будут распределены так же, как и в ветвящихся процессах в разд. 10.4. Аналогия должна нарушиться при $c > 1$, поскольку ветвящиеся процессы могут иметь бесконечную численность популяции, когда $|C(v)|$ будет достигать n. В итоге все те v, для которых ветвящиеся процессы для $C(v)$ не «умирают рано», образуют в совокупности гигантскую компоненту.

Зафиксируем c. Пусть $Y_0^*, Y_1^*, \ldots, T^*, Z_1^*, Z_2^*, \ldots, H^*$ относятся к ветвящемуся процессу из разд. 10.4, а $Y_0, Y_1, \ldots, T, Z_1, Z_2, \ldots, H$ относятся к процессу, порожденному случайным графом. Для любой возможной предыстории $(z_1, \ldots, z_t)$

$$\Pr[H^* = (z_1,\ldots,z_t)] = \prod_{i=1}^{t} \Pr[Z^* = z_i],$$

где Z^* является пуассоновской случайной величиной со средним c, и

$$\Pr[H = (z_1,\ldots,z_t)] = \prod_{i=1}^{t} \Pr[Z_i = z_i],$$

где Z_i имеет биномиальное распределение $B[n-1-z_1-\ldots-z_{i-1},c/n]$. Распределение Пуассона является предельным для биномиального распределения. Когда $m = m(n) \sim n$ и c, i зафиксированы,

$$\lim_{n\to\infty} \Pr\left[B[m,c/n] = z\right] = \lim_{n\to\infty} \binom{m}{z} \left(\frac{c}{n}\right)^z \left(1-\frac{c}{n}\right)^{m-z} = e^{-c}c^z/z!.$$

Отсюда следует, что

$$\lim_{n\to\infty} \Pr\left[H = (z_1,\ldots,z_t)\right] = \Pr[H^* = (z_1,\ldots,z_t)].$$

Предположим, что $c < 1$. Для любого фиксированного t выполнено соотношение $\lim_{n\to\infty} \Pr\left[T = t\right] = \Pr[T^* = t]$. Оценим теперь размер наибольшей компоненты. Для любого t имеем

$$\Pr\left[T > t\right] \leqslant \Pr[Y_t > 0] = \Pr\left[B[n-1,1-(1-p)^t] \geqslant t\right] \leqslant \Pr\left[B[n,tc/n] \geqslant t\right],$$

поскольку $1-(1-p)^t \leqslant tp$ и $n-1<n$. Из неравенства больших уклонений А.1.14 следует, что

$$\Pr\bigl[T>t\bigr]<e^{-\alpha t},$$

где $\alpha=\alpha(c)>0$. Пусть $\beta=\beta(c)$ удовлетворяет неравенству $\alpha\beta>1$. Тогда

$$\Pr\bigl[T>\beta\ln n\bigr]<n^{-\alpha\beta}=o(n^{-1}).$$

Начальную вершину v можно выбрать n способами. Поэтому почти наверное все компоненты имеют размер $O(\ln n)$.

Теперь предположим, что $c>1$. Для любого фиксированного t справедливо соотношение $\lim_{n\to\infty}\Pr[T=t]=\Pr[T^*=t]$. Но что происходит при $T^*=\infty$? При $t=o(n)$ мы можем применить оценки $1-(1-p)^t\sim pt$ и $n-1\sim n$ и, таким образом, из неравенства больших уклонений получить, что вероятность

$$\Pr\bigl[Y_t\leqslant 0\bigr]=\Pr\bigl[B[n-1,1-(1-p)^t]\leqslant t-1\bigr]\sim\Pr\bigl[B[n,tc/n]\leqslant t\bigr]$$

экспоненциально убывает по t. Когда $t=\alpha n$, мы оцениваем $1-(1-p)^t$ через $1-e^{-c\alpha}$. Уравнение $1-e^{-c\alpha}=\alpha$ имеет решение $\alpha=1-y$, где y является вероятностью вырождения, определенной в разд. 10.4.

Поскольку $\alpha<1-y$, то $1-e^{-c\alpha}>\alpha$, следовательно вероятность

$$\Pr\bigl[Y_t\leqslant 0\bigr]\sim\Pr\bigl[B[n,1-e^{-c\alpha}]\leqslant\alpha n\bigr]$$

экспоненциально мала. При $\alpha>1-y$ верно, что $1-e^{-c\alpha}<\alpha$ и $\Pr[Y_t\leqslant 0]\sim 1$. Таким образом, почти всегда $Y_t=0$ при некотором $t\sim(1-y)n$. В основном, $T^*=\infty$ соответствует $T\sim(1-y)n$. Пусть $\varepsilon,\delta>0$ являются произвольно малыми. При более аккуратной оценке можно показать, что существует такое t_0, что для достаточно большого n выполнено неравенство

$$\Pr\bigl[t_0<T<(1-\delta)n(1-y)\ \text{ или }\ T>(1+\delta)n(1-y)\bigr]<\varepsilon.$$

Выберем t_0 достаточно большим, чтобы

$$y-\varepsilon\leqslant\Pr[T^*\leqslant t_0]\leqslant y.$$

Тогда, поскольку $\lim_{n\to\infty}\Pr[T\leqslant t_0]=\Pr[T^*\leqslant t_0]$, то

$$y-2\varepsilon\leqslant\Pr\bigl[T\leqslant t_0\bigr]\leqslant y+\varepsilon,$$
$$1-y-2\varepsilon\leqslant\Pr\bigl[(1-\delta)n(1-y)<T<(1+\delta)n(1-y)\bigr]<1-y+3\varepsilon.$$

Теперь обобщим нашу процедуру, чтобы найти компоненты графов. Начнем со случая $G\sim G(n,p)$. Выберем $v=v_1\in G$ и построим $C(v_1)$ как ранее. Затем удалим $C(v_1)$, выберем $v_2\in G\setminus C(v_1)$ и так далее. На каждой стадии оставшийся граф имеет распределение $G(m,p)$, где m обозначает число вершин. Заметим, что в оставшемся графе не было проверено ни одной пары $\{w,w'\}$, и поэтому он сохраняет свое распределение. Будем говорить, что компонента $C(v)$ *малая*, если $|C(v)|\leqslant t_0$, *гигантская*, если $(1-\delta)n(1-y)<|C(v)|<(1+\delta)n(1-y)$, и *средняя* в остальных случаях. Выберем $s=s(\varepsilon)$ так, что $(y+\varepsilon)^s<\varepsilon$. Заметим, что при малых ε выполнено $s\sim K\ln\varepsilon^{-1}$. Начнем эту процедуру с полного графа и остановимся,

когда будет обнаружена либо гигантская компонента, либо средняя, либо s малых компонент. Так как вплоть до этой стадии были найдены только малые компоненты, на каждом шаге число оставшихся вершин равно $m = n - O(1) \sim n$. Поэтому условные вероятности малой, гигантской и средней компонент остаются асимптотически одинаковыми. Вероятность обнаружения средней компоненты не превосходит $s\varepsilon$. Вероятность обнаружения всех малых компонент не превосходит $(y + \varepsilon)^s \leqslant \varepsilon$. Таким образом, с вероятностью не меньше $1 - \varepsilon'$, где значение $\varepsilon' = (s + 1)\varepsilon$ может быть сделано сколь угодно малым, мы можем найти серию из меньше чем s малых компонент, за которой последует гигантская компонента. Оставшийся граф имеет $m \sim yn$ вершин, и $pm \sim cy = d$ (сопряженный с c элемент, определенный в разд. 10.4). Поскольку $d < 1$, то предыдущий анализ дает максимальные компоненты. Резюмируя, заметим, что почти всегда граф $G(n, c/n)$ имеет гигантскую компоненту размера $\sim (1 - y)n$ и все другие компоненты размера $O(\ln n)$. Более того, принцип двойственности, введенный в разд. 10.4, имеет дискретный аналог.

Дискретный принцип двойственности. Пусть числа d и c, $d < 1 < c$, сопряжены. Структура графа $G(n, c/n)$ без ее гигантской компоненты — это структура графа $G(m, d/m)$, где m — число вершин, не входящих в гигантскую компоненту, и $m \sim ny$.

Малые компоненты графа $G(n, c/n)$ могут быть изучены также со статистической точки зрения. При фиксированном k обозначим через X количество компонент, являющихся деревом размера k. Тогда

$$\mathbf{E}[X] = \binom{n}{k} k^{k-2} \left(\frac{c}{n}\right)^{k-1} \left(1 - \frac{c}{n}\right)^{k(n-k)+\binom{k}{2}-(k-1)}.$$

Здесь мы пользовались нетривиальным фактом, доказанным Кэли, что существует k^{k-2} деревьев на заданном k-множестве. При фиксированных c и k

$$\mathbf{E}[X] \sim n \frac{e^{-ck} k^{k-2} c^{k-1}}{k!}.$$

Поскольку деревья являются строго сбалансированными, используя метод второго момента, получаем, что почти всегда $X \sim \mathbf{E}[X]$. Таким образом, $\sim p_k n$ вершин содержатся в компонентах, являющихся деревьями размера k, где

$$p_k = \frac{e^{-ck}(ck)^{k-1}}{k!}.$$

Аналитически может быть показано, что $p_k = \Pr[T = k]$ в ветвящемся процессе со средним c из разд. 10.4. Пусть Y_k обозначает количество циклов размера k, а Y — количество всех циклов. Тогда

$$\mathbf{E}[Y_k] = \frac{(n)_k}{2k} (\frac{c}{n})^k \sim \frac{c^k}{2k}$$

при фиксированном k. При $c < 1$ математическое ожидание

$$\mathbf{E}[Y] = \sum \mathbf{E}[Y_k] \to \sum_{k=1}^{\infty} \frac{c^k}{2k}$$

и имеет конечный предел, в то время как при $c > 1$ имеет место сходимость $\mathbf{E}[Y] \to \infty$. Даже при $c > 1$ для любого фиксированного k количество k-циклов имеет ограниченное математическое ожидание, и поэтому не влияет на асимптотику числа компонент заданного размера.

10.6. ФАЗОВЫЙ ПЕРЕХОД ИЗНУТРИ

В эволюции случайного графа $G(n,p)$ решающее изменение происходит в окрестности $p = c/n$ при $c = 1$. Малые компоненты в это время быстро объединяются, образуя гигантскую компоненту. Это соответствует ветвящимся процессам, когда число рождений является пуассоновской случайной величиной со средним 1. В них число T организмов будет конечным почти всегда, но, тем не менее, его математическое ожидание будет бесконечным. В последние годы обозначился большой интерес в изучении фазового перехода «изнутри» относительно роста наибольших компонент (см., например, [Luczak (1990)] или монументальную работу [Janson, Knuth, Luczak and Pittel (1993)]). Подходящая параметризация, возможно, неожиданна:

$$p = \frac{1}{n} + \frac{\lambda}{n^{4/3}}.$$

Когда $\lambda = \lambda(n) \to -\infty$, фазовый переход еще не начался. Наибольшие компоненты имеют размер $o(n^{2/3})$, и существует еще много компонент почти наибольшего размера. Когда $\lambda = \lambda(n) \to +\infty$, фазовый переход осуществлен: появилась наибольшая компонента, размер которой $\gg n^{2/3}$, а все остальные компоненты имеют размер $o(n^{2/3})$. Зафиксируем λ и c и обозначим через X количество компонент размера $k = cn^{2/3}$, являющихся деревьями. Тогда

$$\mathbf{E}[X] = \binom{n}{k} k^{k-2} \left(\frac{c}{n}\right)^{k-1} \left(1 - \frac{c}{n}\right)^{k(n-k)+\binom{k}{2}-(k-1)}.$$

Проследите за сокращениями:

$$\binom{n}{k} = \frac{(n)_k}{k!} \sim \frac{n^k e^k}{k^k \sqrt{2\pi k}} \prod_{i=1}^{k-1} \left(1 - \frac{i}{n}\right).$$

При $i < k$ можно записать

$$-\ln\left(1 - \frac{i}{n}\right) = \frac{i}{n} + \frac{i^2}{2n^2} + O\left(\frac{i^3}{n^3}\right),$$

поэтому

$$\sum_{i=1}^{k-1} -\ln(1 - \frac{i}{n}) = \frac{k^2}{2n} + \frac{k^3}{6n^2} + o(1) = \frac{k^2}{2n} + \frac{c^3}{6} + o(1).$$

Заметим, что

$$p^{k-1} = n^{1-k}\left(1+\frac{\lambda}{n^{1/3}}\right)^{k-1},$$

$$(k-1)\ln\left(1+\frac{\lambda}{n^{1/3}}\right) = (k-1)\left(\frac{\lambda}{n^{1/3}} - \frac{\lambda^2}{2n^{2/3}} + O(n^{-1})\right) =$$

$$= \frac{\lambda k}{n^{1/3}} - \frac{\lambda^2 c}{2} + o(1).$$

Также

$$\ln(1-p) = -p + O(n^{-2}) = -\frac{1}{n} - \frac{\lambda}{n^{4/3}} + O(n^{-2})$$

и

$$k(n-k) + \binom{k}{2} - (k-1) = kn - \frac{k^2}{2} + O(n^{2/3}),$$

поэтому

$$\left[k(n-k) + \binom{k}{2} - (k-1)\right]\ln(1-p) = -k + \frac{k^2}{2n} - \frac{\lambda k}{n^{1/3}} + \frac{\lambda c^2}{2} + o(1)$$

и

$$\mathbf{E}[X] \sim \frac{n^k k^{k-2}}{k^k\sqrt{2\pi k}n^{k-1}}e^A,$$

где

$$A = k - \frac{k^2}{2n} - \frac{c^3}{6} + \frac{\lambda k}{n^{1/3}} - \frac{\lambda^2 c}{2} - k + \frac{k^2}{2n} - \frac{\lambda k}{n^{1/3}} + \frac{\lambda c^2}{2} + o(1) =$$

$$= -\frac{c^3}{6} - \frac{\lambda^2 c}{2} + \frac{\lambda c^2}{2} + o(1).$$

Отсюда вытекает, что

$$\mathbf{E}[X] \sim n^{-2/3}e^{-\frac{c^3}{6}-\frac{\lambda^2 c}{2}+\frac{\lambda c^2}{2}}c^{-5/2}(2\pi)^{-1/2}.$$

Для любого фиксированного k имеет место сходимость $\mathbf{E}[X] \to 0$, но если будем суммировать по k от $cn^{2/3}$ до $(c+dc)n^{2/3}$, то появляется множитель $n^{2/3}dc$. Переходя к пределу, получаем интеграл. Для любых фиксированных a, b, λ обозначим через X количество компонент, являющихся деревьями, размера между $an^{2/3}$ и $bn^{2/3}$. Тогда

$$\lim_{n\to\infty}\mathbf{E}[X] = \int_a^b e^{-\frac{c^3}{6}-\frac{\lambda^2 c}{2}+\frac{\lambda c^2}{2}}c^{-5/2}(2\pi)^{-1/2}\,dc.$$

Большие компоненты не всегда являются деревьями. В статье [Wright (1977)] показано, что при фиксированном l существует асимптотически $c_l k^{k-2+\frac{3}{2}l}$ связных графов на k вершинах с $k-1+l$ ребрами, где c_l задается рекуррентно. Асимптотически относительно l имеем $c_l = l^{-l/2(1+o(1))}$. Вычислив $X^{(l)}$, число таких компонент на k вершинах, получим дополнительные множители $c_l k^{\frac{3}{2}l}$ и n^{-l}, которые дают $c_l c^{\frac{3}{2}l}$. При фиксированных a, b, λ, l число $X^{(l)}$ компонент размера от $an^{2/3}$ до $bn^{2/3}$, в которых число ребер больше числа

вершин на $l-1$, удовлетворяет соотношению

$$\lim_{n\to\infty} \mathbf{E}[X^{(l)}] = \int_a^b e^{-\frac{c^3}{6}-\frac{\lambda^2 c}{2}+\frac{\lambda c^2}{2}} c^{-5/2}(2\pi)^{-1/2}(c_l c^{\frac{3}{2}l})\ dc.$$

Пусть X^* — количество всех компонент размера от $an^{2/3}$ до $bn^{2/3}$, тогда

$$\lim_{n\to\infty} \mathbf{E}[X^*] = \int_a^b e^{-\frac{c^3}{6}-\frac{\lambda^2 c}{2}+\frac{\lambda c^2}{2}} c^{-5/2}(2\pi)^{-1/2} g(c)\ dc,$$

где

$$g(c) = \sum_{l=0}^{\infty} c_l c^{\frac{3}{2}l},$$

ряд сходится при всех c (здесь $c_0 = 1$). Компонента размера $\sim cn^{2/3}$ имеет на $l-1$ больше ребер, чем вершин, с вероятностью $c_l c^{\frac{3}{2}l}/g(c)$, не зависящей от λ. Поскольку $\lim_{c\to 0} g(c) = 1$, большинство компонент размера $\varepsilon n^{2/3}$, $\varepsilon \ll 1$, являются деревьями, но с ростом c распределение относительно l монотонно растет.

Подведем *итог* настоящего раздела. Для любого фиксированного λ наибольшие компоненты имеют размер $cn^{2/3}$, где значение константы зависит от компоненты. При $\lambda = -10^6$ существует некоторая положительная предельная вероятность того, что размер наибольшей компоненты больше $10^6 n^{2/3}$, а при $\lambda = +10^6$ существует положительная предельная вероятность того, что размер наибольшей компоненты меньше $10^{-6}n^{2/3}$, хотя обе эти вероятности мизерны. Подынтегральные функции имеют полюс в точке $c = 0$, отражая тот факт, что для любого λ должно существовать много компонент размера около $\varepsilon n^{2/3}$ при достаточно малом $\varepsilon = \varepsilon(\lambda)$. Когда λ является большим отрицательным числом (например -10^6), наибольшая компонента скорее всего будет иметь размер $\varepsilon n^{2/3}$, где ε — малая величина, и будет много компонент почти такого же размера. Число компонент, не являющихся деревьями, будет составлять незначительную долю числа компонент, являющихся деревьями.

Теперь рассмотрим эволюцию $G(n,p)$ относительно λ. Предположим, что при заданном λ существуют компоненты размера $c_1 n^{2/3}$ и $c_2 n^{2/3}$. Когда мы движемся от λ к $\lambda + d\lambda$, эти компоненты объединятся с вероятностью $c_1 c_2 d\lambda$. На компоненты действует «сила притяжения»: вероятность объединения пропорциональна их размерам. Так как с вероятностью $(c_1^2/2)d\lambda$ в компоненте размера $c_1 n^{2/3}$ может появиться новое внутреннее ребро, большие компоненты редко остаются деревьями. В то же время, большие компоненты поглощают другие вершины.

При $\lambda = -10^6$ будем считать, что у нас феодализм. Много маленьких компонент (за́мков) конкурируют друг с другом за возможность стать самой крупной. При увеличении λ компоненты увеличиваются в размерах, и начинают появляться немногие большие компоненты (нации). Большая Франция имеет намного больше шансов увеличиться, чем маленькая Андорра. Наибольшие компоненты стремятся к объединению, и при достижении $\lambda = +10^6$ скорее всего образуется гигантская компонента, Римская Империя. С большой

вероятностью эта компонента больше не конкурирует за превосходство, но она продолжает поглощать меньшие компоненты, пока не придет к полноте — Всему Миру.

10.7. ЗАКОНЫ «НУЛЯ ИЛИ ЕДИНИЦЫ»

В этом разделе мы ограничимся рассмотрением теоретико-графовых свойств, выразимых в формальной теории графов первого порядка. Язык этой теории состоит из переменных $(x, y, z, \ldots)$, которые всегда обозначают вершины графа, символов равенства и смежности $(x = y, x \sim y)$, обычных булевых связок $(\wedge, \neg, \ldots)$ и кванторов общности и существования $(\forall_x, \exists_y)$. Предложения теории конечны. Приведем несколько примеров записи свойств на этом формальном языке. Свойство содержания треугольника:

$$\exists_x \exists_y \exists_z [x \sim y \wedge x \sim z \wedge y \sim z],$$

отсутствия изолированной точки:

$$\forall_x \exists_y [x \sim y],$$

того, что радиус не превосходит двух:

$$\exists_x \forall_y \Big[\neg(y = x) \wedge \neg(y \sim x) \longrightarrow \exists_z [y \sim z \wedge z \sim x]\Big].$$

Для всякого свойства A и произвольных n и p рассмотрим вероятность того, что случайный граф $G(n, p)$ удовлетворяет свойству A. Эта вероятность обозначается

$$\Pr[G(n, p) \models A].$$

Целью данного раздела будет результат, полученный в работах [Glebskii, Kogan, Liagonkii and Talanov (1969)] и [Fagin (1976)] независимо друг от друга, (теорема 10.7.1), а также результат, выписанный в статье [Shelah and Spencer (1988)] (теорема 10.7.2).

Теорема 10.7.1. *Для любого фиксированного p, $0 < p < 1$, и любого свойства A первого порядка*

$$\lim_{n\to\infty} \Pr[G(n, p) \models A] = 0 \quad \textit{или} \quad \lim_{n\to\infty} \Pr[G(n, p) \models A] = 1.$$

Теорема 10.7.2. *Для любого иррационального α, $0 < \alpha < 1$, пусть $p = p(n) = n^{-\alpha}$. Тогда для любого свойства A первого порядка*

$$\lim_{n\to\infty} \Pr[G(n, p) \models A] = 0 \quad \textit{или} \quad \lim_{n\to\infty} \Pr[G(n, p) \models A] = 1.$$

Мы приводим наброски доказательств обеих теорем.

Будем говорить, что функция $p = p(n)$ удовлетворяет закону «нуля или единицы», если приведенное выше равенство верно для всех свойств A первого порядка.

Теорема Глебского—Фагина имеет естественную интерпретацию при $p = 0.5$, поскольку в этом случае $G(n, p)$ предоставляет равную вероятность

каждому (помеченному) графу. Далее теорема утверждает, что любое свойство A первого порядка выполняется либо почти для всех графов, либо почти ни для каких графов. Теорема Шелла—Спенсера может быть интерпретирована в терминах пороговых функций. Из общих результатов разд. 10.1 вытекает, например, что $p = n^{-2/3}$ является пороговой функцией содержания K_4. То есть, если $p \ll n^{-2/3}$, то $G(n,p)$ почти наверное не содержит K_4, если же $p \gg n^{-2/3}$, то $G(n,p)$ почти наверное содержит K_4. В промежуточном случае, скажем, при $p = n^{-2/3}$ вероятность находится строго между 0 и 1. В этом случае она равняется $1-e^{-1/24}$. Будем употреблять следующее (правда, несколько вольное) выражение: при пороговой функции закон «нуля или единицы» не будет выполняться, и поэтому под тем, что $p(n)$ удовлетворяет закону «нуля или единицы», понимается, что $p(n)$ не является пороговой функцией — это является скучным моментом в эволюции случайных графов, по крайней мере, с точки зрения языка первого порядка. Что происходит с эволюцией $G(n,p)$ при $p = n^{-\pi/7}$? Ответ: ничего!

В основе нашего подхода к законам «нуля или единицы» лежит вариант игры Эренфойхт, которую мы сейчас определим. Пусть G и H — два графа с непересекающимися множествами вершин, а t — натуральное число. Рассмотрим совершенную информационную игру, обозначаемую $\mathrm{EHR}[G,H,t]$, с двумя игроками, именуемыми Спойлер и Дупликатор. Игра длится t раундов. Каждый раунд состоит из двух частей. Сначала Спойлер выбирает вершину $x \in V(G)$ или же вершину $y \in V(H)$. Тем самым он выбирает граф, из которого затем выбираются вершины. Затем Дупликатор должен выбрать вершину в другом графе. К концу раунда t выбрано по t вершин из каждого графа. Пусть $x_1,\ldots,x_t$ являются вершинами, выбранными из $V(G)$, а $y_1,\ldots,y_t$ — вершинами, выбранными из $V(H)$, где x_i и y_i — вершины, выбранные в i-м раунде. Дупликатор выигрывает тогда и только тогда, когда индуцированные графы на выбранных вершинах являются изоморфными, т. е. если для всех $1 \leqslant i < j \leqslant t$

$$\{x_i,x_j\} \in E(G) \longleftrightarrow \{y_i,y_j\} \in E(H).$$

Поскольку нет скрытых ходов и ничьих, один из игроков должен иметь выигрышную стратегию. Будем говорить, что этот игрок *выигрывает* в игре $\mathrm{EHR}[G,H,t]$.

Лемма 10.7.3. *Для любого свойства A первого порядка существует такое $t = t(A)$, что для произвольных графов G и H со свойствами $G \models A$ и $H \models \neg A$ Спойлер выигрывает в игре* $\mathrm{EHR}[G,H,t]$.

Полное доказательство потребовало бы формального анализа языка первого порядка, поэтому мы только приведем пример. Пусть A обозначает свойство $\forall_x \exists_y [x \sim y]$ отсутствия изолированной точки, и $t = 2$. Спойлер начинает, выбирая изолированную точку $y_1 \in V(H)$, он может это сделать, поскольку $H \models \neg A$. Дупликатор должен выбрать $x_1 \in V(G)$. Так как $G \models A$, то вершина x_1 не является изолированной, поэтому Спойлер может выбрать $x_2 \in V(G)$, так что $x_1 \sim x_2$, и теперь Дупликатор не может выбрать «дублирующую» вершину y_2.

Теорема 10.7.4. *Функция $p = p(n)$ удовлетворяет закону «нуля или единицы» тогда и только тогда, когда для любого t*

$$\lim_{m,n\to\infty} \Pr\Big[\text{Дупликатор выигрывает EHR}[G(n,p(n)), H(m,p(m)), t]\Big] = 1,$$

где $G(n,p(n))$ и $H(m,p(m))$ — независимые случайные графы с непересекающимися множествами вершин.

Замечание. При любой выбранной паре G и H кто-то должен выиграть в игре EHR$[G,H,t]$. (То есть не существует случайной игры, игра является совершенной.) Когда задано вероятностное распределение по (G,H), существует вероятность того, что игра EHR$[G,H,t]$ будет выиграна Дупликатором, и эта вероятность должна стремиться к единице.

Доказательство. Мы докажем только необходимость. Предположим, что значение $p = p(n)$ не удовлетворяет закону «нуля или единицы». Пусть A удовлетворяет условию

$$\lim_{n\to\infty} \Pr\left[G(n,p(n)) \models A\right] = c,$$

где $0 < c < 1$. Пусть $t = t(A)$ удовлетворяет условию леммы. С предельной вероятностью $2c(1-c) > 0$ ровно один из графов $G(n,p(n))$ и $H(n,p(n))$ будет обладать свойством A, поэтому победит Спойлер, что противоречит предположению. Это не полное доказательство, поскольку, когда закон «нуля или единицы» не выполняется, $\lim_{n\to\infty} \Pr[G(n,p(n)) \models A]$ может не существовать. Если существует подпоследовательность n_i, на которой предел равен $c \in (0,1)$, то мы можем использовать то же доказательство. В противном случае будут существовать две подпоследовательности n_i и m_i, пределы которых будут равны 0 и 1, соответственно. Тогда, устремляя $n, m \to \infty$ по n_i и m_i, соответственно, Спойлер выиграет EHR$[G,H,t]$ с вероятностью, стремящейся к единице. ■

Теорема 10.7.4 является своеобразным мостом между математической логикой и случайными графами. Чтобы доказать, что $p = p(n)$ удовлетворяет закону «нуля или единицы», нам более не нужно знать что-то относительно логики — достаточно найти хорошую стратегию для Дупликатора.

Скажем, что граф G обладает свойством расширения полного уровня s, если для любых различных $u_1,\ldots,u_a,v_1,\ldots,v_b \in G$, где $a+b \leqslant s$, существует такое $x \in V(G)$, что $\{x,u_i\} \in E(G)$, $1 \leqslant i \leqslant a$, и $\{x,v_j\} \notin E(G)$, $1 \leqslant j \leqslant b$. Предположим, что и G и H обладают свойством расширения полного уровня $s-1$. Тогда Дупликатор выигрывает EHR$[G,H,s]$ по следующей простой стратегии. В i-м раунде, при уже выбранных $x_1,\ldots,x_{i-1},y_1,\ldots,y_{i-1}$ и выборе Спойлером, скажем, x_i, Дупликатор просто выбирает y_i, смежное с теми же вершинами $y_j, j < i$, что и x_i — с $x_j, j < i$. Из свойства полного расширения вытекает, что такая вершина y_i существует.

Теорема 10.7.5. *Для любого фиксированного p, $0 < p < 1$, и произвольного s граф $G(n,p)$ почти наверное обладает свойством расширения полного уровня s.*

Доказательство. Для любых различных $u_1, \ldots, u_a, v_1, \ldots, v_b, x \in G$, таких, что $a + b \leqslant s$, определим событие $E_{u_1,\ldots,u_a,v_1,\ldots,v_b,x}$, заключающееся в том, что $\{x, u_i\} \in E(G)$, $1 \leqslant i \leqslant a$, и $\{x, v_j\} \notin E(G)$, $1 \leqslant j \leqslant b$. Тогда

$$\Pr\left[E_{u_1,\ldots,u_a,v_1,\ldots,v_b,x}\right] = p^a(1-p)^b.$$

Теперь положим

$$E_{u_1,\ldots,u_a,v_1,\ldots,v_b} = \bigwedge_x \overline{E_{u_1,\ldots,u_a,v_1,\ldots,v_b,x}},$$

где пересечение берется по $x \neq u_1, \ldots, u_a, v_1, \ldots, v_b$. Эти события попарно независимы по x, так как они задействуют различные ребра. Следовательно,

$$\Pr\left[\bigwedge_x \overline{E_{u_1,\ldots,u_a,v_1,\ldots,v_b,x}}\right] = \left(1 - p^a(1-p)^b\right)^{n-a-b}.$$

Введем обозначение $\varepsilon = \min(p, 1-p)^s$. Получим

$$\Pr\left[\bigwedge_x \overline{E_{u_1,\ldots,u_a,v_1,\ldots,v_b,x}}\right] \leqslant (1-\varepsilon)^{n-s}.$$

Здесь используется то, что ε является фиксированным (зависящим от p и s) положительным числом. Введем обозначение

$$E = \bigvee E_{u_1,\ldots,u_a,v_1,\ldots,v_b},$$

где дизъюнкция берется по всем различным $u_1, \ldots, u_a, v_1, \ldots, v_b \in G$, $a+b \leqslant s$. Существует меньше $s^2 n^s = O(n^s)$ таких возможностей выбора, поскольку мы можем выбрать a и b, а затем вершины. Таким образом,

$$\Pr[E] \leqslant s^2 n^s (1-\varepsilon)^{n-s}.$$

Но

$$\lim_{n\to\infty} s^2 n^s (1-\varepsilon)^{n-s} = 0,$$

и поэтому E почти наверное не выполняется. Итак, $\neg E$, которое в точности утверждает, что $G(n,p)$ обладает свойством расширения полного уровня s, почти наверное выполняется. ■

Мы доказали теорему 10.7.1. Для любого $p \in (0,1)$, произвольного фиксированного s при $m, n \to \infty$ с вероятностью, стремящейся к единице, оба графа $G(n,p)$ и $H(m,p)$ будут обладать свойством расширения полного уровня s и поэтому Дупликатор выиграет $\mathrm{EHR}[G(n,p), H(m,p), s]$.

Почему Дупликатор не может использовать эту стратегию, когда $p = n^{-\alpha}$? Мы проиллюстрируем трудности при помощи простого примера. Пусть $0.5 < \alpha < 1$, и пусть Спойлер и Дупликатор разыгрывают игру из трех ходов на графах G, H. Спойлер подумывает о выборе точки $z \in G$, но не говорит Дупликатору об этом. Вместо этого он выбирает $x_1, x_2 \in G$, смежные с z. Дупликатор просто выбирает точки $y_1, y_2 \in H$, смежные либо несмежные, в зависимости от того, верно ли, что $x_1 \sim x_2$. Но теперь хитрый Спойлер выбирает $x_3 = z$.

Граф $H \sim H(m, m^{-\alpha})$ не обладает свойством расширения полного уровня 2. В частности, большинство пар y_1, y_2 не имеют общего соседа. Если только Дупликатору не повезет, или он не окажется достаточно прозорливым, то он не сможет найти $y_3 \sim y_1, y_2$, и поэтому проиграет. Из этого примера не следует, что Дупликатор проиграет даже при совершенной стратегии (на самом деле, как мы установим ниже, он почти всегда выигрывает при использовании совершенной стратегии), он только показывает, что применяемая стратегия должна быть более сложной.

Начнем доказательство закона «нуля или единицы», теоремы 10.7.2. Зафиксируем иррациональное число $\alpha \in (0,1)$. Пару (R,H), где H является графом на множестве вершин, скажем, $V(H) = \{X_1, \dots, X_r, Y_1, \dots, Y_v\}$, а $R = \{X_1, \dots, X_r\}$ — заданное подмножество $V(H)$, элементы которого называются *корнями*, назовем сетью. Например, (R,H) может состоять из одной вершины Y_1, смежной с корнями X_1 и X_2. Обозначим через $v = v(R,H)$ количество вершин, которые не являются корнями, а через $e = e(R,H)$ — количество ребер, не соединяющих два корня. Мы назовем сеть (R,H) *плотной*, если $v - e\alpha < 0$; *редкой*, если $v - e\alpha > 0$. Из иррациональности α следует, что каждая сеть (R,H) принадлежит одному из этих классов. Будем говорить, что сеть (R,H) *жесткая*, если для всех S, таких, что $R \subseteq S \subset V(H)$, сеть (S,H) является плотной. Назовем сеть (R,H) *надежной*, если для всех S, таких, что $R \subset S \subseteq V(H)$, сеть $(R, H|_S)$ является редкой. Доказать некоторые элементарные свойства этих понятий предлагается в упр. 4. Иногда будем писать (R,S) вместо $(R, H|_S)$, когда из контекста будет понятно, о каком графе H идет речь.

Мы представляем себе сети как абстрактные точки. Будем говорить, что вершины $y_1, \dots, y_v$ графа G образуют (R,H)-*расширение* вершин $x_1, \dots, x_r$ тогда и только тогда, когда X_i смежно с Y_j в H, x_i смежно с y_j в G, а также когда Y_i и Y_j смежны в H, y_i и y_j смежны в G. Заметим, что графу G «позволено» иметь больше ребер чем графу H, и что ребра между корнями «не в счет».

Лемма 10.7.6 (универсальное расширение). *Пусть сеть (R,H), определенная выше, является надежной. Пусть $t \geqslant 0$ — произвольное, но фиксированное целое число. Тогда в графе $G \sim G(n, n^{-\alpha})$ для всех $x_1, \dots, x_r$ почти наверное существуют $y_1, \dots, y_v$, такие, что*

i) вершины $y_1, \dots, y_v$ образуют (R,H)-расширение $x_1, \dots, x_r$.

ii) вершины x_i и y_j являются смежными в G тогда и только тогда, когда X_i и Y_j являются смежными в H; y_i и y_j являются смежными в G тогда и только тогда, когда Y_i и Y_j смежны в H.

iii) если (при $t > 0$) $z_1, \dots, z_u$, $u \leqslant t$, образуют жесткое (R', H')-расширение над $x_1, \dots, x_r, y_1, \dots, y_v$, то ни одна из пар $\{z_k, y_j\}$ не образует ребро.

Пример. Пусть $\alpha \in (\frac{1}{2}, 1)$, $t = 2$, и пусть сеть (R, H) содержит корень X_1, некорневую вершину Y_1 и ребро $\{X_1, Y_1\}$. Заметим, что для сети (R', H'), содержащей два корня X_1 и X_2 с общим соседом Y_1, выполнены равенства $v = 1, e = 2$, и эта сеть является жесткой. Лемма об универсальном расширении в данном случае утверждает, что любая вершина x_1 имеет соседа y_1, такого, что x_1 и y_1 имеют общего соседа z_1.

Доказательство. Из упр. 5 следует, что почти наверное любые $x_1, \dots, x_r$ имеют $\Theta(n^v p^e)$ (R, H)-расширений $y_1, \dots, y_v$. Предположим, что число тех $y_1, \dots, y_v$, которые не являются универсальными, в любом из ограниченного числа возможных способов, будет ограничено меньшей степенью n.

Назовем y *специальным*, если $y \in \mathrm{cl}_{t+v}(x_1, \dots, x_r)$ (как это определяется ниже), иначе — *неспециальным*. Пусть K из леммы о конечном замыкании 10.7.7, приведенной ниже, является почти наверное оценкой специальных y, равномерно распределенных по всем способам выбора вершин x. Расширим сеть (R, H) до (R^+, H^+), добавив K новых корней, но не добавляя новых ребер. Это расширение также является надежным, причем того же типа, что и (R, H), поэтому из упражнения 5 снова следует, что почти наверное каждые $x_1, \dots, x_r, z_1, \dots, z_K$ имеют $\Theta(n^v p^e)$ (R^+, H^+)-расширений $y_1, \dots, y_v$. Допуская вершинам z включать все специальные вершины, мы получаем, что почти наверное каждый набор $x_1, \dots, x_r$ имеет $\Theta(n^v p^e)$ (R, H)-расширений $y_1, \dots, y_v$, где все y_i неспециальные. Оценим теперь сверху число тех неспециальных (R, H)-расширений, которые не удовлетворяют условиям 2 или 3.

Рассмотрим расширение (R, H') с дополнительным ребром (y_i, y_j) или (x_i, y_j). Это расширение не может содержать жесткое подрасширение, поскольку это бы значило, что некоторое y_i является специальным. Из упражнения 4 следует, что оно должно быть надежным расширением. Используя упражнение 5, получаем, что существует $\Theta(n^v p^{e+1}) = o(n^v p^e)$ таких расширений.

Рассмотрим расширения $y_1, \dots, y_v$ и $z_1, \dots, z_u$ из условия 3, где некоторые z_j и y_k смежны. Далее мы можем считать, что вершины z образуют минимальное жесткое расширение над вершинами x и y. Пусть вершины z имеют тип (v_1, e_1) как расширение над x и y, так что $v_1 - e_1\alpha$ отрицательно. Если y и z вместе образуют надежное расширение над x, то будет существовать $\Theta(n^{v+v_1} p^{e+e_1}) = o(n^v p^e)$ таких расширений и, следовательно, не более чем столько способов выбора для вершин y. В противном случае из упражнения 4 следовало бы существование жесткого подрасширения. Оно не имеет пересечений с неспециальными вершинами y. Из минимальности следует, что оно должно в точности совпадать со всеми вершинами z. Если заданы вершины x из леммы о конечном замыкании 10.7.7, то существует $O(1)$ способов выбора вершин z. Тогда y образуют (v, e') расширение над x и y с $e' > e$. Это расширение не имеет жестких расширений (опять же, потому, что вершины y не являются специальными) и, следовательно, является надежным. Снова используя упражнение 5, получаем, что для каждого выбора вершин z существует $\Theta(n^v p^{e'})$ таких y, и поэтому общее число таких вершин y равно $O(n^v p^{e'}) = o(n^v p^e)$.

Во всех случаях число вершин y, которые не удовлетворяют условиям 2 или 3, есть $o(n^v p^e)$. Следовательно, существуют y (на самом деле, большинство неспециальных y) которые являются (R,H)-расширениями и удовлетворяют условиям 2 и 3. ■

Последовательность $X = X_0 \subset X_1 \subset \ldots \subset X_K$, такая, что все сети (X_{i-1}, X_i) являются жесткими, и все $|X_{i+1} - X_i| \leqslant t$, называется *жесткой t-цепью в G*. Максимальное Y, для которого существует жесткая t-цепь (произвольной длины) $X = X_0 \subset X_1 \subset \ldots \subset X_K = Y$, называется *$t$-замыканием X* и обозначается $\mathrm{cl}_t(X)$. Когда не существует жестких t-цепей, мы по определению будем считать, что $\mathrm{cl}_t(X) = X$. Для проверки корректности данного определения заметим (используя упр. 4), что если $X = X_0 \subset X_1 \subset \ldots \subset X_K = Z$ и $X = X_0 \subset Y_1 \subset \ldots \subset Y_L = Y$ являются жесткими t-цепями, то тогда и $X = X_0 \subset X_1 \subset \ldots \subset X_K \subset Z \cup Y_1 \subset \ldots \subset Z \cup Y_L = Z \cup Y$ обладает тем же свойством. Наоборот, минимальное множество, содержащее X, которое не имеет жестких расширений $\leqslant t$ вершин, назовем t-замыканием $\mathrm{cl}_t(X)$. Будем говорить, что $x_1, \ldots, x_r \in G$ и $y_1, \ldots, y_r \in H$ имеют тот же t-тип, если их t-замыкания изоморфны как графы, причем изоморфизм ставит в соответствие каждому x_i соответствующее y_i.

Понятие t-замыкания является ключевым определением, описывающим специальные свойства корней. Предположим, например, что $\alpha \in (\frac{1}{2}, 1)$ и рассмотрим $\mathrm{cl}_1(x_1, x_2)$. Единственное жесткое расширение с $t = 1$ в данном промежутке — это некорневая вершина, смежная с двумя (или более) корнями. Примером 1-типа является следующий: x_1 и x_2 имеют общих соседей y_1 и y_2 и, более того, x_1 и y_1 имеют общего соседа y_3, и между этими вершинами нет других ребер, а никакие другие пары, кроме описанных выше, не имеют общих соседей. Случайно выбранные x_1 и x_2 будут иметь тип: x_1 и x_2 не имеют общих соседей и несмежны.

Теперь мы можем начать описывать стратегию Дупликатора. К концу r-го хода, когда уже будут выбраны $x_1, \ldots, x_r$ и $y_1, \ldots, y_r$ из двух графов, Дупликатор убедится, что эти множества имеют тот же a_r-тип. Мы будем называть совокупность $(a_1, \ldots, a_t)$ *стратегией предварительного просмотра*. Здесь a_r должно зависеть только от t, общего числа ходов в игре, и α. Условимся, что $a_t = 0$, так что в конце игры Дупликатор, если он сможет следовать $(a_1, \ldots, a_t)$ стратегии предварительного просмотра, победит. Однако, если Спойлер выберет, скажем, x_{r+1} так, что не будет существовать соответствующей вершины y_{r+1}, такой, что $x_1, \ldots, x_{r+1}$ и $y_1, \ldots, y_{r+1}$ имеют один и тот же a_{r+1}-тип, то стратегия будет проигрышной, и будем говорить, что Спойлер победил. Значения a_r дают «предварительный просмотр», того, что использует Дупликатор, но перед тем как определять их, нам потребуются некоторые предварительные результаты.

Лемма 10.7.7 (конечное замыкание). *Пусть зафиксированы α, $r > 0$. Положим ε равным минимальному значению выражения $\frac{e\alpha - v}{v}$ по всем целым v и e, таким, что $1 \leqslant v \leqslant t$ и $e\alpha - v > 0$. Пусть K таково, что $r - K\varepsilon < 0$. Тогда в графе $G(n, n^{-\alpha})$ для всех $X \subset G$, $|X| = r$, почти наверное выполнено*

неравенство

$$|\mathrm{cl}_t(X)| \leqslant K + r.$$

Доказательство. Если бы это было не так, то существовала бы жесткая t-цепь $X = X_0 \subset X_1 \subset \ldots \subset X_L = Y$, такая, что $K + r < |Y| < K + r + t$. Полагая, что (X_{i-1}, X_i) имеет тип (v_i, e_i), ограничение G на Y имело бы $r + \sum v_i$ вершин и хотя бы $\sum e_i$ ребер. Но

$$\left(r + \sum v_i\right) - \alpha\left(\sum e_i\right) = r + \sum(v_i - \alpha e_i) \leqslant r - \varepsilon \sum v_i < r - K\varepsilon < 0,$$

и G почти наверное не содержит такого подграфа. ■

Замечание. Оценка на $|\mathrm{cl}_t(X)|$, полученная в этом доказательстве, очень сильно зависит от того, насколько хорошо может быть приближено α рациональными числами, знаменатель которых не превышает t. Часто это является решающим обстоятельством. Если, например, $\frac{1}{2} + \frac{1}{s-1} > \alpha > \frac{1}{2} + \frac{1}{s}$, то почти наверное будут существовать две точки $x_1, x_2 \in G(n, n^{-\alpha})$, у которых есть s общих соседей. Так что $|\mathrm{cl}_1(x_1, x_2)| \geqslant s + 2$.

Теперь определим элементы $a_1, \ldots, a_t$ стратегии предварительного просмотра, используя обратную индукцию. Положим $a_t = 0$. Если к концу игры Дупликатор может убедиться, что 0-типы из $x_1, \ldots, x_t$ и $y_1, \ldots, y_t$ совпадают, то они имеют одинаковые индуцированные подграфы, и он победил. Предположение индукции: пусть определено $b = a_{r+1}$. Определим $a = a_r$ как произвольное целое число, удовлетворяющее условиям

1) $a \geqslant b$;

2) почти наверное $|\mathrm{cl}_b(W)| - r \leqslant a$ для всех множеств W размера $r + 1$.

Теперь мы должны показать, что эта стратегия почти наверное является выигрышной. Пусть $G_1 \sim G(n, n^{-\alpha})$, $G_2 \sim G(m, m^{-\alpha})$ и Дупликатор пытается играть с помощью $(a_1, \ldots, a_t)$ стратегии предварительного просмотра на $\mathrm{EHR}(G_1, G_2, t)$.

Рассмотрим $(r + 1)$-й ход. Пусть, как и выше, $b = a_{r+1}$ и $a = a_r$. Точки $x_1, \ldots, x_r \in G_1$, $y_1, \ldots, y_r \in G_2$ уже выбраны. Положим для удобства $X = \{x_1, \ldots, x_r\}$, $Y = \{y_1, \ldots, y_r\}$. Предположим, что Дупликатор остался в игре до сих пор, так что $\mathrm{cl}_a(X) \cong \mathrm{cl}_a(Y)$, изоморфизм сопоставляет каждому x_i соответствующее y_i. Спойлер выбирает, скажем, $x = x_{r+1} \in G_1$. Положим $X^+ = X \cup \{x\}$ и $Y^+ = Y \cup \{y\}$, где y — не определенный еще контр-ход Дупликатора. Рассмотрим два случая.

Будем говорить, что Спойлер *пошел внутрь*, если $x \in \mathrm{cl}_a(X)$. Тогда, $\mathrm{cl}_b(X^+) \subseteq \mathrm{cl}_a(X)$, поскольку $b \leqslant a$. Дупликатор смотрит на изоморфизм $\Psi : \mathrm{cl}_a(X) \to \mathrm{cl}_a(Y)$ и выбирает $y = \Psi(x)$.

Скажем, что Спойлер *пошел наружу*, если $x \notin \mathrm{cl}_a(X)$. Пусть NEW — те вершины $\mathrm{cl}_b(X^+)$, которые не лежат в $\mathrm{cl}_a(X)$. Имеем $NEW \neq \varnothing$, поскольку $x \in NEW$. Кроме того, $|NEW| \leqslant a$, поскольку $NEW \subseteq \mathrm{cl}_b(X^+) - X$. Рассмотрим NEW как (R, H)-расширение $\mathrm{cl}_a(X)$. Это расширение должно быть

надежным, потому что в противном случае оно бы имело жесткое подрасширение NEW^-, но тогда это подрасширение содержалось бы в $\mathrm{cl}_a(X)$. Дупликатор теперь ходит к G_2 и, применяя лемму об универсальном расширении 10.7.6 с $t = b$, находит (R, H)-расширение $\mathrm{cl}_a(Y)$. То есть он находит инъекцию, сохраняющую ребро, $\Psi : \mathrm{cl}_a(X) \cup NEW \to H$, таким образом расширяя изоморфизм между $\mathrm{cl}_a(X)$ и $\mathrm{cl}_a(Y)$. Дупликатор выбирает $y = \Psi(x)$.

Почему это работает? Положим $NEW' = \Psi(NEW)$ и $CORE = \Psi(\mathrm{cl}_b(X^+))$. Мы можем достичь $\mathrm{cl}_b(X^+)$ при помощи жесткой b-цепи от X^+, а изоморфизм дает ту же цепь от Y^+ до $CORE$, так что $\mathrm{cl}_b(Y^+)$ содержит $CORE$. Но может ли оно содержать дополнительные вершины? Мы используем универсальность, чтобы ответить на этот вопрос отрицательно. Предположим, что существует жесткое расширение $MORE$ над $CORE$ с не более чем b некорневыми вершинами. Расширение $MORE$ не может целиком содержаться в $\Psi[\mathrm{cl}_a(X) \cup NEW]$, так как в этом случае $\Psi^{-1}[MORE]$ содержалось бы в $\mathrm{cl}_b(X^+)$. Пусть $MORE^+$ обозначает множество вершин из $MORE$, лежащих вне $\Psi[\mathrm{cl}_a(X) \cup NEW]$. Тогда $MORE^+$ является жестким расширением $\Psi[\mathrm{cl}_a(X) \cup NEW]$. Из универсальности следует, что $MORE^+$ не имеет смежных вершин из NEW', и, таким образом, будет жестким расширением $\Psi[\mathrm{cl}_a(X)] = \mathrm{cl}_a(Y)$. Так как $a \geqslant b$, a-замыкание множества не может иметь жестких расширений с не более чем b вершинами. Следовательно, $MORE$ не существует.

Первый ход аналогичен тому же принципу, но он несколько проще. Положим $b = a_1$. Пусть a удовлетворяет условиям: $a \geqslant b$ и $a \geqslant |\mathrm{cl}_b(x)|$ при любом x. Спойлер играет $x \in G_1$. На самом деле, не может быть хода внутрь, поскольку $X = \varnothing$ является множеством предыдущих ходов и $\mathrm{cl}_a(\varnothing) = \varnothing$. Дупликатор находит граф $H = \mathrm{cl}_b(x)$, который имеет, скажем, v вершин (включая x) и e ребер. Поскольку H является подграфом G_1, то функция появления H должна появиться до $n^{-\alpha}$. В частности, для любого подграфа H' графа H с v' вершина и e' ребрами не выполняется неравенство $v' - \alpha e' < 0$, и поэтому должно выполняться $v' - \alpha e' > 0$. Тогда выполняются условия теоремы 4.4.5, поэтому G_2 почти наверное содержит $\Theta(m^{e-v\alpha})$ копий H. Рассмотрим произвольный граф H^+, содержащий H вместе с жестким расширением H и с не более чем b вершинами. Такой граф H^+ будет содержать $v + v^+$ вершин и $e + e^+$ ребер, также будет выполняться условие $v^+ - \alpha e^+ < 0$. Тогда математическое ожидание копий H^+ будет равняться $\Theta(m^{e-v\alpha+(v^+-\alpha e^+)})$, что есть $o(m^{e-v\alpha})$. Следовательно, в G_2 найдется копия H, которая не является частью никакого такого H^+. (На самом деле, это универсальное расширение над пустым множеством.) Дупликатор найдет инъекцию, сохраняющую ребра, $\Psi : \mathrm{cl}_b(x) \to G_2$, таким образом представляя такую копию H, и выберет $y = \Psi(x)$.

Итак, нами показано, что $(a_1, \ldots, a_t)$ стратегия предварительного просмотра почти наверное приводит к победе Дупликатора. Отсюда по теореме 10.7.4 следует закон «нуля или единицы». 10.7.2

10.8. УПРАЖНЕНИЯ

1. Покажите, что существует граф на n вершинах, в котором минимальная степень не меньше $n/2$ и размер любого доминирующего множества не меньше $\Omega(\log n)$.

2. Найти пороговую функцию следующего свойства: $G(n,p)$ содержит копию графа, состоящего из полного графа на четырех вершинах и еще одной вершины, соединенной с одной из ее вершин.

3. Пусть X обозначает число циклов в случайном графе $G(n,p)$ с $p = \frac{c}{n}$. Найти точную формулу для $\mathbf{E}[X]$. Найти асимптотику $\mathbf{E}[X]$ при $c < 1$. Найти асимптотику $\mathbf{E}[X]$ при $c = 1$.

4. Здесь мы пишем (R,S) вместо $(R, H|_S)$, где H — некоторый фиксированный граф.

   - Пусть $R \subset S \subset T$. Покажите, что если сети (R,S) и (S,T) плотны, то и (R,T) плотна. Покажите, что если (R,S) и (S,T) являются редкими, то такой же является и (R,T).
   - Пусть $R \subset S$. Покажите, что, если сеть (R,S) является жесткой, то и $(X \cup R, X \cup S)$ является жесткой для любого X.
   - Пусть $R \subset U$ и сеть (R,U) не является редкой. Покажите, что существует T, $R \subset T \subset U$, такое, что сеть (R,T) плотна. Докажите также, что существует S, $R \subset S \subset T$, такое, что сеть (R,S) жесткая.
   - Покажите, что любая сеть (R,T) является либо жесткой, либо редкой, либо существует S, $R \subset S \subset T$, такое, что сеть (R,S) является жесткой, а (S,T) — редкой.

5. Назовем сеть (R,H) *шарнирной*, если она надежная, и не существует множества S, $R \subset S \subset V(H)$, такого, что (S,H) надежная. Обозначим через $N(x_1,\ldots,x_r)$, $x_1,\ldots,x_r \in G$, число (R,H) расширений. Положим $\mu = \mathbf{E}[N] \sim n^v p^e$.

   - Пусть (R,H) является шарнирной. Зафиксируем $x_1,\ldots,x_r \in G$. Используя конструкцию, описанную в разд. 8.5 и, особенно, в теореме 8.5.4, покажите, что
   $$\Pr\Big[|N(x_1,\ldots,x_r) - \mu| > \varepsilon\mu\Big] = o(n^{-r}).$$
   - Выведите, что почти наверное все $N(x_1,\ldots,x_r) \sim \mu$.
   - Покажите, что $N(x_1,\ldots,x_r) \sim \mu$ выполняется для любой надежной сети (R,H), разлагая (R,H) на шарнирные расширения.

ВЕРОЯТНОСТНЫЙ ВЗГЛЯД:

Число подграфов

Граф $G = (V, E)$ на n вершинах имеет 2^n индуцированных подграфов, но некоторые из них наверняка будут изоморфны. Сколько различных подграфов может иметь G? Здесь мы покажем, что существуют графы G с $2^n(1 - o(1))$ различными подграфами. Доказательство, которое мы приводим, довольно грубое. Это типичный случай из числа тех, в которых вероятностный подход дает довольно быстрые ответы на вопросы, к которым трудно подступиться другими методами.

Пусть G — случайный граф на n вершинах, в котором ребро появляется с вероятностью $1/2$. Зафиксируем $S \subseteq V$, $|S| = t$. Для любого взаимно однозначного отображения $\rho : S \to V$, $\rho \neq id$, пусть A_ρ обозначает событие «ρ представляет изоморфизм графов», т. е. для $x, y \in S$ выполняется $\{x, y\} \in E \Leftrightarrow \{\rho x, \rho y\} \in E$. Положим $M_\rho = \{x \in S : \rho x \neq x\}$. Мы разбиваем множество отображений ρ на $g = g(\rho) = |M_\rho|$ частей.

Рассмотрим $g(t - g) + \binom{g}{2}$ пар (x, y), $x, y \in S$, и по крайней мере одну (x, y) в M. Для всех, таких пар, за исключением более $g/2$ пар, выполнено $\{x, y\} \neq \{\rho x, \rho y\}$. (Исключением является случай, когда $\rho x = y, \rho y = x$.) Пусть E_ρ является множеством пар $\{x, y\}$, таких, что $\{x, y\} \neq \{\rho x, \rho y\}$. Определим граф H_ρ множеством вершин которого является E_ρ, и вершина $\{x, y\}$ смежна с $\{\rho x, \rho y\}$. В H_ρ каждая вершина имеет степень не более двух, поскольку $\{x, y\}$ может быть также смежной с $\{\rho^{-1}x, \rho^{-1}y\}$, и поэтому она разлагается в объединение изолированных вершин, цепей и циклов. В каждой такой компоненте есть независимое множество размера не меньше чем одна треть от числа элементов; равенство достигается на треугольнике. Таким образом, существует множество $I_\rho \subseteq E_\rho$, такое, что

$$|I_\rho| \geqslant |E_\rho| \geqslant \frac{g(t - g) + \binom{g}{2} - g/2}{3}.$$

Поэтому пары $\{x, y\}, \{\rho x, \rho y\}$, такие, что $\{x, y\} \in I_\rho$, все различны.

При каждом $\{x, y\} \in I_\rho$ событие $\{x, y\} \in E \Leftrightarrow \{\rho x, \rho y\} \in E$ выполняется с вероятностью $1/2$. Более того, эти события взаимно независимы для всех $\{x, y\} \in I_\rho$, так как они порождены различными парами. Отсюда вытекает оценка

$$\Pr[A_\rho] \leqslant 2^{-|I_\rho|} \leqslant 2^{-(g(t-g)+\binom{g}{2}-g/2)/3}.$$

При заданном g функция ρ определяется условием $\{x : \rho x \neq x\}$ и значениями ρx для этих x, поэтому существует меньше n^{2g} таких ρ. Получаем оценки

$$\sum_{\rho \neq id} \Pr[A_\rho] = \sum_{g=1}^{t} \sum_{g(\rho)=g} \Pr[A_\rho] \leqslant \sum_{g=1}^{t} n^{2g} 2^{-(g(t-g)+\binom{g}{2}-g/2)/3}.$$

Мы приводим грубую оценку

$$g(t-g)+\binom{g}{2}-g/2=g\left(t-\frac{g}{2}-1\right)\geqslant g\left(\frac{t}{2}-1\right),$$

поскольку $g\leqslant t$. Тогда

$$\sum_{\rho\neq id}\Pr[A_\rho]\leqslant\sum_{g=1}^{t}\left[n^2 2^{(-\frac{t}{2}+1)/3}\right]^g.$$

Имеем $2^{\frac{1}{3}-\frac{t}{6}}<n^{-3}$ и $\sum_{\rho\neq id}\Pr[A_\rho]=o(1)$, поскольку (опять же из грубых оценок) $t>50\ln n$. То есть почти наверное не существует изоморфной копии $G|_S$.

Для всех $S\subseteq V$, таких, что $|S|>50\ln n$, обозначим через I_S индикатор события «нет других подграфов, изоморфных $G|_S$». Положим $X=\sum I_S$. Тогда $\mathbf{E}[I_S]=1-o(1)$. Отсюда, используя линейность математического ожидания и то, что существует $2^n(1-o(1))$ таких S, получаем

$$\mathbf{E}[X]=2^n(1-o(1)).$$

Следовательно, существует некоторый граф G, такой, что $X>2^n(1-o(1))$.

11

Сложность схем

Не знание, а процесс изучения, не обладание, а процесс постижения приносит наибольшее наслаждение. Когда я изучил и исчерпал предмет, я оставляю его, чтобы снова идти в темноту. Так странен человек, вечно не удовлетворенный, — если он завершает постройку, то не для того, чтобы жить там в мире, а для того, чтобы начинать новое строительство. Я полагаю, что так должен себя чувствовать завоеватель мира, который, едва завоевав одно королевство, тянет руки к другому.

Карл Фридрих Гаусс

11.1. ПРЕДВАРИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

Булева функция $f = f(x_1, \ldots, x_n)$ от n переменных $x_1, x_2, \ldots, x_n$ — это отображение вида $f : \{0,1\}^n \to \{0,1\}$. В частности, $0, 1, x_1 \wedge \ldots \wedge x_n, x_1 \vee \ldots \vee x_n, x_1 \oplus \ldots \oplus x_n$ обозначают, как обычно, две постоянные функции, функцию *И* (ее значение равно единице тогда и только тогда, когда $x_i = 1$ для всех i), функцию *ИЛИ* (ее значение равно 0 тогда и только тогда, когда $x_i = 0$ для всех i), и *счетчик четности* (ее значение равно 0 тогда и только тогда, когда число переменных x_i, равных 1, четно), соответственно. Через $\overline{f} = f \oplus 1$ обозначается *не f*, т. е. *отрицание* функции f. Функции x_i и $\overline{x}_i$ называются *атомами*. В этом разделе мы рассматриваем проблему эффективного вычисления различных булевых функций. *Схема* — ориентированный, ациклический граф, с особой вершиной без исходящих ребер, называемой *выходом*. Каждая вершина помечена символом одной из базисных булевых функций, зависящей от ее непосредственных предков, а вершины без предков (т. е. вершин без входящих ребер), называемые *входами схемы*, помечены либо символом одной из переменных x_i, либо одной из констант 0 или 1. Для всякого двоичного набора значений входных переменных x_i можно рекурсивно вычислить соответствующее значение функции, реализуемой в произвольной вершине схемы, вычисляя соответствующую базисную функцию, символом которой помечена эта вершина, от уже вычисленных значений ее непосредственных предков. Мы говорим, что схема *вычисляет* функцию $f = f(x_1, \ldots, x_n)$, если при всех $x_i \in \{0,1\}$ соответствующее значение выхода схемы равно $f(x_1, \ldots, x_n)$. Например, на рис. 11.1 показана схема, вычисляющая функцию $f(x_1, x_2, x_3) = (x_1 \oplus (x_2 \wedge x_3)) \wedge x_1$.

Если из каждой вершины в схеме исходит не больше одной дуги (т. е. если соответствующий граф является деревом), то схема называется *формулой*. Если в каждую вершину схемы заходит не больше двух дуг, то схема

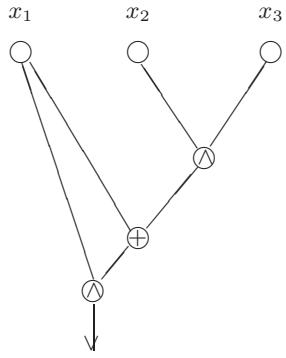


Рис. 11.1

называется[1] *2-схемой*. Таким образом, схема на рис. 11.1 является булевой схемой, но не является формулой. *Сложность* схемы — это число ее вершин (не являющихся входами), а ее *глубина* — это максимальная длина направленного пути в ней. *Сложность* булевой функции *в классе 2-схем* — это размер наименьшей вычисляющей ее 2-схемы. Простой подсчет показывает, что для больших n сложность почти всех функций от n переменных в классе 2-схем не меньше $(1 + o(1))2^n/n$. Это объясняется тем, что число 2-схем сложности s, как легко может быть доказано, меньше $(c_1 s)^s$, в то время как общее число булевых функций, зависящих от n переменных, равно 2^{2^n}. С другой стороны, не известно нелинейных, а тем более — экспоненциальных (по n) нижних оценок сложности для какой-либо «явной» функции в классе 2-схем. Здесь под «явной» мы понимаем **NP**-функцию, т. е. одну из семейства $\{f_{n_i}\}_{i\geqslant 1}$ булевых функций, где f_{n_i} — функция n_i переменных, $n_i \to \infty$, и существует недетерминированная машина Тьюринга, которая по n_i и $x_1, \ldots, x_{n_i}$ может определить за (недетерминированное) полиномиальное (по n_i) время, верно ли, что $f_{n_i}(x_1, \ldots, x_{n_i}) = 1$. (Пример такого семейства — $\frac{n}{2}$-кликовая функция; здесь $n_i = \binom{i}{2}$, n_i переменных $x_1, \ldots, x_{n_i}$ соответствуют ребрам графа на i вершинах, и $f_{n_i}(x_1, \ldots, x_{n_i}) = 1$ тогда и только тогда, когда соответствующий граф содержит клику на не менее чем $i/2$ вершинах.) Любая неполиномиальная нижняя оценка сложности явной функции в классе 2-схем означала бы (помимо прочего), что $\mathrm{P} \neq \mathrm{NP}$, и таким образом решалась бы возможно важнейшая открытая проблема теоретической информатики. К сожалению, наилучшая известная на данный момент нижняя оценка сложности явной функции от n переменных в классе 2-схем — всего лишь $3n$ (см. [Blum (1984)], [Paul (1977)]). Тем не менее, известно несколько нетривиальных нижних оценок для случая, когда мы накладываем определенные ограничения на структуру схем. Большинство известных доказательств этих оценок в значитель-

[1] Авторы используют словосочетание «binary circuit», что в буквальном переводе означает «двоичная схема». Это не соответствует смыслу определения. Для краткости мы используем здесь термин «2-схема». — *Прим. ред.*

ной степени опирается на вероятностные методы. В этой главе мы опишем некоторые из этих результатов. Заметим, что известно множество других красивых результатов о сложности схем; см., например, [Wegener (1987)] и [Karchmer and Wigderson (1990)], однако те, что приведены здесь, не только входят в число важнейших, но и позволяют продемонстрировать элегантные методы, используемые в данной области. Так как большинство результатов этой главы — асимптотические, мы считаем, когда это требуется, что число переменных достаточно велико.

11.2. СЛУЧАЙНЫЕ ОГРАНИЧЕНИЯ И СХЕМЫ ОГРАНИЧЕННОЙ ГЛУБИНЫ

Буквой назовем формулу вида x_i или $\overline{x}_i$, где x_i — символ переменной. Формулы вида $y_1 \wedge \ldots \wedge y_n$ и $y_1 \vee \ldots \vee y_n$, где y_i — буквы, называются соответственно *элементарной конъюнкцией* и *элементарной дизъюнкцией*. Число букв в элементарной конъюнкции (элементарной дизъюнкции) называется ее *рангом*. Булеву формулу G назовем *t-конъюнктивной нормальной формой* (сокращенно, t-КНФ), если она представляет собой конъюнкцию произвольного числа элементарных дизъюнкций ранга, не большего t, т. е. $G = G_1 \wedge \ldots \wedge G_w$, где $G_i = y_{i1} \vee \ldots \vee y_{ia_i}, a_i \leqslant t$ и каждое y_j является буквой. Аналогично, назовем булеву формулу *s-дизъюнктивной нормальной формой* (сокращенно, s-ДНФ), если она может быть записана как дизъюнкция конъюнкций, каждая из которых содержит не более s букв. Импликанта функции f — это элементарная конъюнкция K, удовлетворяющая условию $K \bigvee f = f$. Импликанта K функции f называется *простой импликантой* или *минтермом*, если после удаления любой буквы она перестает быть импликантой функции f. Заметим, что функция является s-ДНФ тогда и только тогда, когда ранг каждого из ее минтермов не превосходит s. *Ограничение* — это отображение ρ множества индексов $\{1, \ldots, n\}$ на множество $\{0, 1, *\}$. Ограничение ρ на функцию $G = G(x_1, \ldots, x_n)$, обозначаемое $G|\rho$, — это булева функция, полученная из G присваиванием каждой переменной x_i значения $\rho(i)$ при условии, что $i \in \rho^{-1}\{0, 1\}$. Каждая x_j при $j \in \rho^{-1}(*)$ остается переменной. Тогда, например, если $G(x_1, x_2, x_3) = (x_1 \wedge x_2) \vee x_3$ и $\rho(1) = 0$, $\rho(2) = \rho(3) = *$, то $G|_\rho = x_3$. Для $0 \leqslant p \leqslant 1$ *случайное p-ограничение* — это случайное ограничение ρ, определенное путем независимого выбора для каждого $1 \leqslant i \leqslant n$ значения $\rho(i)$ согласно следующему распределению:

$$\Pr[\rho(i) = *] = p, \quad \Pr[\rho(i) = 0] = \Pr[\rho(i) = 1] = (1 - p)/2. \tag{11.1}$$

Улучшив результаты [Furst, Saxe and Sipser (1984)], [Ajtai (1983)] и [Yao (1985)], Хостад [Håstad (1988)] доказал следующее утверждение, очень полезное при установке нижних оценок для схем ограниченной глубины.

Лемма 11.2.1 (переключающая лемма). *Пусть $G = G(x_1, \ldots, x_n)$ — t-КНФ, т. е. $G = G_1 \wedge G_2 \wedge \ldots \wedge G_w$, где каждое G_i является дизъюнкцией*

не более чем t букв. Пусть ρ — случайное ограничение, заданное распределением (11.1).

Тогда

$$\Pr[G|_\rho \text{ не является } (s-1)\text{-ДНФ}] = \\ = \Pr[\text{у}G|_\rho \text{ есть минтерм ранга } \geqslant s] \leqslant (5pt)^s.$$

Доказательство. Пусть E_s — событие, состоящее в том, что у $G|_\rho$ есть минтерм ранга не меньше s. Чтобы оценить $\Pr[E_s]$, мы докажем более сильное утверждение: для любой булевой функции F выполняется

$$\Pr\left[E_s \,\middle|\, F|_\rho \equiv 1\right] \leqslant (5pt)^s. \tag{11.2}$$

Здесь мы считаем, что если условие не выполнено, то условная вероятность равна нулю. Утверждение леммы 11.2.1 следует из неравенства (11.2) при $F \equiv 1$. Мы доказываем (11.2) индукцией по w. При $w = 0$ имеем $G \equiv 1$, и неравенство очевидно. Предположим, что (11.2) верно, когда количество G_i меньше w, и докажем это для w. Положим $G = G_1 \wedge G^*$, где $G^* = G_2 \wedge \ldots \wedge G_w$, и пусть E_s^* — событие, состоящее в том, что у $G^*|_\rho$ есть минтерм ранга не меньше s. Заменяя при необходимости некоторые переменные их отрицаниями, мы можем для удобства считать, что $G_1 = \bigvee\limits_{i \in T} x_i$, где $|T| \leqslant t$. Либо $G_1|_\rho \equiv 1$, либо $G_1|_\rho \not\equiv 1$. В первом случае E_s выполняется тогда и только тогда, когда выполняется E_s^*, а значит, по индукции

$$\Pr\left[E_s \middle| F|_\rho \equiv 1, G_1|_\rho \equiv 1\right] = \Pr\left[E_s^* \middle| (F \wedge G_1)|_\rho \equiv 1\right] \leqslant (5pt)^s. \tag{11.3}$$

Случай $G_1|_\rho \not\equiv 1$ требует большего внимания. В этом случае любой минтерм $G|_\rho$ должен содержать хотя бы одну переменную x_i, где $i \in T$, без отрицания. Для не пустого $Y \subseteq T$ и для функции $\sigma : Y \to \{0,1\}$, не равной тождественно нулю, пусть $E_s(Y,\sigma)$ — событие, состоящее в том, что у $G|_\rho$ есть минтерм ранга не меньше s, который содержит букву $x_i^{\sigma(i)}$ для каждого $i \in Y$, и не содержит отрицаний переменных x_j, где $j \in T$. По предыдущему замечанию

$$\Pr\left[E_s \middle| F|_\rho \equiv 1, G_1|_\rho \not\equiv 1\right] \leqslant \sum_{Y,\sigma} \Pr\left[E_s(Y,\sigma) \middle| F|_\rho \equiv 1, G_1|_\rho \not\equiv 1\right]. \tag{11.4}$$

Заметим, что условие $G_1|_\rho \not\equiv 1$ в точности означает, что $\rho(i) \in \{0, *\}$ для всех $i \in T$, а значит, для каждого $i \in T$ выполнено соотношение

$$\Pr\left[\rho(i) = * \middle| G_1|_\rho \not\equiv 1\right] = \frac{p}{p + (1-p)/2} = 2p/(1+p).$$

Следовательно, если $|Y| = y$, то

$$\Pr\left[\rho(Y) = * \middle| G_1|_\rho \not\equiv 1\right] \leqslant \left(\frac{2p}{1+p}\right)^y.$$

Дополнительное условие $F|_\rho \equiv 1$ может только уменьшить эту вероятность. Это может быть показано с помощью FKG-неравенства (см. гл. 6). Это также можно показать напрямую следующим образом. Для любого фиксированного

$\rho' : N \setminus Y \to \{0, 1, *\}$, где $N = \{1, \ldots, n\}$, мы утверждаем, что

$$\Pr\left[\rho(Y) = * \middle| F|_\rho \equiv 1, G_1|_\rho \not\equiv 1, \rho|_{N\setminus Y} = \rho'\right] \leqslant \left(\frac{2p}{1+p}\right)^y.$$

В самом деле, у данного ρ' есть единственное расширение ρ с $\rho(Y) = *$. Если ρ не удовлетворяет вышеупомянутым условиям, то условная вероятность равна нулю. Если удовлетворяет, то этим условиям удовлетворяют также все расширения ρ с $\rho(i) \in \{0, *\}$ при $i \in Y$, так что неравенство выполняется и в этом случае. Так как это верно для всех фиксированных ρ', мы приходим к выводу, что действительно

$$\Pr\left[\rho(Y) = * \middle| F|_\rho \equiv 1, G_1|_\rho \not\equiv 1\right] \leqslant \left(\frac{2p}{1+p}\right)^y \leqslant (2p)^y. \tag{11.5}$$

Пусть при $\rho' : T \to \{0, *\}$ выполняется $\rho(Y) = *$. Рассмотрим все возможные ограничения ρ, для которых $\rho|_T = \rho'$. При этом условии ρ можно рассматривать как случайное ограничение на $N \setminus T$. Событие $F|_\rho \equiv 1$ сводится к событию $\tilde{F}|_{\rho|_{N\setminus T}} \equiv 1$, где $\tilde{F}$ — конъюнкция всех функций, полученных из F заменой значений x_i согласно ρ' для тех $i \in T$, у которых $\rho'(i) = 0$, и перебором всех возможных значений для остальных переменных x_j, где $j \in T$. Если происходит событие $E_s(Y, \sigma)$, то у $G^*|_{\rho\sigma}$ есть минтерм ранга не меньше $s - y$, не содержащий никакой переменной x_i с $i \in T \setminus Y$. Но это происходит тогда и только тогда, когда у $\tilde{G}|_{\rho|_{N\setminus T}}$ есть минтерм ранга не меньше $s - y$, где $\tilde{G}$ — это функция, полученная из G^* заменой значений x_j при $j \in Y$ согласно σ, значений x_i при $i \in T \setminus Y$ и $\rho'(i) = 0$ согласно ρ', и выбрасыванием всех переменных x_k с $k \in T \setminus Y$ и $\rho'(k) = *$. Обозначив это событие $\tilde{E}_{s-y}$, мы можем применить индукцию и получить, что

$$\Pr\left[E_s(Y, \sigma) \middle| F|_\rho \equiv 1, G_1|_\rho \not\equiv 1, \rho|_T = \rho'\right] \leqslant \Pr\left[\tilde{E}_{s-y} \middle| \tilde{F}|_\rho \equiv 1\right] \leqslant (5pt)^{s-y}.$$

Так как любое ρ с $F|_\rho \equiv 1, G_1|_\rho \equiv 1, \rho(Y) = *$ должно удовлетворять $\rho|_T = \rho'$ для некоторого ρ' этого вида, и поскольку событие $E_s(Y, \sigma)$ может произойти только когда $\rho(Y) = *$, мы приходим к выводу, что

$$\Pr\left[E_s(Y, \sigma) \middle| F|_\rho \equiv 1, G_1|_\rho \not\equiv 1, \rho(Y) = *\right] \leqslant (5pt)^{s-y},$$

и согласно неравенству (11.5)

$$\begin{aligned}\Pr\left[E_s(Y, \sigma) \middle| F|_\rho \equiv 1, G_1|_\rho \not\equiv 1\right] &= \Pr\left[\rho(Y) = * \middle| F|_\rho \equiv 1, G_1|_\rho \not\equiv 1\right] \times \\ &\times \Pr\left[E_s(Y, \sigma) \middle| F|_\rho \equiv 1, G_1|_\rho \not\equiv 1, \rho(Y) = *\right] \leqslant (2p)^y (5pt)^{s-y}.\end{aligned}$$

Подставляя это в неравенство (11.4) и используя тот факт, что $|T| \leqslant t$, а также что

$$\sum_{y=1}^{t} (2^y - 1) 2^y / (5^y y!) \leqslant \frac{2}{5} + \sum_{y=2}^{\infty} \frac{(4/5)^y}{y!} = \frac{2}{5} + e^{4/5} - 1 - \frac{4}{5} < 1,$$

получаем:

$$\Pr\left[E_s\middle|F|_\rho \equiv 1, G_1|_\rho \not\equiv 1\right] \leqslant$$

$$\leqslant \sum_{y=1}^{|T|}\binom{|T|}{y}(2^y-1)(2p)^y(5pt)^{s-y} \leqslant (5pt)^s\sum_{y=1}^{t}\frac{t^y}{y!}(2^y-1)\left(\frac{2}{5t}\right)^y =$$

$$= (5pt)^s\sum_{y=1}^{t}(2^y-1)\cdot\frac{2^y}{5^y\cdot y!} \leqslant (5pt)^s.$$

Это вместе с соотношением (11.3) дает неравенство

$$\Pr\left[E_s\middle|F|_\rho \equiv 1\right] \leqslant (5pt)^s,$$

завершая индукцию и доказательство. ■

Подставляя отрицание функции G в лемму 11.2.1 и применяя правила де Моргана, мы легко получим ее двойственную форму. Если G — t-ДНФ и ρ — случайное ограничение, заданное распределением (11.1), то

$$\Pr[G|_\rho \text{ не является } (s-1)\text{-КНФ}] \leqslant (5pt)^s.$$

Теперь мы опишем применение переключающей леммы, дающее нижнюю оценку сложности схем минимальной глубины, реализующих счетчик четности $x_1 \oplus \ldots \oplus x_n$. Мы рассматриваем схемы, в которых вершины упорядочены по уровням, вершины первого уровня — буквы (т. е. переменные или их отрицания), и все остальные элементы — либо ИЛИ, либо И произвольного числа вершин предыдущего уровня. Мы предполагаем, что элементы каждого уровня либо все являются элементами типа И, либо все являются элементами типа ИЛИ, и что уровни меняются всякий раз с типа И на тип ИЛИ и наоборот. Схема такого вида называется $C(s,s',d,t)$-схемой, если она содержит не более s элементов, не более s' из которых выше второго уровня, ее глубина не более d, и число входов каждого элемента второго уровня не более t. Таким образом, например, схема, реализующая счетчик четности путем вычисления дизъюнкции от 2^{n-1} конъюнкций вида $x_1^{\varepsilon_1}\wedge\ldots\wedge x_n^{\varepsilon_n}$, где $(\varepsilon_1,\ldots,\varepsilon_n)$ — всевозможные четные двоичные векторы, и $x_i^{\varepsilon_i} = x_i \oplus \varepsilon_i$, является $C(2^{n-1}+1,1,2,n)$-схемой.

Теорема 11.2.2. *Пусть $f = f(x_1,\ldots,x_n)$ — функция, и пусть C — $C(\infty,s,d,t)$-схема, вычисляющая f, где $s\cdot(\frac{1}{2})^t \leqslant 0.5$. Тогда либо у f, либо у ее дополнения $\overline{f}$ есть минтерм ранга не больше $n - \frac{n}{2\cdot(10t)^{d-2}} + t$.*

Доказательство. Применим к C последовательно $d-2$ раз случайное $1/(10t)$-ограничение. Каждое из этих случайных ограничений, будучи примененным к любой нижней подсхеме глубины 2, преобразует ее по лемме 11.2.1 с вероятностью не меньше $1-(\frac{1}{2})^t$ из t-ДНФ в t-КНФ (или наоборот). Если все эти преобразования успешны, мы можем отождествить новые элементы И с элементами более высокого уровня и получить схему меньшей глубины. Так как общая сложность схемы не превосходит s, и $s(\frac{1}{2})^t \leqslant 0.5$,

мы делаем вывод, что с вероятностью не меньше 0.5 все преобразования успешны, и C преобразуется в $C(\infty, 1, 2, t)$-схему. Каждая переменная x_i независимо остается переменной (т. е. ей не было присвоено значение) с вероятностью $\frac{1}{(10t)^{d-2}}$. Следовательно, количество оставшихся переменных — это биномиальная случайная величина с математическим ожиданием $\frac{n}{(10t)^{d-2}}$ и немного меньшей дисперсией. По стандартным оценкам для биномиального распределения(см. приложение А) вероятность того, что не менее $\frac{n}{2\cdot(10t)^{d-2}}$ переменных остались переменными, больше 0.5. Следовательно, с положительной вероятностью не более $n - \frac{n}{2\cdot(10t)^{d-2}}$ переменных зафиксированы, и у получившегося ограничения f есть $C(\infty, 1, 2, t)$-схема, т. е. ее значение может быть зафиксировано присваиванием значений не более чем t отрицаниям переменных. Это завершает доказательство. ■

Следствие 11.2.3. *Для любого $d \geqslant 2$ не существует*

$$C\left(\infty, \frac{1}{2}\cdot 2^{\frac{1}{10}n^{1/(d-1)}}, d, \frac{1}{10}n^{1/(d-1)}\right)\text{-схемы,}$$

реализующей счетчик четности $f(x_1, \ldots, x_n) = x_1 \oplus \ldots \oplus x_n$.

Доказательство. Предполагая, что такая схема существует, получаем по теореме 11.2.2, что значение f может быть зафиксировано присваиванием значений не более чем $n - \frac{1}{2}n^{1/(d-1)} + \frac{1}{10}n^{1/(d-1)} < n$ переменным. Это неверно, а значит, такой схемы не существует. ■

Оценка в следствии 11.2.3, фактически, близка к наилучшей. Так как каждая $C(s, s', d, t)$-схема может быть преобразована в $C(ts, s, d+1, 2)$-схему (заменой каждого атома на ИЛИ или И двух его копий), из следствия 11.2.3 легко получить, что глубина d любой $C(s, s', d, t)$-схемы полиномиальной сложности, реализующей счетчик четности от n переменных, не меньше $\Omega(\log n/\log\log n)$. Эта нижняя оценка также является оптимальной.

11.3. ЕЩЕ О СХЕМАХ ОГРАНИЧЕННОЙ ГЛУБИНЫ

В предыдущем разделе мы видели, что счетчик четности сложно реализовать с малой глубиной, используя элементы И, ИЛИ и НЕ. Оказывается, что даже если допустить использование элементов-счетчиков четности (вместе с элементами И, ИЛИ и НЕ), все же существуют некоторые относительно простые функции, которые сложно вычислить. Такой результат был впервые доказан Разборовым (1987). Его метод был изменен и усилен Смоленским [Smolensky (1987)]. Для целого $k \geqslant 2$ пусть $\mathrm{Mod}_k(x_1, x_2, \ldots, x_n)$ — булева функция, принимающая значение 1 тогда и только тогда, когда $\Sigma x_i \not\equiv 0 \pmod{k}$. Смоленский показал, что для любых двух степеней p и q различных простых чисел функция Mod_p не может быть вычислена ограниченной по глубине схемой полиномиальной сложности, состоящей из элементов И, ИЛИ, НЕ и Mod_q. Здесь мы приводим частный случай этого результата, в котором $q = 3$ и $p = 2$.

Пусть C — произвольная схема глубины d и сложности s, состоящая из элементов И, ИЛИ, НЕ и Mod_3. Важный факт, доказанный Разборовым, заключается в том, что функция, которую реализует C, может быть довольно хорошо приближена (в зависимости от d и s) полиномом относительно малой степени над полем $GF(3)$. Это доказывается с помощью следующего применения вероятностного метода. Каждый элемент схемы C заменим приближающей полиномиальной операцией, согласно следующим правилам, гарантирующим, что в каждой вершине новой схемы мы вычисляем полином над полем $GF(3)$, принимающий только значения 0 и 1 (если на входы подаются только 0 и 1).

(i) Каждый элемент-отрицание $\overline{y}$ заменяется полиномиальным элементом $(1-y)$.

(ii) Каждый Mod_3-элемент $\mathrm{Mod}_3(y_1,\ldots,y_m)$ заменяется полиномиальным элементом $(y_1+y_2+\ldots+y_m)^2$.

Правило для замены элементов ИЛИ и И немного сложнее. Заметим, что в двух предыдущих случаях (i) и (ii) не было приближенных вычислений; новые элементы вычисляют в точности то же, что и старые, для всех возможных булевых значений переменных. В принципе, это может быть сделано и здесь. Элемент И $y_1\wedge\ldots\wedge y_m$ следует просто заменить произведением $y_1\ldots y_m$. Тогда элемент ИЛИ $y_i\vee\ldots\vee y_m$ может быть вычислен по правилам де Моргана. Так как $y_1\vee\ldots\vee y_m=\overline{(\overline{y}_1\wedge\ldots\wedge\overline{y}_m)}$ и $\overline{y}$ реализуется как $(1-y)$, это дает представление

$$1-(1-y_1)(1-y_2)\ldots(1-y_m). \tag{11.6}$$

Проблема в том, что эта процедура слишком сильно увеличила бы степень наших полиномов. Значит, нам нужно поступить немного хитрее. Пусть ℓ — целое число, которое мы определим позже. По элементу ИЛИ $y_1\vee\ldots\vee y_m$ мы выбираем ℓ случайных подмножеств $I_1,\ldots,I_\ell$ множества $\{1,\ldots,m\}$, где для каждого $1\leqslant i\leqslant\ell$ и для каждого $1\leqslant j\leqslant m$ независимо $\Pr[j\in I_i]=1/2$. Заметим, что для любого фиксированного i, $1\leqslant i\leqslant\ell$, сумма $\Big(\sum\limits_{j\in I_i}y_j\Big)^2$ над полем $GF(3)$ обязательно равна 0, если $y_1\vee\ldots\vee y_m=0$, и равна 1 с вероятностью не меньше $\frac{1}{2}$, если $y_1\vee\ldots\vee y_m=1$. Таким образом, если мы вычислим значение ИЛИ от ℓ выражений $\Big(\sum\limits_{j\in I_i}y_j\Big)^2, 1\leqslant i\leqslant\ell$, то эта функция равна 0, если $y_1\vee\ldots\vee y_m=0$, и равна 1 с вероятностью не меньше $1-(1/2)^\ell$, если $y_1\vee\ldots\vee y_m=1$. Тогда мы вычислим ИЛИ и запишем результат в виде полинома, как это было сделано в формуле (11.6). Это дает

$$1-\prod_{i=1}^{\ell}\left(1-\left(\sum_{j\in I_i}y_j\right)^2\right). \tag{11.7}$$

Следовательно, в нашей новой схеме мы заменим каждый элемент ИЛИ на приближающий полиномиальный элемент вида (11.7). Получив приближение для элемента ИЛИ, мы можем получить соответствующее приближение и для элемента И, применяя правила де Моргана. Так как

$y_1 \wedge \ldots \wedge y_m = \overline{(\overline{y}_1 \vee \ldots \vee \overline{y}_m)}$, мы заменяем каждый элемент И вида $y_1 \wedge \ldots \wedge y_m$ на

$$\prod_{i=1}^{\ell} \left(1 - \left[\sum_{j \in I_i} (1 - y_j) \right]^2 \right). \tag{11.8}$$

Заметим, что степени полиномов (11.7) и (11.8) не превосходят 2ℓ.

В исходной схеме C глубины d и сложности s мы теперь можем заменить все ее элементы нашими приближающими полиномиальными элементами и получить новую схему CP, которая зависит от всех случайных выборов, сделанных при каждой замене каждого элемента И/ИЛИ. Новая схема CP вычисляет полином $P(x_1, \ldots, x_n)$ степени не больше $(2\ell)^d$. Более того, для любых фиксированных булевых значений $x_1, x_2, \ldots, x_n$ вероятность того, что все новые элементы вычисляют то же, что и соответствующие элементы в C, не меньше $1 - s/2^\ell$. Следовательно, ожидаемое число входных значений, на которых $P(x_1, \ldots, x_n)$ совпадает с выходными значениями C, не меньше $2^n(1 - s/2^\ell)$. Таким образом, мы доказали следующее утверждение.

Лемма 11.3.1. *Для любой схемы C глубины d и сложности s от n булевых переменных, использующей элементы НЕ, ИЛИ, И и* Mod_3*, и для любого целого ℓ существует полином $P = P(x_1, \ldots, x_n)$ степени не больше $(2\ell)^d$ над полем $GF(3)$, значение которого совпадает с функцией, реализуемой схемой C, не менее чем на $2^n(1 - s/2^\ell)$ входных наборах.*

Чтобы применить эту лемму для получения нижних оценок сложности любой схемы описанного выше типа, реализующей счетчик четности, нам потребуется следующий дополнительный комбинаторный результат.

Лемма 11.3.2. *Не существует полинома $P(x_1, \ldots, x_n)$ над полем $GF(3)$ степени не больше $\sqrt{n}$, равного счетчику четности от $x_1, \ldots, x_n$ на множестве S из не менее чем $0.9 \cdot 2^n$ различных булевых векторов $(x_1, \ldots, x_n)$.*

Доказательство. Предположим, что это не так, и что $S \subset \{0,1\}^n$, $|S| \geqslant 0.9 \cdot 2^n$ и $P(x_1, \ldots, x_n) = x_1 \oplus \cdots \oplus x_n$ для всех $(x_1, \ldots, x_n) \in S$. Определим полином $Q = Q(y_1, \ldots, y_n)$ как $Q = Q(y_1, \ldots, y_n) = P(y_1 + 2, \ldots, y_n + 2) - 2$ и множество $T = \{(y_1, \ldots, y_n) \in \{1, -1\}^n : (y_1 + 2, \ldots, y_n + 2) \in S\}$, где все суммы берутся по модулю 3. Ясно, что степень многочлена Q не превосходит $\sqrt{n}$, и $Q(y_1, , \ldots, y_n) = \prod_{i=1}^{n} y_i$ для всех $(y_1, \ldots, y_n) \in T$. Пусть теперь $G = G(y_1, \ldots, y_n) : T \to GF(3)$ — произвольная функция. Расширим ее произвольным образом до функции из $(GF(3))^n \to GF(3)$ и запишем эту функцию как полином от n переменных. Очевидно, что любая функция из $(GF(3))^n \to GF(3)$ является полиномом, так как может быть представлена в виде линейной комбинации функций вида $\prod_{i=1}^{n}(y_i - \varepsilon_i)(y_i - \varepsilon_i - 1)$, где $\varepsilon_i \in GF(3)$. Заменим каждое вхождение y_i^2 в этом полиноме на 1, чтобы

получить полилинейный полином $\tilde{G}$, совпадающий с G на T. Теперь заменим каждый одночлен $\prod_{i\in U} y_i$, где $|U| > \frac{n}{2} + \frac{\sqrt{n}}{2}$, на $\prod_{i\notin U} y_i \cdot Q(y_1,\ldots,y_n)$, и этот новый полином заменим на полилинейный полином $\tilde{\tilde{G}}$, снова заменяя каждое y_i^2 на 1. Так как для $y_i \in \{\pm 1\}$, $\prod_{i\notin U} y_i \cdot \prod_{i=1}^{n} y_i = \prod_{i\in U} y_i$, $\tilde{\tilde{G}}$ равно G на T, и его степень не превосходит $\frac{n}{2} + \frac{\sqrt{n}}{2}$. Тем не менее, число возможных $\tilde{\tilde{G}}$ равно $3^{\sum_{i=0}^{\lfloor \frac{n}{2}+\frac{\sqrt{n}}{2}\rfloor}\binom{n}{i}} < 3^{0.88\cdot 2^n}$, в то время как число возможных G равно $3^{|T|} \geqslant 3^{0.9\cdot 2^n}$. Это невозможно, а значит верно утверждение леммы. ■

Следствие 11.3.3. *Не существует схемы глубины d и сложности $s \leqslant \frac{1}{10} 2^{\frac{1}{2}} n^{1/2d}$, реализующей счетчик четности от $x_1, x_2, \ldots, x_n$, состоящей из элементов НЕ, И, ИЛИ и* Mod_3.

Доказательство. Предположим, что это не так, и пусть C — такая схема. Положим $\ell = \frac{1}{2}\cdot n^{1/2d}$. По лемме 11.3.1 существует полином $P = P(x_1,\ldots,x_n)$ над полем $GF(3)$, степень которого не превосходит $(2\ell)^d = \sqrt{n}$, равный счетчику четности от $x_1,\ldots,x_n$ не менее чем на $2^n(1 - \frac{s}{2^{\frac{1}{2}} n^{1/2d}}) \geqslant 0.9\cdot 2^n$ входных значениях. Это противоречит лемме 11.3.2, что завершает доказательство. ■

11.4. МОНОТОННЫЕ СХЕМЫ

Булева функция $f = f(x_1,\ldots,x_n)$ называется *монотонной*, если из условий $f(x_1,\ldots,x_n) = 1$ и $x_i \leqslant y_i$ следует, что $f(y_1,\ldots,y_n) = 1$. *Монотонная 2-схема* — это 2-схема, содержащая только 2-входовые элементы И и ИЛИ. Нетрудно видеть, что функция монотонна тогда и только тогда, когда существует вычисляющая ее монотонная 2-схема. *Сложность* монотонной функции *в классе монотонных схем* — это минимальная сложность вычисляющей ее монотонной 2-схемы. До 1985 г. лучшей известной нижней оценкой для сложности монотонной **NP**-функции от n переменных было $4n$. Существенное улучшение было получено в фундаментальной работе [Разборов (1985)], в которой установлена оценка $n^{\Omega(\log n)}$ для сложности характеристической функции множества клик размера k (равной 1 тогда и только тогда, когда граф содержит клику размера k). Вскоре после этого Андреев (1985), используя аналогичные методы, получил экспоненциальную нижнюю оценку для несколько искусственной **NP**-функции. Алон и Боппана [Alon and Boppana (1987)] усилили комбинаторные соображения Разборова и доказали экспоненциальную нижнюю оценку для сложности кликовой функции в классе монотонных схем. В этом разделе мы опишем особый случай этой оценки, показав, что не существует монотонных схем линейной сложности, определяющих, содержит ли данный граф треугольник. Хотя этот результат гораздо слабее приведенных выше, он хорошо иллюстрирует все вероятностные соображения более сложных доказательств и обходит некоторые комбинаторные тонкости.

Положим $n = \binom{m}{2}$, и пусть $x_1, x_2, \ldots, x_n$ — n булевых переменных, представляющих ребра графа на множестве вершин $\{1, 2, \ldots, m\}$. Пусть $T = T(x_1, \ldots, x_n)$ — монотонная булева функция, чье значение равно 1, если соответствующий граф содержит треугольник. Ясно, что существует монотонная 2-схема сложности $O(m^3)$, вычисляющая функцию T. Поэтому следующая теорема неулучшаема с точностью до константы.

Теорема 11.4.1. *Сложность монотонной схемы для T не меньше $\Omega(m^3/\log^4 m)$.*

Перед тем как приводить доказательство этой теоремы, введем некоторые обозначения и докажем простую лемму. Для любой булевой функции $f = f(x_1, \ldots, x_n)$ положим $A(f) = \{(x_1, \ldots, x_n) \in \{0,1\}^n : f(x_1, \ldots, x_n) = 1\}$. Ясно, что $A(f \vee g) = A(f) \cup A(g)$ и $A(f \wedge g) = A(f) \cap A(g)$. Пусть C — монотонная схема сложности s, вычисляющая функцию $f = f(x_1, \ldots, x_n)$. Ясно, что C дает монотонную прямолинейную программу длины s, вычисляющую f, т. е. последовательность функций $x_1, x_2, \ldots, x_n, f_1, \ldots, f_s$, где $f_s = f$, и каждая f_i, где $1 \leqslant i \leqslant s$, либо ИЛИ, либо И какой-то пары предшествующих функций. Применяя операцию A, получим последовательность $A(C)$ подмножеств $(0,1)^n$: $A_{-n} = A_{x_n}, \ldots, A_{-1} = A_{x_1}, A_1, \ldots, A_s$, где $A_{x_i} = A(x_i), A_s = A(f)$, и каждое A_i, где $1 \leqslant i \leqslant s$, есть либо объединение, либо пересечение какой-то пары предшествующих подмножеств. Заменим последовательность $A(C)$ на *приближающую последовательность* $M(C) : M_{-n} = M_{x_n} = A_{x_n}, \ldots, M_{-1} = M_{x_1} = A_{x_1}, M_1, \ldots, M_s$, определяемую заменой операций объединения и пересечения в $A(C)$ на приближающие операции $\sqcup$ и $\sqcap$ соответственно. Точное определение этих двух операций будет дано позже, причем для всех допустимых M и L будут выполняться включения

$$M \sqcup L \supseteq M \cup L, \quad M \sqcap L \subseteq M \cap L. \tag{11.9}$$

Тогда $M_{x_i} = A_{x_i}$ для всех $1 \leqslant i \leqslant n$, и если для некоторых $1 \leqslant j \leqslant s$ верно $A_j = A_\ell \cup A_k$, то $M_j = M_\ell \sqcup M_k$, в то время как если $A_j = A_\ell \cap A_k$, то $M_j = M_\ell \sqcap M_k$. В первом случае положим $\delta^j_\sqcup = M_j \setminus (M_\ell \cup M_k)$ и $\delta^j_\sqcap = \varnothing$, а во втором случае положим $\delta^j_\sqcap = (M_\ell \cap M_k) \setminus M_j$ и $\delta^j_\sqcup = \varnothing$.

Лемма 11.4.2. *Для всех M_i, членов последовательности $M(C)$, имеют место включения*

$$A_i \setminus \left(\bigcup_{j \leqslant i} \delta^j_\sqcap \right) \subseteq M_i \subseteq A_i \cup \bigcup_{j \leqslant i} \delta^j_\sqcup. \tag{11.10}$$

Доказательство. Применим индукцию по i. При $i < 0$ имеем $M_i = A_i$, а значит формула (11.10) верна. Предположим, что вложение (11.10) выполнено для всех M_j при $j < i$, и докажем это для i. Если $A_i = A_\ell \cup A_k$, то по предположению индукции

$$M_i = M_\ell \cup M_k \cup \delta^i_\sqcup \subseteq A_\ell \cup A_k \cup \bigcup_{j \leqslant i} \delta^j_\sqcup = A_i \cup \bigcup_{j \leqslant i} \delta^j_\sqcup$$

и

$$M_i = M_\ell \sqcup M_k \supseteq M_\ell \cup M_k \supseteq \left(A_\ell \setminus (\bigcup_{j\leqslant \ell} \delta_\sqcap^j)\right) \cup \left(A_k \setminus (\bigcup_{j\leqslant k} \delta_\sqcap^j)\right) \supseteq$$

$$\supseteq A_i \setminus (\bigcup_{j\leqslant i} \delta_\sqcap^j), \tag{11.11}$$

что и требовалось доказать. Случай $A_i = A_\ell \cap A_k$ доказывается аналогично. ■

Лемма 11.4.2 верна при любом выборе операций $\sqcup$ и $\sqcap$, удовлетворяющих соотношениям (11.9). Чтобы доказать теорему 11.4.1, определим эти операции следующим образом. Положим $r = 100\log_2^2 m$. Для любого множества R, содержащего не более r ребер на $V = \{1, 2, \ldots, m\}$, пусть $\lceil R\rceil$ обозначает множество всех графов на V, содержащих по крайней мере одно ребро из R. В частности, $\lceil\varnothing\rceil$ — это пустое множество. Пусть $\lceil *\rceil$ обозначает множество всех графов. Все элементы $M(C)$ будут иметь вид $\lceil R\rceil$ или $\lceil *\rceil$. Заметим, что $A_{x_i} = M_{x_i}$ — это просто множество $\lceil R\rceil$, где R — это множество, состоящее из единственного соответствующего ребра. Для двух множеств R_1 и R_2, каждое из которых содержит не более r ребер, положим $\lceil R_1\rceil \sqcap \lceil R_2\rceil = \lceil R_1 \cap R_2\rceil$, $\lceil R_1\rceil \sqcap \lceil *\rceil = \lceil R_1\rceil$ и $\lceil *\rceil \sqcap \lceil *\rceil = \lceil *\rceil$. Аналогично, если $|R_1 \cup R_2| \leqslant r$, мы определим $\lceil R_1\rceil \sqcup \lceil R_2\rceil = \lceil R_1 \cup R_2\rceil$, в то время как если $|R_1 \cup R_2| > r$, то $\lceil R_1\rceil \sqcup \lceil R_2\rceil = \lceil *\rceil$. Наконец $\lceil *\rceil \sqcup \lceil R_1\rceil = \lceil *\rceil \sqcup \lceil *\rceil = \lceil *\rceil$.

Доказательство теоремы 11.4.1. Теперь мы докажем теорему 11.4.1, показав, что не существует монотонной схемы сложности $s < \binom{m}{3}/2r^2$, вычисляющей функцию T. Действительно, предположим, что это не так, и пусть C — такая схема. Пусть $M(C) = M_{x_n}, \ldots, M_{x_1}, M_1, \ldots, M_s$ — приближающая последовательность длины s, полученная из C описанным выше способом. По лемме 11.4.2

$$A(T) \setminus \left(\bigcup_{j\leqslant s} \delta_\sqcap^j\right) \subseteq M_s \subseteq A(T) \cup \bigcup_{j\leqslant s} \delta_\sqcup^j. \tag{11.12}$$

Рассмотрим два возможных случая.

Случай 1: $M_s = \lceil R\rceil$, где $|R| \leqslant r$. Выберем случайный треугольник Δ на $\{1, 2, \ldots, m\}$. Ясно, что

$$\Pr(\Delta \in M_s) \leqslant r \cdot (m-2)/\tbinom{m}{3} < \frac{1}{2}.$$

Более того, для каждого фиксированного j, $j \leqslant s$,

$$\Pr(\Delta \in \delta_\sqcap^j) \leqslant r^2/\tbinom{m}{3}.$$

Это объясняется тем, что если $\delta_\sqcap^j \neq \varnothing$, то $\delta_\sqcap^j = (\lceil R_1\rceil \cap \lceil R_2\rceil) \setminus \lceil R_1 \cap R_2\rceil$ для некоторых двух множеств R_1, R_2, содержащих не более r ребер. В эту разность входят только треугольники, содержащие ребро из R_1 и ребро из R_2 (и ни одного ребра из обоих). Существует не больше r^2 таких треугольников, из чего следует последнее неравенство. Так как $s < \binom{m}{3}/2r^2$, из двух последних

неравенств получаем, что $\Pr[\Delta \notin M_s$ и $\Delta \notin \bigcup\limits_{j \leqslant s} \delta^j_{\sqcap}] > 0$, а значит, существует такой треугольник Δ. Этот треугольник принадлежит $A(T)$, что противоречит (11.12), поэтому случай 1 невозможен.

Случай 2: $M_s = \lceil * \rceil$. Пусть B — случайный остовный полный двудольный граф на множестве $V = \{1, 2, \ldots, m\}$, полученный раскрашиванием каждой вершины из V случайно и независимо в цвета 0 и 1, в котором две вершины соединены ребром тогда и только тогда, когда они разных цветов. Так как M_s — множество всех графов, $B \in M_s$. Также $B \notin A(T)$, так как он не содержит треугольников. Мы утверждаем, что для любого фиксированного $j, j \leqslant s$,

$$\Pr(B \in \delta^j_{\sqcup}) \leqslant 2^{-\sqrt{r}/2} < \frac{1}{m^5}. \tag{11.13}$$

Действительно, если $\delta^j_{\sqcup} \neq \varnothing$, то $\delta^j_{\sqcup} = \lceil * \rceil \setminus (\lceil R_1 \rceil \cup \lceil R_2 \rceil)$, где $|R_1 \cup R_2| > r$. Рассмотрим граф с множеством ребер $R_1 \cup R_2$. Пусть d — его максимальная степень. По теореме Визинга множество его ребер может быть разбито не более чем на $d + 1$ паросочетаний. Значит, или $d > \frac{\sqrt{r}}{2}$, или размер максимального паросочетания в этом графе не меньше $\sqrt{r}/2$. Следовательно, наш граф содержит множество из $k = \sqrt{r}/2$ ребер $e_1, \ldots, e_k$, которые образуют либо звезду, либо паросочетание. В каждом из этих двух случаев $\Pr[e_i \in B] = \frac{1}{2}$, и эти события попарно независимы. Тогда

$$\Pr[B \notin \lceil R_1 \rceil \cup \lceil R_2 \rceil] \leqslant 2^{-\sqrt{r}/2},$$

и отсюда следует неравенство (11.13). Заметим, что аналогичная оценка может быть получена без теоремы Визинга, если воспользоваться тем, что B не принадлежит $(\lceil R_1 \rceil \cup \lceil R_2 \rceil)$ тогда и только тогда, когда вершины в любой связной компоненте графа с множеством ребер $R_1 \cup R_2$ раскрашены в B в один цвет.

Так как $s < \binom{m}{3}/2r^2 < m^5$, из неравенства (11.13) следует, что существует двудольный граф B, такой что $B \in M_s, B \notin A(T)$ и $B \notin \bigcup\limits_{j \leqslant s} \delta^j_{\sqcup}$. Это противоречит (11.12). Следовательно, случай 2 невозможен, и теорема 11.4.1 доказана. ■

11.5. ФОРМУЛЫ

Напомним, что формула — это схема, в которой число выходов каждого элемента не превосходит 1. В отличие от случая со схемами, есть сверхлинейные нижние оценки минимальной сложности формул, вычисляющих различные явные **NP**-функции над полным двоичным базисом. Для булевой функции $f = f(x_1, \ldots, x_n)$ обозначим через $L(f)$ минимальное число элементов И и ИЛИ в формуле, использующей элементы И, ИЛИ и НЕ и вычисляющей f. В силу правил де Моргана, мы можем считать, что все элементы НЕ входят в первый уровень этой формулы. Мы завершаем эту главу простым результатом Субботовской (1961), который заключается в том, что $L(f) \geqslant \Omega(n^{3/2})$ для счетчика четности $f = x_1 \oplus \ldots \oplus x_n$. Эта оценка была позднее улучшена Храпченко

(1971) до $L(f) = n^2 - 1$. Тем не менее, мы представляем здесь лишь более слабую нижнюю оценку $\Omega(n^{3/2})$ не только потому, что это позволит еще раз продемонстрировать силу относительно простых вероятностных приемов, но и потому, что модификация этого доказательства позволила Андрееву (1987) получить нижнюю оценку $\Omega(n^{5/2}/(\log n)^{O(1)})$ на $L(g)$ для другой **NP**-функции $g = g(x_1, \ldots, x_n)$. Хостад [Håstad (1998)] позднее улучшил эту нижнюю оценку до $\Omega(n^{3-o(1)})$. В настоящее время это лучшая из известных нижняя оценка сложности формулы для **NP**-функции от n переменных над полным базисом.

Метод Субботовской основан на случайных ограничениях, аналогичных использованным в разд. 11.2. Главная лемма заключается в следующем.

Лемма 11.5.1. *Пусть $f = f(x_1, \ldots, x_n)$ — булева функция от n переменных, не являющаяся атомом. Тогда существуют i, $1 \leqslant i \leqslant n$, и $\varepsilon \in \{0, 1\}$, такие, что для функции $g = f(x_1, \ldots, x_{i-1}, \varepsilon, x_{i+1}, \ldots, x_n)$ от $n-1$ переменных, полученной из f заменой $x_i = \varepsilon$, верно следующее неравенство:*

$$(L(g) + 1) \leqslant \left(1 - \frac{3}{2n}\right)(L(f) + 1) \leqslant \left(1 - \frac{1}{n}\right)^{3/2}(L(f) + 1).$$

Доказательство. Зафиксируем формулу F, вычисляющую функцию f, с $l = L(f)$ элементами И и ИЛИ. Формула F может быть представлена как двоичное дерево, каждый из $l + 1$ листьев которого помечен атомом x_i или $\overline{x}_i$. Случайно выберем переменную $x_i, 1 \leqslant i \leqslant n$, согласно равномерному распределению, и присвоим ей случайное двоичное значение $\varepsilon \in \{0, 1\}$. Когда мы заменяем значения ε и $1 - \varepsilon$ на x_i и $\overline{x}_i$ соответственно, число листьев F уменьшается; математическое ожидание числа листьев, удаляемых таким образом, равно $(l+1)/n$. Однако, может произойти дальнейшее сокращение. В самом деле, предположим, что лист помечен символом x_i, и питает, допустим, элемент И $x_i \wedge H$ в F. Мы можем считать, что переменная x_i не входит в подформулу H, так как в противном случае F может быть упрощена заменой $x_i = 1$ в H. Если $x_i = \varepsilon = 0$, то H может быть удалена после того, как мы заменим значение для x_i, тем самым еще сильнее сокращая число листьев. Так, аналогичный эффект получается, если взять элемент ИЛИ вместо И (а также если взять $\overline{x}_i$ вместо x_i), математическое ожидание числа дополнительно удаленных листьев не меньше $(l + 1)/2n$. Следовательно, математическое ожидание числа оставшихся в упрощенной формуле листьев не превосходит $(l + 1)[1 - \frac{3}{2n}]$, что и требовалось доказать.

■

Многократно применяя лемму 11.5.1, получим

Следствие 11.5.2. *Если $f = f(x_1, \ldots, x_n)$ и $L(f) \leqslant (\frac{n}{k})^{3/2} - 1$, то можно так присвоить значения $n - k$ переменным, что получившаяся функция g будет атомом.*

Доказательство. Применение леммы 11.5.1 $n-k$ раз дает функцию g, для которой

$$(L(g)+1) \leqslant \prod_{i=k+1}^{n} \left(1-\frac{1}{i}\right)^{3/2} (L(f)+1) = (k/n)^{3/2}(L(f)+1) \leqslant 1.$$

Значит, g равно либо x_i, либо $\overline{x}_i$ для некоторого i. ■

Следствие 11.5.3. *Для счетчика четности $f = x_1 \oplus \ldots \oplus x_n$ справедливо равенство*

$$L(f) > \left(\frac{n}{2}\right)^{3/2} - 1.$$

11.6. УПРАЖНЕНИЯ

1. Показать, что существует такая константа c, что число булевых 2-схем сложности s не превосходит $(cs)^s$.

2. Пусть f — булева формула над n переменными $x_1, x_2, \ldots, x_n$, где f — конъюнкция произвольного (конечного) числа скобок, каждая скобка — дизъюнкция 10 букв, где каждая буква — это или переменная, или ее отрицание, и пусть каждая переменная входит (с отрицанием или без) не более чем в 10 скобок. Доказать, что f выполнима.

3* Доказать, что существует ограниченная по глубине монотонная схема полиномиальной сложности с n входами $x_1, x_2, \ldots, x_n$, вычисляющая функцию f, принимающую значение 1, если $\sum_{i=1}^{n} x_i \geqslant n/2 + n/\log_2 n$, и 0, если $\sum_{i=1}^{n} x_i \leqslant n/2 - n/\log_2 n$.

ВЕРОЯТНОСТНЫЙ ВЗГЛЯД:

Максимальные антицепи

Семейство $\mathcal{F}$ подмножеств множества $\{1, \ldots, n\}$ называется *антицепью*, если ни одно множество из $\mathcal{F}$ не содержится в другом.

Теорема. *Пусть $\mathcal{F}$ — антицепь. Тогда*

$$\sum_{A \in \mathcal{F}} \left[1 \Big/ \binom{n}{|A|} \right] \leqslant 1.$$

Доказательство. Пусть σ — равновероятно выбранная перестановка на множестве $\{1, \ldots, n\}$. Положим

$$\mathcal{C}_\sigma = \{\{\sigma(j) : 1 \leqslant j \leqslant i\} : 0 \leqslant i \leqslant n\}.$$

(В случаях $i = 0$ и $i = n$ мы получим $\varnothing, \{1, \ldots, n\} \in \mathcal{C}$, соответственно.) Введем случайную величину

$$X = |\mathcal{F} \cap \mathcal{C}_\sigma|.$$

Рассмотрим разложение

$$X = \sum_{A \in \mathcal{F}} X_A,$$

где X_A — индикатор события «$A \in \mathcal{C}$». Тогда

$$\mathbf{E}[X_A] = \Pr[A \in \mathcal{C}_\sigma] = 1 \Big/ \binom{n}{|A|},$$

так как $\mathcal{C}_\sigma$ содержит ровно одно множество размера $|A|$, которое выбирается равновероятно среди всех множеств размера $|A|$. Из линейности математического ожидания следует, что

$$\mathbf{E}[X] = \sum_{A \in \mathcal{F}} 1 \Big/ \binom{n}{|A|}.$$

$\mathcal{C}_\sigma$ образует цепь для *любого* σ, следовательно, каждая пара множеств сравнима. Так как $\mathcal{F}$ — антицепь, *должно* выполняться неравенство $X = |\mathcal{F} \cap \mathcal{C}_\sigma| \leqslant 1$. Следовательно, $\mathbf{E}[X] \leqslant 1$. ■

Следствие (теорема Шпернера). *Пусть $\mathcal{F}$ — антицепь. Тогда*

$$|\mathcal{F}| \leqslant \binom{n}{\lfloor n/2 \rfloor}.$$

Доказательство. Функция $\binom{n}{x}$ достигает максимума при $x = \lfloor n/2 \rfloor$, так что

$$1 \geqslant \sum_{A \in \mathcal{F}} \frac{1}{\binom{n}{|A|}} \geqslant \frac{|\mathcal{F}|}{\binom{n}{\lfloor n/2 \rfloor}}.$$ ■

12

Разброс

Тайна и величие математики состоит не столько в том, что абстрактные теории оказываются полезными на практике, но в чудесном свойстве, что теория, предназначенная для решения одного типа задач, часто дает единственно верный путь решения проблем совершенно иного толка, задач, для которых теория не была предназначена. Такие совпадения происходят настолько часто, что должны отражать саму суть математики.

Жан-Карло Рота

12.1. ОСНОВЫ

Рассмотрим некоторое конечное семейство конечных множеств. Нам нужно раскрасить точки множеств в красный и синий цвета так, что бы каждое из множеств содержало примерно одинаковое количество красных и синих точек. Может оказаться, что наша задача невыполнима, например, если семейство состоит из всех подмножеств некоторого множества Ω. Тогда, независимо от раскраски, найдется монохроматическое множество, состоящее не менее чем из половины точек Ω. В противоположной ситуации, когда множества семейства не имеют общих элементов, задача становится тривиальной. В этом случае можно так раскрасить точки множеств, что все множества будут иметь либо равное количество красных и синих точек, либо их количество будет отличаться на единицу, если мощность множества нечетна. Разброс будет определять, насколько хорошую раскраску мы можем найти.

Дадим формальное определение разброса. Рассмотрим семейство $\mathcal{A}$ подмножеств множества Ω. Отказавшись от использования красного и синего цветов, определим раскраску как отображение

$$\chi : \Omega \longrightarrow \{-1, +1\}.$$

Для каждого $A \subset \Omega$ положим

$$\chi(A) = \sum_{a \in A} \chi(a).$$

Определим *разброс* $\mathcal{A}$ *относительно* χ равенством

$$\operatorname{disc}(\mathcal{A}, \chi) = \max_{A \in \mathcal{A}} |\chi(A)|$$

и *разброс* $\mathcal{A}$ формулой

$$\operatorname{disc}(\mathcal{A}) = \min_{\chi : \Omega \longrightarrow \{-1,+1\}} \operatorname{disc}(\mathcal{A}, \chi).$$

Следующее эквивалентное определение разброса открывает нам его геометрический смысл. Рассмотрим множества $\mathcal{A} = \{S_1, \ldots, S_m\}$, $\Omega = \{1, \ldots, n\}$ и матрицу инцидентности $B = [b_{ij}]$ размерности $m \times n$ такую, что $b_{ij} = 1$, если $j \in S_i$, и $b_{ij} = 0$ иначе. Раскраска χ может быть задана вектором $u = (\chi(1), \ldots, \chi(n)) \in \{-1, +1\}^n$, так что $Bu^T = (\chi(S_1), \ldots, \chi(S_m))$ и

$$\operatorname{disc}(\mathcal{A}) = \min_{u \in \{-1,+1\}^n} |Bu^T|_\infty,$$

где $|v|_\infty$ означает норму L^∞, т. е. максимальное из абсолютных значений координат. Иными словами, обозначив j-й столбец матрицы B (образ точки j) через v_j, получаем, что

$$\operatorname{disc}(\mathcal{A}) = \min |\pm v_1 \pm \ldots \pm v_n|_\infty,$$

где минимум берется по всем 2^n возможным вариантам выбора знаков.

Далее нас будут интересовать верхние оценки разброса. Пусть $\operatorname{disc}(\mathcal{A}) \leqslant K$, тогда и только тогда, когда *существует* раскраска χ, для которой $|\chi(A)| \leqslant K$ для всех $A \in \mathcal{A}$. Естественно будет воспользоваться случайной раскраской.

Теорема 12.1.1. *Пусть $\mathcal{A}$ — семейство из n подмножеств m-множества Ω. Тогда*

$$\operatorname{disc}(\mathcal{A}) \leqslant \sqrt{2m \ln(2n)}.$$

Доказательство. Рассмотрим случайную раскраску $\chi : \Omega \longrightarrow \{-1, +1\}$. Для каждого $A \subset \Omega$ введем случайную величину X_A — индикатор события $|\chi(A)| > \alpha$, где $\alpha = \sqrt{2m \ln(2n)}$. Если $|A| = a$, то $\chi(A)$ имеет распределение $\mathbf{S}_a$, тогда по теореме А.1.1

$$\mathbf{E}[X_A] = \Pr[|\chi(A)| > \alpha] < 2e^{-\alpha^2/2a} \leqslant 2e^{-\alpha^2/2m} = 1/n$$

при нашем выборе α. Обозначим через X число множеств $A \in \mathcal{A}$ таких, что $|\chi(A)| > \alpha$. Тогда

$$X = \sum_{A \in \mathcal{A}} X_A,$$

и из линейности математического ожидания следует, что

$$\mathbf{E}[X] = \sum_{A \in \mathcal{A}} \mathbf{E}[X_A] < |\mathcal{A}|(1/n) = 1.$$

Тогда найдется такая раскраска χ, что $X = 0$. Это означает, что $\operatorname{disc}(\mathcal{A}, \chi) \leqslant \alpha$ и, следовательно, $\operatorname{disc}(\mathcal{A}) \leqslant \alpha$. ■

12.2. ДОСТАТОЧНОСТЬ ШЕСТИ СТАНДАРТНЫХ ОТКЛОНЕНИЙ

В случае, когда семейство $\mathcal{A}$ состоит из n подмножеств n-множества, по теореме 12.1.1 имеем

$$\operatorname{disc}(\mathcal{A}) = O(\sqrt{n \ln(n)}).$$

Следующая теорема дает нам лучшую оценку. Ее доказательство аналогично доказательству основного результата работы [Beck (1981)]. Подход с использованием энтропии был предложен Р. Боппана.

Теорема 12.2.1 [Spencer (1985a)]. *Пусть $\mathcal{A}$ — семейство из n подмножеств n-множества Ω. Тогда*

$$\operatorname{disc}(\mathcal{A}) < 6\sqrt{n}.$$

При случайной раскраске $\chi : \Omega \to \{-1, +1\}$ и $A \in \mathcal{A}$ величина $\chi(A)$ имеет нулевое математическое ожидание и дисперсию, не превосходящую $\sqrt{n}$. Если $|\chi(A)| > 6\sqrt{n}$, то $\chi(A)$ находится на расстоянии не менее шести стандартных отклонений от своего среднего значения. Вероятность этого события очень мала, но, тем не менее, является некоторой положительной константой, а число множеств $A \in \mathcal{A}$ стремится к бесконечности с ростом n. Таким образом, случайная раскраска χ почти всегда *не будет* искомой. Константа 6 (в действительности, 5.32) является результатом специальных вычислений и может быть улучшена. Мы не будем здесь касаться этих вычислений. Вместо этого докажем теорему 12.2.1 с заменой множителя 6 константой 11. Отображение

$$\chi : \Omega \longrightarrow \{-1, 0, +1\}$$

назовем *частичной* раскраской. Если $\chi(a) = 0$, будем говорить, что точка a не раскрашена. Определим $\chi(A)$ так же как и раньше.

Лемма 12.2.2. *Пусть $\mathcal{A}$ — семейство из n подмножеств n-множества Ω. Тогда найдется частичная раскраска χ, такая, что не более $10^{-9}n$ точек останутся не раскрашенными и*

$$|\chi(A)| \leqslant 10\sqrt{n}$$

для всех $A \in \mathcal{A}$.

Значения 10 и 10^{-9} не являются наилучшими. Важно, что это — абсолютные константы. Обозначим для определенности множества семейства $\mathcal{A}$ через $A_1, \ldots, A_n$. Рассмотрим случайную раскраску

$$\chi : \Omega \longrightarrow \{-1, +1\}.$$

Для каждого $1 \leqslant i \leqslant n$ определим

$$b_i = \text{ ближайшее целое число к } \chi(A_i)/(20\sqrt{n}).$$

Например, $b_i = 0$ при $-10\sqrt{n} < \chi(A_i) < 10\sqrt{n}$ и $b_i = -3$ при $-70\sqrt{n} < \chi(A_i) < -50\sqrt{n}$. Из теоремы A.1.1 (также как в доказательстве теоремы 12.1.1) следует, что

$$\begin{aligned}
&\Pr[b_i = 0] > 1 - 2e^{-50},\\
&\Pr[b_i = 1] = \Pr[b_i = -1] < e^{-50},\\
&\Pr[b_i = 2] = \Pr[b_i = -2] < e^{-450},
\end{aligned}$$

и, в общем случае,

$$\Pr[b_i = s] = \Pr[b_i = -s] < e^{-50(2s-1)^2}.$$

Оценим теперь *энтропию* $H[b_i]$. Это важное понятие подробно рассматривается в разд. 14.6. Положим $p_j = \Pr[b_i = j]$, тогда

$$H(b_i) = \sum_{j=-\infty}^{+\infty} (-p_j \log_2(p_j)) \leqslant (1 - 2e^{-50})[-\log_2(1 - 2e^{-50})] +$$
$$+ 2e^{-50}[-\log_2 e^{-50}] + 2e^{-450}[-\log_2 e^{-450}] + \dots .$$

Бесконечная сумма, очевидно, сходится и строго мажорируется своим вторым слагаемым. Вычисления показывают, что

$$H(b_i) \leqslant \varepsilon = 3 \times 10^{-20}.$$

Рассмотрим теперь вектор $(b_1, \dots, b_n)$. Безусловно, между различными b_i могут быть взаимосвязи. В самом деле, если S_i и S_j мало отличаются, то b_i и b_j будут, скорее всего, одинаковыми. Но по утверждению 14.6.2 энтропия субаддитивна. Следовательно,

$$H((b_1, \dots, b_n)) \leqslant \sum_{i=1}^{n} H(b_i) \leqslant \varepsilon n.$$

Если случайная величина Z не принимает ни одно из значений с вероятностью, большей 2^{-t}, то $H(Z) \geqslant t$. Из утверждения, обратного к данному, следует, что найдется вектор $(s_1, \dots, s_n)$ такой, что

$$\Pr[(b_1, \dots, b_n) = (s_1, \dots, s_n)] \geqslant 2^{-\varepsilon n}.$$

Наше вероятностное пространство состоит из 2^n равновероятных раскрасок χ. Следовательно, найдется множество $\mathcal{C}'$, состоящее по крайней мере из $2^{(1-\varepsilon)n}$ раскрасок $\chi : \Omega \longrightarrow \{-1, +1\}$, имеющих одни и те же значения $(b_1, \dots, b_n)$.

Будем рассматривать класс $\mathcal{C}$ всех раскрасок $\chi : \Omega \longrightarrow \{-1, +1\}$ как куб Хэмминга $\{-1, +1\}^n$ с введенной на нем метрикой Хэмминга

$$\rho(\chi, \chi') = |\{a : \chi(a) \neq \chi'(a)\}|.$$

В работе [Kleitman (1966a)] было доказано, что если $\mathcal{D} \subset \mathcal{C}$ и

$$|\mathcal{D}| \geqslant \sum_{i \leqslant r} \binom{n}{i}$$

при $r \leqslant \frac{n}{2}$, то множество $\mathcal{D}$ имеет диаметр по крайней мере $2r$. То есть множество данного размера с минимальным диаметром есть шар. (То, что множество $\mathcal{D}$ имеет диаметр по крайней мере r, доказывается тривиально. Этого было бы достаточно для доказательства леммы 12.2.2 и теоремы 12.2.1 с худшими значениями констант.)

Доказательство. Мы можем положить $r = \alpha n$, если $\alpha < \frac{1}{2}$, и

$$2^{H(\alpha)} \leqslant 2^{1-\varepsilon}.$$

Вычисления показывают, что мы можем взять $\alpha = \frac{1}{2}(1-10^{-9})$. (Из разложения в ряд Тейлора следует, что

$$H\left(\frac{1}{2}-x\right) \sim 1 - \frac{2}{\ln 2}x^2$$

при малых x.) Тогда множество $\mathcal{C}'$ имеет диаметр по крайней мере $n(1-10^{-9})$. Пусть раскраски $\chi_1, \chi_2 \in \mathcal{C}'$ находятся на максимальном расстоянии друг от друга. Положим

$$\chi = \frac{\chi_1 - \chi_2}{2}.$$

Заметим, что χ — частичная раскраска множества Ω. Заметим также, что $\chi(a) = 0$ тогда и только тогда, когда $\chi_1(a) = \chi_2(a)$, что верно для $n - \rho(\chi_1, \chi_2) \leqslant 10^{-9}n$ координат a. И, наконец, наиболее критический момент: для каждого i, $1 \leqslant i \leqslant n$ раскраски χ_1, χ_2 дают одно и тоже значение b_i, а это означает, что величины $\chi_1(A_i)$ и $\chi_2(A_i)$ находятся внутри интервала длины $20\sqrt{n}$. Следовательно,

$$|\chi(A_i)| = \left|\frac{\chi_1(A_i) - \chi_2(A_i)}{2}\right| \leqslant 10\sqrt{n},$$

что и требовалось доказать. ■

Для доказательства теоремы 12.2.1 требуется раскраска всех точек множества, в то время как лемма 12.2.2 оставляет $10^{-9}n$ точек нераскрашенными. Идея дальнейших действий заключается в последовательном применении процедуры раскраски из леммы 12.2.2, которая оставит нераскрашенными, скажем, $10^{-18}n$ точек на втором шаге. Но мы не можем применить лемму 12.2.2 напрямую, так как нарушена симметричность ситуации — у нас на n множеств приходится только $10^{-9}n$ точек.

Лемма 12.2.3. *Пусть $\mathcal{A}$ — семейство из n подмножеств r-множества Ω при $r \leqslant 10^{-9}n$. Тогда найдется частичная раскраска χ множества Ω с не более чем $10^{-40}r$ нераскрашенными точками, такая, что*

$$|\chi(A)| < 10\sqrt{r}\sqrt{\ln(n/r)}$$

для всех $A \in \mathcal{A}$.

Доказательство. Мы опишем доказательство, которое оставляет простор для улучшений. Обозначим через $A_1, \ldots, A_n$ множества семейства $\mathcal{A}$. Рассмотрим случайную раскраску $\chi : \Omega \longrightarrow \{-1, +1\}$. Положим

$$b_i = \text{ ближайшее целое число к } \frac{\chi(A_i)}{20\sqrt{r}\sqrt{\ln(n/r)}}$$

для каждого $1 \leqslant i \leqslant n$. Вероятность того, что $b_i = 1$, менее $(r/n)^{50}$. Энтропия $H(b_i)$ мажорируется соответствующим слагаемым, следовательно,

$$3\left(\frac{r}{n}\right)^{50}\left[-\log_2\left(\left(\frac{r}{n}\right)^{50}\right)\right] < 10^{-100}\frac{r}{n}.$$

Тогда энтропия $(b_1, \ldots, b_n)$ меньше $10^{-100}r$. Найдем пару практически противоположных раскрасок χ_1, χ_2 с одинаковыми b и положим $\chi = (\chi_1 - \chi_2)/2$, как и в предыдущем случае. ■

Доказательство теоремы 12.2.1. Применим лемму 12.2.2 для нахождения χ^1, и затем, последовательно применяя лемму 12.2.3, будем получать раскраски $\chi^2, \chi^3, \ldots$ до тех пор, пока все точки не окажутся раскрашенными. Обозначим через χ итоговую раскраску. Для каждого $A \in \mathcal{A}$ верно, что

$$\chi(A) = \chi^1(A) + \chi^2(A) + \ldots$$

и

$$|\chi(A)| \leqslant 10\sqrt{n} + 10\sqrt{10^{-9}n}\sqrt{\ln 10^9} + \\ + 10\sqrt{10^{-49}n}\sqrt{\ln 10^{49}} + 10\sqrt{10^{-89}n}\sqrt{\ln 10^{89}} + \ldots .$$

При вынесении $\sqrt{n}$ за скобку получаем сходящийся ряд. Отсюда выводим оценку

$$|\chi(A)| \leqslant 11\sqrt{n}$$

с возможностью уточнения. ■

Пусть семейство $\mathcal{A}$ состоит из n подмножеств r-множества и $r < n$. Мы можем последовательно применять лемму 12.2.3 (сначала применив лемму 12.2.2, если $r > 10^{-9}n$) для получения такой раскраски χ, что

$$\operatorname{disc}(\mathcal{A}, \chi) < K\sqrt{r}\sqrt{\ln(n/r)},$$

где K — абсолютная константа. При $r = n^{1-o(1)}$ эта оценка лучше оценки для случайной раскраски из теоремы 12.1.1.

12.3. ЛИНЕЙНЫЙ И НАСЛЕДСТВЕННЫЙ РАЗБРОС

Пусть $\mathcal{A} = \{A_1, \ldots, A_n\}$ и $\Omega = \{1, \ldots, m\}$, $m > n$. Заметим, что утверждение $\operatorname{disc}(\mathcal{A}) \leqslant K$ эквивалентно существованию множества $S = \{j : \chi(j) = +1\}$, такого, что величина $|S \cap A|$ отстоит не более чем на $K/2$ от $|A|/2$ для всех множеств $A \in \mathcal{A}$. Определим *линейный разброс* $\operatorname{lindisc}(\mathcal{A})$ равенством

$$\operatorname{lindisc}(\mathcal{A}) = \max_{p_1,\ldots,p_m \in [0,1]} \min_{\varepsilon_1,\ldots,\varepsilon_m \in \{0,1\}} \max_{A \in \mathcal{A}} \Big|\sum_{i \in A} (\varepsilon_i - p_i)\Big|.$$

Неравенство $\operatorname{lindisc}(\mathcal{A}) \leqslant K$ означает, что при любых заданных $p_1, \ldots, p_m$ найдется их «одновременное округление» $\varepsilon_1, \ldots, \varepsilon_m$, такое, что для множества $S = \{j : \varepsilon_j = 1\}$ величина $|S \cap A|$ отстоит не более чем на K от суммы весов $\sum_{j \in A} p_j$ для всех множеств $A \in \mathcal{A}$. Если положить $p_j = \frac{1}{2}$, то из этой верхней оценки вытекает, что $\operatorname{disc}(\mathcal{A}) \leqslant 2K$. Но из утверждения $\operatorname{lindisc}(\mathcal{A}) \leqslant K$ можно вывести гораздо более сильные следствия. Например, если положить $p_j = \frac{1}{3}$, то из него следует существование множества S, для которого величина $|S \cap A|$ находится на расстоянии не более чем K от $|A|/3$. Линейный разброс, а также описанный ниже наследственный разброс были впервые введены в

работе [Lovász, Spencer and Vesztergombi (1986)]. Для произвольного множества $X \subset \Omega$ обозначим через $\mathcal{A}|_X$ *ограничение* $\mathcal{A}$ по X, т. е. семейство $\{A \cap X : A \in \mathcal{A}\}$. Следующий результат сводит оценку $\operatorname{disc}(\mathcal{A})$ в ситуации, когда точек больше чем множеств, к оценке $\operatorname{lindisc}(\mathcal{A})$ в случае, когда число точек не превышает числа множеств.

Теорема 12.3.1. *Пусть $\mathcal{A}$ — семейство из n подмножеств m-множества и $m \geqslant n$. Тогда, если $\operatorname{lindisc}(\mathcal{A}|_X) \leqslant K$ для каждого подмножества X мощности не более n, то $\operatorname{lindisc}(\mathcal{A}) \leqslant K$.*

Доказательство. Пусть заданы некоторые $p_1, \ldots, p_m \in [0,1]$. Опишем процесс сведения оценки $\operatorname{disc}(\mathcal{A})$ к оценке $\operatorname{lindisc}(\mathcal{A})$. Назовем индекс j *фиксированным*, если $p_j \in \{0,1\}$, иначе назовем его *плавающим*. Обозначим через F множество всех плавающих индексов. Остановим процесс в случае, когда $|F| \leqslant n$. Пусть $|F| > n$, $j \in F$ и y_j — ненулевое решение однородной системы

$$\sum_{j \in A \cap F} y_j = 0, \qquad A \in \mathcal{A}.$$

Такое решение существует, так как система содержит больше переменных ($|F|$) чем уравнений (n), и может быть найдено стандартными методами линейной алгебры. Положим

$$p'_j = p_j + \lambda y_j, \quad j \in F,$$
$$p'_j = p_j, \qquad\quad j \notin F,$$

где λ — действительное число с наименьшим абсолютным значением, такое, что для некоторого $j \in F$ величина p'_j становится равной нулю или единице. Заметим, что

$$\sum_{j \in A} p'_j = \sum_{j \in A} p_j + \lambda \sum_{j \in A \cap F} y_j = \sum_{j \in A} p_j \qquad (*)$$

для всех множеств $A \in \mathcal{A}$. Повторим описанную процедуру с новыми p'_j. На каждой итерации по крайней мере один индекс j становится фиксированным, так что процесс остановится при некоторых $p^*_1, \ldots, p^*_m$. Обозначим через X получившееся множество плавающих индексов j. Имеем $|X| \leqslant n$. По предположению теоремы существует набор $\varepsilon_j, j \in X$, такой, что

$$\left| \sum_{j \in A \cap X} p^*_j - \varepsilon_j \right| \leqslant K, \qquad A \in \mathcal{A}.$$

Положим $\varepsilon_j = p^*_j$ для всех $j \notin X$. Так как равенство $(*)$ выполняется на каждой итерации, то

$$\sum_{j \in A} p^*_j = \sum_{j \in A} p_j.$$

Следовательно,

$$\left| \sum_{j \in A} (p_j - \varepsilon_j) \right| = \left| \sum_{j \in A} (p_j - p^*_j) + \sum_{j \in A \cap X} (p^*_j - \varepsilon_j) \right| \leqslant K$$

для всех $A \in \mathcal{A}$. ■

Определим *наследственный разброс* $\operatorname{herdisc}(\mathcal{A})$ следующим образом:

$$\operatorname{herdisc}(\mathcal{A}) = \max_{X\subseteq\Omega} \operatorname{disc}(\mathcal{A}|_X).$$

Пример. Положим $\Omega = \{1,\dots,n\}$ и пусть семейство $\mathcal{A}$ состоит из всех интервалов $[i,j] = \{i, i+1,\dots,j\}$, где $1 \leqslant i \leqslant j \leqslant n$. Тогда $\operatorname{disc}(\mathcal{A}) = 1$, так как мы можем раскрасить элементы множества Ω поочередно $+1$ и -1. Аналогично $\operatorname{herdisc}(\mathcal{A}) = 1$. Для каждого множества $X \subseteq \Omega$, состоящего из элементов $x_1 < x_2 < \ldots < x_r$, мы можем раскрасить X попеременно $\chi(x_k) = (-1)^k$. Тогда для каждого интервала $[i,j] \in \mathcal{A}$ элементы множества $[i,j] \cap X$ будут раскрашены поочередно в разные цвета.

Теорема 12.3.2. *Справедливо неравенство* $\operatorname{lindisc}(\mathcal{A}) \leqslant \operatorname{herdisc}(\mathcal{A})$.

Доказательство. Пусть $K = \operatorname{herdisc}(\mathcal{A})$. Кроме того, пусть семейство $\mathcal{A}$ состоит из подмножеств множества $\Omega = \{1,\dots,m\}$. Рассмотрим некоторый набор $p_1,\dots,p_m \in [0,1]$. Предположим, что все числа p_i имеют конечную длину записи в двоичной системе счисления. Обозначим через T наименьшее целое число, такое, что $p_i 2^T \in \mathbb{Z}$. Обозначим через J множество индексов i, для которых T-я цифра записи числа p_i в двоичной системе счисления равна единице, т. е. тех, для которых $p_i 2^{T-1} \notin \mathbb{Z}$. Так как $\operatorname{disc}(\mathcal{A}|_J) \leqslant K$, найдется набор $\varepsilon_j \in \{-1,+1\}$, такой, что

$$\Big|\sum_{j\in J\cap A} \varepsilon_j\Big| \leqslant K$$

для всех множеств $A \in \mathcal{A}$. Положим $p_j = p_j^{(T)}$. Пусть

$$p_j^{(T-1)} = \begin{cases} p_j^{(T)} & \text{при } j \notin J,\\ p_j^{(T)} + \varepsilon_j 2^{-T} & \text{при } j \in J.\end{cases}$$

Таким образом, числа $p_j^{(T-1)}$ являются «округлениями» чисел $p_j^{(T)}$ до $(T-1)$-го знака после запятой. Заметим, что все числа $p_j^{(T-1)} 2^{-(T-1)} \in \mathbb{Z}$. Для каждого множества $A \in \mathcal{A}$ верно следующее утверждение:

$$\left|\sum_{j\in A} p_j^{(T-1)} - p_j^{(T)}\right| = \left|\sum_{j\in J\cap A} 2^{-T}\varepsilon_j\right| \leqslant 2^{-T}K.$$

Повторим процедуру и последовательно найдем числа $p_j^{(T-2)},\dots,p_j^{(1)},p_j^{(0)}$. Все числа $p_j^{(0)} 2^{-0} \in \mathbb{Z}$, тогда $p_j^{(0)} \in \{0,1\}$ и

$$\left|\sum_{j\in A} p_j^{(0)} - p_j^{(T)}\right| \leqslant \sum_{i=1}^{T} \left|\sum_{j\in A} p_j^{(i-1)} - p_j^{(i)}\right| \leqslant \sum_{i=1}^{T} 2^{-i}K \leqslant K,$$

что и требовалось доказать.

Что же делать случае произвольных $p_1,\dots,p_m \in [0,1]$? Можно, конечно, сказать, что, по крайней мере для вычислительной техники, все действительные числа имеют конечную длину записи в двоичной системе счисления. Но

мы будем, все же, более аккуратны и рассмотрим функцию

$$f(p_1,\dots,p_m)=\min_{\varepsilon_1,\dots,\varepsilon_m\in\{0,1\}}\max_{A\in\mathcal{A}}\left|\sum_{i\in A}(\varepsilon_i-p_i)\right|,$$

которая является минимумом конечного числа максимальных значений конечного числа непрерывных функций. Следовательно, сама функция также непрерывна. Множество чисел $p_1,\dots,p_m\in[0,1]$, таких, что $p_i2^T\in\mathbb{Z}$ для некоторого T, образует плотное подмножество отрезка $[0,1]$. Так как $f\leqslant K$ на этом плотном множестве, то $f\leqslant K$ на всех наборах $p_1,\dots,p_m\in[0,1]$. ■

Следствие 12.3.3. *Пусть $\mathcal{A}$ — семейство из n подмножеств m-множества. Пусть* $\operatorname{disc}(\mathcal{A}|_X)\leqslant K$ *для всех множеств X мощности не более чем n. Тогда* $\operatorname{disc}(\mathcal{A})\leqslant 2K$.

Доказательство. Для каждого множества $X\subseteq\Omega$, такого, что $|X|\leqslant n$, справедливо неравенство $\operatorname{herdisc}(\mathcal{A}|_X)\leqslant K$. Тогда из теоремы 12.3.2 следует, что $\operatorname{lindisc}(\mathcal{A}|_X)\leqslant K$. По теореме 12.3.1 имеем $\operatorname{lindisc}(\mathcal{A})\leqslant K$. Тогда

$$\operatorname{disc}(\mathcal{A})\leqslant 2\operatorname{lindisc}(\mathcal{A})\leqslant 2K.$$ ■

Следствие 12.3.4. *Для каждого семейства $\mathcal{A}$, состоящего из n множеств произвольной мощности, справедливо неравенство*

$$\operatorname{disc}(\mathcal{A})\leqslant 12\sqrt{n}.$$

Доказательство. Применим теорему 12.2.1 и следствие 12.3.3. ■

12.4. НИЖНИЕ ОЦЕНКИ

Мы приведем два различных доказательства того факта, что с точностью до константы при $\sqrt{n}$ следствие 12.3.4 дает наилучшую из возможных оценок. Матрицей Адамара называется квадратная матрица $H=(h_{ij})$ с элементами $h_{ij}\in\{-1,+1\}$ и попарно ортогональными строками (следовательно, столбцы этой матрицы также попарно ортогональны). Рассмотрим матрицу Адамара H порядка n и вектор $v=(v_1,\dots,v_n)^T$, $v_i\in\{-1,+1\}$. Тогда

$$Hv=v_1c_1+\ldots+v_nc_n,$$

где через c_i обозначен i-й столбец матрицы H. Положим $Hv=(L_1,\dots,L_n)^T$ и, понимая под $|c|$ обычную евклидову норму, имеем

$$L_1^2+\ldots+L_n^2=|Hv|^2=v_1^2|c_1|^2+\ldots+v_n^2|c_n|^2=n+\ldots+n=n^2,$$

так как все векторы c_i взаимно ортогональны. Следовательно, найдется слагаемое $L_i^2\geqslant n$ и

$$|Hv|_\infty=\max(|L_1|,\dots,|L_n|)\geqslant\sqrt{n}.$$

Перенесем этот результат на семейства множеств. Рассмотрим матрицу Адамара H порядка n, в первом столбце и первой строке которой стоят только

единицы. (Заметим, что любая матрица Адамара может быть приведена к такому виду путем умножения соответствующих строк и столбцов на -1.) Обозначим через J единичную матрицу порядка n. Пусть $v_1, \ldots, L_1, \ldots$ обозначают то же, что и прежде. Тогда

$$L_1 + \ldots + L_n = \sum_{i,j=1}^{n} v_j h_{ij} = \sum_{j=1}^{n} v_j \sum_{i=1}^{n} h_{ij} = nv_1 = \pm n,$$

так как сумма элементов первого столбца равна n, а суммы элементов остальных столбцов, ортогональных к первому, равны нулю. Положим $\lambda = v_1 + \ldots + v_n$, тогда $Jv = (\lambda, \ldots, \lambda)$ и

$$(H + J)v = (L_1 + \lambda, \ldots, L_n + \lambda).$$

Имеем

$$|(H + J)v|^2 = \sum_{i=1}^{n}(L_i + \lambda)^2 = \sum_{i=1}^{n}(L_i^2 + 2\lambda L_i + \lambda^2) = n^2 \pm 2n\lambda + n\lambda^2.$$

Будем рассматривать только четные n (для нечетных n, за исключением $n = 1$, матриц Адамара не существует). Тогда λ четно. Квадратичная (относительно λ) функция $n^2 \pm 2n\lambda + n\lambda^2$ достигает минимума в ∓ 1. Тогда при ограничении, что минимум должен достигаться в четном целом λ, получаем, что минимум функции достигается в $\lambda = 0, \mp 2$, и

$$|(H + J)v|^2 \geqslant n^2.$$

Вновь некоторая координата должна быть не меньше $\sqrt{n}$. Положим $H^* = \frac{H+J}{2}$, тогда

$$|H^* v|_\infty \geqslant \sqrt{n}/2.$$

Пусть $\mathcal{A} = \{A_1, \ldots, A_m\}$ — произвольное семейство подмножеств множества $\Omega = \{1, \ldots, n\}$, обозначим через M соответствующую матрицу инцидентности размерности $m \times n$. Раскраска $\chi : \Omega \longrightarrow \{-1, +1\}$ дает нам вектор $v = (\chi(1), \ldots, \chi(n)) \in \{-1, +1\}^n$. Тогда

$$\operatorname{disc}(\mathcal{A}, \chi) = |Mv|_\infty$$

и

$$\operatorname{disc}(\mathcal{A}) = \min_{v \in \{-1,+1\}^n} |Mv|_\infty.$$

Заметим, что элементами матрицы H^* являются числа 0 и 1. Таким образом, верна следующая теорема.

Теорема 12.4.1. *Если существует матрица Адамара порядка $n > 1$, то найдется такое семейство $\mathcal{A}$ из n подмножеств n-множества, что*

$$\operatorname{disc}(\mathcal{A}) \geqslant \sqrt{n}/2.$$

Несмотря на то, что неизвестно точно, для каких n существуют матрицы Адамара (гипотеза Адамара состоит в том, что они существуют для $n = 1, 2$

и всех чисел, кратных четырем; см., например, [Hall (1986)]), известно, что порядки матриц Адамара образуют плотное множество в том смысле, что для любого ε при достаточно больших n найдется матрица Адамара, порядок которой лежит между n и $n(1-\varepsilon)$. Этого факта достаточно для получения из теоремы 12.4.1 асимптотического результата для всех n.

Второе доказательство существования семейства $\mathcal{A}$ с большим разбросом сопряжено с «переворачиванием вероятностного доказательства с ног на голову». Рассмотрим случайную $0,1$ матрицу M порядка n. Зафиксируем вектор $v = (v_1, \ldots, v_n), v_j = \pm 1$ и положим $Mv = (L_1, \ldots, L_n)$. Пусть половина координат $v_j = +1$ и половина равна -1. Тогда

$$L_i \sim B\left(\frac{n}{2}, \frac{1}{2}\right) - B\left(\frac{n}{2}, \frac{1}{2}\right),$$

что приблизительно совпадает с нормальным распределением $N(0, \sqrt{n}/2)$. Выберем константу $\lambda > 0$ таким образом, чтобы

$$\int_{-\lambda}^{\lambda} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt < \frac{1}{2}.$$

Тогда

$$\Pr\left[|L_i| < \lambda\sqrt{n}/2\right] < \frac{1}{2}.$$

Данное неравенство верно и в том случае, когда вектор v не сбалансирован по числу вхождений $+1$ и -1, доказательство этого факта мы опустим. Заметим то важное обстоятельство, что все L_i взаимно независимы. Это следует из независимости выбора элементов матрицы M. Следовательно,

$$\Pr\left[|L_i| < \lambda\sqrt{n}/2 \text{ для всех } 1 \leqslant i \leqslant n\right] < \left(\frac{1}{2}\right)^n.$$

Всего существует 2^n векторов v. Тогда математическое ожидание числа векторов v, для которых $|Mv|_\infty < \lambda\sqrt{n}/2$, меньше $2^n 2^{-n} = 1$. Для некоторой матрицы M эта величина должна равняться нулю, т. е. не должно найтись соответствующих векторов v. По матрице M строим семейство $\mathcal{A}$, такое, что

$$\operatorname{disc}(\mathcal{A}) > \lambda\sqrt{n}/2.$$

12.5. ТЕОРЕМА БЕКА—ФИАЛА

Пусть $\mathcal{A}$ — некоторое семейство множеств. Обозначим через $\deg(\mathcal{A})$ максимальное число множеств из $\mathcal{A}$, одновременно содержащих какой-либо элемент. Следующий результат [Beck and Fiala (1981)] получен с использованием только лишь методов линейной алгебры и формально выбивается из ряда рассматриваемых в этой книге задач. Но мы все-таки рассмотрим его потому, что доказательство замечательно красиво, а также потому, что сам результат во многом отвечает духу этой главы.

Теорема 12.5.1. *Пусть $\mathcal{A}$ — произвольное конечное семейство конечных множеств. Кроме того, пусть* $\deg(\mathcal{A}) \leqslant t$. *Тогда*

$$\operatorname{disc}(\mathcal{A}) \leqslant 2t - 1.$$

Доказательство. Рассмотрим семейство $\mathcal{A} = \{A_1, \ldots, A_m\}$, где все множества $A_i \subseteq \Omega = \{1, \ldots, n\}$. Сопоставим каждому элементу $j \in \Omega$ некоторое значение x_j, которое будет меняться по ходу доказательства. Первоначально пусть все $x_j = 0$. В конечном итоге, все $x_j = \pm 1$. Переменная x_j будет принимать значения из интервала $(-1, +1)$, до тех пор пока не примет одно из значений ± 1, которое и станет конечным. Множество S_i (по определению) имеет значение $\sum_{j \in S_i} x_j$. Назовем элемент j *фиксированным*, если $x_j = \pm 1$, в противном случае назовем его *плавающим*. Назовем множество S_i *безопасным*, если оно содержит менее чем t плавающих точек, в противном случае назовем его *активным*. Отметим важный факт, что так как (по условию) каждая точка содержится не более чем в t множествах, и активное множество содержит более чем t плавающих точек, то активных множеств должно быть меньше чем плавающих точек.

Утверждается, что в каждый момент все активные множества имеют значение ноль. Это выполнено изначально, так как все множества имеют значение ноль. Предположим, что утверждение верно на некотором шаге. Будем рассматривать x_j как переменные для плавающих j и как константы для фиксированных j. Условие, что множество S_i имеет значение ноль, равносильно линейному уравнению для соответствующих переменных. Получаем неопределенную систему линейных уравнений, т. е. содержащую меньше уравнений (активных множеств), чем переменных (плавающих точек). Следовательно, мы можем найти прямую, параметризованную следующим образом:

$$x'_j = x_j + \lambda y_j, \quad \text{для плавающих } j,$$

на которой активные множества сохраняют значение ноль. Возьмем наименьшее значение λ, при котором какое-либо значение x'_j становится равным ± 1, и заменим все x_j на x'_j. (Геометрически это означает перемещение по прямой до точки ее пересечения с кубом в пространстве плавающих переменных.) В результате этого процесса фиксированные переменные остаются фиксированными и, следовательно, безопасные множества остаются безопасными (в то время как некоторые активные множества могут стать безопасными), а требуемое условие продолжает выполняться. К тому же, по крайней мере одна из плавающих точек j становится фиксированной.

Будем повторять описанную процедуру до тех пор пока все точки j не станут фиксированными. (На какой-то момент может вовсе не остаться активных множеств, в этом случае присвоим оставшимся плавающим x_j значения ± 1 произвольным образом.) Рассмотрим произвольное множество S_i. Оно имело значение ноль изначально и сохраняло это значение до тех пор, пока в нем было по крайней мере t плавающих точек. Рассмотрим момент, когда множество S_i становится безопасным. Пусть $1, \ldots, l$ — плавающие точки данного множества. На данный момент значение множества равно нулю. Значения переменных

$y_1, \ldots, y_l$ при переходе к их конечному варианту могут теперь измениться только лишь на величину меньшую двух, так как значения всех переменных принадлежат отрезку $[-1, +1]$. Тогда сумма всех изменений должна быть меньше $2t$. Следовательно, окончательное значение S_i меньше $2t$ и, так как оно должно быть целым числом, не превосходит $2t - 1$. ■

Гипотеза 12.5.2. *Пусть* $\deg(\mathcal{A}) \leqslant t$, *тогда* $\operatorname{disc}(\mathcal{A}) \leqslant K\sqrt{t}$, *где* K — *абсолютная константа.*

Видимо, при доказательстве этой гипотезы имеет смысл сочетать вероятностные методы с методами линейной алгебры. Построение семейства из t подмножеств t-множества, описанное в разд. 12.4, показывает, что в случае, если гипотеза верна, она дает оценку наилучшую из возможных.

12.6. УПРАЖНЕНИЯ

1. Пусть $\mathcal{A}$ — семейство из n подмножеств множества $\Omega = \{1, \ldots, m\}$, где m — четное число. Пусть раскраска $\chi(i)$ точек $1 \leqslant i \leqslant \frac{m}{2}$ принимает значения из множества $\{-1, +1\}$ независимо и равновероятно. Положим $\chi(i + \frac{m}{2}) = -\chi(i)$ для $1 \leqslant i \leqslant \frac{m}{n}$. С использованием понятия случайной раскраски улучшить результат теоремы 12.1.1, показав, что

$$\operatorname{disc}(\mathcal{A}) \leqslant \sqrt{\frac{m}{2}\ln(2n)}.$$

2. Пусть $\vec{v}_1, \ldots, \vec{v}_s \in \mathbb{R}^n$. Пусть числа $x_1, \ldots, x_s \in [-1, +1]$ таковы, что $\sum_{i=1}^{s} x_i\vec{v}_i = \vec{0}$, и не более чем n из них отличны от ± 1. Пусть $\vec{v}_{s+1} \in \mathbb{R}^n$. С использованием линейных методов из разд. 12.5 найти числа $x'_1, \ldots x'_s, x'_{s+1}$ со следующими свойствами:

   - $\sum_{i=1}^{s+1} x'_i\vec{v}_i = \vec{0}$.
   - Все $x'_i \in [-1, +1]$.
   - Не более чем n чисел x'_i отличны от ± 1.
   - $x'_i = x_i$, если $x_i \in \{-1, +1\}$.

   С использованием полученного результата доказать следующее утверждение, полученное впервые Барани и Гринбергом: Пусть $|\cdot|$ — произвольная норма в $\mathbb{R}^n$. Пусть $\vec{v}_1, \ldots, \vec{v}_s \in \mathbb{R}^n$ и все $|v_i| \leqslant 1$. Тогда найдется набор $x_1, \ldots, x_s \in \{-1, +1\}$, такой, что

$$\left|\sum_{i=1}^{t} x_i\vec{v}_i\right| \leqslant 2n$$

   для всех $1 \leqslant t \leqslant s$.

3. Пусть $A_1, \ldots, A_n \subset \Omega = \{1, \ldots, m\}$, где $m \sim n \ln n$. Пусть $|A_i| \leqslant n$. С использованием идей из доказательства теоремы 12.2.1, включая теорему Клейтмана, доказать существование раскраски $\chi : \{1, \ldots, m\} \to \{-1, 0, +1\}$, такой, что $\chi(A_i) = O(\sqrt{n \ln \ln n})$, и $\chi(x) = 0$ для не более чем n точек x. Доказать при помощи теоремы 12.2.1 существование раскраски $\chi : \{1, \ldots, m\} \to \{-1, +1\}$, такой, что $\chi(A_i) = O(\sqrt{n \ln \ln n})$.

ВЕРОЯТНОСТНЫЙ ВЗГЛЯД:

Несбалансированные матрицы

Пусть $B = (b_{ij})$ — матрица размерности $m \times n$, состоящая из элементов $b_{ij} = \pm 1$, положим

$$F[B] = \max_{x_i, y_j = \pm 1} \sum_{i=1}^{m} \sum_{j=1}^{n} x_i y_j b_{ij}.$$

Как и в разд. 2.5, мы можем интерпретировать матрицу B как массив из $m \times n$ лампочек, каждая из которых либо включена ($b_{ij} = +1$), либо выключена ($b_{ij} = -1$). Для каждой строки и каждого столбца есть переключатель, поворот которого включает все выключенные и выключает все включенные лампочки в соответствующей линии. Тогда функция $F[B]$ равна максимально возможному значению разности числа включенных и выключенных лампочек. В разд. 2.5 мы получили нижнюю оценку функции $F[B]$ при $m = n$. Положим теперь $n = 2^m$ и найдем точную нижнюю оценку.

Пусть $n = 2^m$ и A — матрица размерности $m \times n$, столбцами которой являются все возможные векторы длины m из ± 1. Утверждается, что $F[A]$ есть наименьшее значение функции $F[B]$ на множестве всех матриц B размерности $m \times n$.

Для произвольной матрицы B выберем независимо и равновероятно случайные величины $x_1, \ldots, x_m = \pm 1$. Пусть

$$X_j = \sum_{i=1}^{m} x_i b_{ij},$$

$$X = |X_1| + \ldots + |X_n|,$$

тогда

$$F[B] = \max_{y_j = \pm 1} \max_{x_i = \pm 1} \sum_{j=1}^{n} y_j X_j = \max_{x_i = \pm 1} \sum_{j=1}^{n} |X_j| = \max X.$$

Независимо от b_{ij} случайная величина X_i имеет распределение $\mathbf{S}_m$, так что $\mathbf{E}[|X_i|] = \mathbf{E}[|\mathbf{S}_m|]$ и, вследствие линейности математического ожидания,

$$\mathbf{E}[X] = n\mathbf{E}[|\mathbf{S}_m|].$$

При $B = A$ каждый выбор $x_1, \ldots, x_m = \pm 1$ приводит, фактически, к перестановке столбцов в матрице A, так как матрица $(x_i a_{ij})$ также состоит из всех возможных столбцов. Следовательно, величина $X = |X_1| + \ldots + |X_n|$ является константой. Заметим, что $\mathbf{E}[X]$ не зависит от B. Заметим также, что при $\mathbf{E}[X] = \mu$ наименьшее возможное значение $\max X$ достигается, когда величина X есть константа μ. Следовательно, функция $F[B]$ достигает минимума при $B = A$.

13

Геометрия

Мало кто задумывается чаще двух или трех раз в год. Я приобрел мировое имя, думая один или два раза в неделю.

Джордж Бернард Шоу

Выберем случайно и равномерно n точек $P_1, \ldots, P_n$ на единичной окружности. Какова вероятность того, что начало координат окажется внутри выпуклой оболочки этих точек? Есть удивительно простой (и остроумный) способ вычислить эту вероятность. Выберем n случайных пар противоположных точек $Q_1, Q_{n+1} = -Q_1$, $Q_2, Q_{n+2} = -Q_2, \ldots, Q_n, Q_{2n} = -Q_n$ в соответствии с равномерным распределением. Заметим, что с вероятностью 1 все эти пары различны. Далее выберем в качестве P_i с равной вероятностью или точку Q_i, или противоположную ей точку $Q_{n+i} = -Q_i$. Ясно, что эта процедура соответствует случайному выбору точек P_i. Вероятность того, что начало координат *не* окажется внутри выпуклой оболочки точек P_i, при заданных (различных) точках Q_j есть в точности $\frac{x}{2^n}$, где x — количество подмножеств множества точек Q_j, таких, что все точки подмножества содержатся в открытой полуплоскости, определяемой прямой, которая проходит через начало координат и не проходит ни через одну из точек Q_j. Легко видеть, что $x = 2n$. Действительно, перенумеруем точки Q_j так, чтобы порядок их расположения на окружности был следующим: $Q_1, \ldots, Q_n, Q_{n+1}, \ldots, Q_{2n}$, и $Q_{n+i} = -Q_i$. Тогда подмножества, содержащиеся в таких полуплоскостях — это множества вида $\{Q_i, \ldots, Q_{n+i-1}\}$, где все индексы берутся по модулю $2n$. Следовательно, вероятность того, что начало координат содержится внутри выпуклой оболочки n случайно выбранных точек единичной окружности, в точности равна $1 - \frac{2n}{2^n}$. Отметим, что аналогичный результат можно получить, заменив единичную окружность на произвольную ограниченную центрально симметричную область плоскости с центром в начале координат, а также, что этот результат может быть с легкостью обобщен на случай пространств больших размерностей.

Этот результат [Wendel (1962)] показывает, как в некоторых случаях можно с помощью хорошей идеи избавиться от утомительных вычислений. Также на этом примере хорошо видна связь между вероятностью и геометрией. В

последнее время вероятностные методы широко используются в дискретной и вычислительной геометрии. Некоторые из результатов, полученных при помощи вероятностных методов, приведены в данной главе.

13.1. НАИБОЛЬШИЙ УГОЛ МЕЖДУ ТОЧКАМИ В ЕВКЛИДОВОМ ПРОСТРАНСТВЕ

Есть несколько замечательных примеров того, как в самых разных областях комбинаторики вероятностный метод позволял находить очень простые контрпримеры для давно сформулированных гипотез. Следующий результат [Erdős and Fűredi (1983)] является одним из них.

Теорема 13.1.1. *Для каждого $d \geqslant 1$ найдется множество, состоящее не менее чем из $\lfloor \frac{1}{2}(\frac{2}{\sqrt{3}})^d \rfloor$ точек d-мерного евклидова пространства $\mathbb{R}^d$, такое, что все углы, образованные тройками точек этого множества, строго меньше $\pi/2$.*

Эта теорема опровергает старую гипотезу [Danzer and Grünbaum (1962)] о том, что максимальная мощность такого множества не превосходит $2d-1$. Доказано (см. [Danzer and Grünbaum (1962)]), что максимальная мощность множества точек $\mathbb{R}^d$, в котором все углы не превосходят $\pi/2$, есть 2^d.

Доказательство теоремы 13.1.1. Будем выбирать точки множества $X \subset \mathbb{R}^d$ среди вершин d-мерного куба. Как обычно будем рассматривать вершины куба, которые являются двоичными векторами длины d, как характеристические векторы подмножеств d-множества. То есть каждому $0,1$-вектору a длины d соответствует множество $A = \{i : 1 \leqslant i \leqslant d, a_i = 1\}$. Три вершины a, b и c в d-мерном кубе, соответствующие множествам A, B и C, образуют прямой угол в точке c тогда и только тогда, когда

$$A \cap B \subset C \subset A \cup B, \tag{13.1}$$

это является простым следствием теоремы Пифагора. Так как все углы, образованные тройками вершин d-мерного куба, не превосходят $\pi/2$, нам достаточно сконструировать множество X мощности не менее $\lfloor \frac{1}{2}(\frac{2}{\sqrt{3}})^d \rfloor$, никакие три различных элемента которого не удовлетворяют условию (13.1).

Положим $m = \lfloor \frac{1}{2}(\frac{2}{\sqrt{3}})^d \rfloor$ и выберем случайным образом и независимо $2m$ двоичных d-мерных векторов $a_1, \ldots, a_{2m}$ так, чтобы значение каждой координаты каждого вектора независимо выбиралось равным 0 или 1 с вероятностью $\frac{1}{2}$. Для каждой тройки элементов a, b и c выбранного множества вероятность того, что соответствующие этим точкам множества удовлетворяют условию (13.1), в точности равна $(3/4)^d$. Действительно, вложение (13.1) означает, что для всех i $(1 \leqslant i \leqslant d)$ не верно ни то, что $a_i = b_i = 0, c_i = 1$, ни то, что $a_i = b_i = 1, c_i = 0$. Следовательно, вероятность того, что для трех фиксированных индексов i, j и k выбранные нами точки a_i, a_j, a_k образуют прямой угол в точке a_k, равна $(3/4)^d$. Так как количество всевозможных троек, которые могут образовать прямой угол, равно $\binom{2m}{3}3$, то математическое

ожидание числа прямых углов равно

$$\binom{2m}{3} 3(3/4)^d \leqslant m,$$

где последнее неравенство следует из выбора m. Тогда найдется множество X из $2m$ точек, в котором количество прямых углов не превосходит m. Удаляя по одной точке из каждого такого угла, мы получаем множество мощности не менее чем $2m-m=m$, в котором все углы строго меньше $\pi/2$. Заметим, что все оставшиеся точки различны, так как условие (13.1) выполняется тривиально при $A=C$. Теорема полностью доказана. ■

Как было показано Эрдёшем и Фюреди, приведенное выше доказательство легко модифицируется для получения следующего результата.

Теорема 13.1.2. *Для любого $\varepsilon > 0$ существует $\delta > 0$, такое, что для каждого $d \geqslant 1$ найдется множество из не менее чем $(1+\delta)^d$ точек пространства $\mathbb{R}^d$, удовлетворяющее условию, что все углы, определяемые тройками различных его точек, не превосходят $\pi/3+\varepsilon$.*

Мы опустим подробное доказательство этого утверждения.

13.2. ПУСТЫЕ ТРЕУГОЛЬНИКИ, ОПРЕДЕЛЯЕМЫЕ ТОЧКАМИ ПЛОСКОСТИ

Пусть X — множество точек плоскости, находящихся в общем положении (т. е. никакая тройка точек не лежит на одной прямой). Обозначим через $f(X)$ число *пустых треугольников*, т. е. таких, которые не содержат другие точки множества X, определяемых тройками точек множества X. В работе [Katchalski and Meir (1988)] изучалось минимально возможное значение функции $f(X)$ для множеств X мощности n. Положим $f(n)=\min\{f(X)\}$, где минимум берется по всем множествам X, состоящим из n точек плоскости, находящихся в общем положении. Было доказано, что

$$\binom{n-1}{2} \leqslant f(n) < 200n^2.$$

Эти оценки были улучшены в работе [Bárány and Fűredi (1987)], где показано, что при растущих n справедливы соотношения

$$(1+o(1))n^2 \leqslant f(n) \leqslant (1+o(1))2n^2.$$

Для получения верхней оценки была использована вероятностная конструкция, представленная в следующей теореме. Несколько лучший результат был получен в работе [Valtr (1995)].

Теорема 13.2.1. *Пусть $I_1, I_2, \ldots, I_n$ — параллельные единичные отрезки на плоскости:*

$$I_i = \{(x,y) : x=i, 0 \leqslant y \leqslant 1\}.$$

Для каждого i выберем точку p_i независимо и случайным образом из отрезка I_i в соответствии с равномерным распределением. Обозначим через X множество этих n случайно выбранных точек. Тогда математическое ожидание числа пустых треугольников в X не превосходит $2n^2 + O(n \log n)$.

Очевидно, что с вероятностью 1 множество X является множеством точек в общем положении и, следовательно, из теоремы вытекает неравенство $f(n) \leqslant 2n^2 + O(n \log n)$.

Доказательство. Сначала мы оценим вероятность того, что треугольник, определенный точками p_i, p_{i+a} и p_{i+k}, пуст при некоторых фиксированных i, a и $k = a + b \geqslant 3$. Обозначим через $A = (i, x)$, $B = (i + a, y)$ и $C = (i + k, z)$ точки p_i, p_{i+a} и p_{i+k}, соответственно, а через m — расстояние между точкой B и точкой пересечения отрезка AC с отрезком I_{i+a}. Так как все точки p_j, где $i < j < i + k$, были выбраны случайно и равномерно на I_j, вероятность того, что треугольник, определенный точками A, B и C, пуст, равна

$$\begin{aligned}
&\left(1 - \frac{m}{a}\right)\left(1 - 2\frac{m}{a}\right) \ldots \left(1 - (a-1)\frac{m}{a}\right)\left(1 - (b-1)\frac{m}{b}\right) \ldots \left(1 - \frac{m}{b}\right) \leqslant \\
&\quad \leqslant \exp\left(-\frac{m}{a} - 2\frac{m}{a} \ldots - (a-1)\frac{m}{a} - (b-1)\frac{m}{b} \ldots - \frac{m}{b}\right) = \\
&\quad = \exp\left(-\binom{a}{2}\frac{m}{a} - \binom{b}{2}\frac{m}{b}\right) = \exp\left(-(k-2)\frac{m}{2}\right).
\end{aligned}$$

Зафиксируем произвольные точки A и C при случайном выборе точки $p_{i+a} = B$. Вероятность того, что она находится на расстоянии $m \leqslant d$ от точки пересечения отрезка AC с отрезком I_{i+a}, очевидно, не превосходит $2d$ для каждого $d \geqslant 0$. Следовательно, вероятность того, что треугольник, определенный точками p_i, p_{i+a} и p_{i+k}, пуст, не превосходит

$$2\int_{m \geqslant 0} \exp(-(k-2)m/2)dm = 4/(k-2).$$

Отсюда следует, что математическое ожидание общего количества пустых треугольников не превосходит

$$\begin{aligned}
n - 2 + &\sum_{1 \leqslant i \leqslant n-3} \sum_{3 \leqslant k \leqslant n-i} \sum_{1 \leqslant a \leqslant k-1} 4/(k-2) = \\
&= n - 2 + \sum_{3 \leqslant k \leqslant n-1} (n-k)\frac{4(k-1)}{k-2} = \\
&= n - 2 + \sum_{3 \leqslant k \leqslant n-1} (n-k)4/(k-2) + 4 \sum_{3 \leqslant k \leqslant n-1} (n-k) = \\
&= 2n^2 + O(n \log n).
\end{aligned}$$

Теорема полностью доказана. ■

Приведенный выше результат может быть обобщен на случай больших размерностей с помощью схожей вероятностной конструкции. Множество X

из n точек d-мерного евклидова пространства назовем *независимым*, если никакие $d+1$ точек не лежат в гиперплоскости. Симплекс, определенный $d+1$ точками, назовем *пустым*, если он не содержит никакой другой точки X. Через $f_d(X)$ обозначим число пустых симплексов множества X, положим $f_d(n) = \min f_d(X)$, где минимум берется по всем независимым множествам X, состоящим из n точек пространства $\mathbb{R}^d$. В работе [Katchalski and Meir (1988)] было доказано, что $f_d(n) \geqslant \binom{n-1}{d}$. Следующая теорема Барани и Фюреди показывает, что здесь вероятностная конструкция снова дает нам верхнюю оценку, совпадающую с нижней с точностью до постоянного множителя (который зависит от размерности). Мы опустим подробное доказательство.

Теорема 13.2.2. *Найдется такая константа $K = K(d)$, что для каждого выпуклого ограниченного множества $A \subset \mathbb{R}^d$ с непустой внутренней частью выполнено следующее. Если X — случайное множество из n точек, полученное в результате n случайных и независимых выборов точек множества A в соответствии с равномерным распределением, то математическое ожидание числа пустых симплексов множества X не превосходит $K\binom{n}{d}$.*

13.3. ГЕОМЕТРИЧЕСКАЯ РЕАЛИЗАЦИЯ ±1-МАТРИЦ

Пусть $A = (a_{i,j})$ — матрица размера m на n, состоящая из элементов $+1$ и -1. Будем говорить, что матрица A *реализуема* в пространстве $\mathbb{R}^d$, если в $\mathbb{R}^d$ найдется m проходящих через начало координат гиперплоскостей $H_1, \ldots, H_m$, и n точек $P_1, \ldots, P_n$, таких, что для всех i и j точка P_j лежит в положительном полупространстве относительно H_i при $a_{i,j} = +1$, и в отрицательном — при $a_{i,j} = -1$. Обозначим через $d(A)$ минимальную размерность d, при которой A реализуема в $\mathbb{R}^d$, и положим $d(m,n) = \max(d(A))$, где максимум берется по всем матрицам A размера m на n, состоящим из $+1$ и -1. Так как $d(m,n) = d(n,m)$, мы можем ограничиться рассмотрением случая $m \geqslant n$.

Проблема нахождения или оценки $d(m,n)$ (в частности, $d(n,n)$) была поставлена в работе [Paturi and Simon (1984)]. Постановка этого вопроса была обоснована попыткой оценить максимально возможную вероятностную коммуникативную сложность булевой функции без ограничений на ошибки. В работе [Alon, Frankl and Rödl (1985)] было доказано, что $n/32 \leqslant d(n,n) \leqslant (\frac{1}{2} + o(1))n$ при растущем n. Обе оценки были получены путем комбинирования вероятностных и некоторых других идей. В следующей теореме доказывается верхняя оценка, вероятно, наиболее близкая к точной.

Теорема 13.3.1. *Для всех $m \geqslant n$ справедливо неравенство*[1)]

$$d(m,n) \leqslant (n+1)/2 + \sqrt{\frac{n-1}{2}\log m}.$$

[1)] Здесь и далее в этой главе запись $\log a$ означает $\log_2 a$. — *Прим. ред.*

Для доказательства нам потребуются одно определение и две леммы. Пусть $\mathbf{a} = (a_1, \ldots, a_n)$ — вектор из $+1$ и -1. Назовем *числом перемен знака* вектора $\mathbf{a}$ количество индексов i, $1 \leqslant i \leqslant n-1$, таких, что $a_i = -a_{i+1}$. Для каждой матрицы A, состоящей из $+1$ и -1, обозначим через $s(A)$ максимальное число перемен знака в строках матрицы A.

Лемма 13.3.2. *Для каждой матрицы A, состоящей из $+1$ и -1, выполняется неравенство $d(A) \leqslant s(A) + 1$.*

Доказательство. Пусть $A = (a_{i,j})$ — матрица размера m на n, состоящая из $+1$ и -1, положим $s = s(A)$. Пусть $t_1 < t_2 < \ldots < t_n$ — произвольные действительные числа. Определим n точек $P_1, P_2, \ldots, P_n$ пространства $\mathbb{R}^{s+1}$, положив $P_j = (1, t_j, t_j^2, \ldots, t_j^s)$. Эти точки, последние s координат которых представляют собой точки на d-мерной кривой моментов, будут использоваться при реализации матрицы A. Для завершения доказательства нам требуется показать, что каждая строка матрицы A может быть реализована подходящей гиперплоскостью, проходящей через начало координат. Это можно доказать, применив некоторые известные свойства кривой моментов следующим образом. Рассмотрим вектор из $+1$ и -1, соответствующий произвольной строке матрицы A. Пусть число перемен знака этого вектора равно r, где, естественно, $r \leqslant s$. Пусть перемены знака в этом векторе происходят между координатами i_j и $i_j + 1$, $1 \leqslant j \leqslant r$. Выберем произвольные действительные числа $y_1, \ldots, y_r$, где $t_{i_j} < y_j < t_{i_j+1}$ при $1 \leqslant j \leqslant r$. Рассмотрим полином $P(t) = \prod_{j=1}^{r}(t - y_j)$. Так как его степень не превосходит s, найдутся такие действительные числа a_j, что $P(t) = \sum_{j=0}^{s} a_j t^j$. Определим гиперплоскость H в $\mathbb{R}^{s+1}$ как $H = \{(x_0, x_1, \ldots, x_s) \in \mathbb{R}^{s+1} : \sum_{j=0}^{s} a_j x_j = 0\}$. Ясно, что точка $P_j = (1, t_j, \ldots, t_j^s)$ лежит в положительном полупространстве относительно этой гиперплоскости, если $P(t_j) > 0$, и в отрицательном полупространстве, если $P(t_j) < 0$. Из того, что полином P меняет знак только в точках y_j, следует, что гиперплоскость H разделяет точки $P_1, \ldots, P_n$ в соответствии с набором знаков из соответствующей строки матрицы A. Следовательно, правильным образом выбирая ориентацию гиперплоскостей H, мы получаем реализацию матрицы A в пространстве $\mathbb{R}^{s+1}$. ■

Лемма 13.3.3. *Для каждой матрицы A, состоящей из $+1$ и -1, найдется матрица B, полученная из матрицы A путем умножения некоторых столбцов на -1, такая, что $s(B) \leqslant (n-1)/2 + \sqrt{\frac{n-1}{2} \log m}$.*

Доказательство. Выберем для каждого столбца матрицы A случайно, независимо и равновероятно число $\varepsilon \in \{+1, -1\}$. Умножим этот столбец на ε. Обозначим через B случайную матрицу из $+1$ и -1, полученную таким образом. Рассмотрим произвольную фиксированную строку матрицы B. Нетрудно убедиться, что случайная величина, описывающая число перемен знака в этой строке, имеет биномиальное распределение с параметрами $n-1$ и $p = 1/2$. Это следует из того, что независимо от значений элементов матрицы A в этой строке, строка матрицы B является полностью случайным вектором из -1

и 1. С помощью стандартных оценок для биномиального распределения, приведенных в приложении А, получаем, что вероятность того, что это число превосходит $(n-1)/2+\sqrt{\frac{n-1}{2}\log m}$, меньше $1/m$. Следовательно, с ненулевой вероятностью число перемен знака в каждой из m строк не превосходит $(n-1)/2+\sqrt{\frac{n-1}{2}\log m}$, что и требовалось доказать. ■

Доказательство теоремы 13.3.1. Пусть A — произвольная матрица, состоящая из $+1$ и -1. Из леммы 13.3.3 следует, что существует матрица B, полученная из матрицы A путем замены некоторых столбцов на обратные, такая, что $s(B) \leqslant (n-1)/2+\sqrt{\frac{n-1}{2}\log m}$. Отметим, что $d(A) = d(B)$, так как из любой реализации одной из этих матриц точками и гиперплоскостями, проходящими через начало координат, можно получить реализацию другой матрицы путем замены точек, соответствующих измененным столбцам, на противоположные им. Тогда по лемме 13.3.2

$$d(A) = d(B) \leqslant s(B)+1 \leqslant (n+1)/2+\sqrt{\frac{n-1}{2}\log m},$$

что и требовалось доказать. ■

Применяя теорему о шести стандартных отклонениях (в общем виде), сформулированную в конце разд. 12.2, нетрудно показать, что оценка из леммы 13.3.3 (а следовательно, и теоремы 13.3.1) может быть улучшена до $n/2+O(\sqrt{n\log(m/n)})$. Можно также показать, что если n и m возрастают так, что m/n^2 стремится к бесконечности и $(\log_2 m)/n$ стремится к нулю, то для почти всех матриц A размера m на n, состоящих из ± 1, выполняется равенство $d(A) = (\frac{1}{2}+o(1))n$.

13.4. ε-СЕТИ И VC-РАЗМЕРНОСТИ РАНЖИРОВАННЫХ ПРОСТРАНСТВ

Пусть $f = f(n,\varepsilon)$ — минимальное число, такое, что каждое множество X из n точек плоскости содержит подмножество S из не более чем f точек, такое, что каждый треугольник, содержащий не менее εn точек X, содержит по крайней мере одну точку множества S. Наша задача — найти f. Как мы убедимся в этом разделе, существует абсолютная константа c, такая, что $f(n,\varepsilon) \leqslant \frac{c}{\varepsilon}\log(1/\varepsilon)$, и эта оценка верна для всех n. Этот несколько неожиданный результат на самом деле является очень частным случаем теоремы, доказанной в работе [Вапник и Червоненкис (1971)], и обобщенной затем в работе [Haussler and Welzl (1987)]. У этой теоремы есть множество интересных применений в вычислительной геометрии и математической статистике. Для изложения этого результата нам понадобятся несколько определений. *Ранжированное пространство* (range space) — это пара $S = (X,R)$, где X — множество (конечное или бесконечное) и R — семейство (конечное или бесконечное) подмножеств множества X. Элементы множества X будем называть *точками*, а элементы семейства R — *интервалами*. Пусть A — подмножество множества X. *Проекцией* семейства

R на множество A назовем семейство $P_R(A) = \{r \cap A : r \in R\}$. В случае, если проекция содержит все подмножества множества A, будем говорить, что множество A *расщеплено*. Размерностью *Вапника—Червоненкиса* (или *VC-размерностью* от Vapnik—Chervonenkis) пространства S, обозначаемую как $\mathrm{VC}(S)$, назовем максимальную мощность расщепленного подмножества множества X. В случае, если существуют расщепленные подмножества сколь угодно большой мощности, положим $\mathrm{VC}(S) = \infty$.

Количество интервалов в любом конечном ранжированном пространстве с заданным количеством точек и заданной VC-размерностью не может быть слишком большим. Для произвольных целых чисел $n \geqslant 0$ и $d \geqslant 0$ определим функцию $g(d, n)$ равенством

$$g(d, n) = \sum_{i=0}^{d} \binom{n}{i}.$$

Отметим, что для всех $n, d \geqslant 1$ справедливо соотношение $g(d, n) = g(d, n-1) + g(d-1, n-1)$. Следующая комбинаторная лемма была доказана независимо Перлесом и Шелла, в работе [Sauer (1972)] и, в немного более слабой форме, Вапником и Червоненкисом.

Лемма 13.4.1. *Пусть (X, R) — ранжированное пространство* VC*-размерности d, $|X| = n$. Тогда $|R| \leqslant g(d, n)$.*

Доказательство. Воспользуемся индукцией по $n + d$. Утверждение тривиальным образом выполняется при $d = 0$ и $n = 0$. Предположив, что оно верно для n и $d - 1$, а также для $n - 1$ и $d - 1$, докажем, что утверждение также верно для n и d. Пусть $S = (X, R)$ — ранжированное пространство VC-размерности d на n точках. Пусть $x \in X$, рассмотрим два ранжированных пространства $S - x$ и $S \setminus x$, определенных следующим образом: $S - x = (X - \{x\}, R - x)$, где $R - x = \{r - \{x\} : r \in R\}$; $S \setminus x = (X - \{x\}, R \setminus x)$, где $R \setminus x = \{r \in R : x \notin r, r \cup \{x\} \in R\}$. Очевидно, что VC-размерность пространства $S - x$ не превосходит d. Также легко показать, что VC-размерность пространства $S \setminus x$ не превосходит $d - 1$. Следовательно, по предположению индукции,

$$|R| = |R - x| + |R \setminus x| \leqslant g(d, n-1) + g(d-1, n-1) = g(d, n),$$

что и требовалось доказать. ■

Несложно проверить, что оценка, приведенная в предыдущей лемме, является точной для всех возможных значений n и d. Пусть (X, R) — ранжированное пространство VC-размерности d, $A \subset X$, тогда VC-размерность пространства $(A, P_R(A))$, очевидно, не превосходит d. Из предыдущей леммы вытекает следующее утверждение.

Следствие 13.4.2. *Пусть (X, R) — ранжированное пространство VC-размерности d. Тогда $|P_R(A)| \leqslant g(d, |A|)$ для каждого конечного подмножества A множества X.*

Ранжированные пространства с конечными VC-размерностями естественным образом возникают в дискретной и вычислительной геометрии. Примером может служить пространство $S = (\mathbb{R}^d, H)$, множество точек которого составляют все точки d-мерного евклидова пространства, а множеством интервалов является множество всех (открытых) полупространств. Любое множество из $d+1$ аффинно независимых точек расщеплено в этом пространстве, и по теореме Радона никакое множество из $d+2$ точек не будет расщепленным. Следовательно, $\mathrm{VC}(S) = d+1$. Как было показано в работе [Dudley (1978)], в случае, если пространство (X, R) имеет конечную VC-размерность, то это верно и для пространства (X, R_k), где R_k есть множество всех булевых комбинаций, образованных не более чем k интервалами из R. В частности, из следствия 13.4.2 легко может быть получено следующее утверждение.

Следствие 13.4.3. *Пусть (X, R) — ранжированное пространство VC-размерности $d \geqslant 2$, а (X, R_h) — ранжированное пространство на множестве X, в котором $R_h = \{(r_1 \cap \ldots \cap r_h) : r_1, \ldots r_h \in R\}$. Тогда* $\mathrm{VC}(X, R_h) \leqslant 2dh\log(dh)$.

Доказательство. Пусть A — произвольное подмножество множества X мощности n. Из следствия 13.4.2 вытекает, что $|P_R(A)| \leqslant g(d,n) \leqslant n^d$. Так как каждый интервал пространства $P_{R_h}(A)$ представляет собой пересечение h интервалов пространства $P_R(A)$, то $|P_{R_h}(A)| \leqslant \binom{g(d,n)}{h} \leqslant n^{dh}$. Тогда при $n^{dh} < 2^n$ множество A не может быть расщепленным. Но это неравенство выполняется для всех $n \geqslant 2dh\log(dh)$ при $dh \geqslant 4$. ■

Как было показано выше, ранжированное пространство, множество точек которого составляют все точки евклидова пространства $\mathbb{R}^d$, и множеством интервалов которого является множество всех полупространств, имеет VC-размерность $d+1$. Отсюда и из последнего следствия вытекает, что ранжированное пространство $(\mathbb{R}^d, C_h)$, где C_h — множество всех выпуклых d-многогранников с h гранями, имеет VC-размерность, не превосходящую $2(d+1)h\log((d+1)h)$.

Интересным свойством ранжированных пространств с конечной VC-размерностью является то, что каждое конечное подмножество такого множества содержит относительно мало хороших образцов в смысле, раскрытом ниже. Пусть (X, R) — ранжированное пространство и A — конечное подмножество множества X. Для каждого $0 \leqslant \varepsilon \leqslant 1$ множество $B \subset A$ назовем *ε-образцом* множества A, если для всех $r \in R$ справедливо неравенство

$$||A \cap r|/|A| - |B \cap r|/|B|| \leqslant \varepsilon.$$

Аналогично, множество $N \subset A$ назовем *ε-сетью* множества A, если каждый интервал $r \in R$, такой, что $|r \cap A| > \varepsilon|A|$, содержит по крайней мере одну точку из N.

Заметим, что каждый ε-образец множества A также является ε-сетью, а обратное неверно. Тем не менее, оба понятия определяют подмножества множества A, которые дают приблизительное представление о части свойств множества A, касающихся интервалов. Наша цель — показать существование

маленьких ε-сетей или ε-образцов конечных множеств в некоторых ранжированных пространствах. Отметим, что если (X, R) — ранжированное пространство с бесконечной VC-размерностью, то для каждого n найдется расщепленное подмножество A множества X мощности n. Очевидно, что каждая ε-сеть (и, следовательно, любой ε-образец) такого множества A должна содержать не менее $(1-\varepsilon)n$ точек, т. е. почти все точки множества A. Таким образом, при бесконечных VC-размерностях не существует маленьких сетейобразцов. Тем не менее, оказывается, что в случае конечных VC-размерностей всегда находятся очень маленькие сети и образцы. Следующая теорема была доказана в работе [Вапник и Червоненкис (1971)].

Теорема 13.4.4. *Пусть (X, R) — произвольное ранжированное пространство VC-размерности, не превосходящей d, A — конечное подмножество множества X и $\varepsilon, \delta > 0$. Тогда существует положительная константа c такая, что если s не меньше минимума из чисел $|A|$ и*

$$\frac{c}{\varepsilon^2}\left(d\log\frac{d}{\varepsilon}+\log\frac{1}{\delta}\right),$$

то случайное подмножество B множества A, такое, что $|B| = s$, является ε-образцом множества A с вероятностью не менее $1-\delta$.

С использованием аналогичных идей в работе [Haussler and Welzl (1987)] была доказана следующая теорема.

Теорема 13.4.5. *Пусть (X, R) — ранжированное пространство VC-размерности d, A — конечное подмножество множества X, пусть, кроме того, $0 < \varepsilon, \delta < 1$. Обозначим через N множество, полученное путем случайного и независимого выбора m элементов из множества A, где*

$$m \geqslant \max\left(\frac{4}{\varepsilon}\log\frac{2}{\delta}, \frac{8d}{\varepsilon}\log\frac{8d}{\varepsilon}\right). \tag{13.2}$$

Тогда N является ε-сетью множества A с вероятностью не менее $1-\delta$.

Таким образом, из того, что A является конечным подмножеством ранжированного пространства конечной VC-размерности d, следует, что для любого $\varepsilon > 0$ множество A содержит как ε-сеть, так и ε-образец, размеры которых не превосходят значения некоторой функции, зависящей от ε и d, независимо от мощности A! Результат, касающийся треугольников, упомянутый в первом разделе этой главы, таким образом, следует из теоремы 13.4.5. Отсюда, используя рассуждения, приведенные после следствия 13.4.3, мы получаем, что ранжированное пространство, интервалами которого являются все треугольники на плоскости, имеет конечную VC-размерность. Как было показано в работе [Pach and Woeginger (1990)], существуют примеры, когда при фиксированном δ зависимость m от $1/\varepsilon$ не может быть линейной, но до сих пор не известны естественные геометрические примеры, на которых проявлялся бы этот феномен. В работе [Komlós, Pach and Woeginger (1992)] содержится более короткая формулировка последней теоремы.

Доказательства теорем 13.4.4 и 13.4.5 очень похожи. Мы приводим только доказательство теоремы 13.4.5, так как присутствующие в нем вычисления проще, и предлагаем читателю попробовать модифицировать его для получения доказательства теоремы 13.4.4.

Доказательство теоремы 13.4.5. Пусть (X, R) — ранжированное пространство VC-размерности d, A — подмножество множества X и $|A| = n$. Пусть m удовлетворяет условию (13.2), и множество $N = (x_1, \ldots, x_m)$ получено путем независимого случайного выбора m элементов из множества A. (Разумеется, элементы множества N не обязательно все различны). Обозначим через E_1 следующее событие:

$$E_1 = \{\exists r \in R : |r \cap A| > \varepsilon n, r \cap N = \varnothing\}.$$

Нам достаточно показать, что вероятность события E_1 не превосходит δ. Для этого мы проведем дополнительный случайный выбор и определим другое событие. Независимо от множества N определим множество $T = (y_1, \ldots, y_m)$, полученное путем случайного и независимого выбора m элементов из множества A. Определим событие E_2 следующим образом:

$$E_2 = \left\{\exists r \in R : |r \cap A| > \varepsilon n, r \cap N = \varnothing, |r \cap T| \geqslant \frac{\varepsilon m}{2}\right\}.$$

(Так как элементы множества T не обязательно все различны, то обозначение $|r \cap T|$ понимается здесь в смысле $|\{i : 1 \leqslant i \leqslant m, y_i \in r\}|$. Аналогично определим величины $|r \cap N|$ и $|r \cap (N \cup T)|$.)

Утверждение 13.4.6. *Справедливо неравенство* $\Pr[E_2] \geqslant \frac{1}{2}\Pr[E_1]$.

Доказательство. Достаточно оценить условную вероятность $\Pr[E_2|E_1]$ и показать, что она не менее $1/2$. Предположим, что произошло событие E_1. Тогда найдется такое $r \in R$, что $|r \cap A| > \varepsilon n$ и $r \cap N = \varnothing$. Условная вероятность, упомянутая выше, очевидно, не меньше вероятности того, что для этого фиксированного r верно $|r \cap T| \geqslant \frac{\varepsilon m}{2}$. С другой стороны, величина $|r \cap T|$ есть биномиальная случайная величина с математическим ожиданием pm и дисперсией $p(1-p)m \leqslant pm$, где $p = |r \cap A|/|A| > \varepsilon$. Тогда из неравенства Чебышёва следует, что

$$\Pr\left[|r \cap T| < \frac{\varepsilon m}{2}\right] \leqslant \frac{\varepsilon m}{(\varepsilon m/2)^2} = \frac{4}{\varepsilon m} \leqslant 1/2,$$

где последнее неравенство следует из соотношения (13.2). Таким образом, наше утверждение верно. ■

Утверждение 13.4.7. *Справедливо неравенство*

$$\Pr[E_2] \leqslant g(d, 2m)2^{-\frac{\varepsilon m}{2}}.$$

Доказательство. Случайный выбор множеств N и T может быть описан следующим эквивалентным предыдущему способом. Сформируем сначала множество $N \cup T = (z_1, \ldots, z_{2m})$, сделав $2m$ случайных независимых выборов элементов из множества A. Затем выберем случайным образом m элементов

z_i, которые образуют множество N (оставшиеся элементы z_j, естественно, образуют множество T). Для каждого интервала $r \in R$, такого, что $|r \cap A| > \varepsilon n$, обозначим через E_r событие, что $|r \cap T| > \frac{\varepsilon m}{2}$ и $r \cap N = \varnothing$. Ключевым моментом является то, что в случае, если для двух интервалов $r, r' \in R$ выполнено $|r \cap A| > \varepsilon n$ и $|r' \cap A| > \varepsilon n$, и если $r \cap (N \cup T) = r' \cap (N \cup T)$, то два события E_r и $E_{r'}$ будут совпадать при условии выбора определенного множества $N \cup T$. Действительно, возникновение события E_r зависит только от пересечения $r \cap (N \cup T)$. Следовательно, для каждого фиксированного множества $N \cup T$ число различных событий E_r не превосходит числа различных множеств проекции $P_{N \cup T}(R)$. Так как VC-размерность множества X равна d, из следствия 13.4.2 вытекает, что это число не превосходит $g(d, 2m)$.

Оценим теперь вероятность возникновения события E_r при заданном выборе множества $N \cup T$. Эта вероятность не превосходит

$$\Pr\left[r \cap N = \varnothing \,\middle|\, |r \cap (N \cup T)| > \frac{\varepsilon m}{2}\right].$$

Положим $l = |r \cap (N \cup T)|$. Так как выбор элементов множества N среди элементов множества $N \cup T$ не зависит от выбора самого множества $N \cup T$, указанная выше условная вероятность в точности равна

$$\frac{(2m-l)(2m-l-1)\dots(m-l+1)}{2m(2m-1)\dots(m+1)} = \frac{m(m-1)\dots(m-l+1)}{2m(2m-1)\dots(2m-l+1)} \leqslant 2^{-l} \leqslant 2^{-\frac{\varepsilon m}{2}}.$$

Так как всего может произойти не более чем $g(d, 2m)$ различных событий E_r, вероятность возникновения по крайней мере одного из них при заданном выборе $N \cup T$ не превосходит $g(d, 2m)2^{-\frac{\varepsilon m}{2}}$. Так как эта оценка верна при условии любого выбора множества $N \cup T$, вероятность возникновения события E_2 не превосходит $g(d, 2m)2^{-\frac{\varepsilon m}{2}}$, что и требовалось доказать. ■

Из утверждений 13.4.6 и 13.4.7 следует, что $\Pr[E_1] \leqslant 2g(d, 2m)2^{-\frac{\varepsilon m}{2}}$. Для завершения доказательства теоремы остается показать, что если m удовлетворяет неравенству (13.2), то

$$2g(d, 2m)2^{-\frac{\varepsilon m}{2}} \leqslant \delta.$$

Мы приводим доказательство для случая $d \geqslant 2$. В случае $d = 1$ вычисления еще проще. Так как $g(d, 2m) \leqslant (2m)^d$, достаточно показать, что

$$2(2m)^d \leqslant \delta 2^{\frac{\varepsilon m}{2}},$$

т. е. что

$$\frac{\varepsilon m}{2} \geqslant d \log(2m) + \log \frac{2}{\delta}.$$

Из неравенства (13.2) следует

$$\frac{\varepsilon m}{4} \geqslant \log \frac{2}{\delta},$$

тогда достаточно показать, что

$$\frac{\varepsilon m}{4} \geqslant d \log(2m).$$

Справедливость последнего неравенства для некоторого значения m влечет за собой его справедливость для всех больших значений m. Следовательно, нам

достаточно проверить, что неравенство верно при $m = \frac{8d}{\varepsilon}\log\frac{8d}{\varepsilon}$, т. е.

$$2d\log\frac{8d}{\varepsilon} \geqslant d\log(\frac{16d}{\varepsilon}\log\frac{8d}{\varepsilon}).$$

Последнее неравенство эквивалентно неравенству $\frac{4d}{\varepsilon} \geqslant \log\frac{8d}{\varepsilon}$, которое, конечно же, верно. Теорема полностью доказана. ■

Теоремы 13.4.4 и 13.4.5 используются для построения эффективных структур данных в различных задачах вычислительной геометрии. Рассмотрим следующий тривиальный пример (утверждение вытекает из теоремы 13.4.4): для любого $\varepsilon > 0$ найдется такая константа $c = c(\varepsilon)$, что для каждого n и каждого множества A из n точек плоскости существует структура данных размера $c(\varepsilon)$, которая позволяет для некоторого заданного треугольника на плоскости оценить число точек множества A, находящихся в этом треугольнике, с точностью до аддитивной погрешности εn. Этой структурой данных будет являться множество координат точек, составляющих ε-образец множества A, рассматриваемое как подмножество ранжированного пространства, интервалами которого являются все треугольники плоскости. Более сложные структуры данных, построения которых опираются на две приведенные выше теоремы, можно найти в работе [Haussler and Welzl (1987)].

13.5. ДВОЙСТВЕННАЯ ФУНКЦИЯ РАСЩЕПЛЕНИЯ И РАЗБРОС

Двойственной функцией расщепления h ранжированного пространства $S = (X, R)$ назовем такую функцию h, отображающую множество целых чисел на себя, что $h(g)$ есть максимальное число атомов в диаграмме Венна для g множеств семейства R, где максимум берется по всем возможным наборам из g множеств семейства R. Несложно показать, что если VC-размерность пространства S равна d, то $h(g) \leqslant O(g^{2^{d+1}-1})$, но в геометрических приложениях обычно оказывается, что лучше оценивать эту функцию напрямую для каждого конкретного случая.

В работе [Matoušek, Welzl and Wernisch (1993)] было доказано, что если двойственная функция расщепления ранжированного пространства $S = (X, R)$ удовлетворяет условию $h(g) \leqslant O(g^t)$, A — произвольное множество из n точек ранжированного пространства и $\mathcal{F}$ — проекция $P_R(A)$ семейства R на множество A, то разброс семейства $\mathcal{F}$ удовлетворяет неравенству

$$\operatorname{disc}(\mathcal{F}) \leqslant O(n^{\frac{1}{2}-\frac{1}{2t}}\sqrt{\log n}). \tag{13.3}$$

Это неравенство позволяет доказывать нетривиальные оценки во многих геометрических задачах, в частности, с помощью этого результата можно улучшить тривиальные оценки, вытекающие из теоремы 12.1.1 гл. 12. Распространено мнение, что в большей части этих геометрических задач можно избавиться от множителя $\sqrt{\log n}$. Тем не менее, в общем случае это не так, что было доказано в работе [Matoušek (1997)] (для $t = 2, 3$), и позже — в работе [Alon, Rónyai and Szabó (1999)] для всех t.

Доказательство утверждения (13.3) основывается на замечательном результате из работы [Chazelle and Welzl (1989)], улучшенном в статье [Haussler (1995)]. Несколько проще доказывать это утверждение с добавочным логарифмическим множителем. Это доказательство мы здесь и изложим. Более подробно этот вопрос рассмотрен в работе [Pach and Agarwal (1995)].

Пусть $\mathcal{F}$ — семейство подмножеств конечного множества A. Далее будут рассматриваться графы, ребрами которых являются (неупорядоченные) пары точек множества A. Для каждого множества $F \in \mathcal{F}$ и пары точек $x, y \in A$ будем говорить, что ребро $\{x, y\}$ *прокалывает* множество F, если F содержит ровно одну из двух точек x и y. Следующая теорема была доказана в работе [Chazelle and Welzl (1989)]. Результат был улучшен на логарифмический множитель в статье [Haussler (1995)].

Теорема 13.5.1. *Пусть $(A, \mathcal{F})$ — конечное ранжированное пространство, $|A| = n$, и пусть его двойственная функция расщепления h удовлетворяет условию $h(g) \leqslant cg^t$ для некоторых фиксированных $c, t > 0$. Тогда найдутся константа $C = C(c, t)$ и гамильтонов путь на множестве A, такие, что каждое множество F семейства $\mathcal{F}$ прокалывается не более чем $Cn^{1-1/t} \log n$ ребрами гамильтонова пути.*

Для доказательства этой теоремы нам понадобится следующая лемма.

Лемма 13.5.2. *Пусть $(A, \mathcal{F}), n, h, t$ и c удовлетворяют условиям, сформулированным в теореме 13.5.1, B — конечное подмножество мощности $p > 1$ множества A, $\mathcal{G}$ — семейство из m (не обязательно различных) множеств семейства $\mathcal{F}$. Тогда в множестве B найдутся две различные точки x, y, такие, что ребро xy прокалывает не более чем $\frac{bm \log p}{p^{1/t}}$ множеств семейства $\mathcal{G}$, где $b = b(c)$.*

Доказательство. целое число, такое, что $cg^t \leqslant p - 1$, тогда $g = \lfloor (\frac{p-1}{c})^{1/t} \rfloor$. Обозначим через L случайное семейство из g множеств семейства $\mathcal{G}$, каждое из которых было выбрано случайно и независимо (возможно, с повторениями), среди всех m множеств семейства $\mathcal{G}$ в соответствии с равномерным распределением. Диаграмма Венна всех множеств семейства L разбивает множество B на не более чем $h(g) \leqslant cg^t < p$ атомов, следовательно, найдутся две различных точки x, y множества B, которые находятся в одном и том же атоме. Для завершения доказательства остается показать, что с ненулевой вероятностью для каждой пары точек множества B, которые прокалывают более чем $(bm \log p)/p^{1/t}$ множеств семейства $\mathcal{G}$, по крайней мере одно из этих множеств содержится в семействе L (а значит, пара не содержится в соответствующем атоме диаграммы Венна). Всего таких пар $\binom{p}{2}$, и для каждой из них вероятность того, что L не содержит ни одного из множеств семейства $\mathcal{G}$, которые они прокалывают, не превосходит

$$\left(1 - \frac{b \log p}{p^{1/t}}\right)^g \leqslant e^{-\frac{b \log p}{p^{1/t}} \lfloor (\frac{p-1}{c})^{1/t} \rfloor},$$

что меньше $1/p^2$ при соответствующем выборе константы $b = b(c)$. ■

Доказательство теоремы 13.5.1. Заметим, что, если d — это VC-размерность заданного пространства, то найдется расщепленное множество D мощности d. Несложно видеть, что среди расщепляющих D множеств найдутся $g = \lceil \log_2 d \rceil$ таких, что никакие две точки множества D не принадлежат одному и тому же атому их диаграммы Венна. Тогда $d \leqslant c(\lceil \log_2 d \rceil)^t$, откуда следует, что $d \leqslant 2^{c' t \log t}$, где $c' = c'(c)$. По лемме 13.4.1 отсюда вытекает, что общее количество интервалов R не превосходит $n^{2^{c' t \log t}}$.

Докажем далее, что существует остовное дерево на множестве A, удовлетворяющее утверждению теоремы 13.5.1, а затем покажем, как заменить это дерево гамильтоновым путем. Из леммы 13.5.2 при $B_0 = A, p_0 = n$ и $\mathcal{G}_0 = \mathcal{F}, m_0 = |\mathcal{G}_0|$ $(\leqslant n^{2^{c' t \log t}})$ вытекает, что найдется пара точек x_0, y_0 множества A, такая, что ребро $x_0 y_0$ не прокалывает более $\frac{b \log n}{n^{1/t}} m_0$ множеств семейства $\mathcal{G}$. Пусть $\mathcal{G}_1$ — семейство множеств, полученное из $\mathcal{G}$ дублированием всех множеств семейства $\mathcal{G}$, проколотых ребром $x_0 y_0$. Положим $B_1 = B - x_0, p_1 = n - 1$, $m_1 = |\mathcal{G}_1| \leqslant m_0(1 + \frac{b \log n}{n^{1/t}})$. Вновь применяя лемму 13.5.2, на этот раз для B_1 и $\mathcal{G}_1$, получаем другую пару $x_1 y_1$. Положим $B_2 = B_1 - x_1, p_2 = p_1 - 1 = n - 2$, и обозначим через $\mathcal{G}_2$ семейство множеств, полученное из $\mathcal{G}_1$ дублированием всех множеств семейства $\mathcal{G}_1$, проколотых ребром $x_1 y_1$, $m_2 = |\mathcal{G}_2|$. По утверждению леммы $m_2 \leqslant m_1(1 + \frac{b \log n}{(n-1)^{1/t}})$. Повторяя этот процесс, мы получим последовательность $x_0 y_0, x_1 y_1, \ldots, x_{n-1} y_{n-1}$ ребер графа на множестве A, последовательность подмножеств $B_0 = A, B_1, \ldots, B_{n-1}$, где каждое множество B_i получено из предыдущего исключением точки x_{i-1}, и последовательность семейств $\mathcal{G}_0, \mathcal{G}_1, \ldots, \mathcal{G}_{n-1}$, где

$$|\mathcal{G}_{n-1}| \leqslant m_0 \prod_{i=0}^{n-1} \left(1 + \frac{b \log n}{(n-i)^{1/t}}\right) \leqslant$$
$$\leqslant n^{2^{c' t \log t}} e^{b \log n \sum_{i=0}^{n-1} (n-i)^{-1/t}} \leqslant 2^{b' n^{1-1/t} \log n} \tag{13.4}$$

для соответствующего $b' = b'(c, t)$.

Заметим, что ребра $x_i y_i$ образуют остовное дерево на множестве A. Ключевым моментом доказательства является то, что если множество $\mathcal{F}$ проколото s ребрами, то оно s раз дублировалось в процессе построения $\mathcal{G}_{n-1}$, откуда следует, что $2^s \leqslant |\mathcal{G}_{n-1}|$ и, следовательно, $s \leqslant b' n^{1-1/t} \log n$. Остается заменить остовное дерево гамильтоновым путем. Для этого заменим каждое ребро дерева двумя параллельными ребрами и выделим эйлеров цикл в получившемся графе (в котором степени всех вершин четны). Мы получили последовательность $x_0, x_1, x_2, \ldots, x_{2n-2} = x_0$ точек множества A, такую, что каждая соседняя пара элементов последовательности есть ребро дерева и каждое ребро встречается в последовательности дважды. Подпоследовательность этой последовательности, полученная удалением повторных вхождений точек множества A, является гамильтоновым путем, и несложно проверить, что каждое множество семейства $\mathcal{F}$ проколото не более чем $2b' n^{1-1/t} \log n$ ребрами этого пути. Теорема полностью доказана. ■

Следующий результат представляет собой простое следствие теоремы 13.5.1. Как было упомянуто ранее, ее утверждение может быть улучшено на множитель $\sqrt{\log n}$.

Теорема 13.5.3. *Пусть $(A, \mathcal{F})$ — конечное ранжированное пространство, $|A| = n$, и пусть его двойственная функция расщепления h такова, что $h(g) \leqslant cg^t$ для некоторых фиксированных $c, t > 0$. Тогда найдется константа $C' = C'(c,t)$, такая, что разброс семейства $\mathcal{F}$ удовлетворяет неравенству*

$$\operatorname{disc}(\mathcal{F}) \leqslant C' n^{\frac{1}{2}-\frac{1}{2t}} \log n.$$

Доказательство. Не ограничивая общности рассуждений, мы можем предположить, что число точек множества A четно (в противном случае, удалим одну точку). Из теоремы 13.5.1 следует, что найдется гамильтонов путь $x_1 x_2 \ldots x_n$ на множестве A, такой что каждое множество семейства $\mathcal{F}$ проколото не более чем $Cn^{1-1/t} \log n$ ребрами этого пути. Обозначим через $f : A \mapsto \{-1, 1\}$ случайную раскраску множества A, такую, что для каждого i, $1 \leqslant i \leqslant n/2$, случайно и независимо пара вершин окрашивается либо как $f(x_{2i-1}) = 1, f(x_{2i}) = -1$, либо как $f(x_{2i-1}) = -1, f(x_{2i}) = 1$ с равной вероятностью. Зафиксируем множество $F \in \mathcal{F}$ и отметим, что вклад каждой пары $x_{2i-1}x_{2i}$ в сумму $\sum_{x_j \in F} f(x_j)$ равен нулю, если ребро $x_{2i-1}x_{2i}$ не прокалывает F, и равен $+1$ или -1 в противном случае. Отсюда следует, что эта сумма (в обозначениях теоремы A.1.1) имеет распределение $\mathbf{S}_r$ для некоторого $r \leqslant Cn^{1-1/t} \log n$. Таким образом, вероятность того, что абсолютное значение этой величины не меньше α, может быть оценена как $2e^{-\alpha^2/2r}$. Как было показано в первом разделе доказательства теоремы 13.5.1, общее число множеств семейства $\mathcal{F}$ не превосходит $n^{2^{c't\log t}}$. Тогда вероятность того, что существует множество $F \in \mathcal{F}$, для которого сумма $\sum_{x_j \in F} f(x_j)$ превосходит $C' n^{\frac{1}{2}-\frac{1}{2t}} \log n$, меньше 1 при соответствующем выборе константы $C' = C'(c,t)$. ■

Ранжированное пространство, множеством точек которого является произвольное множество точек плоскости, и множество интервалов которого составляют все круги плоскости, имеет двойственную функцию расщепления, равную $O(g^2)$. Из приведенной выше теоремы следует, что можно так раскрасить произвольные n точек плоскости в красный и синий цвета, что абсолютное значение разности между числом красных и числом синих точек внутри любого круга не будет превосходить $n^{1/4+o(1)}$. Аналогичные результаты можно получить для многих других ранжированных пространств.

13.6. УПРАЖНЕНИЯ

1. Пусть A — множество из n точек плоскости, обозначим через $\mathcal{F}$ семейство пересечений множества A со всеми открытыми треугольниками плоскости. Доказать, что разброс семейства $\mathcal{F}$ не превосходит $n^{1/4+o(1)}$.

2. Доказать, что n различных точек плоскости определяют не более $O(n^{4/3})$ единичных отрезков.

ВЕРОЯТНОСТНЫЙ ВЗГЛЯД:

Эффективная упаковка

Пусть множество $C \subset \mathbb{R}^n$ имеет ограниченную риманову меру $\mu = \mu(C) > 0$. Обозначим через $N(C, x)$ максимальное число непересекающихся копий множества C, которые можно разместить в кубе со стороной x. Определим константу упаковки следующим образом:

$$\delta(C) = \mu(C) \lim_{x\to\infty} N(C, x)x^{-n},$$

т. е. как максимальную долю пространства, которая может быть заполнена непересекающимися копиями множества C. Следующая теорема улучшает результат, описанный в гл. 3, разд. 3.4.

Теорема. *Пусть множество C ограничено, выпукло и центрально-симметрично относительно начала координат. Тогда*

$$\delta(C) \geqslant 2^{-(n-1)}.$$

Доказательство. Зафиксируем $\varepsilon > 0$. Пусть $\mu = \mu(C) = 2 - \varepsilon$. Для каждого действительного z обозначим через C_z вектор $(z_1, \ldots, z_{n-1}) \in \mathbb{R}^{n-1}$, такой, что $(z_1, \ldots, z_{n-1}, z) \in C$, и пусть $\mu(C_z)$ — обычная мера $(n-1)$-мерного вектора C_z. Из измеримости по Риману следует, что

$$\lim_{\gamma\to 0} \sum_{m\in\mathbb{Z}} \mu(C_{m\gamma})\gamma = \mu(C).$$

Обозначим через K целое число, для которого выполняется неравенство

$$\sum_{m\in\mathbb{Z}} \mu(C_{mK^{-(n-1)}})K^{-(n-1)} < 2$$

и, следовательно, все точки множества C имеют координаты меньшие $K/2$.

Для $1 \leqslant i \leqslant n-1$ обозначим через $v_i \in \mathbb{R}^n$ вектор, все координаты которого равны нулю, а i-я координата равна K. Пусть

$$v = (z_1, \ldots, z_{n-1}, K^{-(n-1)}),$$

где действительные числа $z_1, \ldots, z_{n-1}$ выбраны независимо в соответствии с равномерным распределением на интервале $[0, K)$. Обозначим через Λ_v решетку образованную векторами v:

$$\Lambda_v = \{m_1v_1 + \ldots + m_{n-1}v_{n-1} + mv : m_1, \ldots, m_{n-1}, m \in \mathbb{Z}\} =$$
$$= \{(mz_1 + m_1K, \ldots, mz_{n-1} + m_{n-1}K, mK^{-(n-1)} : m_1, \ldots, m_{n-1}, m \in \mathbb{Z}\}.$$

Обозначим через $\theta(x)$ число $x' \in (-\frac{K}{2}, \frac{K}{2}]$, для которого при некотором $m \in \mathbb{Z}$ выполнено $x - mK = x'$. Существует только одно такое число. Для $m \in \mathbb{Z}$ обозначим через A_m событие «некоторая сумма $m_1v_1 + \ldots + m_{n-1}v_{n-1} + mv \in C$».

Так как все координаты точек множества C меньше $K/2$, событие A_m происходит тогда и только тогда, когда

$$(\theta(mz_1), \ldots, \theta(mz_{n-1}), mK^{-(n-1)}) \in C,$$

что верно лишь в том случае, если $(\theta(mz_1), \ldots, \theta(mz_{n-1})) \in C_{mK^{-(n-1)}}$. Из независимости и равномерности выбора чисел z_i на интервале $[0, K)$ вытекает независимость и равномерность распределения величин $\theta(z_i)$ на интервале $(-\frac{K}{2}, \frac{K}{2}]$, следовательно,

$$\Pr[A_m] = K^{-(n-1)}\mu(C_{mK^{-(n-1)}}).$$

Суммируя по всем положительным m и используя центральную симметрию, получаем, что

$$\sum_{m>0} \Pr[A_m] < \frac{1}{2} \sum_{m\in\mathbb{Z}} K^{-(n-1)}\mu(C_{mK^{-(n-1)}}) < \frac{1}{2}2 = 1.$$

Следовательно, *существует* вектор v, для которого ни одно из событий A_m, $m > 0$ не произошло. Из центральной симметрии следует, что A_m и A_{-m} обозначают одно и то же событие. Следовательно, не произошло ни одно из событий A_m при $m \neq 0$. При $m = 0$ все точки $m_1v_1 + \ldots + m_{n-1}v_{n-1} = K(m_1, \ldots, m_{n-1}, 0)$, за исключением начала координат, лежат вне множества C. Для этого v

$$\Lambda_v \cap C = \{0\}.$$

Рассмотрим множество копий $C + 2w, w \in \Lambda_v$. Пусть

$$z = c_1 + 2w_1 = c_2 + 2w_2 \text{ при } c_1, c_2 \in C, w_1, w_2 \in \Lambda_v.$$

Тогда $(c_1 - c_2)/2 = w_2 - w_1$. Из выпуклости и центральной симметрии следует, что $(c_1 - c_2)/2 \in C$. Так как разность $w_2 - w_1 \in \Lambda_v$, а значит, равна нулевому вектору, то $c_1 = c_2$ и $w_1 = w_2$. Следовательно, данное множество копий образует упаковку пространства $\mathbb{R}^n$. Из того, что $\det(2\Lambda_v) = 2^n \det(\Lambda_v) = 2^n$, следует, что эта упаковка имеет плотность равную $2^{-n}\mu = 2^{-n}(2 - \varepsilon)$. Поскольку $\varepsilon > 0$ выбирался произвольно, то $\delta(C) \geqslant 2^{-(n-1)}$. ■

14

Коды, игры и энтропия

— Скажи, зачем ты приехал в Касабланку, Рик?
— Я приехал на воды.
— Воды, какие воды? Касабланка находится посреди пустыни.
— Я был неверно проинформирован.

Клод Рэйнс Хэмфри Богарту
в фильме «Касабланка»

14.1. КОДЫ

Предположим, требуется передать некоторое сообщение в виде двоичной последовательности по каналу с помехами. С вероятностью p каждый бит может быть принят неправильно, т. е. p — это вероятность того, что посланный ноль будет принят как единица, а посланная единица будет принята как ноль. Величина p является параметром канала, и влиять на него мы не можем. Посланные биты всегда принимаются, иногда с ошибкой. Будем считать, что события, заключающиеся в неверном приеме битов, взаимно независимы. Рассмотрим пример такой ситуации. Пусть, скажем, $p = 0.1$.

Как добиться высокой надежности системы? Простым решением будет посылать каждый бит трижды. Для декодирования принятых трех бит будем применять правило голосования. Тогда вероятность неправильной расшифровки будет равна в нашем случае $3p^2 + p^3 = 0.031$. Мы пожертвовали скоростью передачи данных, она теперь составляет $1/3$ от исходной, зато повысили надежность. Если посылать каждый бит пять раз и использовать правило голосования для декодирования, то вероятность неправильной расшифровки понизится до 0.01051, но скорость передачи также понизится. Она составит $1/5$ от исходной. Ясно, что мы можем сделать вероятность неправильной расшифровки сколь угодно малой, но, по-видимому, за это придется поплатиться стремящейся к нулю скоростью передачи данных. Фундаментальная теорема теории информации, принадлежащая Клоду Шеннону, заключается в том, что такой обмен не обязателен: существуют коды со скоростью передачи данных, стремящейся к некоторой положительной константе (зависящей от p), и со стремящейся к нулю вероятностью ошибки при передаче.

Схему кодирования составляют два целых числа m, n, функция $f\colon \{0,1\}^m \to \{0,1\}^n$, которую называют *функцией кодирования*, и функция $g\colon \{0,1\}^n \to \{0,1\}^m$, называемая *функцией декодирования*. Эта запись означает, что сообщение (или часть сообщения) $x \in \{0,1\}^m$ будет закодирована

и послана как $f(x)$, а полученное сообщение $y \in \{0,1\}^n$ будет декодировано как $g(y)$. Скорость передачи данных в такой схеме определяется отношением m/n. Пусть $E = (e_1, \ldots, e_n)$ — случайный вектор, в котором $\Pr[e_i = 1] = p$, $\Pr[e_i = 0] = 1 - p$, а случайные величины e_i взаимно независимы. Определим вероятность правильной передачи сообщения как $\Pr[g(f(x) + E) = x]$ ($+$ означает сложение векторов по $\bmod 2$). Подразумевается, что случайная величина x имеет равномерное распределение на множестве $\{0,1\}^m$ и не зависит от E.

Ключевую роль в дальнейшем будет играть *функция энтропии*

$$H(p) = -p\log_2 p - (1-p)\log_2(1-p),$$

определенная для всех $p \in (0,1)$. Функция энтропии фигурирует в асимптотической формуле

$$\binom{n}{pn} = \frac{n^n e^{-n}}{(pn)^{pn}e^{-pn}((1-p)n)^{(1-p)n}e^{-(1-p)n}}(1+o(1))^n = 2^{n(H(p)+o(1))},$$

справедливой для каждого фиксированного p. При $p \in (0, 0.5)$ мы будем пользоваться оценкой

$$\sum_{i \leqslant pn} \binom{n}{i} \leqslant (1+pn)\binom{n}{pn} = 2^{n(H(p)+o(1))}.$$

Теорема 14.1.1 (теорема Шеннона). *Пусть $p \in (0, 0.5)$ — некоторое фиксированное число. Для любого сколь угодно малого $\varepsilon > 0$ существует схема кодирования со скоростью передачи данных, превосходящей $1 - H(p) - \varepsilon$, и вероятностью ошибки при передаче меньшей, чем ε.*

Доказательство. Пусть $\delta > 0$ таково, что $p + \delta < 0.5$ и $H(p+\delta) < H(p) + \varepsilon/2$. Для достаточно больших n положим $m = n(1 - H(p) - \varepsilon)$, обеспечивая требуемую скорость передачи данных. Зададим случайным образом функцию $f\colon \{0,1\}^m \to \{0,1\}^n$, выбирая каждое значение $f(x)$ независимо в соответствии с равномерным распределением. При заданной функции f определим декодирующую функцию $g\colon \{0,1\}^n \to \{0,1\}^m$ следующим образом. Положим $g(y) = x$, если x — это единственный вектор, для которого расстояние между $f(x)$ и y не превосходит $n(p+\delta)$. В качестве расстояния используется метрика Хэмминга ρ, где $\rho(y, y')$ равно числу координат, в которых отличаются векторы y и y'. Если такой вектор x не существует или существует более чем один такой вектор, то будем считать декодирование некорректным.

Рассмотрим две ситуации, в которых декодирование оказывается некорректным. Во-первых, вектор $f(x)+E$ может оказаться на расстоянии, большем чем $n(p+\delta)$ от $f(x)$. Заметим, что расстояние от $f(x) + E$ до $f(x)$ равно весу вектора E, имеющему биномиальное распределение $B(n,p)$, а значит, такое событие происходит с вероятностью $o(1)$, фактически с экспоненциально малой вероятностью. Во-вторых, может случиться, что существует $x' \neq x$, такой, что $f(x') \in S$, где S — множество векторов y', находящихся на расстоянии не

более $n(p+\delta)$ от $f(x)$. При заданных значениях $f(x)$ и E, величина $f(x')$ попрежнему равномерно распределена на множестве $\{0,1\}^n$, а следовательно, такое событие происходит с вероятностью $|S|2^{-n}$ для каждого отдельного x'. Тогда общая вероятность возникновения этого события не превосходит

$$2^m|S|2^{-n} < 2^{-n(\frac{\varepsilon}{2}+o(1))} = o(1).$$

Таким образом, суммарная вероятность того, что декодирование некорректно, равна $o(1)$ и, фактически, экспоненциально мала. Для достаточно больших n она меньше ε.

Среднее значение вероятности некорректного декодирования x по всем возможным выборам f меньше ε. Следовательно, найдется такая функция f (а значит, и схема кодирования) с вероятностью неверного декодирования меньшей ε. ■

Теорема Шеннона касается в высшей степени практической стороны передачи информации, тем самым особенно четко проявляются недостатки вероятностного подхода. Как найти подходящую схему кодирования? Если схема кодирования существует, то как обеспечить быстрое кодирование-декодирование сообщений?

Групповой код — это схема кодирования, в которой отображение $f : \{0,1\}^m \to \{0,1\}^n$ линейно, т. е. $f(0) = 0$ и $f(x+x') = f(x) + f(x')$, все вычисления проводятся по модулю 2. (Другими словами, множество значений функции f есть подгруппа группы $\{0,1\}^n$.) Групповым кодам уделяется особое внимание, в частности, из-за простоты кодирования.

Теорема 14.1.2. *Пусть зафиксировано некоторое значение $p \in (0, 0.5)$. Для любого сколь угодно малого $\varepsilon > 0$ существует групповой код со скоростью передачи данных, превосходящей $1 - H(p) - \varepsilon$, и вероятностью ошибки при передаче меньшей ε.*

Доказательство. Для всякого $1 \leqslant i \leqslant m$ обозначим через u_i двоичный вектор длины m, содержащий единицу на i-й позиции и нули на всех остальных. Выберем значения $f(u_1), \ldots, f(u_m)$ случайно и независимо, определим функцию f на остальных наборах, положив

$$f(\varepsilon_1 u_1 + \ldots + \varepsilon_m u_m) = \varepsilon_1 f(u_1) + \ldots + \varepsilon_m f(u_m).$$

Повторим доказательство теоремы Шеннона до момента оценки вероятности того, что $f(x) + E$ находится на расстоянии не более чем $n(p+\delta)$ от $f(x)$. Положим $z = x - x' = \varepsilon_1 u_1 + \ldots + \varepsilon_m u_m$, где все операции проводятся по модулю 2. Так как $x \neq x'$, то $z \neq 0$. Перенумеруем для определенности координаты так, чтобы $\varepsilon_m = 1$. Из линейности следует, что $f(z) = f(x) - f(x')$. Оценим вероятность $\Pr[f(z) \in S]$, где S — это множество векторов, находящихся на расстоянии не более $n(p+\delta)$ от нулевого вектора. При любых фиксированных ε_i и $f(u_i)$, $i < m$, у $f(z)$ остается еще слагаемое $f(u_m)$, которое независимо и равномерно распределено. Следовательно, $f(z)$ распределено равномерно. Тогда $\Pr[f(z) \in S] = |S|2^{-n}$ и доказательство завершается так же, как доказательство теоремы Шеннона. ■

14.2. ИГРА ЛЖЕЦА

Пол пытается узнать число $x \in \{1, \ldots, n\}$ у лживой Кэрол, сопротивляющейся этому. Он может задать q вопросов вида «Верно ли, что $x \in S$?», где S может быть произвольным множеством. Вопросы задаются последовательно, и i-й вопрос Пола может зависеть от предыдущих ответов. Кэрол может лгать, но она не может солгать больше k раз. При каких n, q, k Пол сможет определить число x?

При $k = 0$ Пол может выиграть в том и только том случае, когда $n \leqslant 2^q$. При значениях $n = 100$, $q = 10$, $k = 1$ мы получаем замечательную детскую игру. Кэрол вряд ли можно назвать пассивным наблюдателем. Она может осуществлять стратегию, препятствующую Полу. Под этим мы подразумеваем, что Кэрол не загадывала x заранее, но ее ответы согласуются по крайней мере с одним x. В конце игры, если ее ответы согласуются с более чем одним x, то она выиграла. Таким образом, игра, называемая (n, q, k)-игрой лжеца, является игрой с полной информацией без скрытых ходов и без ничьих. Следовательно, либо у Пола, либо у Кэрол есть выигрышная стратегия. Но у кого?

Опишем эквивалентную игру, называемую *игрой лжеца с фишками*. Имеется доска с позициями $0, 1, \ldots, k$ и n фишек, пронумерованных числами $1, \ldots, n$ и находящихся на позиции k. Игра длится q раундов. В каждом раунде Пол выбирает множество фишек S. Кэрол может *либо* передвинуть все фишки из дополнения $\overline{S}$ множества S на одну позицию влево, *либо* передвинуть все фишки из множества S на одну позицию влево. (Будем считать, что позиция $i - 1$ находится слева от позиции i. Фишки, передвинутые на одну позицию влево с позиции 0, убираются с доски.) По окончании q раундов Кэрол выигрывает, если на доске находится более одной фишки, а Пол выигрывает, если на доске останется одна или ноль фишек. Присутствие фишки i на позиции j означает, что при ответе $x = i$ Кэрол солгала $k-j$ раз. Выбор Полом множества S соответствует вопросу: «Верно ли, что $x \in S$?». Перемещение Кэрол фишек из $\overline{S}$ соответствует ответу «да», перемещение фишек из S соответствует ответу «нет». (В игре лжеца с фишками Кэрол может удалить все фишки с доски, в то время как в игре лжеца ответы должны согласовываться по крайней мере с одним x. Но если Кэрол уберет все фишки с доски, то она автоматически проиграет. Следовательно, это различие не влияет на определение победителя.)

В игре лжеца с фишками нет необходимости помещать сначала все фишки на позицию k. Для $x_0, \ldots, x_k \geqslant 0$ определим $[(x_0, \ldots, x_k), q]$-игру лжеца с фишками так же, как и q-раундовую игру, описанную выше, но с размещением x_i фишек на позиции i вначале. Что, в свою очередь, соответствует игре лжеца, в которой для Кэрол есть x_i чисел, относительно которых можно солгать не более i раз.

Обозначим через $B(q, j)$ вероятность того, что при q подбрасываниях монеты герб выпадает не более j раз. Мы, конечно же, располагаем точной формулой

$$B(q, j) = 2^{-q} \sum_{i=0}^{j} \binom{q}{i}.$$

Теорема 14.2.1. *При условии*

$$\sum_{i=0}^{k} x_i B(q,i) > 1$$

Кэрол выигрывает $[(x_0,\ldots,x_k),q]$*-игру лжеца с фишками.*

Следствие 14.2.2. *При условии*

$$n > \frac{2^q}{\sum_{i=0}^{k}\binom{q}{i}}$$

Кэрол выигрывает (n,q,k)*-игру лжеца.*

Доказательство теоремы 14.2.1. Зафиксируем стратегию Пола. Пусть теперь игра Кэрол случайна! То есть в каждом раунде, после того как Пол выбрал множество фишек S, Кэрол подбрасывает монету. Если выпадает герб, то она передвигает фишки из $\overline{S}$ на одну позицию влево, если выпадает решка, то она передвигает на одну позицию влево фишки из S. Для каждой фишки c обозначим через I_c индикатор, описывающий присутствие или отсутствие фишки c на доске в конце игры. Обозначим через $X = \sum I_c$ число фишек, остающихся на доске в конце игры. Рассмотрим некоторую фишку c. В каждом раунде Пол мог выбрать $c \in S$ или $c \notin S$, но независимо от этого фишка c передвигается влево с вероятностью $\frac{1}{2}$. Пусть фишка c находится на позиции i в начале игры. Она останется на доске в конце игры, тогда и только тогда, когда во всех q раундах она передвигалась влево не более i раз. Тогда $\mathbf{E}[I_c]$ — вероятность этого события — в точности равна $B(q,i)$. Из линейности математического ожидания следует, что $\mathbf{E}[X] = \sum_{i=0}^{k} x_i B(q,i)$. Из условий теоремы вытекает, что $\mathbf{E}[X] > 1$. Но тогда событие $X > 1$ должно происходить с ненулевой вероятностью. То есть с ненулевой вероятностью Кэрол выиграет.

Никакая стратегия не поможет Полу всегда выигрывать. Но мы имеем дело с игрой с полной информацией без ничьих, следовательно, у кого-то должна быть выигрышная стратегия. Этот кто-то — не Пол, следовательно, это должна быть Кэрол. ■

Изложенное выше доказательство хорошо иллюстрирует присутствие магического элемента в вероятностном подходе. У Кэрол есть выигрышная стратегия, но как ее найти? Переход от вероятностного доказательства существования к явному построению называется *дерандомизацией* и будет детально рассмотрен в следующей главе. Здесь же мы приведем явную стратегию. При остающихся до конца игры l ходах и числом фишек y_i на позиции i, определим *вес игровой ситуации* как сумму $\sum_i y_i B(l,i)$. Заметим, что эта сумма представляет собой математическое ожидание $\mathbf{E}[Y]$, где Y — это число фишек, которые останутся на доске, если Кэрол будет играть случайно. Явную стратегию Кэрол описывает следующее правило: всегда делать ход так, чтобы максимизировать вес игровой ситуации.

Рассмотрим произвольную игровую ситуацию с весом W и некоторый ход Пола S. Обозначим через W^y, W^n веса, которые получатся при перемещении

Кэрол всех фишек из $\overline{S}$ или всех фишек из S соответственно. Утверждается, что $W = \frac{1}{2}(W^y + W^n)$. Подтверждением этого может служить то, что, вследствие линейности веса, проверка этого равенства сводится к рассмотрению случая с одной фишкой, равенство для которого следует из тождества $B(l, j) = \frac{1}{2}(B(l-1, j) + B(l-1, j-1))$. В действительности, нам не требуются никакие вычисления. Кэрол играет случайным образом, следовательно, вероятность каждого из ее решений в l-м раунде равна $1/2$. Тогда $\mathbf{E}[Y]$ — это среднее арифметическое двух условных математических ожиданий.

В начале игры по условию теоремы 14.2.1 вес игровой ситуации больше единицы. Явная стратегия Кэрол не дает весу уменьшиться, следовательно, к концу игры вес по-прежнему будет больше единицы. Но в конце игры вес равен числу оставшихся на доске фишек, а Кэрол побеждает, если это число больше единицы.

Утверждения, обратные к теореме 14.2.1 и ее следствию, не верны. Рассмотрим игру лжеца при $n = 5$, с $q = 5$ вопросами и $k = 1$ возможностью солгать. Она соответствует $[(0, 5), 5]$-игре лжеца с фишками. Отметим, что $B(5, 1) = 6/32$ и $5(6/32) \leqslant 1$. Но Кэрол все-равно выиграет с помощью своей стратегии. Проблема в том, что у Пола нет хорошего первого хода. Предположим, что он выберет какие-то две фишки в качестве множества S (спросит: «Верно ли, что $x \leqslant 2$?» в игре лжеца). Тогда Кэрол передвинет две фишки влево (ответит «да»), создав ситуацию $(2, 3)$ при четырех остающихся вопросах. Кэрол теперь обязательно выиграет, так как $2B(4, 0) + 3B(4, 1) = 17/16 > 1$. Несложно проверить, что все остальные ходы Пола также приводят к поражению. Проблема здесь в том, что Пол, находясь в ситуации с весом $W \leqslant 1$, не смог найти ход, приводящий к ситуации $W^y \leqslant 1$ и $W^n \leqslant 1$.

14.3. ИГРА «ПОСТОЯННАЯ ДОЛЖНОСТЬ»

Декан Пол пытается добиться постоянной должности для одного из своих преподавателей, но ему противодействует скупая Кэрол, ректор университета. Всего есть k уровней, предшествующих постоянной должности, которые мы пометим числами $1, \ldots, k$, 1-й уровень — наивысший, а 0-й означает постоянную должность. Каждый преподаватель представлен фишкой в нашей игре. Игра «постоянная должность» с параметрами $(x_1, \ldots, x_k)$ начинается с x_i фишками на i-м уровне при $1 \leqslant i \leqslant k$, а на нулевом уровне фишек нет. Каждый год Пол представляет Кэрол некоторое множество фишек S. Кэрол может сделать следующее:

- повысить в должности все фишки из множества S и уволить остальные *или*
- повысить в должности все фишки из дополнения $\overline{S}$ множества S и уволить фишки из S.

Повышение фишки в должности означает ее передвижение с уровня i на уровень $i-1$. Увольнение фишки означает удаление ее из игры. Пол побеждает, если какая-то фишка достигла уровня 0. Драконовские правила повышения в

должности гарантируют, что игра закончится через k лет, и либо выиграет Пол, либо, успешно удалив из игры все фишки, выиграет Кэрол.

Теорема 14.3.1. *Кэрол выигрывает игру «постоянная должность» с параметрами* $(x_1, \ldots, x_k)$ *при* $\sum_i x_i 2^{-i} < 1$.

Доказательство. Зафиксируем некоторую стратегию Пола. Пусть игра Кэрол будет случайной! То есть в каждом раунде, после того как Пол выберет множество фишек S, Кэрол подбрасывает монету — если выпадет герб, то она передвигает все фишки из множества $\overline{S}$ на одну позицию влево, а если выпадет решка, то на одну позицию влево передвигаются фишки из множества S. Для каждой фишки c обозначим через I_c индикатор события, что фишка c достигла уровня 0. Пусть $X = \sum I_c$ — число фишек, достигших уровня 0 к концу игры. Рассмотрим некоторую фишку c. В каждом раунде Пол мог решить, что $c \in S$, или $c \notin S$, но в любом случае фишка c передвигается влево с вероятностью $\frac{1}{2}$. Пусть фишка c находится на позиции i в начале игры. Она останется на доске к концу игры, только если первые i подбрасываний Кэрол монеты приводили к повышению в должности фишки c. В таком случае $\mathbf{E}[I_c]$, вероятность такого события, есть 2^{-i}. В силу линейности математического ожидания $\mathbf{E}[X] = \sum_{i=1}^{k} x_i 2^{-i}$. Из условия теоремы получаем, что $\mathbf{E}[X] < 1$. Тогда событие $X < 1$ должно происходить с ненулевой вероятностью. То есть с ненулевой вероятностью Кэрол должна выигрывать.

Никакая стратегия Пола не позволяет ему всегда выигрывать. Но мы имеем дело с игрой с полной информацией без ничьих, так что у кого-то должна быть выигрышная стратегия. Если это не Пол, то это должна быть Кэрол. ■

Как и в игре лжеца, мы можем дерандомизировать приведенное выше доказательство и привести явную стратегию для Кэрол. Пусть на i-й позиции находится y_i фишек. Определим *вес игровой ситуации* суммой $\sum_i y_i 2^{-i}$. Заметим, что эта величина равна $\mathbf{E}[Y]$, где Y — это количество фишек, которые достигнут 0-го уровня, если Кэрол будет играть случайно. Явная стратегия Кэрол заключается в выборе хода, минимизирующего вес игровой ситуации. Рассмотрим произвольную игровую ситуацию с весом W и некоторый ход Пола S. Обозначим через W^y, W^n веса ситуаций, возникающих, если Кэрол передвинет все фишки из множества $\overline{S}$ или все фишки из множества S, соответственно. Как и в игре лжеца, $W = \frac{1}{2}(W^y + W^n)$. По условию теоремы, в начале игры вес был меньше единицы. Явная стратегия Кэрол не дает весу увеличиться, следовательно, к концу игры вес по-прежнему будет меньше единицы. Фишка, находящаяся на позиции 0, добавила бы к весу игровой ситуации единицу, следовательно, на позиции 0 нет ни одной фишки, а это означает, что Кэрол выиграла.

В игре лжеца достаточные условия выигрыша Кэрол не являются одновременно и необходимыми, так как в любом случае у Пола может не оказаться хорошего хода. Тем не менее, справедлива следующая интересная лемма.

Лемма 14.3.2. *Если вес игровой ситуации не меньше единицы, то она может быть разделена на две игровые ситуации, вес каждой из которых не меньше* $\frac{1}{2}$.

Доказательство. Обязательно найдется позиция i, на которой находятся не менее двух фишек. В противном случае, вес был бы меньше единицы. Если на позиции 1 находятся две фишки, то отнесем их к разным игровым ситуациям. Если две фишки находятся на позиции $i > 1$, склеим их и будем рассматривать как одну суперфишку на позиции $i - 1$. Заметим, что мы получили игровую ситуацию с тем же весом, но с меньшим количеством фишек. Далее применяем индукцию по числу фишек. ■

Теорема 14.3.3. *Пол выигрывает игру «постоянная должность» с параметрами* $(x_1, \ldots, x_k)$ *при* $\sum x_i 2^{-i} \geqslant 1$.

Доказательство. Вес игровой ситуации в начале игры не меньше единицы. Пол применяет последнюю лемму и разделяет фишки на два подмножества так, чтобы каждому из них соответствовала игровая ситуация веса не менее $\frac{1}{2}$. Одно из этих подмножеств Пол выберет в качестве множества S. Кэрол передвинет все фишки одного из подмножеств на одну позицию влево, удваивая их вес, и мы снова получаем ситуацию с весом не менее единицы. Таким образом, мы никогда не получим ситуацию веса меньше единицы. Следовательно, игра не может закончиться удалением всех фишек с доски (у этой ситуации вес равен нулю), а, значит, игра закончится победой Пола. ■

14.4. ИГРА «БАЛАНСИРОВКА ВЕКТОРОВ»

Игра «балансировка векторов» — это игра с полной информацией, в которой участвуют два игрока — Предлагающий и Решающий. У этой игры есть параметр $n \geqslant 1$, и нас будет интересовать асимптотика по n. Игра длится n раундов, в каждом из которых мы имеем дело с векторами пространства $\mathbb{R}^n$. Имеется вектор игровой ситуации $P \in \mathbb{R}^n$, изначально равный нулю. Каждый раунд состоит из двух частей. Вначале Предлагающий выбирает вектор $v \in \{-1, +1\}^n$. Затем Решающий заменяет вектор P либо вектором $P+v$, либо вектором $P - v$. К концу n-го раунда выигрыш Предлагающего равен $|P|_\infty$, максимальному из абсолютных значений координат вектора P. Обозначим через $\mathrm{VAL}(n)$ значение этой игры для Предлагающего, а через $\mathbf{S}_n$ обозначим, как обычно, сумму n независимых случайных величин, принимающих с равной вероятностью значения из множества $\{1, -1\}$.

Теорема 14.4.1. *Если* $\Pr[|\mathbf{S}_n| > \alpha] < n^{-1}$, *то* $\mathrm{VAL}(n) \leqslant \alpha$.

Доказательство. Будем считать, что Предлагающий выиграл, если в конце игры $|P|_\infty > \alpha$. Пусть Решающий определяет, заменить ли ему вектор P вектором $P + v$ или вектором $P - v$, подбрасывая монету. Через x_i обозначим i-ю координату окончательного значения вектора игровой ситуации P.

Обозначим через W_i событие «$|x_i| > \alpha$», положим $W = \vee W_i$. Тогда событие W означает победу Предлагающего. Случайная величина x_i имеет распределение $\mathbf{S}_n$ независимо от стратегии Предлагающего. Тогда

$$\Pr[W] \leqslant \sum_{i=1}^{n} \Pr[|\mathbf{S}_n| > \alpha] < 1.$$

Предлагающий не может всегда выигрывать, следовательно, всегда выигрывать будет Решающий. ■

Следствие 14.4.2. *Справедливо соотношение* $\mathrm{VAL}(n) = O(\sqrt{n \ln n})$.

Для получения нижней оценки $\mathrm{VAL}(n)$ требуется найти выигрышную стратегию для Предлагающего, т. е. стратегию, которая позволит ему выигрывать независимо от действий Решающего. Мы не можем ограничиться рассмотрением случая, когда Решающий играет случайно. Решающий — противник Предлагающего и может придерживаться любой стратегии. Тем не менее, справедлив следующий результат, касающийся Решающего, который играет случайно.

Теорема 14.4.3. *Если* $\Pr[|\mathbf{S}_n| > \alpha] > cn^{-1/2}$, *где* c — *абсолютная константа, то* $\mathrm{VAL}(n) > \alpha$.

Следствие 14.4.4. *Справедливо соотношение* $\mathrm{VAL}(n) = \Omega(\sqrt{n \ln n})$ *и, следовательно,* $\mathrm{VAL}(n) = \Theta(\sqrt{n \ln n})$.

Доказательство теоремы 14.4.3. Положим для $x \in \mathbb{Z}, 0 \leqslant i \leqslant n$,

$$w_i(x) = \Pr[|x + \mathbf{S}_{n-i}| > \alpha].$$

Для $P = (x_1, \ldots, x_n)$ положим $w_i(P) = \sum_{1 \leqslant j \leqslant n} w_i(x_j)$. Пусть P — вектор игровой ситуации в конце i-го раунда, тогда величину $w_i(P)$ можно интерпретировать как ожидаемое количество координат, абсолютная величина которых будет больше чем α к концу игры, если Решающий будет играть случайно. В начале игры $w_0(P) = w_0(0) > c\sqrt{n}$ по условию теоремы. Стратегия Предлагающего заключается в следующем. При заданном в конце i-го раунда векторе игровой ситуации P Предлагающий так должен выбрать вектор $v \in \{-1, +1\}^n$, чтобы значения $w_{i+1}(P - v)$ и $w_{i+1}(P + v)$ различались мало.

Случайная величина $x + \mathbf{S}_{n-i}$ превращается в величину $x + 1 + \mathbf{S}_{n-i-1}$ или $x - 1 + \mathbf{S}_{n-i-1}$ в зависимости от первого подбрасывания монеты. Тогда для всех i, x верно равенство

$$w_i(x) = \frac{1}{2}[w_{i+1}(x + 1) + w_{i+1}(x - 1)].$$

Пусть $P = (x_1, \ldots, x_n)$, $v = (v_1, \ldots, v_n)$. Для $1 \leqslant j \leqslant n$ положим

$$\Delta_j = w_{i+1}(x_j + 1) - w_{i+1}(x_j - 1).$$

Тогда

$$w_{i+1}(P + v) - w_{i+1}(P - v) = \sum_{j=1}^{n} v_j \Delta_j,$$

и для $\varepsilon = \pm 1$ имеем

$$w_{i+1}(P + \varepsilon v) = w_i(P) + \frac{1}{2}\varepsilon \sum_{j=1}^{n} v_j \Delta_j.$$

Оценим $|\Delta_j|$. Заметим, что

$$\Delta_j = \Pr[\mathbf{S}_{n-i-1} = y] - \Pr[\mathbf{S}_{n-i-1} = z],$$

где y — это единственное целое число той же четности, что и $n-i-1$, в интервале $(\alpha - (x_j + 1), \alpha - (x_j - 1)]$, а z — это единственное целое число той же четности, что и $n-i-1$, в интервале $[-\alpha - (x_j + 1), -\alpha - (x_j - 1))$. Пусть

$$g(m) = \max_s \Pr[\mathbf{S}_m = s] = \binom{m}{\lfloor m/2 \rfloor} 2^{-m} \sim \sqrt{\frac{2}{\pi m}},$$

тогда $|\Delta_j| \leqslant g(n-i-1)$ для всех j.

Стратегия Предлагающего будет заключаться в следующем. Так переупорядочить координаты, чтобы $|\Delta_1| \geqslant \ldots \geqslant |\Delta_n|$, а затем последовательно выбрать значения $v_1, \ldots, v_n \in \{-1, +1\}$ так, чтобы величины $v_i\Delta_i$ и $v_1\Delta_1 + \ldots + v_{i-1}\Delta_{i-1}$ имели разные знаки. (Значения v_i выбираются произвольно в случае $i = 1$ или при равной нулю второй величине.) Тогда

$$|v_1\Delta_1 + \ldots + v_n\Delta_n| \leqslant |\Delta_1| \leqslant g(n-i-1).$$

Обозначим через P^i вектор игровой ситуации в конце i-го раунда, а через v — вектор, выбранный Предлагающим для $(i+1)$-го раунда. Независимо от выбора Решающим $\varepsilon = \pm 1$ будут выполнены соотношения

$$w_{i+1}(P^{i+1}) = w_{i+1}(P^i + \varepsilon v) \geqslant w_i(P^i) - \frac{1}{2}\left|\sum_{j=1}^{n} v_j \Delta_j\right| \geqslant w_i(P^i) - \frac{1}{2}g(n-i-1).$$

Таким образом,

$$w_n(P^n) \geqslant w_0(P^0) - \frac{1}{2}\sum_{i=0}^{n-1} g(n-i-1).$$

С помощью несложных вычислений можно убедиться в том, что значение указанной выше суммы асимптотически стремится к $(8n/\pi)^{1/2}$. При $c > (2/\pi)^{1/2}$ имеем $w_n(P^n) > 0$. Но $w_n(P^n)$ — это *число координат* вектора финального состояния $P = P^n$, абсолютные величины которых больше α. Стратегия Предлагающего гарантирует, что это число будет больше нуля, а, значит, и его победу. ■

14.5. НЕАДАПТИВНЫЕ АЛГОРИТМЫ

Модифицируем игру «балансировка векторов» из разд. 14.4, предоставив Предлагающему выбирать вектор, состоящий из нулей и единиц, а не из ± 1. Обозначим через $\mathrm{VAL}^*(n)$ значение модифицированной игры. С использованием оценок для $\mathrm{VAL}(n)$, можно показать, что $\mathrm{VAL}^*(n) = \Theta(\sqrt{n \ln n})$.

В гл. 12 было показано, что каждое семейство из n подмножеств $S_1, \ldots, S_n$ n-множества $\{1, \ldots, n\}$ имеет разброс, равный $O(\sqrt{n})$. То есть существует раскраска $\chi : \{1, \ldots, n\} \to \{-1, +1\}$ такая, что $|\chi(S_i)| \leqslant c\sqrt{n}$ при всех значениях i. Из доказательства этого утверждения нельзя получить эффективный алгоритм нахождения такой раскраски. В самом деле, пока не известно, существует ли полиномиальный алгоритм решения такой задачи. Зададимся вопросом существования *неадаптивного* или *интерактивного* алгоритма в следующем смысле. Вместо получения всей информации о семействе $S_1, \ldots, S_n$ сразу, мы будем последовательно получать информацию о каждой отдельной точке. На j-м шаге алгоритм получает информацию о том, какие из множеств S_i содержат точку j, или, иными словами, j-й столбец матрицы инцидентности. В этот момент алгоритм должен решить, как раскрасить точку j, причем впоследствии эта раскраска не может быть изменена. Как можно оценить сверху величину $\max|\chi(S_i)|$ для такого алгоритма? Будем считать, что точки нам предлагает наш соперник. Будем рассматривать эти точки как векторы-столбцы матрицы инцидентности. Предлагающего будем рассматривать как наиболее неудобного соперника, а Решающего — как алгоритм. Тогда требуемой оценкой будет $\mathrm{VAL}^*(n)$.

Требование неадаптивности алгоритма одновременно и сильнее и слабее требования его полиномиальности. Тем не менее, приведенный выше результат косвенно подтверждает гипотезу об отсутствии полиномиального алгоритма нахождения раскраски, при которой $|\chi(S_i)| \leqslant c\sqrt{n}$ для всех значений i.

14.6. ЭНТРОПИЯ

Пусть X — случайная величина, принимающая значения на множестве S. Обозначим через $\Pr[X = x]$ вероятность того, что случайная величина X принимает значение x. *Бинарная энтропия* случайной величины X, обозначаемая через $H(X)$, определяется следующим образом:

$$H(X) = \sum_{x \in S} \Pr[X = x] \log_2 \left(\frac{1}{\Pr[X = x]} \right).$$

Пусть Y — другая случайная величина, принимающая значения на множестве T, и (X, Y) — случайная величина, принимающая значения на множестве $S \times T$ в соответствии с совместным распределением случайных величин X и Y. Тогда *условной энтропией* случайной величины X при заданном значении Y называется

$$H(X|Y) = H(X, Y) - H(Y).$$

В этом разделе мы приведем некоторые простые свойства энтропии и несколько замечательных примеров применения энтропии в комбинаторике и геометрии. Интуитивно ясно, что энтропия случайной величины характеризует количество информации, содержащейся в ней. Отсюда становится ясен «физический» смысл четырех утверждений следующей леммы. Формальное доказательство, приведенное ниже, опирается на свойства функций $\log z$ и $z \log z$, где через $\log z$ здесь обозначен двоичный логарифм $\log_2 z$.

Лемма 14.6.1. *Пусть X, Y и Z — случайные величины, принимающие значения на множествах S, T и U соответственно. Тогда*

i) $H(X) \leqslant \log_2 |S|$;
ii) $H(X,Y) \geqslant H(X)$;
iii) $H(X,Y) \leqslant H(X) + H(Y)$;
iv) $H(X|Y,Z) \leqslant H(X|Y)$.

Доказательство.

i) Так как функция $\log z$ вогнута, то из неравенства Йенсена следует, что

$$H(X) = \sum_{i\in S} \Pr\big[X=i\big] \log\left(\frac{1}{\Pr[X=i]}\right) \leqslant$$
$$\leqslant \log\left(\sum_{i\in S} \Pr\big[X=i\big] \frac{1}{\Pr[X=i]}\right) = \log|S|.$$

ii) Из монотонности функции $\log z$ при всех $z > 0$ следует, что

$$H(X,Y) = \sum_{i\in S}\sum_{j\in T} \Pr\big[X=i, Y=j\big] \log\left(\frac{1}{\Pr[X=i,Y=j]}\right) \geqslant$$
$$\geqslant \sum_{i\in S}\sum_{j\in T} \Pr\big[X=i, Y=j\big] \log\left(\frac{1}{\Pr[X=i]}\right) =$$
$$= \sum_{i\in S} \Pr\big[X=i\big] \log\left(\frac{1}{\Pr[X=i]}\right) = H(X). \tag{14.1}$$

iii) По определению

$$H(X) + H(Y) - H(X,Y) =$$
$$= \sum_{i\in S}\sum_{j\in T} \Pr\big[X=i, Y=j\big] \log\left(\frac{\Pr[X=i,Y=j]}{\Pr[X=i]\Pr[Y=j]}\right) =$$
$$= \sum_{i\in S}\sum_{j\in T} \Pr\big[X=i\big]\Pr[Y=j] f(z_{ij}), \tag{14.2}$$

где $f(z) = z\log z$ и $z_{ij} = \frac{\Pr[X=i,Y=j]}{\Pr[X=i]\Pr[Y=j]}$. Так как функция $f(z)$ выпукла, то из неравенства Йенсена следует, что последняя величина не меньше

$$f\left(\sum_{i\in S}\sum_{j\in T} \Pr[X=i]\Pr[Y=j] z_{ij}\right) = f(1) = 0.$$

iv) Заметим, что

$$H(X|Y) = H(X,Y) - H(Y) =$$
$$= \sum_{i\in S}\sum_{j\in T} \Pr[X=i, Y=j] \log\left(\frac{\Pr[Y=j]}{\Pr[X=i,Y=j]}\right).$$

Аналогично,

$$H(X|Y,Z)=\sum_{i\in S}\sum_{j\in T}\sum_{k\in U}\Pr[X=i,Y=j,Z=k]\log\left(\frac{\Pr[Y=j,Z=k]}{\Pr[X=i,Y=j,Z=k]}\right).$$

Следовательно,

$$\begin{aligned}H(X|Y)-H(X|Y,Z)=\\ =\sum_{i\in S}\sum_{j\in T}\sum_{k\in U}\Pr[X=i,Y=j,Z=k]\cdot\log\left(\frac{\Pr[Y=j]\Pr[X=i,Y=j,Z=k]}{\Pr[X=i,Y=j]\Pr[Y=j,Z=k]}\right)=\\ =\sum_{i\in S}\sum_{j\in T}\sum_{k\in U}\frac{\Pr[X=i,Y=j]\Pr[Y=j,Z=k]}{\Pr[Y=j]}f(z_{ijk}),\end{aligned}$$

где $f(z)=z\log z$ и

$$z_{ijk}=\frac{\Pr[Y=j]\Pr[X=i,Y=j,Z=k]}{\Pr[X=i,Y=j]\Pr[Y=j,Z=k]}.$$

Из выпуклости функции $f(z)$ и того, что

$$\sum_{i\in S}\sum_{j\in T}\sum_{k\in U}\frac{\Pr[X=i,Y=j]\Pr[Y=j,Z=k]}{\Pr[Y=j]}=1,$$

вытекает, что последняя величина не меньше

$$f\left(\sum_{i\in S}\sum_{j\in T}\sum_{k\in U}\frac{\Pr[X=i,Y=j]\Pr[Y=j,Z=k]}{\Pr[Y=j]}z_{ijk}\right)=f(1)=0. \quad \blacksquare$$

Приведенное ниже простое, но полезное утверждение о том, что энтропия субаддитивна, уже применялось в гл. 12, разд. 12.2.

Предложение 14.6.2. *Пусть $X=(X_1,\ldots,X_n)$ — случайная величина, принимающая значения на множестве $S=S_1\times S_2\times\ldots\times S_n$, где каждая координата X_i вектора X является случайной величиной, принимающей значения на множестве S_i. Тогда*

$$H(X)\leqslant\sum_{i=1}^{n}H(X_i).$$

Доказательство. Утверждение следует по индукции из третьего пункта леммы 14.6.1. $\blacksquare$

Приведенное выше утверждение используется в работе [Kleitman, Shearer and Sturtevant (1981)] для получения нескольких интересных приложений в экстремальной теории конечных множеств, включая верхнюю оценку для максимально возможной мощности семейства k-множеств, в котором ни одно из пересечений двух множеств не содержится в третьем. Основная идея работы [Kleitman et al. (1981)] может быть проиллюстрирована с помощью следующего очень простого следствия последнего утверждения.

Следствие 14.6.3. *Пусть $\mathcal{F}$ — некоторое семейство подмножеств множества $\{1,2,\dots,n\}$. Обозначим через p_i долю множеств семейства $\mathcal{F}$, содержащих элемент i. Тогда*

$$|\mathcal{F}| \leqslant 2^{\sum_{i=1}^{n} H(p_i)},$$

где $H(y) = -y\log_2 y - (1-y)\log_2(1-y)$.

Доказательство. Поставим в соответствие каждому множеству $F \in \mathcal{F}$ его характеристический вектор $v(F)$, представляющий собой двоичный вектор длины n. Пусть $X = (X_1,\dots,X_n)$ — случайная величина, принимающая значения на множестве $\{0,1\}^n$, где $\Pr[X = v(F)] = 1/|\mathcal{F}|$ для всех $F \in \mathcal{F}$. Ясно, что $H(X) = |\mathcal{F}|(\frac{1}{|\mathcal{F}|}\log|\mathcal{F}|) = \log|\mathcal{F}|$, и так как $H(X_i) = H(p_i)$ для всех $1 \leqslant i \leqslant n$, требуемый результат следует из утверждения 14.6.2. ■

Следующее интересное обобщение утверждения 14.6.2 было получено Ширером (см. [Chung, Frankl, Graham and Shearer (1986)]). Пусть $X = (X_1,\dots,X_n)$ — случайная величина, принимающая значения на множестве $S = S_1 \times S_2 \times \dots \times S_n$, где каждая координата X_i — случайная величина, принимающая значения на множестве S_i. Для подмножества I множества $\{1,2,\dots,n\}$, обозначим через $X(I)$ случайную величину $(X_i)_{i\in I}$.

Предложение 14.6.4. *Пусть $X = (X_1,\dots,X_n)$ и S обозначают то же, что и прежде. Пусть $\mathcal{G}$ — семейство подмножеств множества $\{1,\dots,n\}$, и каждый элемент $i \in \{1,\dots,n\}$ принадлежит не менее чем k множествам семейства $\mathcal{G}$. Тогда*

$$kH(X) \leqslant \sum_{G\in\mathcal{G}} H(X(G)).$$

Доказательство. Применим индукцию по k. При $k = 1$ заменим каждое множество $G \in \mathcal{G}$ его подмножеством так, чтобы получить семейство $\mathcal{G}'$, члены которого образуют разбиение множества $\{1,\dots,n\}$. Из второго пункта леммы 14.6.1 следует, что $\sum_{G\in\mathcal{G}} H(X(G)) \geqslant \sum_{G'\in\mathcal{G}'} H(X(G'))$, из третьего пункта леммы 14.6.1 следует, что $\sum_{G'\in\mathcal{G}'} H(X(G')) \geqslant H(X)$, а, значит, и требуемое утверждение для $k = 1$.

Пусть утверждение верно для $k-1$, докажем его для k ($\geqslant 2$). Если семейство $\mathcal{G}$ содержит множество $G = \{1,\dots,n\}$, то требуемое утверждение следует из предположения индукции. В противном случае рассмотрим два множества G_1, G_2, принадлежащих семейству $\mathcal{G}$. Воспользовавшись четвертым пунктом леммы 14.6.1, получаем, что

$$H(X(G_1 \setminus G_2)|X(G_1 \cap G_2), X(G_2 \setminus G_1)) \leqslant H(X(G_1 \setminus G_2)|X(G_1 \cap G_2)),$$

откуда следует неравенство

$$H(X(G_1 \cup G_2)) - H(X(G_2)) \leqslant H(X(G_1)) - H(X(G_1 \cap G_2)).$$

Следовательно, $H((X(G_1 \cup G_2)) + H(X(G_1 \cap G_2)) \leqslant H(X(G_1)) + H(X(G_2))$. Тогда мы можем модифицировать семейство $\mathcal{G}$, заменив в нем множества G_1 и G_2 их объединением и пересечением, при этом значение суммы $\sum_{G\in\mathcal{G}} H(X(G))$

может только уменьшиться. После конечного числа таких преобразований мы получим семейство $\mathcal{G}$, содержащее множество $\{1,\dots,n\}$, а этот случай мы уже рассматривали. Утверждение полностью доказано. ■

Следствие 14.6.5. *Пусть $\mathcal{F}$ — семейство векторов из $S_1 \times S_2 \dots \times S_n$. Пусть $\mathcal{G} = \{G_1, G_2, \dots G_m\}$ — семейство подмножеств множества $N = \{1,2,\dots,n\}$, и пусть каждый элемент $i \in N$ принадлежит не менее чем k множествам семейства $\mathcal{G}$. Для каждого $1 \leqslant i \leqslant m$ обозначим через $\mathcal{F}_i$ семейство проекций членов семейства $\mathcal{F}$ на множество G_i. Тогда*

$$|\mathcal{F}|^k \leqslant \prod_{i=1}^{m} |\mathcal{F}_i|.$$

Доказательство. Пусть $X = (X_1,\dots,X_n)$ — случайная величина, принимающая значения на множестве $\mathcal{F}$, где $\Pr[X = F] = 1/|\mathcal{F}|$ для всех $F \in \mathcal{F}$. Из утверждения 14.6.4 следует, что

$$kH(X) \leqslant \sum_{i=1}^{m} H(X(G_i)).$$

Требуемое утверждение вытекает теперь из того, что $H(X) = \log_2 |\mathcal{F}|$, и неравенства $H(X(G_i)) \leqslant \log_2 |\mathcal{F}_i|$, следующего из первого пункта леммы 14.6.1. ■

Объем любого d-измеримого подмножества в $\mathbb{R}^n$ может быть оценен объемом аппроксимирующей это множество фигуры, составленной из образованных достаточно частой решеткой кубов, поэтому последний результат имеет и геометрическое применение. Нижеследующее утверждение доказывается другим способом в работе [Loomis and Whitney (1949)].

Следствие 14.6.6. *Пусть B — измеримое тело в n-мерном евклидовом пространстве. Обозначим через $\mathrm{Vol}(B)$ его (n-мерный) объем, а через $\mathrm{Vol}(B_i)$ обозначим $(n-1)$-мерный объем проекции тела B на гиперплоскость, полученную фиксацией i-й координаты. Тогда*

$$(\mathrm{Vol}(B))^{n-1} \leqslant \prod_{i=1}^{n} \mathrm{Vol}(B_i).$$ ■

При $S_i = \{0,1\}$ для всех i из следствия 14.6.5 получаем следующее утверждение относительно систем множеств.

Следствие 14.6.7 [Chung et al. (1986)]. *Пусть N — конечное множество и $\mathcal{F}$ — семейство подмножеств множества N. Пусть $\mathcal{G} = \{G_1, \dots G_m\}$ — семейство подмножеств множества N, и пусть каждый элемент множества S принадлежит не менее чем k множествам семейства $\mathcal{G}$. Для каждого $1 \leqslant i \leqslant m$ положим $\mathcal{F}_i = \{F \cap G_i \colon F \in \mathcal{F}\}$. Тогда*

$$|\mathcal{F}|^k \leqslant \prod_{i=1}^{m} |\mathcal{F}_i|.$$ ■

В заключение данной главы приведем следующее применение последнего результата, описанное в работе [Chung et al. (1986)].

Следствие 14.6.8. *Пусть $\mathcal{F}$ — семейство графов на помеченном множестве вершин $\{1, 2, \dots t\}$, и пусть для каждой пары графов семейства $\mathcal{F}$ найдется треугольник, содержащийся в каждом из них. Тогда*

$$|\mathcal{F}| < \frac{1}{4} 2^{\binom{t}{2}}.$$

Доказательство. Обозначим через N множество всех $\binom{t}{2}$ неупорядоченных пар вершин из множества $T = \{1, 2 \dots, t\}$, будем рассматривать $\mathcal{F}$ как семейство подмножеств множества N. Обозначим через $\mathcal{G}$ семейство всех подмножеств множества N, которые состоят из множеств ребер объединения двух полных графов на непересекающихся почти равномощных множествах вершин из множества T. Обозначим через

$$s = \binom{\lceil t/2 \rceil}{2} + \binom{\lfloor t/2 \rfloor}{2}$$

число ребер такого объединения, а через m — мощность семейства $\mathcal{G}$. Из соображений симметрии следует, что каждое ребро множества N принадлежит ровно $k = \frac{sm}{\binom{t}{2}}$ графам семейства $\mathcal{G}$. Ключевым моментом доказательства является то, что у любых двух графов семейства $\mathcal{F}$ должно быть не менее одного общего ребра в каждом из графов $G \in \mathcal{G}$, так как их пересечение содержит треугольник (а дополнение графа G не содержит треугольников). Следовательно, в обозначениях следствия 14.6.7, мощность каждого семейства $\mathcal{F}_i$ не превосходит 2^{s-1}. Таким образом, получаем неравенство

$$|\mathcal{F}|^{sm/\binom{t}{2}} \leqslant (2^{s-1})^m,$$

откуда следует, что

$$|\mathcal{F}| \leqslant 2^{\binom{t}{2} - \binom{t}{2}/s},$$

так как $s < \binom{t}{2}/2$, отсюда вытекает требуемое утверждение. ■

Симонович и Шош предположили, что если семейство $\mathcal{F}$ удовлетворяет условиям последнего следствия, то

$$|\mathcal{F}| \leqslant \frac{1}{8} 2^{\binom{t}{2}}.$$

Причем, если эта оценка верна, то она является наилучшей. Эта гипотеза до сих пор не подтверждена и не опровергнута. Представляется правдоподобной гипотеза, что существует такая абсолютная константа $\varepsilon > 0$, что для каждого фиксированного графа H, не являющегося звездным лесом (т. е. лесом, каждая связная компонента которого представляет собой звезду), верно следующее. Пусть $\mathcal{F}$ — семейство графов на помеченном множестве вершин $\{1, 2, \dots t\}$, и пусть для каждых двух графов семейства $\mathcal{F}$ найдется копия графа H,

содержащаяся в каждом из них. Тогда

$$|\mathcal{F}| < \left(\frac{1}{2} - \varepsilon\right) 2^{\binom{t}{2}}.$$

Вопрос об истинности данного утверждения также является открытым, но, тем не менее, несложно показать, что оно верно для каждого графа H, хроматическое число которого не меньше 3, а также, что оно не выполняется для произвольного звездного леса H.

14.7. УПРАЖНЕНИЯ

1. Пусть в $(x_1, \ldots, x_k)$-игре «постоянная должность» из разд. 14.3 задачей Пола будет максимизировать число преподавателей, получивших постоянную должность, в то время как задачей Кэрол является минимизировать эту величину. Обозначим через v это число при игре с полной информацией. Доказать, что $v = \lfloor \sum_{i=1}^{k} x_i 2^{-k} \rfloor$.

2. Пусть $A_1, \ldots, A_n \subseteq \{1, \ldots, m\}$ при $\sum_{i=1}^{n} 2^{-|A_i|} < 1$. Пол и Кэрол по очереди выбирают различные вершины из множества $\{1, \ldots, m\}$. Пол выбирает первым, процесс продолжается пока не выберут все вершины. Кэрол выигрывает, если она выберет все вершины какого-либо из множеств A_i. Пол выигрывает, если не выиграет Кэрол. Найти выигрышную стратегию Пола.

3. Пусть $\mathcal{F}$ — семейство графов на помеченном множестве вершин $\{1, 2, \ldots 2t\}$, пусть для каждых двух графов семейства $\mathcal{F}$ найдется паросочетание из t ребер, содержащееся в каждом из них. Доказать, что

$$|\mathcal{F}| \leqslant 2^{\binom{2t}{2} - t}.$$

4. (Неравенство Хана.) Пусть $X = (X_1, \ldots, X_m)$ — случайная величина, а $H(X)$ обозначает ее энтропию. Для каждого подмножества I множества $\{1, 2, \ldots, m\}$ обозначим через $X(I)$ случайную величину $(X_i)_{i \in I}$. Для $1 \leqslant q \leqslant m$ положим

$$H_q(X) = \frac{1}{\binom{m-1}{q-1}} \sum_{Q \subset \{1, \ldots m\}, |Q| = q} H(X(Q)).$$

   Доказать, что

$$H_1(X) \geqslant H_2(X) \geqslant \ldots \geqslant H_m(X) = H(X).$$

5. Пусть $X_i = \pm 1$, $1 \leqslant i \leqslant n$, — независимые равномерно распределенные случайные величины и $\mathbf{S}_n = \sum_{i=1}^{n} X_i$. Пусть $0 \leqslant p \leqslant \frac{1}{2}$. Доказать, что

$$\Pr[\mathbf{S}_n \geqslant (1-2p)n] \leqslant 2^{H(p)n} 2^{-n},$$

   путем вычисления границы Чернова для $\min_{\lambda \geqslant 0} E[e^{\lambda \mathbf{S}_n}] e^{-\lambda(1-2p)n}$. (Случай $p = 0$ потребует небольшой модификации метода, тем не менее, результат будет таким же.)

ВЕРОЯТНОСТНЫЙ ВЗГЛЯД:

Экстремальные графы

Пусть T и B — непересекающиеся множества мощности m, а G — двудольный граф с долями T и B. Предположим, что G не содержит 4-циклов. Сколько ребер может содержать граф G? Этот вопрос относится к экстремальной теории графов. Удивительно, но для некоторых значений m мы можем дать точный ответ.

Пусть $m = n^2 + n + 1$ и существует проективная плоскость P порядка n (а, следовательно, содержащая m точек). Поставим в соответствие множеству T множество точек проективной плоскости P, а множеству B — прямые на плоскости P, положим $G = G_P$ и пусть вершина $t \in T$ будет смежной с вершиной $b \in B$, только если точка t принадлежит прямой b на плоскости P. Так как две точки не могут принадлежать двум различным прямым, граф G_P не содержит 4-циклов. Утверждается, что граф G_P содержит наибольшее количество ребер среди всех графов G без 4-циклов. Более того, каждый граф G, не содержащий 4-циклов, и содержащий такое количество ребер может быть записан в виде $G = G_P$.

Пусть G не содержит 4-циклов. Выберем случайным образом два различных элемента $b_1, b_2 \in B$. Для каждой вершины $t \in T$ обозначим через $D(t)$ множество вершин $b \in B$, смежных с t, через $d(t) = |D(t)|$ обозначим степень вершины t. Обозначим через I_t индикатор события «вершина t смежна с вершинами b_1, b_2». Тогда

$$\mathbf{E}[I_t] = \Pr[b_1, b_2 \in D(t)] = \binom{d(t)}{2} \Big/ \binom{m}{2}.$$

Положим

$$X = \sum_{t \in T} I_t.$$

Таким образом, случайная величина X равна числу вершин $t \in T$, смежных с вершинами b_1, b_2. Тогда $X \leqslant 1$, т. е. каждая пара вершин b_1, b_2 имеет по крайней мере одного общего соседа. (Неравенство $X \leqslant 1$ эквивалентно утверждению, что граф G не содержит 4-циклов.) Из линейности математического ожидания следует, что

$$\mathbf{E}[X] = \sum_{t \in T} \mathbf{E}[I_t] = \sum_{t \in T} \binom{d(t)}{2} \Big/ \binom{m}{2}.$$

Обозначим через $\overline{d} = m^{-1} \sum_{t \in T} d(t)$ среднее значение степени вершин графа G. Выпуклость функции $\binom{y}{2}$ влечет за собой неравенство

$$\sum_{t \in T} \binom{d(t)}{2} \Big/ \binom{m}{2} \geqslant m \binom{\overline{d}}{2} \Big/ \binom{m}{2},$$

обращающееся в равенство, только если у всех вершин $t \in T$ степени совпадают. Заметим, что

$$1 \geqslant \max X \geqslant \mathbf{E}[X] \geqslant m\binom{\overline{d}}{2} \Big/ \binom{m}{2}.$$

При $G = G_P$ все степени $d(x) = \overline{d}$ (каждая прямая содержит ровно $n+1$ точку), и всегда верно, что $X = 1$ (две точки определяют ровно одну прямую). Таким образом, приведенные выше неравенства обращаются в равенства и

$$1 = m\binom{\overline{d}}{2} \Big/ \binom{m}{2}.$$

Любой граф с большим количеством ребер будет иметь среднюю степень, строго большую $\overline{d}$, а следовательно, не будет выполняться условие $1 \geqslant m\binom{\overline{d}}{2}/\binom{m}{2}$, и граф будет содержать 4-цикл.

Пусть граф G содержит такое же количество вершин, что и G_P, и не содержит 4-циклов. Приведенные выше неравенства должны тогда обратиться в равенства, и всегда должно выполняться $X = 1$. Построим геометрический объект, соответствующий данному графу. Вершинам из множества T поставим в соответствие точки, а каждое множество вершин, смежных с некоторой вершиной $b \in B$, пусть определяет прямую. Так как $X = 1$, каждые две точки определяют единственную прямую. Поменяв местами множества T и B, получаем, что каждые две прямые определяют единственную точку. Тогда получившийся объект — проективная плоскость.

15

Дерандомизация

Математика естественна. Никто не смог бы выдумать математическую вселенную. Она существовала всегда, ждала своих исследователей; она великолепна, она фантастична.

Джон Конвей

Как уже было сказано в гл. 1, с помощью вероятностного метода часто можно получить эффективный вероятностный алгоритм для алгоритмических задач из самых разных областей. В некоторых случаях эти алгоритмы могут быть дерандомизированы и превращены в детерминированные. В этой главе мы обсудим несколько таких случаев.

15.1. МЕТОД УСЛОВНЫХ ВЕРОЯТНОСТЕЙ

Простое применение вероятностного метода дает нам следующее утверждение, являющееся частным случаем теоремы 2.3.1.

Предложение 15.1.1. *Для каждого целого числа n найдется раскраска ребер полного графа K_n в два цвета, при которой число одноцветных копий графа K_4 не превосходит $\binom{n}{4} \cdot 2^{-5}$.*

Действительно, $\binom{n}{4} \cdot 2^{-5}$ — это математическое ожидание числа одноцветных копий графа K_4 при случайной реберной 2-раскраске графа K_n, следовательно, раскраска из утверждения 15.1.1 существует.

Можем ли мы *детерминированным* образом найти такую раскраску за полиномиальное от n время? Опишем процедуру, решающую эту задачу. Она представляет собой частный случай техники, называемой *методом условных вероятностей.*

Прежде всего нам нужно определить весовую функцию для частично раскрашенного графа K_n. При заданной раскраске части ребер графа K_n в красный и синий цвета определим для каждой копии K графа K_4, содержащейся в графе K_n, ее вес $w(K)$ следующим образом. Если в подграфе K найдутся два ребра, раскрашенные в разные цвета, то $w(K) = 0$. Если ни одно из ребер подграфа K не раскрашено, то $w(K) = 2^{-5}$, а если $r \geqslant 1$

ребер подграфа K имеют один и тот же цвет, то $w(K) = 2^{r-6}$. Определим также общий вес W частично раскрашенного графа K_n как сумму $\sum w(K)$, где суммирование ведется по всем копиям K графа K_4 в K_n. Отметим, что вес каждой копии K графа K_4 равен вероятности того, что эта копия окажется одноцветной в случае, если все бесцветные ребра графа K_n будут раскрашены случайно и независимо в красный и синий цвета. Тогда в силу линейности математического ожидания общий вес W представляет собой математическое ожидание числа одноцветных копий графа K_4 в раскраске, полученной путем случайного завершения частичной раскраски графа K_n.

Опишем теперь процедуру нахождения раскраски, описанной в формулировке утверждения 15.1.1. Упорядочим произвольным образом все $\binom{n}{2}$ ребер графа K_n. Построим требуемую 2-раскраску, присваивая в выбранном порядке каждому из ребер красный или синий цвет. Пусть ребра $e_1, \ldots, e_{i-1}$ уже раскрашены и требуется присвоить цвет ребру e_i. Пусть W — текущий вес графа K_n, определяемый частичной раскраской c ребер $e_1, \ldots, e_{i-1}$. Обозначим через W_{red} вес графа K_n, определяемый частичной раскраской, полученной из раскраски c при раскраске ребра e_i в красный цвет, а через W_{blue} — вес графа K_n, определяемый частичной раскраской, полученной из раскраски c при раскраске ребра e_i в синий цвет. По определению веса W (это следует также из свойств математического ожидания)

$$W = \frac{W_{\text{red}} + W_{\text{blue}}}{2}.$$

Цвет ребра e_i выбирается так, чтобы минимизировать получающийся вес, т. е. при $W_{\text{red}} \leqslant W_{\text{blue}}$ раскрашиваем ребро e_i в красный цвет, в противном случае — раскрашиваем его в синий. Из приведенного выше неравенства следует, что в нашем алгоритме вес не возрастает с течением времени. Так как начальное значение веса равно $\binom{n}{4}2^{-5}$, его конечное значение не превосходит этого числа. Процесс останавливается только когда все ребра раскрашены, в этот момент вес равен числу одноцветных копий графа K_4. Следовательно, описанный алгоритм создает детерминированным образом и за полиномиальное время реберную 2-раскраску графа K_n, удовлетворяющую требованиям утверждения 15.1.1.

Опишем теперь метод условных вероятностей в более общем виде. Неявным образом этот метод возник в работе [Erdős and Selfridge (1973)], более явные примеры его применения можно найти в работах [Spencer (1987)] и [Raghavan (1988)]. Рассмотрим некоторое вероятностное пространство. Пусть, для простоты, оно симметрично и состоит из 2^l точек, соответствующих двоичным векторам длины l. Пусть $A_1, \ldots, A_s$ — некоторое множество событий и $\sum_{i=1}^{s} \Pr[A_i] = k$. Таким образом, k — это математическое ожидание числа произошедших событий A_i, а, следовательно, найдется точка пространства $(\varepsilon_1, \ldots, \varepsilon_l)$, в которой происходит не более k событий. Наша задача — найти такую точку детерминированным образом.

Для каждого набора $(\varepsilon_1, \ldots, \varepsilon_{j-1})$ и для каждого события A_i условная вероятность

$$\Pr[A_i \mid \varepsilon_1, \ldots, \varepsilon_{j-1}]$$

события A_i при заданных значениях $\varepsilon_1, \ldots, \varepsilon_{j-1}$, очевидно, равна среднему арифметическому двух условных вероятностей, соответствующих двум возможным значениям ε_j. То есть

$$\Pr[A_i \mid \varepsilon_1, \ldots, \varepsilon_{j-1}] = \frac{\Pr[A_i \mid \varepsilon_1, \ldots, \varepsilon_{j-1}, 0] + \Pr[A_i \mid \varepsilon_1, \ldots, \varepsilon_{j-1}, 1]}{2}.$$

Следовательно,

$$\sum_{i=1}^{s} \Pr[A_i \mid \varepsilon_1, \ldots, \varepsilon_{j-1}] =$$

$$= \frac{\sum_{i=1}^{s} \Pr[A_i \mid \varepsilon_1, \ldots, \varepsilon_{j-1}, 0] + \sum_{i=1}^{s} \Pr[A_i \mid \varepsilon_1, \ldots, \varepsilon_{j-1}, 1]}{2} \geqslant$$

$$\geqslant \min \left\{ \sum_{i=1}^{s} \Pr\left[A_i \mid \varepsilon_1, \ldots, \varepsilon_{j-1}, 0\right], \sum_{i=1}^{s} \Pr\left[A_i \mid \varepsilon_1, \ldots, \varepsilon_{j-1}, 1\right] \right\}.$$

Таким образом, если значения ε_j выбираются поочередно таким образом, чтобы минимизировать значение суммы $\sum_{i=1}^{s} \Pr[A_i \mid \varepsilon_1, \ldots, \varepsilon_j]$, то это значение не может увеличиваться. Так как значение этой суммы изначально равно k, оно не превосходит k и в конце процесса. Но по завершении процесса значения всех ε_j зафиксированы, следовательно, значение суммы равно числу событий A_i произошедших в точке $(\varepsilon_1, \ldots, \varepsilon_l)$, а это означает, что наш алгоритм выдал нужный результат.

Отметим, что требования симметричности вероятностного пространства и того, чтобы оно состояло из 2^l точек, могут быть ослаблены. Описанный выше алгоритм эффективен, если значение l не слишком велико (в комбинаторных задачах так обычно и бывает), и, что более важно, если условные вероятности $\Pr[A_i \mid \varepsilon_1, \ldots, \varepsilon_j]$ могут быть эффективно вычислены для каждого из событий A_i и для каждого из наборов $\varepsilon_1, \ldots, \varepsilon_j$. Это безусловно так для примера, рассмотренного в утверждении 15.1.1. Тем не менее, есть множество интересных задач, для которых упомянутые требования не выполняются. Технический трюк, полезный в такого рода ситуациях, называется *введением функций пессимистических оценок*. Впервые этот прием был использован в работе [Raghavan (1988)]. Рассмотрим вновь описанное выше симметричное вероятностное пространство, состоящее из 2^l точек, и события $A_1, \ldots, A_s$ в нем. Пусть для каждого события A_i и для каждого числа $0 \leqslant j \leqslant l$ имеется эффективно вычислимая функция $f_j^i(\varepsilon_1, \ldots, \varepsilon_j)$. Предположим также, что

$$f_{j-1}^i(\varepsilon_1, \ldots, \varepsilon_{j-1}) \geqslant \min \{f_j^i(\varepsilon_1, \ldots, \varepsilon_{j-1}, 0), f_j^i(\varepsilon_1, \ldots, \varepsilon_{j-1}, 1)\}, \tag{15.1}$$

и что функция f_j^i мажорирует условные вероятности для события A_i, т. е.

$$f_j^i(\varepsilon_1, \ldots, \varepsilon_j) \geqslant \Pr[A_i \mid \varepsilon_1, \ldots, \varepsilon_j]. \tag{15.2}$$

(Было бы достаточно предположить справедливость аналогичных неравенств для сумм по i правых и левых частей последнего неравенства, но в данном случае нас устраивает и такой вариант.) Таким образом, если в начале процесса

$\sum_{i=1}^{s} f_0^i \leqslant t$, и на каждом шаге значение ε_j выбирается так, чтобы минимизировать сумму $\sum_{i=1}^{s} f_j^i(\varepsilon_1 \ldots, \varepsilon_j)$, то в итоге мы получим точку $(\varepsilon_1, \ldots, \varepsilon_l)$, для которой $\sum_{i=1}^{s} f_l^i(\varepsilon_1, \ldots, \varepsilon_l) \leqslant t$. Таким образом, число событий A_i, происходящих в этой точке, не превосходит t. Функции f_j^i, упомянутые выше, называют функциями пессимистических оценок.

Этот метод позволяет получать эффективные алгоритмы в ряде случаев, когда не известен эффективный способ вычисления требуемых условных вероятностей. Примером тому служит следующая теорема, схожая с некоторыми результатами из гл. 12 и 14.

Теорема 15.1.2. *Пусть $(a_{ij})_{i,j=1}^{n}$ — матрица размера $n \times n$, состоящая из действительных чисел, такая, что $-1 \leqslant a_{ij} \leqslant 1$ для всех i, j. Тогда за полиномиальное время можно найти такие значения $\varepsilon_1, \ldots, \varepsilon_n \in \{-1, 1\}$, что для каждого i, $1 \leqslant i \leqslant n$, выполняется неравенство*

$$\left| \sum_{j=1}^{n} \varepsilon_j a_{ij} \right| \leqslant \sqrt{2n \ln(2n)}.$$

Доказательство. Рассмотрим симметричное вероятностное пространство, состоящее из 2^n точек, соответствующих 2^n возможным векторам $(\varepsilon_1, \ldots, \varepsilon_n) \in \{-1, 1\}^n$. Положим $\beta = \sqrt{2n \ln(2n)}$ и обозначим через A_i событие, заключающееся в том, что $|\sum_{j=1}^{n} \varepsilon_j a_{ij}| > \beta$. Покажем, что метод условных вероятностей при использовании соответствующих пессимистичных оценок позволяет эффективно определить точку пространства, в которой не происходит ни одно из событий A_i.

Положим $\alpha = \beta/n$ и определим функцию $G(x)$ как

$$G(x) = \operatorname{ch}(\alpha x) = \frac{e^{\alpha x} + e^{-\alpha x}}{2}.$$

Путем сопоставления соответствующих рядов Тейлора несложно убедиться в том, что для всех действительных чисел x

$$G(x) \leqslant e^{\frac{\alpha^2 x^2}{2}},$$

причем неравенство обращается в строгое, когда x и α одновременно отличны от нуля. Не составляет труда также проверить, что для всех действительных чисел x и y справедливо соотношение

$$G(x)G(y) = \frac{G(x+y) + G(x-y)}{2}.$$

Определим теперь функции f_p^i, которые будут выполнять роль наших пессимистичных оценок. Для каждого $1 \leqslant i \leqslant n$ и для каждого набора $\varepsilon_1, \ldots, \varepsilon_p \in \{-1, 1\}$ положим

$$f_p^i(\varepsilon_1, \ldots, \varepsilon_p) = 2e^{-\alpha\beta} G\left(\sum_{j=1}^{p} \varepsilon_j a_{ij} \right) \prod_{j=p+1}^{n} G(a_{ij}).$$

Несомненно, эти функции могут быть эффективно вычислены. Остается убедиться в том, что они удовлетворяют условиям (15.1), (15.2), и в том, что сумма $\sum_{i=1}^{n} f_0^i$ меньше 1. Подтверждением тому служат следующие результаты.

Утверждение 15.1.3. *Для каждого* $1 \leqslant i \leqslant n$ *и для каждого набора* $\varepsilon_1, \ldots, \varepsilon_{p-1} \in \{-1, 1\}$ *справедливо неравенство*

$$f_{p-1}^i(\varepsilon_1, \ldots, \varepsilon_{p-1}) \geqslant \min\ \{f_p^i(\varepsilon_1, \ldots, \varepsilon_{p-1}, -1), f_p^i(\varepsilon_1, \ldots, \varepsilon_{p-1}, 1)\}.$$

Доказательство. Положим $v = \sum_{j=1}^{p-1} \varepsilon_j a_{ij}$. Из определения функций f_p^i и свойств функции G вытекает, что

$$\begin{aligned} f_{p-1}^i(\varepsilon_1, \ldots, \varepsilon_{p-1}) &= 2e^{-\alpha\beta} G(v) G(a_{ip}) \prod_{j=p+1}^{n} G(a_{ij}) = \\ &= 2e^{-\alpha\beta} \frac{G(v - a_{ip}) + G(v + a_{ip})}{2} \prod_{j=p+1}^{n} G(a_{ij}) = \\ &= \frac{f_p^i(\varepsilon_1, \ldots, \varepsilon_{p-1}, -1) + f_p^i(\varepsilon_1, \ldots, \varepsilon_{p-1}, 1)}{2} \geqslant \\ &\geqslant \min\ \{\, f_p^i(\varepsilon_1, \ldots, \varepsilon_{p-1}, -1), f_p^i(\varepsilon_1, \ldots, \varepsilon_{p-1}, 1)\}, \end{aligned}$$

а это и требовалось доказать. ■

Утверждение 15.1.4. *Для каждого* $1 \leqslant i \leqslant n$ *и для каждого набора* $\varepsilon_1, \ldots, \varepsilon_{p-1} \in \{-1, 1\}$ *справедливо неравенство*

$$f_{p-1}^i(\varepsilon_1, \ldots, \varepsilon_{p-1}) \geqslant \Pr(A_i \mid \varepsilon_1, \ldots, \varepsilon_{p-1}).$$

Доказательство. Пусть v означает то же, что и в доказательстве утверждения 15.1.3. Тогда

$$\begin{aligned} \Pr[A_i \mid \varepsilon_1, \ldots, \varepsilon_{p-1}] &\leqslant \Pr[v + \sum_{j \geqslant p} \varepsilon_j a_{ij} > \beta] + \Pr[-v - \sum_{j \geqslant p} \varepsilon_j a_{ij} > \beta] = \\ &= \Pr[e^{\alpha(v + \sum_{j \geqslant p} \varepsilon_j a_{ij})} > e^{\alpha\beta}] + \Pr[e^{-\alpha(v + \sum_{j \geqslant p} \varepsilon_j a_{ij})} > e^{\alpha\beta}] \leqslant \\ &\leqslant e^{\alpha v} e^{-\alpha\beta} \mathbf{E}[e^{\alpha(\sum_{j \geqslant p} \varepsilon_j a_{ij})}] + e^{-\alpha v} e^{-\alpha\beta} \mathbf{E}[e^{-\alpha(\sum_{j \geqslant p} \varepsilon_j a_{ij})}] = \\ &= 2e^{-\alpha\beta} G(v) \prod_{j \geqslant p} G(a_{ij}) = f_{p-1}^i(\varepsilon_1, \ldots, \varepsilon_{p-1}). \end{aligned}$$

Утверждение 15.1.4 доказано. ■

Осталось показать, что $\sum_{i=1}^{n} f_0^i < 1$. Действительно, из свойств функции G, а также из выбора α и β следует, что

$$\begin{aligned} \sum_{i=1}^{n} f_0^i = \sum_{i=1}^{n} 2e^{-\alpha\beta} \prod_{j=1}^{n} G(a_{ij}) &\leqslant \sum_{i=1}^{n} 2e^{-\alpha\beta} \prod_{j=1}^{n} e^{\frac{\alpha^2 a_{ij}^2}{2}} \leqslant \\ &\leqslant \sum_{i=1}^{n} 2e^{-\alpha\beta} e^{\frac{\alpha^2 n}{2}} = 2ne^{\frac{\alpha^2 n}{2} - \alpha\beta} = 2ne^{-\frac{\alpha^2 n}{2}} = 1. \end{aligned}$$

Более того, первое неравенство является строгим, если не все $a_{ij} = 0$, а второе неравенство является строгим, если не все $a_{ij}^2 = 1$. Теорема полностью доказана. ■

15.2. d-НЕЗАВИСИМЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ В ПРОСТРАНСТВАХ МАЛОГО РАЗМЕРА

Класс сложности NC представляет собой, грубо говоря, множество тех задач, которые могут быть решены за полилогарифмическое (от размера входных данных) время на полиномиальном числе параллельно работающих процессоров. При исследовании этого класса используются несколько моделей параллельных вычислений. Наиболее распространенной является модель EREW (=Exclusive Read, Exclusive Write) PRAM, в которой различные процессоры не могут обмениваться данными с одной и той же ячейкой памяти одновременно. Более подробно эта модель рассмотрена в работе [Karp and Ramachandran (1990)].

Обозначим через n размер входных данных. Некоторые операции могут быть легко выполнены в классе NC. Например, можно записать содержимое ячейки c в $m = n^{O(1)}$ ячеек за время $O(\log n)$, с использованием, скажем, m процессоров. Для того, чтобы сделать это, рассмотрим полное двоичное дерево с m листьями и поставим в соответствие каждой из его внутренних вершин процессор, а каждому процессору назначим по одной ячейке памяти. Сначала процессор, соответствующий корню дерева, считывает содержимое ячейки c и записывает его в две ячейки смежных с ним процессоров. Затем, оба этих процессора параллельно считывают информацию из своих ячеек и передают ее дальше. В общем случае, на i-м шаге все процессоры, находящиеся на расстоянии $i-1$ от корневого, параллельно считывают информацию, записанную на предыдущем шаге в их ячейки, и записывают ее в ячейки соседей снизу. Ясно, что процесс, как и требовалось, закончится за время $O(\log m)$. (В действительности, можно показать, что для этой цели достаточно $O(m/\log m)$ процессоров, но мы не будем оптимизировать здесь это число.)

Аналогичный прием можно использовать для сложения m чисел на m процессорах за время $O(\log m)$. Будем считать, что упомянутые числа лежат на листьях полного двоичного дерева с m листьями, а на i-м шаге каждый из процессоров, находящийся на расстоянии i от листьев, вычисляет сумму чисел, находящихся в ячейках его соседей снизу. Ясно, что в корне дерева окажется значение интересующей нас суммы за время $O(\log m)$.

Вернемся теперь к задаче реберной раскраски полного графа K_n, рассмотренной в утверждении 15.1.1. Из сказанного выше вытекает, что задача *проверки* условия, что заданная реберная раскраска содержит не более $\binom{n}{4}2^{-5}$ одноцветных копий графа K_4, принадлежит классу NC. Это означает, что такая проверка может быть выполнена за время $(\log n)^{O(1)}$ (в действительности, за время $O(\log n)$) при использовании $n^{O(1)}$ процессоров. Действительно, построим сначала $\binom{n}{4}$ копий данной раскраски. Затем назначим каждому процессору по копии графа K_4 в K_n, которую он будет проверять на монохроматичность (все эти проверки могут вестись параллельно, так как у нас есть достаточное количество копий раскраски). Затем суммированием соответствующих индикаторов получим число процессоров, которым достались одноцветные копии. Ясно, что этот процесс потребует $n^{O(1)}$ параллельных процессоров и завершится за время $O(\log n)$.

Итак, мы можем *проверить* в NC, удовлетворяет ли заданная раскраска графа K_n условиям утверждения 15.1.1. Возникает вопрос: принадлежит ли классу NC задача нахождения такой раскраски? Метод, описанный в предыдущем разделе, не годится, так как ребра в нем раскрашиваются по очереди, а, значит, такая последовательная процедура займет время $\Omega(n^2)$. Тем не менее, оказывается, что эта задача принадлежит NC, и мы можем найти раскраску с нужными свойствами, применив метод, основывающийся на технике, впервые предложенной в работе [Joffe (1974)], и развитой впоследствии многими исследователями. Этот метод позволяет преобразовывать вероятностные алгоритмы, работа которых зависит только от d-независимых, а не от полностью независимых случайных выборов (для некоторой константы d) в детерминированные (а во многих случаях и в параллельные) алгоритмы. Изложенный здесь подход аналогичен подходу из работы [Alon, Babai and Itai (1986)], но для упрощения модели мы будем рассматривать случайные величины, которые принимают только два значения 0 и 1, каждое с равной вероятностью.

Основная идея заключается в замене экспоненциально большого пространства элементарных событий пространством полиномиального размера. Если случайная величина в таком пространстве принимает какое-то значение с ненулевой вероятностью, то мы можем найти точку, где это происходит, простым перебором всех точек пространства. Это можно сделать без каких-либо потерь во времени с использованием полиномиального числа параллельных процессоров. Заметим, что для задачи реберной раскраски, рассмотренной в утверждении 15.1.1, достаточно 6-независимости случайных величин, соответствующих цветам ребер, ведь в этом случае вероятность «оказаться монохроматической» равна 2^{-5} для каждой из копий K_4, а это дает нам требуемое математическое ожидание числа одноцветных копий. Следовательно, в этом конкретном случае достаточно построить вероятностное пространство размера $n^{O(1)}$ из $\binom{n}{2}$ случайных величин, каждая из которых принимает значения 0 и 1 с вероятностью $1/2$, такое, что каждые 6 случайных величин в нем взаимно независимы.

Пространства исходов малого размера с большим количеством d-независимых $0,1$-случайных величин, могут быть получены из любых линейных кодов, исправляющих ошибки, с подходящими параметрами. Способ, описанный ниже, основывается на двоичных кодах Боуза—Чоудхури—Хоквингема (БЧХ-кодах) (см., например, [MacWilliams and Sloane (1977)]).

Теорема 15.2.1. *Пусть $n = 2^k - 1$ и $d = 2t + 1$. Тогда найдется симметричное вероятностное пространство Ω размера $2(n+1)^t$ и d-независимые случайные величины $y_1, \ldots, y_n$, определенные над пространством Ω, каждая из которых принимает значения 0 и 1 с вероятностью $1/2$.*

Пространство и случайные величины могут быть явным образом построены при заданном представлении поля $F = GF(2^k)$ в виде k-мерной алгебры над полем $GF(2)$.

Доказательство. Пусть $x_1, \ldots, x_n$ — n ненулевых элементов поля F, представленных в виде вектор-столбцов длины k над полем $GF(2)$. Обозначим

через H следующую матрицу размера $1+kt$ на n над полем $GF(2)$:

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^3 & x_2^3 & \dots & x_n^3 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ x_1^{2t-1} & x_2^{2t-1} & \dots & x_n^{2t-1} \end{pmatrix}.$$

Это матрица проверки четности расширенного двоичного БЧХ-кода длины n с кодовым расстоянием $2t+2$. Хорошо известно, что любые $d = 2t+1$ столбцов матрицы H линейно независимы над полем $GF(2)$. Для полноты изложения мы доказываем следующую лемму.

Лемма 15.2.2. *Любые $d = 2t+1$ столбцов матрицы H линейно независимы над полем $GF(2)$.*

Доказательство. Пусть $J \subset \{1, 2, \dots, n\}$ — некоторое $(2t+1)$-элементное подмножество множества индексов столбцов матрицы H. Пусть $\sum_{j \in J} z_j H_j = 0$, где H_j обозначает j-й столбец матрицы H и $z_j \in GF(2)$. Требуется показать, что $z_j = 0$ для всех $j \in J$. По нашему предположению

$$\sum_{j \in J} z_j x_j^i = 0 \tag{15.3}$$

для $i = 0$ и всех нечетных i, удовлетворяющих условию $1 \leqslant i \leqslant 2t-1$. Пусть число $a = 2^b \cdot l$ нечетно при $l \leqslant 2t-1$. Для $i = l$ возведем равенство (15.3) в квадрат b раз. Учитывая, что $(u+v)^2 = u^2 + v^2$ по модулю 2, а также, что $z_j = z_j^2$ для всех j, поскольку все z_j равны либо 0, либо 1, получаем, что равенство (15.3) верно для $i = a$. Следовательно, равенство (15.3) верно для всех i, $0 \leqslant i \leqslant 2t$. Таким образом, мы имеем однородную систему $2t+1$ линейных уравнений с $2t+1$ неизвестными. Матрица этой системы представляет собой матрицу Вандермонда, следовательно, она невырождена. Верно, что единственным решением этой системы является тривиальное решение: $z_j = 0$ при всех $j \in J$, что и требовалось доказать. ■

Возвращаясь к доказательству теоремы, положим $\Omega = \{1, 2, \dots, 2(n+1)^t\}$, и рассмотрим $(0,1)$-матрицу $A = (a_{ij}), i \in \Omega, 1 \leqslant j \leqslant n$, в которой $2(n+1)^t = 2^{kt+1}$ строк представляют собой все возможные линейные комбинации (над полем $GF(2)$) строк матрицы H. Пространство исходов Ω дополнено теперь равномерной вероятностной мерой, определим случайные величины y_j равенствами $y_j(i) = a_{ij}$ для всех $i \in \Omega, 1 \leqslant j \leqslant n$.

Остается показать, что случайные величины y_j являются d-независимыми, и каждая из них принимает значения 0 и 1 с равной вероятностью. Для этого нам нужно показать, что для любого набора J из не более чем d столбцов матрицы A строки матрицы $A_J = (a_{ij}), i \in \Omega, j \in J$, размера $|\Omega|$ на $|J|$ содержат одинаковое количество копий каждого из $2^{|J|}$ $(0,1)$-векторов длины

$|J|$. Заметим, что по лемме 15.2.2 столбцы соответствующей подматрицы H_J матрицы H линейно независимы. Число строк матрицы A_J, совпадающих с некоторым произвольно заданным вектором, равно количеству образующих этот вектор линейных комбинаций строк матрицы H_J. Это количество равно числу решений системы из $|J|$ линейно независимых линейных уравнений с $kt+1$ неизвестными, которое равно $2^{kt+1-|J|}$, независимо от вектора свободных коэффициентов. Теорема полностью доказана. ■

Теорема 15.2.1 предлагает нам эффективный способ построения для каждых фиксированных чисел d и n пространства элементарных событий размера $O\left(n^{\lfloor d/2 \rfloor}\right)$ и n d-независимых случайных величин в нем, каждая из которых принимает значения 0 и 1 с равной вероятностью. В частности, мы можем использовать такое пространство размера $O\left(\binom{n}{2}^3\right) = O(n^6)$ для нахождения раскраски, как в утверждении 15.1.1, в NC. Несколько других применений теоремы 15.2.1 содержится в работе [Alon et al. (1986)].

Возникает естественный вопрос: можно ли уменьшить размер $O(n^{\lfloor d/2 \rfloor})$? Далее мы покажем, что этот размер является наименьшим из возможных, с точностью до постоянного множителя (зависящего от d).

Назовем случайную величину *почти константой*, если она принимает некоторое значение с вероятностью 1. Определим функцию $m(n, d)$ следующим образом:

$$m(n,d) = \sum_{j=0}^{d/2} \binom{n}{j}, \text{ если } d \text{ четно,}$$

и

$$m(n,d) = \sum_{j=0}^{(d-1)/2} \binom{n}{j} + \binom{n-1}{(d-1)/2}, \text{ если } d \text{ нечетно.}$$

Отметим, что $m(n,d) = \Omega(n^{\lfloor d/2 \rfloor})$ для каждого фиксированного числа d.

Предложение 15.2.3. *Если случайные величины $y_1, \ldots, y_n$, определенные на пространстве исходов Ω, являются d-независимыми и ни одна из них не является почти константой, то $|\Omega| \geqslant m(n,d)$.*

Заметим, что мы не предполагаем здесь ни симметричности пространства Ω, ни того, что величины y_j являются $(0,1)$-переменными.

Доказательство. Не ограничивая общности рассуждений, мы можем полагать, что математическое ожидание каждой из переменных y_j равно 0 (иначе можно было бы заменить переменную y_j переменной $y_j - \mathbf{E}[y_j]$). Для каждого подмножества S множества $\{1, \ldots, n\}$, положим $\alpha_S = \prod_{j \in S} y_j$. Отметим, что поскольку ни одна из переменных y_j не является почти константой, и так как эти случайные величины являются d-независимыми, для всех множеств S мощности не более d выполнено неравенство

$$\mathbf{E}[\alpha_S \alpha_S] = \prod_{j \in S} \mathrm{Var}[y_j] > 0. \tag{15.4}$$

Аналогично, для всех множеств S и T, таких, что $|S \cup T| \leqslant d$ и $S \neq T$, имеем

$$\mathbf{E}[\alpha_s \alpha_T] = \prod_{j \in S \cap T} \mathrm{Var}[y_j] \prod_{j \in S \cup T \setminus (S \cap T)} \mathbf{E}[y_j] = 0. \tag{15.5}$$

Пусть $m = m(n, d)$ и подмножества $S_1, \ldots S_m$ множества $\{1, \ldots, n\}$ таковы, что мощность объединения любых двух из них не превосходит d. (Возьмем все подмножества мощности не более $d/2$ и, если d нечетно, добавим все подмножества мощности $(d+1)/2$, содержащие 1.)

Покажем, что m функций α_{S_j} (рассматриваемых как действительные векторы длины $|\Omega|$) линейно независимы. Отсюда будет следовать, что $|\Omega| \geqslant m = m(n, d)$, что и требуется.

Для доказательства линейной независимости положим $\sum_{j=1}^{m} c_j \alpha_{S_j} = 0$. Умножая равенство на α_{S_i} и вычисляя математическое ожидание с использованием соотношения (15.5), получаем

$$0 = \sum_{j=1}^{m} c_j \mathbf{E}[\alpha_{S_j} \alpha_{S_i}] = c_i \mathbf{E}[\alpha_{S_i} \alpha_{S_i}].$$

Отсюда и из (15.4) следует, что $c_i = 0$ для всех i. Таким образом, линейная независимость доказана, а, значит, доказано и требуемое утверждение. ■

Последнее утверждение показывает, что размер пространства исходов с n d-независимыми невырожденными случайными величинами может быть полиномиальным от n, только если значение d фиксировано. Тем не менее, как было показано в работе [Naor and Naor (1990)], если требовать только *почти* d-независимости случайных величин, то размер может быть полиномиальным даже при $d = \Omega(\log n)$. Такие пространства элементарных событий со случайными величинами, которые могут быть явным образом построены несколькими способами, имеют множество интересных применений, в которых достаточно почти d-независимости. Более подробно этот вопрос рассматривается в работах [Naor and Naor (1990)] и [Alon, Goldreich, Håstad and Peralta (1990)].

15.3. УПРАЖНЕНИЯ

1. Пусть множества $A_1, \ldots, A_n \subseteq \{1, \ldots, m\}$ таковы, что $\sum_{i=1}^{n} 2^{1-|A_i|} < 1$. Доказать, что существует 2-раскраска $\chi : \{1, \ldots, m\} \to \{0, 1\}$, при которой ни одно из множеств A_i не является монохромным. При $m = n$ описать детерминированный алгоритм нахождения такой раскраски χ за полиномиальное время.

2. Найти детерминированный алгоритм, строящий при заданном n за полиномиальное от n время семейство $\mathcal{F}$ из n^{10} подмножеств множества $N = \{1, 2, \ldots, n\}$, такое, что мощность каждого из множеств $F \in \mathcal{F}$ не превосходит $10 \log_2 n$ и такое, что для каждого семейства $\mathcal{G}$ из n подмножеств мощности $n/2$ множества N найдется множество $F \in \mathcal{F}$, имеющее непустое пересечение со всеми множествами семейства $\mathcal{G}$.

ВЕРОЯТНОСТНЫЙ ВЗГЛЯД:

Число пересечений, инцидентности, суммы и произведения

Этот раздел мы начнем с изложения одного простого утверждения из теории графов, имеющего вероятностное доказательство. Затем будут рассмотрены некоторые его замечательные следствия, касающиеся комбинаторной геометрии и комбинаторной теории чисел. Вариации большей части этих, на первый взгляд, не связанных между собой утверждений, уже доказывались ранее, причем, доказательства эти были намного более громоздки. До появления новых версий доказательств, приведенных здесь, единственным намеком на то, что между этими утверждениями может быть какая-то связь, являлся тот факт, что соавтором всех первых версий доказательств был А. Семереди.

Вложением графа $G = (V, E)$ в плоскость называют его планарное представление, в котором каждой вершине соответствует точка плоскости, а каждому ребру $\{u, v\}$ соответствует кривая, соединяющая точки, соответствующие вершинам u и v. *Числом пересечений* такого вложения называют число пар пересекающихся кривых, соответствующих парам неинцидентных ребер. *Число пересечений* $\mathrm{cr}(G)$ графа G — это минимально возможное число пересечений его вложения в плоскость. Следующая теорема была доказана в работе [Ajtai, Chvátal, Newborn and Szemerédi (1982)] и, независимо, Лейтоном. Ниже мы приводим ее очень короткое вероятностное доказательство.

Теорема 1. *Число пересечений произвольного графа $G = (V, E)$ без петель и кратных ребер, такого, что $|E| \geqslant 4|V|$, не меньше $|E|^3/(64|V|^2)$.*

Доказательство. Из формулы Эйлера следует, что у планарного n-вершинного графа без петель и кратных ребер может быть не более $3n$ ребер, тогда число пересечений любого простого n-вершинного графа с m ребрами должно быть не меньше $m - 3n$. Пусть $G = (V, E)$ — граф, вложенный в плоскость с $t = \mathrm{cr}(G)$ пересечениями, и $|E| \geqslant 4|V|$. Обозначим через H случайный подграф графа G, полученный случайным и независимым выбором вершин графа G, таким, что каждая вершина оказывалась в графе H с вероятностью p (значение p будет выбрано позже). Математическое ожидание числа вершин графа H равно $p|V|$, математическое ожидание числа его ребер равно $p^2|E|$, а математическое ожидание числа пересечений в заданном его вложении равно $p^4 t$, откуда следует, что математическое ожидание числа пересечений не превосходит $p^4 t$. Таким образом, $p^4 t \geqslant p^2|E| - 3p|V|$, а, значит,

$$\mathrm{cr}(G) = t \geqslant \frac{|E|}{p^2} - 3\frac{|V|}{p^3}.$$

Не делая попытки оптимизировать постоянный множитель, подставим число $p = 4|V|/|E|$ $(\leqslant 1)$, для получения требуемого результата. ■

В работе [Székely (1997)] было показано, что с помощью этого результата можно получить удивительно простое доказательство теоремы Семереди и Троттера, относящейся к комбинаторной геометрии. Исходное доказательство намного сложнее.

Теорема 2. *Пусть P — множество, состоящее из n различных точек плоскости, а L — множество, состоящее из m различных прямых. Тогда число инцидентностей между элементами множеств P и L (т. е. количество пар (p,l), где $p \in P$, $l \in L$ и $p \in l$) не превосходит $c \cdot (m^{2/3}n^{2/3} + m + n)$ для некоторой абсолютной константы c.*

Доказательство. Обозначим число инцидентностей через I. Пусть $G = (V, E)$ — граф на множестве вершин P, в котором две вершины смежны, только если они являются соседними точками на некоторой прямой из множества L. Ясно, что, $|V| = n$ и $|E| = I - m$. Заметим, что граф G уже вложен в плоскость, причем ребрам соответствуют отрезки прямых из множества L. В этом вложении каждое пересечение дает точку, принадлежащую двум прямым из множества L, следовательно, $\mathrm{cr}(G) \leqslant \binom{m}{2} \leqslant m^2/2$. По теореме 1 либо $I - m = |E| < 4|V| = 4n$ и $I \leqslant m + 4n$, либо

$$\frac{m^2}{2} \geqslant \mathrm{cr}(G) \geqslant \frac{(I-m)^3}{64n^2}$$

и $I \leqslant (32)^{1/3}m^{2/3}n^{2/3} + m$. В обоих случаях $I \leqslant 4(m^{2/3}n^{2/3} + m + n)$, что и требовалось доказать. ■

Аналогичным образом можно показать, что максимальное число инцидентностей между n точками и m единичными окружностями на плоскости не превосходит $O(m^{2/3}n^{2/3}+m+n)$, отсюда следует, что число единичных расстояний, определяемых n точками плоскости не превосходит $O(n^{4/3})$. Указанная выше верхняя оценка числа инцидентностей точек и прямых точна (если не учитывать постоянный множитель), но в то же время старая гипотеза Эрдёша гласит, что число единичных расстояний, определяемых n точками плоскости, не превосходит $c_\varepsilon n^{1+\varepsilon}$ для любого $\varepsilon > 0$. Тем не менее, оценка $O(n^{4/3})$ — наилучшая из известных верхних оценок. Она была получена Спенсером, Семереди и Троттером гораздо более сложным способом.

В работе [Elekes (1997)] было найдено несколько применений для теоремы 2 в аддитивной теории чисел. Доказательства также получились очень простыми. Примером может служить следующая теорема.

Теорема 3. *Для любых трех множеств действительных чисел A, B и C мощности s верно, что*

$$|A \cdot B + C| = |\{ab + c : \ a \in A, b \in B, c \in C\}| \geqslant \Omega(s^{3/2}).$$

Доказательство. Положим $R = A \cdot B + C$, $|R| = r$ и пусть

$$P = \{(a,t) : a \in A, t \in R\}, \qquad L = \{y = bx + c : b \in B, c \in C\}.$$

Таким образом, P — множество из $n = sr$ точек плоскости, L — множество из $m = s^2$ прямых на плоскости, и каждая прямая $y = bx + c$ из множества L инцидентна s точкам из множества P, т. е. всем точкам $\{(a, ab + c) : a \in A\}$. Следовательно, по теореме 2, $s^3 \leqslant 4(s^{4/3}(sr)^{2/3} + sr + s^2)$, а, значит, $r \geqslant \Omega(s^{3/2})$, что и требовалось доказать. ■

Аналогично можно показать, что для каждого множества A из n действительных чисел выполнено либо $|A+A| \geqslant \Omega(n^{5/4})$, либо $|A \cdot A| \geqslant n^{5/4}$. Это сильно улучшает и упрощает соответствующий результат Эрдёша и Семереди.

Приложение $\mathcal{A}$

Оценки для больших уклонений

A.1. ОЦЕНКИ ДЛЯ БОЛЬШИХ УКЛОНЕНИЙ

В этом приложении приводятся некоторые базовые оценки для больших уклонений, которые часто оказываются полезными при применении вероятностного метода. Наше изложение не опирается на внешние источники. Бóльшая часть результатов может быть найдена или легко получена из основополагающей работы [Chernoff (1952)]. В случае, если мы руководствуемся асимптотическими соображениями, неравенства доказываются для всех значений параметров из заданной области. Следующий результат, несмотря на то, что он описывает частный случай, хорошо раскрывает основную идею этого приложения.

Теорема A.1.1. *Пусть* $X_i, 1 \leqslant i \leqslant n, -$ *взаимно независимые случайные величины, такие, что*

$$\Pr[X_i = +1] = \Pr[X_i = -1] = \frac{1}{2}.$$

Положим

$$\mathbf{S}_n = X_1 + \ldots + X_n$$

и пусть $a > 0$. *Тогда*

$$\Pr[\mathbf{S}_n > a] < e^{-a^2/2n}.$$

Замечание. Для больших n из центральной предельной теоремы вытекает, что случайная величина $\mathbf{S}_n$ имеет приблизительно нормальное распределение с нулевым средним значением и стандартным отклонением равным $\sqrt{n}$. В частности, при любом фиксированном u имеем:

$$\lim_{n\to\infty} \Pr[\mathbf{S}_n > u\sqrt{n}] = \int_{t=u}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt,$$

причем, можно явным образом показать, что последняя величина меньше $e^{-u^2/2}$. Наше доказательство, подчеркнем это еще раз, верно для всех значений n и $a > 0$.

Нам потребуется неравенство Маркова: если Y — произвольная неотрицательная случайная величина и $\alpha > 0$, то

$$\Pr[Y > \alpha \mathbf{E}[Y]] < 1/\alpha.$$

Доказательство теоремы А.1.1. Зафиксируем значения n, a и пусть, на короткое время, значение $\lambda > 0$ произвольно. Для $1 \leqslant i \leqslant n$ имеем:

$$\mathbf{E}[e^{\lambda X_i}] = (e^{\lambda} + e^{-\lambda})/2 = \operatorname{ch}(\lambda).$$

Нам потребуется неравенство

$$\operatorname{ch}(\lambda) \leqslant e^{\lambda^2/2},$$

справедливое для всех значений $\lambda > 0$, оно следует из леммы А.1.5 при $\alpha = 0$. (Это неравенство может быть получено более простым способом — путем почленного сравнения рядов Тейлора этих функций.) В силу определения величины $\mathbf{S}_n$ справедливо равенство

$$e^{\lambda \mathbf{S}_n} = \prod_{i=1}^{n} e^{\lambda X_i}.$$

Так как величины X_i взаимно независимы, то взаимно независимы и случайные величины $e^{\lambda X_i}$, следовательно, математическое ожидание произведения равно произведению математических ожиданий:

$$\mathbf{E}[e^{\lambda \mathbf{S}_n}] = \prod_{i=1}^{n} \mathbf{E}[e^{\lambda X_i}] = [\operatorname{ch}(\lambda)]^n < e^{\lambda^2 n/2}.$$

Отметим, что $\mathbf{S}_n > a$, только если $e^{\lambda \mathbf{S}_n} > e^{\lambda a}$, и применим неравенство Маркова:

$$\Pr[\mathbf{S}_n > a] = \Pr[e^{\lambda \mathbf{S}_n} > e^{\lambda a}] < \mathbf{E}[e^{\lambda \mathbf{S}_n}]/e^{\lambda a} \leqslant e^{\lambda^2 n/2 - \lambda a}.$$

Чтобы оптимизировать неравенство, положим $\lambda = a/n$. Тогда $\Pr[\mathbf{S}_n > a] < e^{-a^2/2n}$, как и утверждалось. ■

Из соображений симметрии получаем

Следствие А.1.2. *Пусть выполнены условия теоремы А.1.1. Тогда*

$$\Pr[|\mathbf{S}_n| > a] < 2e^{-a^2/2n}.$$

Далее мы будем иметь дело с случайными величинами X, распределенными следующим образом.

Предположение А.1.3.

$$p_1, \ldots, p_n \in [0, 1],$$
$$p = (p_1 + \ldots + p_n)/n;$$

$X_1, \ldots, X_n$ взаимно независимы и имеют следующее распределение вероятностей:

$$\Pr[X_i = 1 - p_i] = p_i,$$
$$\Pr[X_i = -p_i] = 1 - p_i;$$
$$X = X_1 + \ldots + X_n.$$

Замечание. Ясно, что $\mathbf{E}[X] = \mathbf{E}[X_i] = 0$. При всех $p_i = 1/2$ величина X имеет то же распределение, что и случайная величина $\mathbf{S}_n/2$. Если все $p_i = p$, то X имеет распределение $B(n, p) - np$, где через $B(n, p)$ обозначено биномиальное распределение.

Теорема А.1.4. *При предположении А.1.3 и $a > 0$ имеем*

$$\Pr[X > a] < e^{-2a^2/n}.$$

Лемма А.1.5. *Для всех действительных чисел α, β, таких, что $|\alpha| \leqslant 1$, справедливо неравенство*

$$\operatorname{ch}(\beta) + \alpha \operatorname{sh}(\beta) \leqslant e^{\beta^2/2 + \alpha\beta}.$$

Доказательство. Неравенство очевидно при $\alpha = +1$, $\alpha = -1$ или $|\beta| \geqslant 100$. Если бы лемма была неверна, то функция

$$f(\alpha, \beta) = \operatorname{ch}(\beta) + \alpha \operatorname{sh}(\beta) - e^{\beta^2/2 + \alpha\beta}$$

достигала бы положительного глобального максимума во внутренней части прямоугольника

$$R = \{(\alpha, \beta) : |\alpha| \leqslant 1, |\beta| \leqslant 100\}.$$

Приравняв частные производные к нулю, получаем

$$\operatorname{sh}(\beta) + \alpha \operatorname{ch}(\beta) = (\alpha + \beta) e^{\beta^2/2 + \alpha\beta},$$
$$\operatorname{sh}(\beta) = \beta e^{\beta^2/2 + \alpha\beta},$$

следовательно, $\operatorname{th}(\beta) = \beta$, а это верно только при $\beta = 0$. Но $f(\alpha, 0) = 0$ для всех α. Полученное противоречие доказывает лемму. ■

Лемма А.1.6. *Для всех $\theta \in [0, 1]$ и всех λ справедливо неравенство*

$$\theta e^{\lambda(1-\theta)} + (1 - \theta) e^{-\lambda\theta} \leqslant e^{\lambda^2/8}.$$

Доказательство. Положив $\theta = (1 + \alpha)/2$ и $\lambda = 2\beta$, сводим лемму А.1.6 к лемме А.1.5. ■

Доказательство теоремы А.1.4. Пусть, на короткое время, значение $\lambda > 0$ произвольно. По лемме А.1.6

$$\mathbf{E}[e^{\lambda X_i}] = p_i e^{\lambda(1-p_i)} + (1 - p_i) e^{-\lambda p_i} \leqslant e^{\lambda^2/8}.$$

Тогда

$$\mathbf{E}[e^{\lambda X}] = \prod_{i=1}^{n} \mathbf{E}[e^{\lambda X_i}] \leqslant e^{\lambda^2 n/8}.$$

Применяя неравенство Маркова, получаем

$$\Pr[X > a] = \Pr[e^{\lambda X} > e^{\lambda a}] < \mathbf{E}[e^{\lambda X}]/e^{\lambda a} \leqslant e^{\lambda^2 n/8 - \lambda a}.$$

Чтобы оптимизировать неравенство, положим $\lambda = 4a/n$. Получаем

$$\Pr[X > a] < e^{-2a^2/n},$$

как и утверждалось. ■

Из соображений симметрии получаем

Следствие А.1.7. *При предположении А.1.3 и $a > 0$ справедливо неравенство*

$$\Pr[|X| > a] < 2e^{-2a^2/n}.$$

При предположении А.1.3 и при произвольном λ имеем:

$$\mathbf{E}[e^{\lambda X}] = \prod_{i=1}^{n} \mathbf{E}[e^{\lambda X_i}] = \prod_{i=1}^{n} [p_i e^{\lambda(1-p_i)} + (1-p_i)e^{-\lambda p_i}] =$$
$$= e^{-\lambda pn} \prod_{i=1}^{n} [p_i e^{\lambda} + (1-p_i)].$$

При фиксированном λ функция

$$f(x) = \ln[xe^{\lambda} + 1 - x] = \ln[Bx + 1], \text{где} B = e^{\lambda} - 1,$$

вогнута, тогда (по неравенству Йенсена)

$$\sum_{i=1}^{n} f(p_i) \leqslant nf(p).$$

Потенцируя обе части неравенства, получаем

$$\prod_{i=1}^{n} [p_i e^{\lambda} + (1-p_i)] \leqslant [pe^{\lambda} + (1-p)]^n,$$

таким образом, мы доказали следующий результат.

Лемма А.1.8. *При предположении А.1.3 справедливо неравенство*

$$\mathbf{E}[e^{\lambda X}] < e^{-\lambda pn}[pe^{\lambda} + (1-p)]^n.$$

Теорема А.1.9. *При предположении А.1.3 и $a > 0$ неравенство*

$$\Pr[X \geqslant a] < e^{-\lambda pn}[pe^{\lambda} + (1-p)]^n e^{-\lambda a}$$

справедливо для всех $\lambda > 0$.

Доказательство. $\Pr[X > a] = \Pr[e^{\lambda X} > e^{\lambda a}] < \mathbf{E}[e^{\lambda X}]/e^{\lambda a}$. Применим теперь лемму А.1.8. ■

Замечание. При заданных p, n, a оптимальное значение λ в теореме А.1.9 находится с помощью элементарных вычислений и оказывается равным

$$\lambda = \ln\left[\left(\frac{1-p}{p}\right)\left(\frac{a+np}{n-(a+np)}\right)\right].$$

Это значение часто оказывается слишком громоздким в применении. Мы используем близкое к оптимальному значение λ для получения удобочитаемых результатов.

Положив $\lambda = \ln[1+a/pn]$ и используя неравенство $(1+a/n)^n \leqslant e^a$, получаем из теоремы А.1.9 следующий результат.

Следствие А.1.10. *Справедливо неравенство*

$$\Pr[X \geqslant a] < e^{a-pn\ln(1+a/pn)-a\ln(1+a/pn)}.$$

Теорема А.1.11. *Имеет место неравенство*

$$\Pr[X \geqslant a] < e^{-a^2/2pn+a^3/2(pn)^2}.$$

Доказательство. При $u = a/pn$ применим к следствию А.1.10 неравенство

$$\ln(1+u) \geqslant u - u^2/2,$$

справедливое для всех $u \geqslant 0$. ■

Если все $p_i = p$, то величина X имеет дисперсию, равную $np(1-p)$. При $p = o(1)$ и $a = o(pn)$ эта оценка отражает аппроксимацию распределения величины X нормальным распределением с дисперсией $\sim np$. Оценка из теоремы А.1.11 достигает минимума при $a = 2pn/3$. Для $a > 2pn/3$ у нас есть следующая простая оценка:

$$\Pr[X > a] \leqslant \Pr[X > 2pn/3] < e^{-2pn/27},$$

которую можно улучшить описанным ниже способом.

Теорема А.1.12. *Для* $\beta > 1$ *справедливо соотношение*

$$\Pr[X \geqslant (\beta-1)pn] < [e^{\beta-1}\beta^{-\beta}]^{pn}.$$

Доказательство. Подставим соответствующее значение параметра в следствие А.1.10. ■

Величину $X + pn$ можно интерпретировать как количество успехов в n независимых испытаниях с вероятностью успеха в i-м испытании, равной p_i.

Теорема А.1.13. *При предположении А.1.3 и $a > 0$ справедливо неравенство*

$$\Pr[X < -a] < e^{-a^2/2pn}.$$

Заметим, что здесь нельзя воспользоваться «симметричностью», так как при этом p и $1-p$ поменяются местами.

Доказательство. Пусть, на короткое время, значение $\lambda > 0$ произвольно. Из рассуждений, предшествующих лемме А.1.8 следует, что

$$\mathbf{E}[e^{-\lambda X}] \leqslant e^{\lambda pn}[pe^{-\lambda} + (1-p)]^n.$$

Тогда, аналогично теореме А.1.9,

$$\Pr[X < -a] = \Pr[e^{-\lambda X} > e^{\lambda a}] < e^{\lambda pn}[pe^{-\lambda} + (1-p)]^n e^{-\lambda a}.$$

Применив неравенство $1 + u \leqslant e^u$, верное для всех значений u, получаем

$$pe^{-\lambda} + (1-p) = 1 + (e^{-\lambda} - 1)p < e^{p(e^{-\lambda}-1)}$$

и

$$\Pr[X < -a] \leqslant e^{\lambda pn + np(e^{-\lambda}-1) - \lambda a} = e^{np(e^{-\lambda}-1+\lambda) - \lambda a}.$$

Применим неравенство

$$e^{-\lambda} \leqslant 1 - \lambda + \lambda^2/2,$$

верное для всех значений $\lambda > 0$. (Отметим, что аналогичное неравенство $e^{\lambda} \leqslant 1 + \lambda + \lambda^2/2$ *неверно* для $\lambda > 0$, поэтому этот метод, примененный для оценки $\Pr[X > a]$, требует поправочного слагаемого, такого же, как в теореме А.1.11.) Итак,

$$\Pr[X < -a] \leqslant e^{np\lambda^2/2 - \lambda a}.$$

Чтобы оптимизировать неравенство, положим $\lambda = a/np$, получим $\Pr[X < -a] < e^{-a^2/2pn}$, как и утверждалось. ■

Часто оказывается полезным следующий результат.

Следствие А.1.14. *Пусть Y — сумма взаимно независимых индикаторов, $\mu = \mathbf{E}[Y]$. Для всех $\varepsilon > 0$ выполнено соотношение*

$$\Pr[|Y - \mu| > \varepsilon\mu] < 2e^{-c_\varepsilon \mu},$$

где величина $c_\varepsilon > 0$ зависит только от ε.

Доказательство. Применим теоремы А.1.12, А.1.13 при $Y = X + pn$ и

$$c_\varepsilon = \min[-\ln(e^\varepsilon(1+\varepsilon)^{-(1+\varepsilon)}), \varepsilon^2/2].$$ ■

Асимметричность случаев $\Pr[X < a]$ и $\Pr[X > a]$, которую демонстрируют теоремы А.1.12, А.1.13, действительно существует. Нормальное распределение с нулевым средним и дисперсией np аппроксимирует распределение случайной величины X относительно точно при оценке $\Pr[X < a]$ для любого значения a, а также при оценке $\Pr[X > a]$ с $a = o(np)$. Но когда значения a и np становятся сравнимыми, или когда $a \gg np$, пуассоновское поведение начинает

превалировать, и значение $\Pr[X > a]$ *не может* быть достаточно точно оценено с помощью нормального распределения.

В завершение приведем два результата о больших уклонениях, касающихся распределений, а не сумм индикаторов.

Теорема А.1.15. *Пусть случайная величина P имеет пуассоновское распределение со средним значением μ. Для любого $\varepsilon > 0$ справедливы неравенства*

$$\Pr[P \leqslant \mu(1-\varepsilon)] \leqslant e^{-\varepsilon^2\mu/2},$$
$$\Pr[P \geqslant \mu(1+\varepsilon)] \leqslant \left[e^{\varepsilon}(1+\varepsilon)^{-(1+\varepsilon)}\right]^{\mu}.$$

Доказательство. Для любого s имеем

$$\Pr[P = s] = \lim_{n\to\infty} \Pr\left[B\left(n, \frac{\mu}{n}\right) = s\right].$$

Применим теоремы А.1.12, А.1.13. ■

Теорема А.1.16. *Пусть случайные величины X_i, $1 \leqslant i \leqslant n$, взаимно независимы и $\mathbf{E}[X_i] = 0$, $|X_i| \leqslant 1$. Положим $S = X_1 + \ldots + X_n$. Тогда*

$$\Pr[S > a] < e^{-a^2/2n}.$$

Доказательство. Положим $\lambda = a/n$, как и в доказательстве теоремы А.1.1. Пусть

$$h(x) = \frac{e^{\lambda} + e^{-\lambda}}{2} + \frac{e^{\lambda} - e^{-\lambda}}{2}x.$$

Для $x \in [-1, 1]$ выполнено неравенство $e^{\lambda x} \leqslant h(x)$ ($y = h(x)$ — это хорда выпуклой кривой $y = e^{\lambda x}$, проходящая через точки $x = \pm 1$). Тогда

$$\mathbf{E}[e^{\lambda X_i}] \leqslant \mathbf{E}[h(X_i)] = h(\mathbf{E}[X_i]) = h(0) = \operatorname{ch}\lambda.$$

Далее действуем как в доказательстве теоремы А.1.1. ■

Теорема А.1.17. *Пусть $\mathbf{E}[X] = 0$ и расстояние между любыми двумя значениями случайной величины X не превосходит единицы. Тогда для всех $\lambda \geqslant 0$*

$$\mathbf{E}[e^{\lambda X}] \leqslant e^{\lambda^2/8}.$$

Доказательство. Зафиксируем значение $b \in [-\frac{1}{2}, \frac{1}{2}]$, для которого случайная величина $X \in [\frac{-1+b}{2}, \frac{+1+b}{2}]$. Пусть $y = h(x)$ — прямая, пересекающая кривую $y = e^{\lambda x}$ в точках $(\pm 1 + b)/2$. Так как функция $e^{\lambda x}$ выпукла, то $e^{\lambda x} \leqslant h(x)$ для всех $x \in [\frac{-1+b}{2}, \frac{+1+b}{2}]$. Тогда

$$\mathbf{E}[e^{\lambda X}] \leqslant \mathbf{E}[h(X)] = h(\mathbf{E}[X]) = h(0).$$

Имеем $h(0) = e^{\lambda b/2}[\operatorname{ch}(\lambda/2) - b\operatorname{sh}(\lambda/2)]$, а эта величина не превосходит $e^{\lambda^2/8}$ по лемме А.1.5. ■

Теорема А.1.18. *Пусть случайные величины X_i, $1 \leqslant i \leqslant n$, взаимно независимы, $\mathbf{E}[X_i] = 0$ и расстояние между любыми двумя значениями каждой из случайных величин X_i не превосходит единицы. (Тем не менее, значения различных X_i, X_j могут находиться на большем расстоянии.) Положим $S = X_1 + \ldots + X_n$. Тогда*

$$\Pr[S > a] < e^{-2a^2/n}.$$

Доказательство. По теореме А.1.17 $\mathbf{E}[e^{\lambda S}] = \prod_{i=1}^{n} \mathbf{E}[e^{\lambda X_i}] \leqslant e^{n\lambda^2/8}$. Тогда для всех $\lambda \geqslant 0$ имеем

$$\Pr[S > a] = \Pr[e^{\lambda S} \geqslant e^{\lambda a}] \leqslant \exp\left[\frac{n\lambda^2}{8} - \lambda a\right].$$

Положим $\lambda = 4a/n$. ■

Мы, в целом, руководствовались идеей, что если случайная величина X имеет нулевое математическое ожидание и дисперсию σ^2, то вероятность $\Pr[X \geqslant a\sigma]$ должна себя вести как $e^{-a^2/2}$. Но существуют ситуации, когда это утверждение совершенно не верно. Рассмотрим предположение А.1.3 при всех $p_i = 1/n$, тогда $X = P_n - 1$, где величина P_n имеет биномиальное распределение $B(n, 1/n)$ сходящееся к P, пуассоновскому распределению со средним значением равным единице. Тогда $\mathbf{E}[X] = 0$ и $\mathrm{Var}[X] \sim 1$. Для произвольного фиксированного a имеем $\Pr[X = a] \to \frac{1}{e(a+1)!}$, что намного больше $e^{-a^2/2}$. Вместе с этим предостережением мы приводим условия, при которых упомянутое суждение асимптотически верно, если значение a невелико.

Теорема А.1.19. *Для каждого $C > 0$ и $\varepsilon > 0$ существует $\delta > 0$, такое, что верно следующее. Пусть X_i, $1 \leqslant i \leqslant n$, где n произвольно, — независимые случайные величины, $\mathbf{E}[X_i] = 0$, $|X_i| \leqslant C$ и $\mathrm{Var}[X_i] = \sigma_i^2$. Положим $X = \sum_{i=1}^{n} X_i$ и $\sigma^2 = \sum_{i=1}^{n} \sigma_i^2$, так что $\mathrm{Var}[X] = \sigma^2$. Тогда для $0 < a \leqslant \delta\sigma$*

$$\Pr[X > a\sigma] < e^{-\frac{a^2}{2}(1-\varepsilon)}.$$

Доказательство. Положим $\lambda = a/\sigma$, так что $0 \leqslant \lambda \leqslant \delta$. Тогда

$$\mathbf{E}[e^{\lambda X_i}] = \sum_{k=0}^{\infty} \mathbf{E}\left[\frac{\lambda^k}{k!} X_i^k\right] = 1 + \frac{\lambda^2}{2}\sigma_i^2 + \sum_{k=3}^{\infty} \frac{\lambda^k}{k!} \mathbf{E}[X_i^k].$$

Так как $|X_i^k| \leqslant C^{k-2} X_i^2$, выполнено

$$\mathbf{E}[X_i^k] \leqslant \mathbf{E}[|X_i^k|] \leqslant C^{k-2}\mathbf{E}[X_i^2] = C^{k-2}\sigma_i^2.$$

Для $k \geqslant 3$ оценим $\frac{2}{k!} \leqslant \frac{1}{(k-2)!}$, так что

$$\mathbf{E}[e^{\lambda X_i}] \leqslant 1 + \frac{\lambda^2}{2}\sigma_i^2 \left[1 + \sum_{k=3}^{\infty} \frac{(C\lambda)^{k-2}}{(k-2)!}\right] = 1 + \frac{\lambda^2}{2}\sigma_i^2 e^{\lambda C}.$$

Выберем значение δ так, чтобы оно удовлетворяло неравенству $e^{C\delta} \leqslant 1+\varepsilon$. Так как $\lambda \leqslant \delta$, то

$$\mathbf{E}[e^{\lambda X_i}] \leqslant 1 + \frac{\lambda^2}{2}\sigma_i^2(1+\varepsilon) < \exp\left[\frac{\lambda^2}{2}\sigma_i^2(1+\varepsilon)\right].$$

Это неравенство верно для всех X_i, следовательно,

$$\mathbf{E}[e^{\lambda X}] = \prod_{i=1}^{n} \mathbf{E}[e^{\lambda X_i}] < \exp\left[\frac{\lambda^2}{2}\sigma^2(1+\varepsilon)\right]$$

и

$$\Pr[X > a\sigma] \leqslant \mathbf{E}[e^{\lambda X}]e^{-\lambda a\sigma} < e^{-\frac{a^2}{2}(1-\varepsilon)}. \qquad \blacksquare$$

А.2. УПРАЖНЕНИЯ

1. *Числом Хайоша* графа G называют максимальное число k вершин графа G, каждая пара которых может быть соединена путем так, что все $\binom{k}{2}$ путей попарно не пересекаются по внутренним вершинам (а также ни одна из этих k вершин не является одновременно внутренней точкой одного пути и конечной точкой другого). Существует ли граф, хроматическое число которого превосходит удвоенное число Хайоша?

2. Для двух подмножеств A и B множества $\mathbb{Z}_m$ вычетов по модулю m и для некоторого числа $g \in Z_m$ положим

$$s(A,B,g) = |\{(a,b) : a \in A, b \in B, a+b = g\}|.$$

Для разбиения множества $\mathbb{Z}_m$ на два непересекающихся множества $\mathbb{Z}_m = A \cup B$, $A \cap B = \varnothing$ положим

$$c(A,B) = \max_{x \in \mathbb{Z}_m} |s(A,A,x) + s(B,B,x) - 2s(A,B,x)|.$$

Доказать, что для каждого нечетного числа m существует разбиение множества $\mathbb{Z}_m$ на два непересекающихся множества A и B, такое, что $c(A,B) = O(\sqrt{m \log m})$.

3. Пусть N — стандартная нормальная случайная величина с функцией распределения $f(t) = (2\pi)^{-1/2}e^{-t^2/2}$. Найти точное значение $\mathbf{E}[e^{\lambda N}]$. Для $a > 0$ найти точное значение $\min_{\lambda > 0} \mathbf{E}[e^{\lambda N}]e^{-\lambda a}$. Из этого будет следовать оценка Чернова для $\Pr[N \geqslant a]$. Как она соотносится с истинным значением $\Pr[N \geqslant a]$ при больших a?

ВЕРОЯТНОСТНЫЙ ВЗГЛЯД:

Графы, свободные от треугольников, содержат большие независимые множества

Обозначим через $\alpha(G)$ размер наибольшего из независимых множеств графа G. Хорошо известно, что для каждого графа G на n вершинах, степень которых не превосходит d, $\alpha(G) \geqslant n/(d+1)$. В работе [Ajtai, Komlós and Szemerédi (1980)] было показано, что в случае, когда граф G свободен от треугольников, эта оценка может быть улучшена на логарифмический множитель и $\alpha(G) \geqslant cn\log d/d$, где c — некоторая абсолютная положительная константа. В работе [Shearer (1983)] упрощено доказательство и улучшено значение константы до наилучшего значения: $c = 1 + o(1)$. Мы приводим здесь очень короткое доказательство этого факта и не пытаемся оптимизировать значение константы c. Доказательство основано на технике из работы [Shearer (1995)] и ее модификации из работы [Alon (1996)].

Утверждение. *Пусть n-вершинный граф $G = (V, E)$ свободен от треугольников и степень его вершин не превосходит $d \geqslant 1$. Тогда*

$$\alpha(G) \geqslant \frac{n\log d}{8d}.$$

Здесь и далее через $\log d$ *обозначен двоичный логарифм* $\log_2 d$.

Доказательство. При $d < 16$ утверждение следует из тривиальной оценки $\alpha(G) \geqslant n/(d+1)$, а, значит, мы можем далее считать, что $d \geqslant 16$. Пусть W — случайное независимое множество вершин графа G, выбранное равномерно среди всех независимых множеств графа G. Для каждой вершины $v \in V$ определим случайную величину $X_v = d|\{v\} \cap W| + |N(v) \cap W|$, где $N(v)$ обозначает множество всех смежных с v вершин. Утверждается, что математическое ожидание случайной величины X_v удовлетворяет условию $\mathbf{E}[X_v] \geqslant \frac{\log d}{4}$.

Докажем это. Обозначим через H подграф G, индуцированный множеством вершин $V \setminus (N(v) \cup \{v\})$. Зафиксируем независимое множество S в графе H и обозначим через X множество всех вершин из $N(v)$, не смежных с вершинами из S, пусть $|X| = x$. Достаточно показать, что условное математическое ожидание

$$\mathbf{E}[X_v|\ W \cap V(H) = S] \geqslant \frac{\log d}{4} \tag{1}$$

для каждого фиксированного S. При условии, что $W \cap V(H) = S$, существует ровно $2^x + 1$ возможностей для выбора множества W: одна — для случая $W = S \cup \{v\}$, и 2^x — для случая, когда $v \notin W$ и W представляет собой объединение множества S и подмножества множества X. Отсюда следует, что условное математическое ожидание из неравенства (1) равно $\frac{d}{2^x+1} + x2^{x-1}/(2^x+1)$. Для проверки того, что последняя величина не меньше $\log d/4$, достаточно

заметить, что в противном случае $x \geqslant 1$ и $2^x(\log d - 2x) > 4d - \log d$. Тогда $\log d > 2x \geqslant 2$, а, следовательно, $4d - \log d < \sqrt{d}(\log d - 2)$, что неверно для всех $d \geqslant 16$. А значит,

$$\mathbf{E}[X_v |\ W \cap V(H) = S] \geqslant \frac{\log d}{4},$$

как и утверждалось.

Из линейности математического ожидания следует, что математическое ожидание величины $\sum_{v \in V} X_v$ не меньше $\frac{n \log d}{4}$. С другой стороны, ясно, что значение этой суммы не превосходит $2d|W|$, так как каждая вершина $u \in W$ вносит вклад d в слагаемое X_u и вклад, равный своей степени в графе G (последняя не превосходит d), в сумму всех остальных слагаемых X_v. Тогда математическое ожидание мощности множества W не меньше $\frac{n \log d}{8d}$, а, следовательно, в графе найдется независимое множество, мощность которого не меньше, чем это математическое ожидание, что и требовалось доказать. ■

Число Рамсея $r(3, k)$ — это минимальное число r, такое, что в любом графе с не менее чем r вершинами найдется либо треугольник, либо независимое множество мощности k. Асимптотическое поведение этой функции изучалось на протяжении последних пятидесяти лет. Выяснилось, что $r(3, k) = \Theta(k^2/\log k)$. Нижняя оценка была получена в недавней работе [Kim (1995)], основанной на изящной вероятностной конструкции, сопровождающейся тридцатью страницами вычислений. Способ построения такого графа неизвестен, наибольший явным образом построенный свободный от треугольников граф без независимых множеств мощности k описан в работе [Alon (1994)], и он имеет лишь $\Theta(k^{3/2})$ вершин. Точная верхняя оценка для функции $r(3, k)$, доказанная в работе [Ajtai et al. (1980)], представляет собой очень простое следствие последнего утверждения.

Теорема [Ajtai et al. (1980)]. *Существует абсолютная константа b, такая, что*

$$r(3, k) \leqslant bk^2/\log k$$

для всех $k > 1$.

Доказательство. Пусть $G = (V, E)$ — свободный от треугольников граф на $8k^2/\log k$ вершинах. Если G содержит вершину степени не менее k, то ее граница содержит независимое множество мощности k. В противном случае из утверждения следует, что граф G содержит независимое множество мощности не менее $\frac{8k^2}{\log k}\frac{\log k}{8k} = k$. Таким образом, в любом случае $\alpha(G) \geqslant k$, что и требовалось доказать. ■

Приложение

B

Пол Эрдёш

Работа с Полом Эрдёшем была подобна прогулке в горах. Каждый раз, когда я думала, что мы достигли цели и заслужили отдых, Пол указывал на следующую вершину и мы отправлялись к ней.

Фэн Чан

B.1. ТРУДЫ

Пол Эрдёш был самым плодовитым математиком двадцатого столетия, им было написано более 1500 работ и более 490 в соавторстве. Приведенный ниже крайне субъективный список включает в себя только часть тех работ, в которых создан и сформирован вероятностный метод. Аббревиатуры **MR** и **Zbl.** представляют собой ссылки на обзоры в Math Reviews и Zentralblatt, соответственно. Ссылки на главы или разделы относятся к соответствующим частям этой книги.

- A combinatorial problem in geometry, *Compositio Math* **2** (1935), 463–470 (with George Szekeres) **Zbl.** 12,270.
  Работа, написанная, когда Эрдёш был еще юношей. Эта жемчужина содержит повторное открытие теоремы Рамсея и теоремы о монотонной подпоследовательности. Многие авторы отмечают, что эта работа сыграла ключевую роль в переходе Эрдёша к более комбинаторной точке зрения на математику.
- Some remarks on the theory of graphs, *Bull. Amer. Math. Soc.* **53** (1947), 292–294, **MR** 8,479d; **Zbl** 32,192.
  Трехстраничная работа, с которой «начинается» вероятностный метод. В ней приводится экспоненциальная нижняя оценка для диагональных чисел Рамсея $R(k,k)$ (разд. 1.1).
- The Gaussian law of errors in the theory of additive number theoretic functions, *Amer. J. Math.* **62** (1940), 738–742 (with Mark Kac) **MR** 2,42c; **Zbl.** 24,102.
  Доказано, что количество простых делителей числа x, выбранного равномерно из множества $\{1,\ldots,n\}$, имеет асимптотически нормальное распределение. Продемонстрированная связь между теорией вероятностей и теорией чисел в свое время казалась поразительной (разд. 4.2).

- Problems and results in additive number theory, *Colloque sur la Théorie des Nombres, Bruxelles, 1955*, 127–137, George Thone, Liège; Masson and Cie, Paris, 1956; **MR** 18,18a; **Zbl.** 73,31.

  Здесь случайные подмножества используются для доказательства существования такого множества целых чисел, что каждое число n может быть представлено в виде $n = x + y$ по крайней мере одним, но не более чем $c \ln n$ способами. Решена проблема, поставленная Эрдёшу Зидоном в 1930-х годах. Эта проблема и далее продолжает интересовать Эрдёша: см., например, [Erdős and Tetali (1990)] (разд. 8.6).
- On a combinatorial problem, *Nordisk. Mat. Tidskr.* **11** (1963), 220–223 **MR** 28# 4068; **Zbl.** 122,248.
  On a combinatorial problem II, *Acta. Math. Acad. Sci. Hungar.* **15** (1964), 445–447; **MR** 29# 4700; **Zbl.** 201,337.
  Свойство **B**. Вероятностное доказательство того, что любые $m < 2^{n-1}$ множеств размера n могут быть 2-раскрашены так, что ни одно множество не является монохроматическим, несмотря на то, что существуют $cn^2 2^n$ множеств размера n, которые не могут быть так раскрашены (разд. 1.3).
- On the evolution of random graphs, *Magyar. Tud. Akad. Mat. Kutató Int. Közl.* **5** (1960), 17–61 (with Alfred Rényi); **MR** 23# A2338; **Zbl.** 103,163.
  Очень редко целый раздел математики удается исследовать в одной работе. Однако, для случайных графов произошло именно так (гл. 10).
- Graph theory and probability, *Canad. J. Math.* **11** (1959), 34–38; **MR** 21# 876; **Zbl.** 84,396.
  Вероятностным методом доказывается существование графов с произвольно большим обхватом и хроматическим числом. Работа убедительно продемонстрировала силу метода, поскольку несмотря на то, что данная проблема вызывала большой интерес, построить такие графы не удавалось (разд. «Вероятностный взгляд», следующий за гл. 3).
- Graph theory and probability II, *Canad. J. Math.* **13** (1961), 346–352 **MR** 22# 10925; **Zbl.** 97,391.
  Доказано существование n-вершинного свободного от треугольников графа без независимых множеств размера $cn^{1/2} \ln n$, и, следовательно, того, что число Рамсея $R(3, k) = \Omega(k^2 \ln^{-2} k)$. Это технический *tour de force*, очень тонко использующий вероятностный метод, особенно, если учитывать раннюю дату публикации.
- On circuits and subgraphs of chromatic graphs, *Mathematika* **9** (1962), 170–175; **MR** 25 # 3035; **Zbl.** 109,165.
  Работа опровергает мнение, что хроматическое число является чисто локальным свойством. Эрдёш доказывает существование n-вершинного графа, который не может быть k-раскрашен, но каждые εn вершин которого могут быть 3-раскрашены (разд. «Вероятностный взгляд», следующий за гл. 8).
- On a combinatorial game, *J. Combinatorial Theory Ser. A* **14** (1973), 298–301 (with John Selfridge) **MR** 48# 5655; **Zbl.** 293,05004.

Игроки по очереди выбирают вершины, причем второй игрок пытается помешать первому получить выигрышное множество. Использованный здесь метод весовой функции, по существу, является вероятностным и представляет собой одно из первых применений дерандомизации (разд. 15.1).

В.2. ГИПОТЕЗЫ

Гипотезы всегда играли существенную роль в математической жизни Пола Эрдёша. Вот самые интересные на наш взгляд.

- Обязательно ли множество целых чисел положительной плотности содержит арифметические прогрессии произвольной длины? Финитизация проблемы приводит к следующей задаче. Пусть S является подмножеством множества первых n целых чисел, $|S| \leqslant \varepsilon n$. Верно ли, что для всех k и всех $\varepsilon > 0$ существует такое число n_0, что при $n \geqslant n_0$ множество S обязательно содержит арифметическую прогрессию длины k? Эта гипотеза впервые была выдвинута Эрдёшем и Тураном в 1930 г. Она была решена (положительно) Семереди в 1970 г. Обозначим через $F(k, \varepsilon)$ минимальное число n_0, удовлетворяющее приведенным выше условиям. Определение скорости роста функции F остается весьма интригующей проблемой, последние результаты на эту тему были получены Говерсом.
- Будем говорить, что три различных множества S, T, U образуют Δ-систему, если $S \cap T = S \cap U = T \cap U$. Обозначим через $F(n)$ минимальное число m, такое, что среди m произвольных n-множеств какие-то три обязательно сформируют Δ-систему. Эрдёш и Радо показали, что функция $F(n)$ существует, и получили верхнюю оценку $F(n) \leqslant 2^n n!$. Эрдёш предположил, что $F(n) < C^n$ для некоторой константы C.
- Какова асимптотика функции Рамсея $R(k, k)$? В частности, каково значение c предела $\lim_k R(k, k)^{1/k}$, если он существует? Классическая работа Эрдёша 1947 г. показывает, что $c \geqslant \sqrt{2}$. Результат $c \leqslant 4$ следует из доказательства теоремы Рамсея, но дальнейших уточнений значения c нет уже полвека, хотя были результаты, касающиеся членов меньшего порядка.
- Обозначим через $r_S(n)$ число решений уравнения $n = x + y$ для $x, y \in S$. Существует ли множество положительных целых чисел S, такое, что $r_S(n) > 0$ для всех n, за исключением, быть может, конечного числа, при условии, что функция $r_S(n)$ ограничена некоторой константой K? Работа Эрдёша 1955 г. года, ссылка на которую приведена выше, описывает множество S, такое, что $r_S(n) = \Theta(\ln n)$.
- Пусть $m(n)$, как и в разд. 1.3, обозначает минимальный размер семейства из n-множеств, которое не может быть 2-раскрашено без образования монохроматического множества. Какова асимптотика функции $m(n)$? В 1963 и 1964 гг. Эрдёш получает оценки $\Omega(2^n) \leqslant m(n) = O(2^n n^2)$. Впоследствии, Радхакришнан и Шринивасан (см. разд. 3.5) улучшают нижнюю оценку до $\Omega(2^n (n/\ln n)^{1/2})$.

- Рассмотрим $2^{n-2}+1$ точек плоскости, среди которых никакие три не лежат на одной прямой. Должны ли всегда некоторые n из них образовывать выпуклое множество? Эта проблема была поставлена в работе Эрдёша и Секереша 1935 г., ссылка на которую приведена выше.
- Обозначим через $m(n,k,l)$ размер наибольшего семейства таких k-элементных подмножеств n-множества, что никакое l-множество не содержится более чем в одном из них. Простой подсчет показывает, что $m(n,k,l) \leqslant \binom{n}{l}/\binom{k}{l}$. В 1963 г. Эрдёш и Ханани предположили, что для фиксированного $l<k$ эта оценка асимптотически верна, т. е. отношение функций $m(n,k,l)$ и $\binom{n}{l}/\binom{k}{l}$ стремится к единице при $n\to\infty$. Эрдёш обладал удивительной способностью ставить очень трудные, но все же разрешимые задачи. Справедливость этой гипотезы была доказана Войтехом Рёдлем в 1985 г. (см. разд. 4.7). Асимптотика разности $\binom{n}{l}/\binom{k}{l} - m(n,k,l)$ до сих пор остается неизвестной.

В.3. ОБ ЭРДЁШЕ

О жизни и математике Пола Эрдёша написано множество книг и статей. Следующие три из них заслуживают отдельного упоминания:

- *The Mathematics of Paul Erdős* (Ron Graham and Jarik Nešetřil, eds.), Berlin: Springer-Verlag, 1996. (Vols. I and II)
- *Combinatorics, Paul Erdős is Eighty* (D. Miklós, V. T. Sós, T. Szönyi, eds.), Bolyai Soc. Math. Studies, Vol. I (1990) and Vol. II (1993).
- *Erdős on Graphs — His Legacy of Unsolved Problems*, Fan Chung and Ron Graham, A. K. Peters, 1998.

Среди множества работ, написанных математиками, мы отметим

- László Babai, In and out of Hungary: Paul Erdős, his friends, and times. In *Combinatorics, Paul Erdős is Eighty* (listed above), Vol. II, 7–93.
- Béla Bollobás, Paul Erdős — Life and work, in *The Mathematics of Paul Erdős* (listed above), Vol. II, 1–42.
- A. Hajnal, Paul Erdős' Set theory, in *The Mathematics of Paul Erdős* (listed above), Vol. II, 352–393.
- János Pach, Two places at once: a remembrance of Paul Erdős, *Math Intelligencer*, Vol. 19 (1997), no. 2, 38–48.

Появились следующие популярные биографии Эрдёша:

- *The Man Who Loved Only Numbers*, Paul Hoffman, Hyperion (New York), 1998.
- *My Brain is Open - The Mathematical Journies of Paul Erdős*, Bruce Schechter, Simon & Schuster (New York), 1998.

И, наконец, Джорж Сисери снял документальный фильм *N is a Number, A Portrait of Paul Erdős*, выпущенный издательством А. К. Петерса, позволяющий увидеть и услышать Эрдёша на лекции и среди друзей, доказывающего и высказывающего гипотезы.

В.4. ДЯДЮШКА ПОЛ

Пол Эрдёш умер в сентябре 1996 года в возрасте 83 лет. Его теоремы и гипотезы пронизывают всю эту книгу. Эта речь, произнесенная Джоэлем Спенсером на заседании Американского Математического Общества в январе 1997[1)]*, представляет собой попытку выразить тот особый дух, который мы и многие другие восприняли от этого необыкновенного человека.*

Пол Эрдёш был искателем, искателем математической истины. Место Пола в математическом пантеоне будет предметом бурных дебатов, поскольку в этой изысканной атмосфере у него был уникальный стиль. Лучше всего об этом сказал ныне покойный Эрнст Штраус на праздновании 70-летия Эрдёша.

> В наш век, когда в математике так сильно доминируют «создатели теорий», он оставался принцем тех, кто решает проблемы, и абсолютным королем тех, кто их ставит. Один из моих друзей, сильный математик, жаловался мне, что «Эрдёш дает нам только следствия великих метатеорем, которые остаются не сформулированными в его голове». Это высказывание во многом верно, однако я не думаю, чтобы Эрдёш мог или хотел отказаться от этих следствий и сконцентрироваться на формулировании метатеорем. Во многом Пол Эрдёш является Эйлером нашего времени. Как «специальные» проблемы, которые решил Эйлер, открыли путь к аналитической и алгебраической теории чисел, топологии, комбинаторике, функциональным пространствам и т. д., так и методы и результаты работ Эрдёша уже дают нам возможность увидеть очертания великих новых дисциплин, таких как комбинаторная и вероятностная теория чисел, комбинаторная геометрия, вероятностная и трансфинитная комбинаторика, теория графов, а также многих других, которым еще только предстоит возникнуть из его идей.

Штраус, работавший ассистентом Альберта Эйнштейна, отметил, что Эйнштейн выбрал физику, а не математику, поскольку опасался, что бесплодно растратит силы на решение многих замечательных и завораживающих математических задач, не находя центральных проблем. Штраус продолжил:

> Эрдёш последовательно и успешно нарушал все рекомендации Эйнштейна. Он поддавался соблазну каждой встретившейся ему красивой задачи, и очень многие из них поддавались ему. Это убеждает меня в том, что в поисках истины всегда есть место и Дон Жуану, как Эрдёш, и Сэру Галахэду, как Эйнштейн.

Я полагаю, и это определенно предвзятое мнение, что наследие Пола будет сильнейшим в дискретной математике. Интерес Пола к этой области возник с замечательного доклада Джорджа Секереша в 1935 году, однако в полную силу он проявился только после Второй Мировой войны. Думаю, есть две

[1)] Перепечатано с разрешения Bulletin of the American Mathematical Society.

основные причины расцвета дискретной математики в последние полстолетия. Первая — это компьютер. Как чудесно, что этот физический объект привел к таким интригующим математическим задачам. Вторая причина, с полным уважением ко многим другим, заключается в деятельности Пола Эрдёша с его знаменитым лозунгом «Доказывать и ставить проблемы!». Теория Рамсея, экстремальная теория графов, случайные графы. Сколько башен в нашем математическом замке были построены из кирпичиков-теорем Пола, и украшены его многочисленными и всегда прозорливыми гипотезами.

Вероятностный метод, предмет моих собственных исследований, безусловно, мог бы быть назван методом Эрдёша. Начало ему было положено в 1947 году трехстраничной статьей в Бюллетене Американского Математического Общества. Пол доказал существование графа, имеющего определенное Рамсеевское свойство, без фактического его конструирования. Говоря современным языком, он показал, что соответствующим образом определенный случайный граф обладает этим свойством с ненулевой вероятностью и, следовательно, такой граф должен существовать. В течение следующих двадцати лет работы Пола были «гласом вопиющего в пустыне», коллеги восхищались его поразительными результатами, однако приятие методологии шло медленно. Тем не менее, Пол упорно продолжал идти своим путем, он всегда руководствовался чувством математической эстетики, которому полностью доверял, и сегодня метод широко применяется, как в дискретной математике, так и в теоретической информатике.

Не может быть никаких разногласий относительно того, какое влияние оказала личность Пола на формирование современной математики. Пол Эрдёш был самым вдохновляющим человеком из всех, кого я когда-либо встречал. Я начал работать с Полом в конце 1960-х, в бурное время, когда девиз «подумай сам» витал в воздухе. Но тогда как другие только говорили об этом, для Пола это было образом жизни. Для него не существовало понятия «работа» — он работал постоянно. У него не было дома — мир был его домом. Имущество было для него помехой, деньги — скукой. Он жил в системе доверия, беспрерывно путешествуя от центра к центру, осеняя своим математическим вдохновением.

Что привело столь многих из нас в его круг? Как объяснить радость общения с этим прекрасным человеком? Почему мы любим рассказывать истории об Эрдёше? Я много думал об этом и считаю, что все сводится к вере. Мы, математики, знаем о красоте нашего предмета и мы верим в его трансцендентность. Бог создал целые числа, остальное уже работа Человека.

Математическая истина неизменна, она лежит вне физической реальности. Когда мы, к примеру, показываем, что сумма двух степеней числа n никогда не равна степени числа n для $n \geqslant 3$, мы прикасаемся к Истине. В этом наша вера, это дает нам глубинную мотивацию. И все же, наши попытки описать эту веру нашим нематематическим друзьям сродни проповеди о Всемогущем атеисте. Пол воплощал эту веру в математическую истину. Его огромный талант и энергия были полностью посвящены Храму математики. Он не питал никаких сомнений относительно важности, абсолютности его поисков. Видеть его веру

означало уверовать самому. В религиозном мире лучше бы смогли понять особые качества личности Пола. Мы знали его как дядюшку Пола.

Я надеюсь, что краеугольный камень, если хотите, теологии Пола просуществует долго. Я имею в виду Книгу. Книга состоит из всех математических теорем. Для каждой теоремы в Книге только одно доказательство. Это наиболее эстетичное, наиболее ясное доказательство, то, что Пол называл доказательством Книги. И когда одна из бесчисленных проблем Пола решалась, но решалась «некрасиво», Пол был очень счастлив поздравить доказавшего, но добавлял: «А теперь, давайте поищем доказательство Книги». Эта платоновская идея была очевидна для нас, входящих в его круг. Математика уже была там, нам оставалось только открыть ее.

Интенсивность и самоотверженность поиска истины были описаны Хорхе Луисом Борхесом в рассказе «Вавилонская библиотека». Рассказ ведется от лица работника этой библиотеки, бескрайние просторы которой содержали всю мудрость вселенной. Он бродит по ее бесчисленным коридорам в поисках того, что Пол Эрдёш назвал бы Книгой. Он объявляет во всеуслышание:

> Мне не кажется невероятным, что на какой-то книжной полке вселенной стоит всеобъемлющая книга; молю неведомых богов, чтобы человеку — хотя бы одному, хоть через тысячи лет! — удалось найти и прочесть ее. Если почести, и мудрость, и счастье не для меня, пусть они достанутся другим. Пусть существует небо, даже если мое место в аду. Пусть я буду попран и уничтожен, но хотя бы на миг, хотя бы в одном существе твоя огромная Библиотека будет оправдана.

Летом 1985 года я отвез Пола в лагерь, который многие из нас с любовью вспоминают как Лагерь Желтой Свиньи — математический лагерь для талантливых студентов Хемпширского колледжа. Это был прекрасный день — студенты обожали дядюшку Пола, и для Пола было наивысшим удовольствием находиться в обществе пытливых молодых умов. В моем вступлении к его лекции, я говорил о Книге, но я допустил ошибку, описывая ее, как книгу, которую «держит Бог». Пол начал свою лекцию с мягкой поправки, которую я никогда не забуду. «Вы не обязаны верить в Бога», сказал он, «но вы должны верить в Книгу.»

Литература

Андреев А. Е. (1985). О методе получения нижних оценок сложности индивидуальных монотонных функций, *Доклады Академии Наук СССР* **282**(5): 1033–1037.

Андреев А. Е.(1987). О методе получения более чем квадратичных эффективных нижних оценок сложности π-схем, *Вестник Московского Университета, Серия 1, Математика, Механика, (1)* : 70–73.

Брегман Л. М. (1973). Некоторые свойства неотрицательных матриц и их перманентов, *Доклады Академии Наук* **14**: 945–949.

Вапник В. Н., Червоненкис А. Я. (1971). О равномерной сходимости относительных частот к их вероятностям, *Теория вероятностей и ее применения* **16**: 264–280.

Глебский Ю. В., Коган Д. И., Легонький М. И. и Таланов В. А. (1969). Область и степень реализуемости формул ограниченного исчисления предикатов, *Кибернетика* **5**: 142–154.

Маргулис Г. А. (1973). Явные конструкции расширителей, *Проблемы Передачи Информации* **9**, 4: 71–80.

Маргулис Г. А. (1988). Явные теоретико-групповые конструкции комбинаторных схем и их применения к постороению синтезу расширителей и суперконцентраторов, *Проблемы Передачи Информации* **24**, 1: 51–60.

Разборов А. А. (1985). Нижние оценки монотонной сложности некоторых булевых функций, *Доклады Академии Наук СССР* **281**(4): 798–801.

Разборов А. А. (1987). Нижние оценки сложности схем ограниченной глубины над полным базисом, содержащем функцию логического сложения, *Математические Заметки* **41**(4): 598–607.

Субботовская Б. А. Реализация линейных функций формулами в базисе $\bigvee, \bigwedge, -$, *Доклады Академии Наук СССР* **136**,(3), 110–112.

Храпченко В. М. (1971). Об одном методе получения нижних оценок сложности П-схем, *Математические Заметки* **10**, (1): 83–92.

Ahlswede, R. and Daykin, D. E. (1978). An inequality for the weights of two families of sets, their unions and intersections, *Z. Wahrscheinl. V. Geb* **43**: 183–185.

Aho, A. V., Hopcroft, J. E. and Ullman, J. D. (1974). *The Design and Analysis of Computer Algorithms*, Addison Wesley, Reading, MA.

Ajtai, M. (1983). Σ_1^1-formulae on finite structures, *Annals of Pure and Applied Logic* **24**: 1–48.

Ajtai, M., Chvátal, V., Newborn, M. M. and Szemerédi, E. (1982). Crossing-free subgraphs, *Theory and practice of combinatorics, North Holland Math. Stud.* **60**: 9–12.

Ajtai, M., Komlós, J. and Szemerédi, E. (1980). A note on Ramsey numbers, *J. Combinatorial Theory, Ser. A* **29**: 354–360.

Ajtai, M., Komlós, J. and Szemerédi, E. (1983). Sorting in $c \log n$ parallel steps, *Combinatorica* **3**: 1–19.

Ajtai, M., Komlós, J. and Szemerédi, E. (1987). Deterministic simulation in LOGSPACE, *Proc.* 19-*th annual ACM STOC*, New York, pp. 132–140.

Akiyama, J., Exoo, G. and Harary, F. (1981). Covering and packing in graphs IV: Linear arboricity, *Networks* **11**: 69–72.

Alon, N. (1986a). Eigenvalues and expanders, *Combinatorica* **6**: 83–96.

Alon, N. (1986b). Eigenvalues, geometric expanders, sorting in rounds and Ramsey Theory, *Combinatorica* **6**: 207–219.

Alon, N. (1988). The linear arboricity of graphs, *Israel J. Math* **62**: 311–325.

Alon, N. (1990a). The maximum number of Hamiltonian paths in tournaments, *Combinatorica* **10**: 319–324.

Alon, N. (1990b). Transversal numbers of uniform hypergraphs, *Graphs and Combinatorics* **6**: 1–4.

Alon, N. (1994). Explicit Ramsey graphs and orthonormal labelings, *The Electronic J. Combinatorics* **1**: 8 pp. R12.

Alon, N. (1996). Independence numbers of locally sparse graphs and a Ramsey type problem, *Random Structures and Algorithms* **9**: 271–278.

Alon, N. and Boppana, R. B. (1987). The monotone circuit complexity of boolean functions, *Combinatorica* **7**: 1–22.

Alon, N. and Chung, F. R. K. (1988). Explicit construction of linear sized tolerant networks, *Discrete Math.* **72**: 15–19.

Alon, N. and Frankl, P. (1985). The maximum number of disjoint pairs in a family of subsets, *Graphs and Combinatorics* **1**: 13–21.

Alon, N. and Kleitman, D. J. (1990). Sum-free subsets, *in: A tribute to Paul Erdős* (A. Baker, B. Bollobás and A. Hajnál, eds.), Cambridge Univ. Press, Cambridge, England, pp. 13–26.

Alon, N. and Krivelevich, M. (1997). The concentration of the chromatic number of random graphs, *Combinatorica* **17**: 303–313.

Alon, N. and Linial, N. (1989). Cycles of length 0 modulo k in directed graphs, *J. Combinatorial Theory, Ser. B* **47**: 114–119.

Alon, N. and Milman, V. D. (1984). Eigenvalues, expanders and superconcentrators, *Proc.* 25-*th Annual FOCS*, IEEE, New York, pp. 320–322. See also: N. Alon and V. D. Milman, λ_1, isoperimetric inequalities for graphs and superconcentrators, J. Combinatorial Theory, Ser. B, 38, 1985, 73–88.

Alon, N., Babai, L. and Itai, A. (1986). A fast and simple randomized parallel algorithm for the maximal independent set problem, *J. of Algorithms* **7**: 567–583.

Alon, N., Frankl, P. and Rödl, V. (1985). Geometrical realization of set systems and probabilistic communication complexity, *Proc.* 26-*th FOCS*, IEEE, New York, pp. 277–280.

Alon, N., Goldreich, O., Håstad, J. and Peralta, R. (1990). Simple constructions of almost k-wise independent random variables, *Proc. 31-st FOCS, St. Louis*, IEEE, New York, pp. 544–553.

Alon, N., Kim, J. H. and Spencer, J. H. (1997). Nearly perfect matchings in regular simple hypergraphs, *Israel J. Math* **100**: 171–187.

Alon, N., Rónyai, L. and Szabó, T. (1999). Norm-graphs: variations and applications, *J. Combinatorial Theory, Ser. B* **76**: 280–290.

Baik, J., Deift, P. and Johansson, K. (1999). On the distribution of the length of the longest increasing subsequence of random permutations, *J. Amer. Math. Soc.* **12**: 1119–1178.

Bárány, I. and Fűredi, Z. (1987). Empty simplices in Euclidean spaces, *Canad. Math. Bull.* **30**: 436–445.

Beck, J. (1978). On 3-chromatic hypergraphs, *Disc. Math.* **24**: 127–137.

Beck, J. (1981). Roth's estimate of the discrepancy of integer sequences is nearly optimal, *Combinatorica* **1**: 319–325.

Beck, J. (1991). An algorithmic approach to the Lovász local lemma. I., *Random Structures and Algorithms* **2**: 343–365.

Beck, J. and Fiala, T. (1981). Integer-making theorems, *Disc. Appl. Math.* **3**: 1–8.

Bernstein, S. N. (1912). Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités, *Comm. Soc. Math. Kharkov* **13**: 1–2.

Blum, N. (1984). A boolean function requiring $3n$ network size, *Theoretical Computer Science* **28**: 337–345.

Bollobás, B. (1965). On generalized graphs, *Acta Math. Acad. Sci. Hungar.* **16**: 447–452.

Bollobás, B. (1985). *Random Graphs*, Academic Press., New York.

Bollobás, B. (1988). The chromatic number of random graphs, *Combinatorica* **8**: 49–55.

Bollobás, B. and Erdős, P. (1976). Cliques in random graphs, *Math. Proc. Camb. Phil. Soc.* **80**: 419–427.

Boppana, R. B. and Spencer, J. H. (1989). A useful elementary correlation inequality, *J. Combinatorial Theory, Ser. A* **50**: 305–307.

Chazelle, B. and Welzl, E. (1989). Quasi-optimal range searching in spaces of finite VC-dimension, *Discrete and Computational Geometry* **4**: 467–489.

Chernoff, H. (1952). A measure of the asymptotic efficiency for tests of a hypothesis based on the sum of observations, *Ann. Math. Stat.* **23**: 493–509.

Chung, F. R. K., Frankl, P., Graham, R. L. and Shearer, J. B. (1986). Some intersection theorems for ordered sets and graphs, *J. Combinatorial Theory, Ser. A* **43**: 23–37.

Chung, F. R. K., Graham, R. L. and Wilson, R. M. (1989). Quasi-random graphs, *Combinatorica* **9**: 345–362.

Cohen, A. and Wigderson, A. (1989). Dispersers, deterministic amplification, and weak random sources, *Proc. 30-th IEEE FOCS*, IEEE, New York, pp. 14–19.

Danzer, L. and Grünbaum, B. (1962). Űber zwei Probleme bezűglich konvexer Kőrper von P. Erdős und von V. L. Klee, *Math. Z.* **79**: 95–99.

de la Vega, W. F. (1983). On the maximal cardinality of a consistent set of arcs in a random tournament, *J. Combinatorial Theory, Ser. B* **35**: 328–332.

Dudley, R. M. (1978). Central limit theorems for empirical measures, *Ann. Probab.* **6**: 899–929.

Elekes, G. (1997). On the number of sums and products, *Acta Arith.* **81**: 365–367.

Erdős, P. (1947). Some remarks on the theory of graphs, *Bull. Amer. Math. Soc.* **53**: 292–294.

Erdős, P. (1956). Problems and results in additive number theory, *Colloque sur la Théorie des Nombres (CBRM, Bruselles)* pp. 127–137.

Erdős, P. (1959). Graph theory and probability, *Canad. J. Math.* **11**: 34–38.

Erdős, P. (1962). On circuits and subgraphs of chromatic graphs, *Mathematika* **9**: 170–175.

Erdős, P. (1963a). On a combinatorial problem, I, *Nordisk Mat. Tidskr.* **11**: 5–10.

Erdős, P. (1963b). On a problem of graph theory, *Math. Gaz.* **47**: 220–223.

Erdős, P. (1964). On a combinatorial problem II, *Acta Math. Acad. Sci. Hungar.* **15**: 445–447.

Erdős, P. (1965a). Extremal problems in number theory, *Proc. Symp. Pure Math. (AMS)* **VIII**: 181–189.

Erdős, P. (1965b). On extremal problems of graphs and generalized graphs, *Israel J. Math.* **2**: 189–190.

Erdős, P. and Fűredi, Z. (1983). The greatest angle among n points in the d-dimensional Euclidean space, *Annals of Discrete Math.* **17**: 275–283.

Erdős, P. and Hanani, H. (1963). On a limit theorem in combinatorial analysis, *Publ. Math. Debrecen* **10**: 10–13.

Erdős, P. and Kac, M. (1940). The Gaussian law of errors in the theory of additive number theoretic functions, *Amer. J. Math.* **62**: 738–742.

Erdős, P. and Lovász, L. (1975). Problems and results on 3-chromatic hypergraphs and some related questions, *in: Infinite and Finite Sets* (A. Hajnal et al., eds.), North-Holland, Amsterdam, pp. 609–628.

Erdős, P. and Moon, J. W. (1965). On sets of consistent arcs in a tournament, *Canad. Math. Bull.* **8**: 269–271.

Erdős, P. and Rényi, A. (1960). On the evolution of random graphs, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5**: 17–61.

Erdős, P. and Selfridge, J. L. (1973). On a combinatorial game, *J. Combinatorial Theory, Ser. A* **14**: 298–301.

Erdős, P. and Spencer, J. H. (1991). Lopsided Lovász local lemma and latin transversals, *Discrete Appl. Math.* **30**: 151–154.

Erdős, P. and Tetali, P. (1990). Representations of integers as the sum of k terms, *Random Structures and Algorithms* **1**: 245–261.

Fagin, R. (1976). Probabilities in finite models, *J. Symbolic Logic* **41**: 50–58.

Fishburn, P. (1992). Correlation in partially ordered sets, *Discrete Applied Math.* **39**: 173–191.

Fortuin, C. M., Kasteleyn, P. W. and Ginibre, J. (1971). Correlation inequalities on some partially ordered sets, *Comm. Math. Phys.* **22**: 89–103.

Fűredi, Z. (1988). Matchings and covers in hypergraphs, *Graphs and Combinatorics* **4**: 115–206.

Frankl, P. and Wilson, R. M. (1981). Intersection theorems with geometric consequences, *Combinatorica* **1**: 357–368.

Frankl, P., Rödl, V. and Wilson, R. M. (1988). The number of submatrices of given type in a Hadamard matrix and related results, *J. Combinatorial Theory, Ser. B* **44**: 317–328.

Furst, M., Saxe, J. and Sipser, M. (1984). Parity, circuits and the polynomial hierarchy, *Mathematical Systems Theory* **17**: 13–27.

Graham, R. L. and Spencer, J. H. (1971). A constructive solution to a tournament problem, *Canad. Math. Bull.* **14**: 45–48.

Graham, R. L., Rothschild, B. L. and Spencer, J. H. (1990). *Ramsey Theory*, second edition, Wiley, New York.

Halberstam, H. and Roth, K. F. (1983). *Sequences*, second edition, Springer Verlag, Berlin.

Hall, M. (1986). *Combinatorial Theory*, second edition, Wiley, New York.

Harper, L. (1966). Optimal numberings and isoperimetric problems on graphs, *J. Combinatorial Theory* **1**: 385–394.

Harris, T. E. (1960). Lower bound for the critical probability in a certain percolation process, *Math. Proc. Cambridge Phil. Soc.* **56**: 13–20.

Haussler, D. (1995). Sphere packing numbers for subsets of the boolean n-cube with bounded Vapnik-Chervonenkis dimension, *J. Combinatorial Theory, Ser. A* **69**: 217–232.

Haussler, D. and Welzl, E. (1987). ε-nets and simplex range queries, *Discrete and Computational Geometry* **2**: 127–151.

Håstad, J. (1988). Almost optimal lower bounds for small depth circuits, *in Advances in Computer Research* (S. Micali ed.), JAI Press, Chapter 5: Randomness and Computation, pp. 143–170.

Håstad, J. (1998). The shrinkage exponent of De Morgan formulas is 2, *SIAM J. Comput.* **27**: 48–64.

Janson, S. (1990). Poisson approximation for large deviations, *Random Structures and Algorithms* **1**: 221–230.

Janson, S. (1998). New versions of Suen's correlation inequality, *Random Structures and Algorithms* **13**: 467–483.

Janson, S., Knuth, D., Luczak, T. and Pittel, B. (1993). The birth of the giant component, *Random Structures and Algorithms* **4**: 233–358.

Janson, S., Luczak, T. and Rucinski, A. (2000). *Random Graphs*, Wiley, New York.

Joffe, A. (1974). On a set of almost deterministic k-independent random variables, *Ann. Probab.* **2**: 161–162.

Kahn, J. (1996). Asymptotically good list colorings, *J. Combinatorial Theory, Ser. A* **73**: 1–59.

Karchmer, M. and Wigderson, A. (1990). Monotone circuits for connectivity require super-logarithmic depth, *SIAM J. Disc. Math.* **3**: 255–265.

Karp, R. M. (1990). The transitive closure of a random digraph, *Random Structures and Algorithms* **1**: 73–94.

Karp, R. M. and Ramachandran, V. (1990). Parallel algorithms for shared memory machines, *in: Handbook of Theoretical Computer Science* (J. Van Leeuwen, ed.), Vol. A, Chapter 17, Elsevier, New York, pp. 871–941.

Katchalski, M. and Meir, A. (1988). On empty triangles determined by points in the plane, *Acta Math. Hungar.* **51**: 323–328.

Katona, G. O. H. (1972). A simple proof of the Erdős-Ko-Rado theorem, *J. Combinatorial Theory, Ser. B* **13**: 183–184.

Kim, J. and Vu, V. (2000). Concentration of multivariate polynomials and its applications, *Combinatorica* **20**(3): 417–434.

Kim, J. H. (1995). The Ramsey number $R(3,t)$ has order of magnitude $t^2/\log t$, *Random Structures and Algorithms* **7**: 173–207.

Kleitman, D. J. (1966a). On a combinatorial problem of Erdős, *J. Combinatorial Theory* **1**: 209–214.

Kleitman, D. J. (1966b). Families of non-disjoint subsets, *J. Combinatorial Theory* **1**: 153–155.

Kleitman, D. J., Shearer, J. B. and Sturtevant, D. (1981). Intersection of k-element sets, *Combinatorica* **1**: 381–384.

Kolountzakis, M. N. (1999). An effective additive basis for the integers, *Discrete Mathematics* **145**: 307–313.

Komlós, J., Pach, J. and Woeginger, G. (1992). Almost tight bounds on epsilon-nets, *Discrete Comput. Geom.* **7**: 163–173.

Komlós, J., Pintz, J. and Szemerédi, E. (1982). A lower bound for Heilbronn's problem, *J. London Math. Soc.* **25**(2): 13–24.

Loomis, L. H. and Whitney, H. (1949). An inequality related to the isoperimetric inequality, *Bull. Amer. Math. Soc.* **55**: 961–962.

Lovász, L., Spencer, J. H. and Vesztergombi, K. (1986). Discrepancy of set systems and matrices, *Europ. J. Comb.* **7**: 151–160.

Lubotzky, A., Phillips, R. and Sarnak, P. (1986). Explicit expanders and the Ramanujan conjectures, *Proc. 18-th ACM STOC*, pp. 240–246. See also: A. Lubotzky, R. Phillips and P. Sarnak, Ramanujan graphs, Combinatorica 8, 1988, 261–277.

Luczak, T. (1990). Component behavior near the critical point of the random graph process, *Random Structures and Algorithms* **1**: 287–310.

Luczak, T. (1991). A note on the sharp concentration of the chromatic number of random graphs, *Combinatorica* **11**: 295–297.

MacWilliams, F. J. and Sloane, N. J. A. (1977). *The Theory of Error Correcting Codes*, North Holland, Amsterdam.

Mani-Levitska, P. and Pach, J. (1988). Decomposition problems for multiple coverings with unit balls, manuscript.

Marica, J. and Schonheim, J. (1969). Differences of sets and a problem of Graham, *Canad. Math. Bull.* **12**: 635–637.

Matoušek, J. (1997). On discrepancy bounds via dual shatter function, *Mathematika* **44**(1): 42–49.

Matoušek, J., Welzl, E. and Wernisch, L. (1993). Discrepancy and approximation for bounded VC dimension, *Combinatorica* **13**: 455–466.

Matula, D. W. (1976). The largest clique size in a random graph, Technical report, Southern Methodist University, Dallas.

Maurey, B. (1979). Construction de suites symétriques, *Compt. Rend. Acad. Sci. Paris* **288**: 679–681.

Milman, V. D. and Schechtman, G. (1986). *Asymptotic Theory of Finite Dimensional Normed Spaces, Lecture Notes in Mathematics*, Vol. 1200, Springer Verlag, Berlin and New York.

Moon, J. W. (1968). *Topics on Tournaments*, Holt, Reinhart and Winston, New York.

Nakayama, A. and Peroche, B. (1987). Linear arboricity of digraphs, *Networks* **17**: 39–53.

Naor, J. and Naor, M. (1990). Small-bias probability spaces: efficient constructions and applications, *Proc. 22-nd annual ACM STOC*, ACM Press, pp. 213–223.

Nilli, A. (1991). On the second eigenvalue of a graph, *Discrete Mathematics* **91**: 207–210.

Pach, J. and Agarwal, P. K. (1995). *Combinatorial Geometry*, Wiley, New York.

Pach, J. and Woeginger, G. (1990). Some new bounds for epsilon-nets, *Proc. 6-th Annual Symposium on Computational Geometry*, ACM Press, New York, pp. 10–15.

Paturi, R. and Simon, J. (1984). Probabilistic communication complexity, *Proc. 25-th FOCS*, IEEE, New York, pp. 118–126.

Paul, W. J. (1977). A $2.5n$ lower bound on the combinational complexity of boolean functions, *SIAM Journal on Computing* **6**: 427–443.

Pinsker, M. (1973). On the complexity of a concentrator, 7-*th Internat. Teletraffic Conf.*, Stockholm, pp. 318/1–318/4.

Pippenger, N. and Spencer, J. H. (1989). Asymptotic behaviour of the chromatic index for hypergraphs, *J. Combinatorial Theory, Ser. A* **51**: 24–42.

Rabin, M. O. (1980). Probabilistic algorithms for testing primality, *J. Number Theory* **12**: 128–138.

Radhakrishnan, J. and Srinivasan, A. (2000). Improved bounds and algorithms for hypergraph two-coloring, *Random Structures and Algorithms* **16**: 4–32.

Raghavan, P. (1988). Probabilistic construction of deterministic algorithms: approximating packing integer programs, *J. of Computer and Systems Sciences* **37**: 130–143.

Ramsey, F. P. (1929). On a problem of formal logic, *Proc. London Math. Soc.* **30**(2): 264–286.

Rödl, V. (1985). On a packing and covering problem, *European Journal of Combinatorics* **6**: 69–78.

Sauer, N. (1972). On the density of families of sets, *J. Combinatorial Theory, Ser. A* **13**: 145–147.

Schrijver, A. (1978). A short proof of Minc's conjecture, *J. Combinatorial Theory, Ser. A* **25**: 80–83.

Shamir, E. and Spencer, J. H. (1987). Sharp concentration of the chromatic number in random graphs $G_{n,p}$, *Combinatorica* **7**: 121–130.

Shearer, J. B. (1983). A note on the independence number of triangle-free graphs, *Discrete Math* **46**: 83–87.

Shearer, J. B. (1985). On a problem of Spencer, *Combinatorica* **5**: 241–245.

Shearer, J. B. (1995). On the independence number of sparse graphs, *Random Structures and Algorithms* **7**: 269–271.

Shelah, S. and Spencer, J. H. (1988). Zero-one laws for sparse random graphs, *J. Amer. Math. Soc.* **1**: 97–115.

Shepp, L. A. (1982). The XYZ-conjecture and the FKG-inequality, *Ann. of Probab.* **10**: 824–827.

Smolensky, R. (1987). Algebraic methods in the theory of lower bounds for boolean circuit complexity, *Proceedings of the 19-th ACM STOC*, ACM Press, New York, pp. 77–82.

Spencer, J. H. (1977). Asymptotic lower bounds for Ramsey functions, *Disc. Math.* **20**: 69–76.

Spencer, J. H. (1985a). Six standard deviations suffice,, *Trans. Amer. Math. Soc.* **289**: 679–706.

Spencer, J. H. (1985b). Probabilistic methods, *Graphs and Combinatorics* **1**: 357–382.

Spencer, J. H. (1987). *Ten Lectures on the Probabilistic Method*, SIAM, Philadelphia.

Spencer, J. H. (1990a). Threshold functions for extension statements, *J. Combinatorial Theory, Ser. A* **53**: 286–305.

Spencer, J. H. (1990b). Counting extensions, *J. Combinatorial Theory, Ser. A* **55**: 247–255.

Spencer, J. H. (1995). Asymptotic packing via a branching process, *Random Structures and Algorithms* **7**: 167–172.

Suen, W. C. (1990). A correlation inequality and a Poisson limit theorem for nonoverlapping balanced subgraphs of a random graph, *Random Structures and Algorithms* **1**: 231–242.

Székely, L. (1997). Crossing numbers and hard Erdős problems in discrete geometry, *Combin. Probab. Comput.* **6**: 353–358.

Szele, T. (1943). Kombinatorikai vizsgálatok az irányitott teljes gráffal kapcsolatban, *Mat. Fiz. Lapok 50* pp. 223–256. For a German translation see: T. Szele, *Publ. Math. Debrecen* **13**, 1966, 145–168.

Talagrand, M. (1996). Concentration of measures and isopermetric inequalites in product spaces, *Publications Mathematiques de l'I. H. E. S.* **81**: 73–205.

Tanner, R. M. (1984). Explicit construction of concentrators from generalized N-gons, *SIAM J. Alg. Disc. Meth.* **5**: 287–293.

Tarjan, R. E. (1983). *Data Structures and Network Algorithms*, SIAM, Philadelphia.

Thomason, A. (1987). Pseudo-random graphs, *Annals of Discrete Math.* **33**: 307–331.

Turán, P. (1934). On a theorem of Hardy and Ramanujan, *J. London Math Soc.* **9**: 274–276.

Turán, P. (1941). On an extremal problem in graph theory, *Mat. Fiz. Lapok* **48**: 436–452.

Valtr, P. (1995). On the minimum number of empty polygons in planar point sets, *Studia Sci. Math. Hungar.* **30**: 155–163.

Wegener, I. (1987). *The Complexity of Boolean Functions*, Wiley-Teubner, New York.

Weil, A. (1948). *Sur les courbes algébriques et les variétés qui s'en déduisent*, Actualités Sci. Ind, no. 1041. iv+85pp.

Wendel, J. G. (1962). A problem in geometric probability, *Math. Scand.* **11**: 109–111.

Wright, E. M. (1977). The number of connected sparsely edged graphs, *J. Graph Theory* **1**: 317–330.

Yao, A. C. (1985). Separating the polynomial-time hierarchy by oracles, *Proceedings of the* 26*-th Annual IEEE FOCS*, IEEE, New York, pp. 1–10.

Литература, добавленная при переводе

Гаврилов Г. П., Сапоженко А. А., Задачи и упражнения по дискретной математике. М.: Физматлит, 2004 г.

Гончаров В. Л., О распределении циклов в перестановках, *Докл. АН СССР* **35**, No 9 — 1942 — С. 299–301.

Гончаров В. Л., Из области комбинаторики, *Изв. АН СССР, сер. матем.* **8**, No 1 — 1944 — С. 3–48.

Колчин В. Ф., Севастьянов Б. А., Чистяков В. П. Случайные размещения, — М.: Наука, 1976 г.

Комбинаторный анализ. Задачи и упражнения// Под ред. К. А. Рыбникова, — М.: Наука, 1982 г.

Нигматуллин Р. Г. Метод наискорейшего спуска в задачах на покрытие// Вопросы точности и эффективности вычислительных алгоритмов. *Вып.* **5**, *Труды симпозиума*, Киев, 1969, — С. 116–126.

Сапоженко А. А., О сложности д. н. ф., получаемых с помощью градиентного алгоритма // *Сб. Дискретный анализ — Новосибирск — Вып.* **21**, — 1972 — С. 62–71.

Сапоженко А. А., О наибольшей длине тупиковых д. н. ф. у почти всех функций алгебры логики// *Математические заметки* — 1968 — т. 4, No.6, — С. 649–658.

Харари Ф. Теория графов. М.: Мир, 1973 г.

Ширяев А. Н. Вероятность. М.: Физматлит, 1989 г.

Яблонский С. В., Введение в дискретную математику. М.: Физматлит, 1986 г.

Предметный указатель

NC, 280
NC, 281, 283
NP (недетерминированное полиномиальное время), 208, 216, 219–220
VC-размерность, 244–250, 252
ε-образец, 246–247, 250, 280
ε-сеть, 244, 246–247

Автоморфизм, 65, 69, 174, 180
Алгоритм, 19–20, 23–24, 36, 39–40, 51, 96–97, 99, 156–157, 162, 165–166, 266, 275–276, 278
  вероятностный или рандомизированный, 50, 54, 57, 96–97, 165, 275, 281
  детерминированный, 97, 284
  жадный, 49, 54, 57
  неадаптивный или интерактивный, 265–266
  проверки простоты, 165
  типа Монте-Карло, 165
Антицепь, 222

Биномиальная
  случайная величина, 248
Биномиальное
  распределение, 189, 244, 257, 289, 294
Блок-схема, 73

Вектор
  несбалансированный, 233
Выпуклая оболочка, 238, 242, 246, 254
Выпуклое тело, 48, 128–130

Гигантская компонента, 184, 188, 194
Гиперграф, 57, 86, 89, 133
  2-раскраска, 97
  однородный, 50, 54, 58, 73–74, 78, 86
  подгиперграф, 90
  покрытие, 73–74, 77
  порожденный, 74–77
  свойство B, 49, 53, 86
Гипотеза
  Адамара, 232
  Данцера и Грюнбаума, 239
  линейной древесности, 90–93
    ориентированная, 91
  Минка, 42, 80
  Райвала и Сэндса, 110
  Рамануджана, 162
  Римана, 160
  Селе, 80
  Симоновича и Шош, 271
  Хейлбронна, 47
  Эрдёша, 286, 300
  Эрдёша и Секереша, 301
  Эрдёша—Ханани, 57, 73, 78, 301
Граф
  вершинно-транзитивный, 174–175
  гамильтонов, 109
  гамильтонов путь, 32, 40, 80, 102
  зависимости, 97
  квазислучайный, 157, 167–168, 172–173
  клика, 66–67, 70–71, 113–114, 119, 216
  Кэли, 160, 162
    явное построение, 160
  матрица смежности, 160
  независимое множество, 46, 58–59, 91–92, 98, 100, 113, 156–157, 184, 205, 296–297, 299
  обхват, 59, 92–94, 299
    ориентированный, 92–94
  планарный, 102, 109, 285
  реберная связность, 23–24
  с корнем, 134
  сбалансированный, 67–69, 180
    строго, 67, 69, 180–181, 191
  цикл, 59
Графа линейная древесность, 90–91
  ориентированная, 91

Группа
матриц, 160, 162
подгруппа, 258
симметрическая, 117
фактор, 162
Групповой код, 258

Дисперсия, 37, 60–61, 63, 75, 78–79, 123–125, 170, 225, 248, 291–292, 294
Доминирующее множество, 21, 23–24, 204

Евклидова норма, 41, 78, 231
Евклидово
пространство, 89, 239, 242, 246, 270
расстояние, 129

Законы нуля или единицы, 195, 197, 203

Игра Эренфойхт, 196

Клика, 208
в графе, 132–133, 156–157
Клики максимальный размер, 181–182
Ковариация, 63
Код
двоичный БЧХ, 281–282
Компактность, 87, 90

Латинская трансверсаль, 95
Лемма
Клейтмана, 102, 107
локальная Ловаса, 19, 45–46, 83, 86–91, 93–96, 99, 101
симметричный случай, 85
переключающая, 209
Лес, 90, 271
звездный, 271–272
линейный, 90
ориентированный, 91, 93–94
Линейное расширение
частичного порядка, 109–111

Мартингал, 115–118, 120–122, 125–126, 131–132
подбрасывания монеты, 117
проявления вершин, 116–118, 120
проявления ребер, 115–119
Мартингальный процесс Дуба, 116
Математическое ожидание, 35–36, 41, 53, 60, 63–64, 78–79, 93, 107, 115–116, 136, 170, 192, 213, 225, 248, 288, 294, 296–297
линейность, 22, 32, 42, 44–46, 49, 57, 62, 66, 68, 75, 125, 127, 206, 222, 224, 237, 260, 262, 273, 276, 297
условное, 53, 115–117, 124–125, 261, 296
Матрица
Адамара, 231–233
смежности
графа, 160, 163–164, 167–168
турнира, 80
Метод второго момента, 60–61, 66, 72–73, 191
Метод Рёдля, 73
Метод Субботовской, 220
Метрика
Хэмминга, 126, 128
Метрическое пространство, 244
Множества попарно непересекающиеся, 91, 100
Множество
частично упорядоченное, 104, 109–111
Монохроматический, 35–36, 40, 44–45, 50–51, 53, 86–87, 90, 97, 99, 275, 299–300
Монохроматический граф, 276, 280–281

Независимое множество
в графе, 46, 58–59, 91–92, 98, 100, 113, 156–157, 184, 205, 296–297, 299
в евклидовом пространстве, 242
Непересекающиеся
клики, 119
семейства, 89
Неравенство, 64, 81, 85, 90, 93–94, 98, 103–104, 106–109, 111, 117, 121, 124, 126, 130, 159–161, 163–164,

173–174, 184, 211, 218–220, 240, 246, 248–250, 274, 276, 278–279, 287–288, 290–292, 295
FKG, 102, 106–108, 110–111, 210
Ацумы, 117–118, 120, 122, 126, 131–132
больших уклонений, 117
Бонферрони, 143
Гёльдера, 130
изопериметрическое, 117, 126
Йенсена, 267, 290
корреляционное, 102, 107, 109
Коши—Буняковского, 159, 163–164, 172
Маркова, 77, 123, 288, 290
мартингальное, 123, 134
Талаграна, 127–128, 130–132
Хана, 272
Чебышёва, 60, 63, 65, 72, 75–77, 136, 139, 248
Янсона, 108, 137–138, 180–181, 183
обобщенное, 133, 139, 183
Нормальное распределение, 61, 63–64, 233, 291–293, 295, 298

Объединение, 104–105, 110
Объединительно-неизбыточный, 105
Орграф зависимости, 84–85, 88, 92, 94–95
Отклонение
стандартное, 60–61, 93, 136, 224–225, 244, 287

Перекраска, 49
случайная, 49
Пересечение, 104–105, 110
Перманент, 42, 80
Плотность
графа, 67, 69
линейного расширителя, 160
множества, 300
упаковки пространства $\mathbb{R}^d$, 255
Покрытие, 73
гиперграфа, 73–74, 77
графа, 90
минимальный размер, 73
пространства $\mathbb{R}^d$, 89–90
неразложимое, 89
разложимое, 89
Полиномиальное время, 19
Проективная плоскость, 273
Простое число, 40, 48, 61–64, 79, 94, 158, 162, 165–166, 172, 213, 298
Путь, 164–166, 168
случайный, 166

Разброс, 223–224, 233, 250, 253, 266
линейный, 228
наследственный, 228, 230
семейства, 223
Размер максимальной клики, 121
Разрез, 23–24
Рамсея
граф, 156–157
теория, 303
число, 44–45, 58, 87–88, 297–299
Ранжированное пространство, 245–248, 250–251, 253
Раскраска, 21, 35–36, 44–46, 50, 86–87, 90, 93–94, 97, 99, 121, 219, 223–228, 232, 266, 275–276, 280–281, 283
гиперграфа, 97
случайная, 21, 44–45, 49, 87, 93, 101
Распределение, 21, 108, 115, 117–118, 132, 178, 184–185
биномиальное, 56, 93, 136, 189, 213, 243–244, 257, 289, 294
нормальное, 61, 63–64, 125, 233, 291–293, 295, 298
Пуассона, 55–56, 185–186, 189, 293–294
равномерное, 78, 87, 92–93, 95, 101, 128, 165, 220, 238, 241–242, 251
Расширитель, 160–162, 173
линейный, 160
плотность, 160
явное построение, 160, 166
Реберная связность, 24, 109
Решетка, 48, 104, 110, 254
дистрибутивная, 104–106, 110
подрешетка, 105

Свидетель, 165–166
Свойство графов, 109
Случайная величина, 22, 32, 38, 41, 60–61, 63, 78–79, 115, 123, 125, 129, 131, 170, 185, 222, 263,

266–270, 272, 281–284, 287–288, 294, 296
d-независимость, 280–284
биномиальная, 93, 136, 213, 243, 248
индикаторная, 22, 32, 44, 46, 61, 68, 70, 75–76, 113, 206, 222, 224, 226, 243, 260, 262, 273, 292–293
ковариация, 61
разложение, 32, 61, 97
стандартная нормальная, 295
Случайное блуждание, 115, 174
Случайный путь, 166
Собственное значение, 160
графа, 167
матрицы, 162
регулярного графа, 161–162, 164, 166
симметричной матрицы, 160–161, 163–164, 167
Собственный вектор
симметричной матрицы, 161, 163–165
Среднее, 55–56, 61, 118, 125, 153, 185
Среднее значение, 127, 132–133, 287, 292–294
Среднее
геометрическое, 42–43
Стратегия предварительного просмотра, 201
Схема, 207–208, 212–216, 218–219, 221, 299
2-схема, 208, 216–217, 221
булева, 208
глубина, 208
кодирования, 256–258
монотонная, 216–218
ограниченной глубины, 209, 213, 221
подсхема, 212
сложность, 207–209, 216–217
сортировки, 160
Счетчик четности, 207, 212–213, 215, 219, 221

Тактическая конфигурация, 73
Теорема
XYZ-теорема, 110
Альсведе—Дайкина, 103–104, 106, 108
Бека-Фиала, 233
Брегмана, 81
Вейерштрасса, 136
Вейля, 162
Визинга, 219
Кёнига—Холла, 92
Клейтмана, 236
Пифагора, 239
Радона, 246
Рамсея, 298, 300
Тихонова, 87
Турана, 46–47, 113
Фробениуса—Перрона, 164
Шеннона, 257–258
Эрдёша—Ко—Радо, 31
Тройки Штейнера, 73
Турнир, 20–21, 32–33, 80–81, 83, 157–158, 160
квадратичных вычетов, 157–158, 172–173
явное построение, 21, 160

Уклонение
большое, 62, 117, 146–147, 190, 287, 293
стандартное, 186
Упаковка, 48–49, 54, 57–58, 73, 254
жадная, 54
константа, 254
пространства $\mathbb{R}^d$, 255
случайная, 54
число, 57, 73
Упаковки константа, 48
Условие Липшица, 122–123, 125–126, 131–132
вершинное, 118, 120
реберное, 118–119

Фазовый переход, 184–185, 192
Формула
булева, 209
включений-исключений, 143
Стирлинга, 62, 81
Функция
log-супермодулярная, 106–108, 111
булева, 207–210, 213, 216–217, 219–220, 242
сложность, 208
кликовая, 208

липшицева, 131
пессимистических оценок, 277–278
пороговая, 66–69, 143, 180, 203–204
Рамсея, 300
энтропии, 257

Характер квадратичного вычета, 158
Хроматическое число, 59, 116, 118, 135, 183, 272, 295, 299
Хэмминга метрика, 226, 257

Цепь, 222
жесткая, 201–203
Маркова, 185

Число
клик, 70
пересечений, 285

Энтропия, 225–228, 256, 266, 268, 272
бинарная, 266
случайной величины, 266
условная, 266

Явное построение, 156–157, 160, 162, 260, 284, 297
линейного расширителя, 160
расширителя, 166
турнира, 21, 160

Именной указатель

Адамар, 231–233
Акияма, 90
Алон, 80, 123
Альсведе, 102–104, 106, 108
Андреев, 216
Ацума, 117–118, 120, 122, 126, 131–132

Бабай, 301
Барани, 235, 242
Бек, 52, 96–97, 233
Боллобаш, 118, 178, 301
Бонферрони, 143
Боппана, 225
Брегман, 42, 80–81
Буняковский, 159, 163–164, 172

Вапник, 245
Вапник и Червоненкис, 244, 247
Вейерштрасс, 136
Вейль, 160, 162, 173
Визинг, 219
Ву, 133–134

Гёльдер, 130
Гончаров, 19
Гринберг, 235
Грэхем, 160, 301
Грюнбаум, 239

Дайкин, 102–104, 106, 108
Данцер, 239
Де ла Вега, 157
Дуб, 116

Ексо, 90

Зидон, 299

Игуса, 162

Йенсен, 267, 290

Кан, 73
Каро, 113
Кац, 298
Кёниг, 92
Ким, 89, 123, 133–134
Клейтман, 102, 107–108, 236
Ко, 31
Комлош, 47
Конвей, 83
Коши, 159, 163–164, 172
Кэли, 160, 162

Липшиц, 118–120, 122–123, 125–126, 131–132
Ловас, 45–46, 83–84, 86
Лучак, 178

Мани-Левицка, 89
Маргулис, 162, 166
Марков, 77, 123, 288, 290
Матула, 23–24
Миклош, 301
Милман, 121
Минк, 42, 80
Мори, 117, 121

Нешетржиль, 301
Нигматуллин, 23

Пач, 89, 301
Перлес, 245
Пинскер, 160
Пинц, 47
Пиппенджер, 73–74
Пифагор, 239
Поддерюгин, 23
Пуассон, 56, 185, 189
Пэли, 172

Рабин, 165
Радо, 31, 300
Радон, 246
Разборов, 214, 216
Райвал, 110
Рамануджан, 61, 162
Рамсей, 44–45, 58, 87–88, 156–157, 297–300, 303
Реньи, 69, 178–179, 184, 299
Рёдль, 57, 73–74, 78, 127, 301
Риман, 160, 254
Рот, 150
Ручински, 178

Саймонс, 115
Сапоженко, 23, 65
Секереш, 21, 298, 301–302
Селе, 19, 33, 80
Селфридж, 299
Семереди, 47, 285–286, 300
Сёньи, 301
Симонович, 271
Смоленский, 213
Спенсер, 73, 95, 123, 160, 286, 302
Стирлинг, 62, 81
Субботовская, 220
Сэндс, 110

Троттер, 285–286
Туран, 46–47, 63, 113, 300

Уэй, 113

Фиала, 233
Фишборн, 110
Франкл, 73–74
Фюреди, 240, 242

Хайнал, 301
Халберстам, 150
Хан, 272
Ханани, 57, 73, 78, 301
Харари, 90
Харди, 44, 61, 64
Харрис, 102
Хейлбронн, 47
Холл, 92
Хостад, 220
Хоффман, 301

Чан, 298, 301
Чебышёв, 60, 63, 65, 72, 75–77, 136, 139, 174, 248
Червоненкис, 245
Чернов, 272, 295

Шелла, 245
Шеннон, 256–258
Шехтман, 121
Ширер, 85, 269
Шош, 271, 301
Штейнер, 73
Шютте, 20, 160

Эйлер, 302
Эйнштейн, 302
Эйхлер, 162
Эрдёш, 18–19, 31, 57, 60, 69, 73, 78, 84, 86–87, 89, 95, 156, 178–179, 184, 240, 286, 298–304
Эренфойхт, 196

Янсон, 108, 133, 137–139, 153, 178, 180–181, 183

Минимальные системные требования определяются соответствующими требованиями программ Adobe Reader версии не ниже 11-й либо Adobe Digital Editions версии не ниже 4.5 для платформ Windows, Mac OS, Android и iOS; экран 10"

Учебное электронное издание

Алон Нога
Спенсер Джоэл

ВЕРОЯТНОСТНЫЙ МЕТОД
Учебное пособие

Ведущий редактор *М. С. Стригунова*
Художник *Н. В. Зотова*
Технический редактор *Е. Денюкова*
Оригинал-макет подготовлен *О. Г. Лапко* в пакете LaTeX 2_ε

Подписано к использованию 20.08.23.
Формат 145×225 мм

Издательство «Лаборатория знаний»
125167, Москва, проезд Аэропорта, д. 3
Телефон: (499) 157-5272
e-mail: info@pilotLZ.ru, http://www.pilotLZ.ru

Открытие того, что детерминированные утверждения могут быть доказаны с помощью вероятностных соображений, позволило уже в первой половине XX в. получить ряд замечательных утверждений из анализа, теории чисел, комбинаторики и теории информации. Вскоре стало ясно, что метод, который сейчас называется вероятностным, является весьма мощным инструментом получения результатов в математике.

Главная цель монографии — изложение идей вероятностного подхода к решению задач дискретной математики. Авторы придерживаются известного тезиса о том, что пример учит лучше, чем теория. Подбор примеров в книге отвечает самым высоким требованиям целесообразности и вкуса, а некоторые из них являются избранными шедеврами. По существу, это — мастер-класс двух маэстро для лиц, заинтересованных в освоении вероятностных методов.

Книга будет полезна специалистам в области дискретной математики (комбинаторики, теории сложности, приложений теории вероятностей), студентам, аспирантам и преподавателям соответствующих дисциплин.