Лекция «Замечательные идеи комбинаторики»

Поздняков Сергей Николаевич

запись конспекта: Кацер Евгений

дата лекции: 21 декабря 2018 г.

Глава II. Комбинаторика

00:04

Комбинаторика — одна из областей, вход в которую очень короткий. То есть с одной стороны проблемы могут быть объяснены на пальцах, а с другой стороны они могут быть до сих пор не решены. В комбинаторике очень интересно проводить компьютерный эксперимент, пытаться вывести какие-то общие формулы.

Теперь вообще по комбинаторике: как она связана с программированием, с курсом дискретной математики. Во-первых, есть просто комбинаторные алгоритмы, когда вы начинаете что-то перебирать: то нужно выбирать все подмножества какого-то множества, то вам нужно перебирать перестановки, то перестановки, в которых часть элементов повторяются и так далее (эти генераторы и генерацию таких комбинаторных объектов мы будем разбирать через неделю). Во-вторых, совершенно классические задачи (подсчет чего-нибудь). А что в программировании интересно считать? Трудоемкость – сколько времени требуется на выполнение алгоритма, если размер входных данных n (граф с n вершинами, число с п цифрами и др). На семинаре часто упоминаются классическая работа Кнута — «Искусство программирования». В ней последний четвертый том посвящен комбинаторным задачам. И еще хочется сказать, что есть одно свойство у комбинаторики: когда человек что-то изобретает, и были проведены в свое время даже исследования таким математиком Адамаром: он провел анкету среди известных ученых, чтобы узнать как появляется новое, как человек открывает то, что до него никогда не было, оказывается, что процесс открытия связан с напряженной работой и актуализацией тех знаний которые у человека есть.

У меня одна девушка решила по одной из тем коллоквиума взять похожую тему, но попроще, ну просто с ней разобраться. Мы с ней уже два раза виделись, я ее спрашиваю: «Ну как, вы разобрались?» Она говорит: «Нет, я хочу сама придумать этот алгоритм.» Вот это вот правильный подход: надо сначала исчерпать все свои возможности по решению этой задачи. И вот тогда, когда эти возможности исчерпаны, вы запускаете некие механизмы в подсознании. Ученые рассказывают, что происходило после такой неуспешной предварительной деятельности, они бросали эту задачу, кто-то ехал отдыхать, кто-то переключался на другую работу. И вдруг неожиданно всплывало решение этой задачи. Кто-то даже пытался последить за собой, как же это получается. В некоторых наблюдениях говорят, что вот у ученого в голове сталкивались какие-то различные идеи, и вдруг они образовали хорошую комбинацию.

Я хочу сказать, что на каком-то уровне, ну вот сейчас всех интересует

искусственный интеллект и нейронные сети, и если вы будете пытаться смоделировать какие-то интеллектуальные процессы, то вы поймете, что на каком-то уровне наступает, грубо говоря, перебор вариаций, и все вот это искусство заключается в том, что ученые, которые делали открытия, думая над этой задачей, создавали такой богатый контекст, что, когда вот эти ситуации сталкивались, они сразу видели красиво или некрасиво. Это значит не формально: она может подойти или не может. И вот эта эстетическая оценка, которая формируется за счет того, что вы об этом думаете, вы над этим работаете, вы все что можно собираете вокруг этой задачи, резко уменьшает объем комбинаций, потому что такой просто тупой перебор, вы можете подумать про шахматы, там, например, если просто перебирать очень быстро (экспоненциально) растут комбинации. Я даже могу привести такое рассуждение Галилея, он открыл телескоп. Вот он стал рассуждать, как может быть устроен телескоп: простые стекла там быть не могут, значит они выпуклые или вогнутые, если два выпуклых комбинировать, ничего не получится, если два вогнутых, тоже ничего не получится, значит надо выпуклое и вогнутое стекла комбинировать, и подобрать расстояние. Вот такими комбинаторными суждениями Галилей открыл для себя, как устроен телескоп.

Комбинаторные задачи еще интересны тем, что каждая задача в комбинаторике предполагает свою модель. Их можно как-то классифицировать, собрать несколько похожих задач, но это как раз и интересно то, что вы можете к каждой задаче применять эти общие принципы и строить новую модель. Вот сейчас я и хочу продемонстрировать вам такие простые идеи. Когда-то меня попросили прочесть лекцию в Пакистане на английском языке, я думал, что же я смогу прочитать, я выбрал как раз комбинаторику. Прочитал лекцию, которая называется «Brilliant ideas of combinatorics», поэтому значит первая часть лекции это будет вот эти «Brilliant», замечательные идеи комбинаторики.

08:37

Прежде чем начать эти замечательные идеи я еще раз обращусь к тому, насколько в комбинаторике просто поставить нерешенную задачу. Вот как раз опять же в четверг Николай Николаевич Васильев упомянул Владимира Игоревича Арнольда, этого очень известного нашего ученого, и Николай Николаевич говорит, что вот Владимир Игоревич ставил свои задачи так: он выбирал область, где существуют какие-то перестановки, и изучал перестановки в этой области. Николай Николаевич не пояснил что за этим стоит, а вот я могу пояснить: вот есть, например, прямолинейная дорога и ее пересекает какая-то речка:

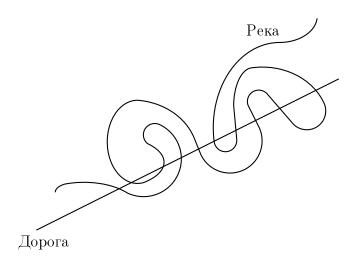


Рис. 1: Пример дороги и реки

Вы можете мостики (пересечение дороги и реки) пронумеровать в порядке того, как вы проходите, когда едете по дороге:

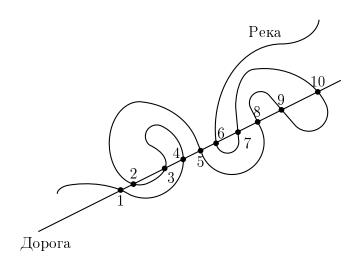


Рис. 2: Пример дороги и реки

Получилось 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. А потом плыть по этой речке в том же направлении и смотреть в каком порядке проходят мостики. Получится 1, 4, 3, 2, 5, 8, 9, 10, 7, 6. Вот у нас получилась перестановка 10 чисел. Понятно, что не любую перестановку можно изобразить меандром (меандр — форма реки). Так вот эта задача о количестве меандров. У меня как-то пару лет назад или может быть три года два студента независимо взялись решать эту задачку численно, и в какой-то момент

нам показалось, что они поставили уже мировой рекорд, потому что количество их так быстро возрастает, но я не помню до 10 по моему, нет они уже до 14 и до 15 дошли, а дальше уже не смогли. Я им говорю: «Давайте вы уже статью может напишите». Тогда они стали изучать литературу более внимательно, нашли что уже кто-то их обогнал, поэтому 14 и 15 не были лучшим результатом, но они вот сами до него дошли.

Видите, задачка очень понятно поставлена и до сих пор не решена, то есть нет формулы, есть только какие-то оценки сверху и снизу для количества. Я думаю, что надо посмотреть, наверное, у Арнольда еще есть задачи на перестановки. Я поговорю об этом с Васильевым. Это к тому, что начинается период, когда надо брать тему альтернативного экзамена, и я прошу вас всех до мая определиться, потому что май уже — это минимум, которых необходим, чтобы выполнить проектную работу.

12:02

Хорошо, теперь перейдем к замечательным идеям комбинаторики:

1. Правило умножения. Вы знаете, что и в алгебре, и в программировании есть двумерные массивы, матрицы. И матрицы естественно использовать для того, чтобы классифицировать какие-то объекты по 2 параметрам. Допустим у вас есть множество букв — $\{a,b,c\}$ и есть множество чисел — $\{1,2,3,4\}$, тогда в таблице естественным образом можно перечислить их комбинации (точнее пары составленные из букв и цифр):

| | 1 | 2 | 3 | 4 |
|---|------------|------------|------------|------------|
| a | a1 | a2 | a3 | a4 |
| b | <i>b</i> 1 | <i>b</i> 2 | <i>b</i> 3 | <i>b</i> 4 |
| c | <i>c</i> 1 | c2 | c3 | c4 |

Вы знаете, наверное, что такая конструкция называется «Декартово произведение двух множеств». То есть, если буквы обозначить множеством А от слова Alphabet, а цифры — множеством D от слова Digit, то получится что вот эта таблица есть не что иное, как произведение вот этих двух множеств:

$$A = \{a, b, c\}; D = \{1, 2, 3, 4\}$$

Произведение декартово, то есть мы рассматриваем пары чисел:

$$A \times D = \{(a_i, d_i) | a_i \in A, d_i \in D\}$$

Так определяется декартово произведение. Понятно, что если вам нужно посчитать количество элементов вот этого множества $|A \times D|$... Вот эти палочки ($| \ | \ |$) в комбинаторике обозначают мощность множества или количество элементов, если это множество конечно. С мощностью вы скорее всего в анализе сталкивались, что вот есть бесконечные множества, например, натуральные числа, есть рациональные числа есть вещественные числа, есть комплексные числа. Так вот мощность обладает интересными свойствами для бесконечных множеств. Для конечных все понятно: если мы умножим количество элементов A на количество элементов D, то мы получим сколько элементов произведения:

$$|A \times D| = |A| \cdot |D|$$

15:17

А вот, когда у нас множество бесконечное, появляется очень любопытный эффект. У нас был математик и популяризатор — Наум Яковлевич Виленкин, он написал книжку «Комбинаторика», а потом написал книжку «Рассказы о множествах». И в «Рассказах о множествах» он вот эти эффекты с мощностью множеств облек в фабулу одного из произведений Станислава Лема — «Звездные дневники Ийона Тихого». Фабула такая: есть отель в котором бесконечное количество номеров, и вот туда прилетает Ийон Тихий, а все номера заняты. Но администратор делает очень просто: он жильца из первого номера переселяется во второй, из второго в третий, из третьего в четвёртый, а Ийона Тихого помещают в первый. После него прилетает делегация из ста человек. Понятно, что нужно делать: надо первого в сто первый, второго в сто второй, а делегацию поселить на освободившиеся места. Но потом прилетает делегация из бесконечного количества человек, тогда администратор селит первого во второй, второго в четвертый, третьего в шестой, К-ого в 2К-ый. Все нечетные номера освобождаются, и администратор селит в них бесконечную делегацию. Но после этого прилетает бесконечное число бесконечных делегаций. Что делать тогда? Тогда администратор строит похожую на нашу табличку, в которой строка — номер делегации, а столбец — номер члена делегации. Только теперь эта таблица не конечная, а бесконечная:

Таблица 1: Таблица соответствует каждому члену делегации

| № дел. | 1 | 2 | 3 | |
|--------|---|---|---|---|
| 1 | | | | |
| 2 | | | | |
| 3 | | | | |
| : | : | : | : | ٠ |

Давайте теперь расселим всех членов делегации. Селить их будем по диагоналям:

Таблица 2: Таблица соответствует каждому члену делегации

| № дел. | 1 | 2 | 3 | |
|--------|----|---------|--------------|---|
| 1 | №1 | N_{2} | $N^{\circ}4$ | |
| 2 | №3 | №5 | | |
| 3 | №6 | | | |
| i : | i | : | ÷ | ٠ |

Это не единственный способ расселения.

18:22

Упражнение. В какой номер будет поселен j-ый член i-ой делегации?

Формулы для генерации очень понадобится, поэтому вот это неплохая достаточно простая задача. Кстати, это обозначает почему четных чисел столько же, сколько натуральных, почему рациональных чисел столько же, сколько целых. Потому что №3 можно считать одной второй. Вот мы их пронумеровали, то есть мы сделали взаимно-однозначное соответствие одного множества и второго. Но для конечных множеств понятно, что если у нас есть 10 юнош и 10 девушек, начинается танец и все заняты — значит их одинаково. В общем случае это то же самое, если вы между двумя множествами нашли взаимно однозначное-соответствие, то мощности этих множеств одинаковы.

Почему вещественных чисел больше чем рациональных я не буду рассказывать. Это к сегодняшней лекции отношения не имеет. Вот через полгода, когда мы будем с вами обсуждать теории алгоритмов, там это будет очень важно, потому что как раз есть задача, которых множество задач эквивалентно множеству всех чисел, а множество алгоритмов эквивалентно множеству натуральных чисели, и поэтому понятно, что есть задачи для которых нет алгоритма. Такие будем приложения вот этой теории с вами обсуждать.

Вот это мы с вами все обсуждаем самый простой принцип умножения, но, кстати, его можно и применить к какой-нибудь простой задачке.

20:34

Задача. Например, сколько существует различных автомобильных номеров?

Решение. Ну, вот я каждый год эту задачку задаю и каждый год забываю, как устроен номер. Сначала идет буква, потом три цифры, потом две буквы. То есть множество номеров можно записать в таком виде:

$$A \times D \times D \times D \times A \times A$$

Множества A и D будут больше, чем в предыдущем примере. Теперь, если мне нужно найти количество этих номеров, я могу воспользоваться той же самой идеей. Ну понятно, что у меня будет уже таблица не квадратная $A \times D \times D \times D \times A \times A$, не кубическая $(A \times D \times D) \times D \times A \times A$, не четырехмерно кубическая $(A \times D \times D \times A \times A)$, не пятимерно $(A \times D \times D \times A \times A)$, шестимерный куб будет, не куб, а параллелепипед потому что неравные стороны. Но все равно его объем, в данном случае, или количество элементов считается по такой же формуле:

$$A \times D \times D \times D \times A \times A = |A| \cdot |D| \cdot |D| \cdot |D| \cdot |A| \cdot |A|$$

В правой части выражения мы перемножаем числа. В левой части выражения множества менять местами нельзя, иначе другой набор будет, а вот в правой |A| — количество букв, |D| — количество цифр, поэтому числа я могу переставлять:

$$|A|\cdot |D|\cdot |D|\cdot |D|\cdot |A|\cdot |A|=|A|^3\cdot |B|^3$$

Как вы думаете сколько там букв можно брать? 20 больше, меньше? Мне считать не хочется. 20 в куб легче возводить и 10 в куб:

$$|A|^3 \cdot |B|^3 \approx 20^3 \cdot 10^3 = 8000000$$

Теперь давайте следующую задачу. Мы ее назовем, если тут было правило умножения или принцип умножения, то тут будет правило сложения или принцип сложение. Но я лучше не правило, а принцип напишу, потому что эта идея достаточно общая.

23:18

2. Принцип сложения. Вы уже слышали, наверно, есть такой тезис «Divide and conquer» — «Разделяй и властвуй». Вот, в частности, принцип сложения относится к той же идее. Если вы не можете решить трудную задачу, разбейте ее на части и решайте эти части. Но на самом деле этот процесс может быть итеративный, и у нас будут получаться так называемые рекуррентные формулы, когда мы не можем написать сразу, как надо делать вычисления, но каждый шаг, переход от одного к другому, мы можем записать. Давайте решим такую задачу: пусть нам нужно рубль разменять монетками 50 копеечными, 10 копеечными. Пятикопеечную уже, наверное, не найти, но мы их тоже учтем.

24:04

Задача. Сколькими способами можно разменять 1 рубль монетами достоинством: 50 копеек, 10 копеек, 5 копеек. Ну, я уж единички брать не буду, а то слишком будет утомительно считать.

Решение. Понятно, что такое размен монет: если вам дали там 50 копеек и пять раз по 10 копеек и если вам сначала дадут 10, потом 50, а потом еще четыре монеты по 10, то это то же самое. То есть нас не интересует порядок. Нас интересует набор монет, только из каких монет он составлен. А в каком порядке нам эти монеты выдали, мы учитывать не будем. Давайте обозначим, ну допустим, D_{100} — количество разменов одного рубля, который равен 100 копеек. И вот теперь все способы мы разобьем на два типа: есть там 50 копеечная монета или нет. Значит, если 50 копеечная монета есть: вот мы ее положили, и нам остается разменять 50 копеек:

$$D_{100} = D_{50} + \dots$$

Теперь случай когда нет 50 копеечных. Значит мне нужно разменять рубль:

$$D_{100} = D_{50} + D'_{100}$$
, где

 D'_{100} — количество разменов 1 рубля монетами достоинством 10 копеек и 5 копеек. То есть 50 копеек я уже использовать не могу. Значит вы согласны с тем, что я правильно разложил на сумму? Значит еще раз: если 50 копеек есть, то мне осталось разменять 50, то есть я беру размениваю, потом каждому набору добавляю еще 50 копеечную монету. С этого их количество не изменится. Либо 50 копеек там просто нет, и я свожу задачу к более простой, где меньше вариантов выбора разменной монеты. Теперь 50 значит. У меня два варианта: либо там одна 50 копеечная монета, и мне остается разменять ноль:

$$D_{100} = D_{50} + D'_{100} = D_0 + \dots$$

Понятно какое число D_0 естественно присвоить — один, потому что это означает, что два раза по 50 копеек. Значит либо там нет монет по 50 копеек, и тогда мы аналогично раскладываем, либо есть.

$$D_{100} = D_{50} + D'_{100} = (\underbrace{D_0}_{1} + D'_{50}) + \dots$$

Дальше, соответственно, разложим D'_{100} . Либо там есть десятикопеечная монета, либо там нет 10 копеечных монет. Если десятикопеечная монета есть, нам остается разложить 90, если десятикопеечной монеты нет, то мы раскладываем 100:

$$\mathrm{D}_{100} = \mathrm{D}_{50} + \mathrm{D}_{100}' = (\underbrace{\mathrm{D}_0}_{1} + \mathrm{D}_{50}') + (\mathrm{D}_{90} + \mathrm{D}_{100}''),$$
где

 D_{100}'' — количество разменов 1 рубля монетами по 5 копеек. Сколько таких вариантов? Один! Это двадцать монет, но размен всего один. Значит после этого у меня получается:

$$D_{100} = D_{50} + D'_{100} = \left(\underbrace{D_0}_{1} + D'_{50}\right) + \left(D_{90} + \underbrace{D''_{100}}_{1}\right) = 2 + D'_{50} + D'_{90}$$

Теперь мы раскладываем аналогичную D'_{50} . Либо есть там 10 копеек, либо нет. Если есть, остается 40, все 50 мы раскладываем по 5 копеек. Но опять же это всего будет один способ. То же самое D'_{90} . Либо два десятка там есть, тогда остается 80, либо все разбиваем на пяти копеечную:

$$2 + D_{50}' + D_{90}' = 2 + (D_{40}' + \underbrace{D_{50}''}_{1}) + (D_{80}' + \underbrace{D_{90}''}_{1}) = \dots$$

И вот теперь, наверное, можно поставить многоточие и сообразить сколько шагов мы так сделали. Значит было 50, потом стало 40 и еще добавилась единичка, 30 еще единичка 20, 10, 0, да? Посчитали сколько там добавятся единичек или не посчитали? Еще 4, наверное. То есть $D'_{40} + D''_{50} = 4 + 1 = 5$ [скорее всего здесь 3 + 1, так как в дальнейшем мы пишем $D'_0 + D''_{10}$] и плюс еще 2 в начале = 7. И еще за эти шаги $D'_{80} + D''_{90}$ добавит 5.

$$2 + (D'_{40} + \underbrace{D''_{50}}_{1}) + (D'_{80} + \underbrace{D''_{90}}_{1}) = 12 + \underbrace{D'_{0}}_{1} + \underbrace{D''_{10}}_{1} + D'_{40} + \underbrace{D''_{50}}_{1} =$$

$$= 15 + D'_{40}$$

 D'_{40} даст нам еще 5, получится:

$$15 + D'_{40} = 20$$

Мне кажется, что я где-то ошибся, чего-то у меня ощущение, что их то ли 18, то ли 19. На самом деле $D'_{40}=4$, поэтому:

$$15 + D'_{40} = 19$$

Я предлагаю вам просто перебрать все эти наборы после лекции и проверить, что мы нигде в вычислениях не ошиблись.

Вы видите, что ту формулу, которую мы получили или получали, они называются рекуррентными — мы выражаем одни величины через другие. Если понадобится вам выразить не рубль, разложить. Это называется разбиение, не рубль разбить, а там N разбить на слагаемые, то вы сделаете аналогично и выведите общую формулу. Кстати, вот это пусть будет заданием.

34:31

Задача. Обобщить предыдущую задачу на случай размена не одного рубля, а N копеек. Можно, конечно, N рублей, но тогда более будет частного вида задача.

Теперь я хочу вам показать один прием, который позволяет нам использовать разные способы рассуждения для того, чтобы вывести различные комбинаторные свойства, формулы и так далее. Сравнение — третий способ. Сравнение различных способов решения.

35:27

3. Сравнение различных способов решения. Вы знаете уже, что когда мы с вами теорему Безу изучали и применили ее к задаче интерполяции, то мы нашли связь между алгеброй и анализом. А сейчас мы будем внутри комбинаторики искать такие связи.

Задача. Вот смотрите какую я задаю задачку. Вот здесь сидит N человек. Мне нужно из N человек выбрать команду и из этой команды выбрать одного капитана. Сколькими способами это можно сделать?

Решение. Я не знаю все ли проходили в школе, но есть такое замечательное число, которое называется C_n^k — это число способов выбора k предметов из n различных. То есть, если я хочу выбирать из вас команду из 10 человек, а вас допустим 120, это будет C_{120}^{10} . Я мог бы это перевести на более формальный математический язык и сказать, что C_n^k — число k элементных подмножеств множества из n элементов. Это будет то же самое. Давайте использовать. Теперь будем считать, что вы не проходили, что такое C_n^k и вы слышите впервые, что я сказал, и не знаете никакой формулы, как C_n^k вычисляется. Можем ли мы вот из ничего извлечь такую формулу? Значит вот давайте вернемся к принципу умножения и к той задаче, которую я вам только что поставил. Вот допустим мы действуем так: сначала мы выбрали команду из k человек, а потом в этой команде выберем капитана. Сколькими способами я могу выбрать капитана? k человек в команде. Любой человек в команде может быть капитаном. Понятно, что k и C_n^k — независимые комбинации, да? Поэтому я могу перемножить:

 kC_n^k

Вот получается у меня один спуск. Могу ли я рассуждать по-другому? Могу сначала выбрать капитана вообще из всех. Сколькими способами я могу выбрать капитана? N способами. А потом к нему добавить команду, но раз я уже капитана выбрал, он уже не считается.

Значит у меня осталось n-1 человека. И сколько мне надо выбрать из них? k-1:

$$nC_{n-1}^{k-1}$$

Но мы решали одну и туже задачу, поэтому и ответ должен получиться тот же самый:

$$kC_n^k = nC_{n-1}^{k-1}$$

А теперь разделим обе части на k:

$$C_n^k = \frac{n}{k} C_{n-1}^{k-1}$$

Видите, мы получили рекуррентную формулу, которая позволяет нам, если нам нужно найти более сложное число, свести это к более простому. Мы всегда можем до вычитать до того, пока k-1 не станет равным 1, например. Сколькими способами 100 человек можно выбрать одного? 100.

$$C_n^1 = n$$

Поэтому вот достаточно такой простой формулы, чтобы потом вычислить любое такое выражение. Но можно и немножко продолжить этот процесс. Можно теперь применить то же самое правило. Тогда получится:

$$C_n^k = \frac{n}{k} \cdot C_{n-1}^{k-1} = \frac{n}{k} \cdot \frac{n-1}{k-1} \cdot C_{n-2}^{k-2}$$

И так мы будем идти. По смыслу k всегда не будет превышать n: я же не могу из ста человек выбрать 200. Это будет 0, поэтому все разумные случаи предполагают, что k не больше n, поэтому k кончится раньше, поэтому, если я напишу многоточие в выражении, то конец у меня будет вот такой:

$$C_n^k = \frac{n}{k} \cdot \frac{n-1}{k-1} \cdot \ldots \cdot \frac{1}{1}$$

В последней дроби внизу у меня будет единичка, а что будет наверху? Ответ из зала: n-k. Все согласны? А вот смотрите, как

легко проверить: разница между числителем и знаменателем в первой дроби n-k, во второй естественно та же самая, мы же вычли по единичке из числителя и из знаменателя, значит и в конце эта разница должна быть той же самой. Это инвариант цикла. То, о чем мы уже говорили. В последней дроби какой инвариант цикла? n-k-1, значит наверху... То есть инвариант цикла n-k, а наверху что должно стоять? n-k+1, чтобы после вычитания из числителя знаменателя получилось n-k:

$$C_n^k = \frac{n}{k} \cdot \frac{n-1}{k-1} \cdot \ldots \cdot \frac{n-k+1}{1}$$

Получившееся выражение в западных учебниках обозначают вот так:

$$\binom{n}{k}$$

Ну, понятно, что берем $\frac{n}{k}$, а потом от числителя и знаменателя вычитаем по единичке, пока это сохраняет смысл. Поэтому, если вы будете читать Кнута, например, то вы увидите вместо C_n^k вот такое. Мне почему-то приятнее работать с C_n^k . Может, потому что я в детстве прочитал Виленкина и на всю жизнь запомнил. Может както западное обозначение не похоже на число, да как это скобочки. Да, когда мы видим C_n^k , то под ним мы разумеем какое число, а в западном обозначении как бы вот что-то такое непривычное. Но не знаю. Но полезно, конечно, пользоваться любыми обозначениями.

Следующая задача связана с самым главным, наверное, приемом в комбинаторике — это взаимно-однозначное соответствие.

42:59

4. Взаимно-однозначное соответствие. Вот опять же если взять вот этот доклад, который был в четверг, там такие диаграммы Юнга рассматривались, а потом в них там ставились циферки, но их можно было рассматривать как последовательное построение этой диаграммы и каждое это число, номер, означало какой по счету это клеточка поставлена. Но потом этой же таблице был сопоставлен путь в некотором графе и вместо того, чтобы считать в таблице, можно считать пути в графе или наоборот. Я вам задам задачку. Есть такое журнал квант, до сих пор который, в 1970 году появился,

и тогда количество людей, интересовавшихся математикой и читавших журнал, было таким, что этот журнал был самый известный в мире, самый лучший журнал по математике для школьников. Так вот я вам дам сейчас задачу из кванта для младших школьников. Вы знаете что такое счастливый билет? Трамвайный, да? Ну например 12754, что надо написать? 1, потому что сумма первых трех цифр равняется 10 и сумма последних трех цифр тоже должна ровняться 10. Так вот оказывается, что счастливых билетов столько же, сколько билетов суммой цифр 27:

$$\underbrace{127}_{\sum=10}\underbrace{541}_{\sum=10}\longleftrightarrow\underbrace{239436}_{\sum=27}$$

45:32

Задача. Вам нужно придумать взаимно-однозначное соответствие, чтобы каждому счастливому сопоставить билет суммой цифр 27 и наоборот.

Я еще раз повторю это для тех, кто будет видео слушать, что есть подозрение, что все-таки не 19, а 18 способов размена, и что где-то вычислениях была сделана ошибка, поэтому ее надо найти каждому отдельно.

Решение. А теперь перейдем к этой задаче есть у кого-нибудь идея или мне показать решение? Идея вот какая: если мы одну из троек, первую или вторую, дополним до 9, то есть вместо этих трех цифр (обведены в квадрат), мы напишем разности:

$$\begin{array}{c}
999 \\
127 \overline{\smash)541} \longleftrightarrow \underline{127458} \\
458 \overline{\smash)5227}
\end{array}$$

Получившийся номер будет иметь сумму цифр 27 согласны? Ну, понятно, смотрите, значит номер 541 с 458 дает три девятки, но сумма цифр 127 и 541 одинаковы, поэтому 127 с 458 тоже дадут три девятки. Ну, в точности наоборот, если вы теперь из трех девяток отнимите 458, естественно, вернетесь к исходному. Поэтому, раз это правило работает и туда и сюда, мы установили взаимное однозначное соответствие. Видите, какие простые и в то же время трудные задачи легко получаются в комбинаторике.

Но и еще один интересный принцип или интересную идею мы рассмотрим, которая связана с различными языками описания комбинаторных объектов.

47:47

5. Различные языки описания комбинаторных объектов. Кстати сказать, если вы будете читать разные учебники, обычно авторы стараются изложить какую-то новую идею на основе предыдущей. Вот они выбирают один язык и на нём всё излагают. На самом деле это для комбинаторики не всегда удобно: иногда чуть-чуть заменишь язык и становится все понятным, прозрачным. Даже есть такая максима, что любое утверждение можно сделать понятным, если его правильно переформулировать. Так что один из способов в чемто разобраться — найти способ представления этой идеи, который бы вам стал очевидным.

Давайте рассмотрим такую задачу: вот у вас допустим празднуются чей-то день рождения и одного человека посылают за пирожными. В магазине 5 видов пирожных, каждого из которых достаточно много, чтобы можно было купить хотя бы все пирожные одного вида. Нам нужно купить 10.

49:26

Задача. Сколькими способами можно купить 10 пирожных, если имеется 5 сортов.

Не буду писать чего 5 сортов и так понятно. Ну вот, если начнете пробовать перебирать, то окажется, что это довольно трудно. Правило умножения здесь не будет работать, потому что порядок не важен. Вот мы купили, например, 10 эклеров, другой вариант — 9 эклеров и какой-нибудь буше, как тут правило перемножения работает? Если бы каждые пирожные покупались независимо, тогда, да. У каждого пирожного было бы пять способов, вы бы эту 10 раз перемножили пятерку, получили бы 5^{10} и сказали: «Вот столько способов». Но потом бы вам кто-нибудь сказал: «Хорошо, вот вы взяли первое пирожное эклер, второе — буше, а теперь наоборот — первое буше, а второе эклер. Это что у вас разные способы получились?» Вы бы сказали: « Ай-ай-ай, я видимо очень много раз посчитал одно и тоже» и стали бы думать как это сделать. Но дело в том, что, если бы вы каждый способ посчитали одинаковое количество раз как разные, потом бы на него просто поделили. Так иногда числа сочетаний сводят к другой задаче, а потом их объединяют все, которые перестановками получились друг из друга. Но тут когда у вас 9 пирожных одного типа, десятое другого понятно, что там переставить их можно 10 способами: вот это последнее пирожное можно первым взять, вторым, третьим, а если у вас, например, 5 пирожных одного, 5 другого, то их уже гораздо больше способов переставить между собой, и причем не очень очевидных. Так что я попробовал как бы побывать в вашей шкуре и посмотреть как можно было бы начать думать. И вот достаточно быстро получается, что задачка не так просто решается. И вот я вам покажу способ поиска другого языка.

Решение. Сейчас мы придумаем код для пирожных: \oslash —не нолик, это будет у меня пирожное, но на самом деле это нолик. А вот это |, которая на самом деле единица, а $\oslash \oslash |$ будут пирожные 1 сорта. Ну, что я там называл, эклеры?



А может стоять вот эта | сразу за $\oslash \oslash |$. Это что значит? Это значит, что буше, например, которое мы считаем вторым по счету, отсутствует в нашей покупке:



Потом может стоять там еще. Какие вы еще пирожные знаете? <u>Ответ из зала</u>: берлинер. О, я даже не знаю:



Так, хорошо, потом сколько у нас сортов то всего... Пять: раз, два, три. Какой ещё сорт назовете? <u>Ответ из зала</u>: тарталетка. Современные студенты знают о пирожных больше предыдущего курса, там они вообще, кроме булочки с маком, ничего не знали:



Хорошо, значит у нас раз, два, три, четыре и еще пятый. Или нет, уже палочек уже хватит, наставил. Четыре палочки уже поставил, значит один сорт пирожного остался. Ну, еще напрягитесь, какие

еще бывают? <u>Ответ из зала</u>: картошка. Хорошо, это хоть я знаю. И сколько пирожных у нас осталось? 4? Чтобы 10 получилось.



Ну вот, согласны, что это один из видов покупки, да? И вот я теперь перейду на язык двоичных наборов: вместо пирожных я поставлю нули, а вместо границ между сортами я поставлю единички:

И получается, что вот этой покупке я присвоил вот такой двоичный код. Но при этом мы договорились, что порядок пирожных он всегда сохранится. Эклер — это сорт номер один, буше — сорт номер два, берлинер — сорт номер три, тарталетка — сорт номер четыре, картошка — сорт номер пять, иначе, конечно, не расшифровать. Ну, давайте я возьму любой другой набор. Из скольки единичек? Из четырех и скольких нулей? Из десяти. Почему единичек 4? Потому что они делят все на 5 частей. Например:

11000100001000

Что это за покупка? Значит ноль эклеров, ноль буше, три берлинера, четыре тарталетки и три картошки. Так что согласны, что это взаимно-однозначное соответствие.

55:52

А теперь задачу я ставлю так: сколько существует бинарных наборов (двоичных наборов) на русском или на нерусском, наборов из четырех единиц и 10 нулей.

Это опять же я решаю, как команду выбираю. Вот у меня есть всего 14 позиций для цифр или 14-значный набор. Я выбираю 4 позиции и ставлю на них единички, а на остальные ставлю 0, и получается C из скольки? C_{14}^4 . 14 мест из них 4 занимают единички. Могу выбрать позиции для нулей

$$C_{14}^4 = C_{14}^{10}$$

Обратите внимание, что опять бесплатно я получил формулу, которую вы, наверное, знаете:

$$C_n^k = C_{n-k}^k$$

Когда вы k человек выбрали, n-k осталось. Вместо того, чтобы выбирать кого я беру в команду, я могу выбирать, кого не беру. Это то же самое.

Кстати сказать, тут могут быть интересные вариации: допустим я предлагаю купить так, чтобы каждого пирожного было бы хотя бы по единичке. Вы можете сказать, что эту задачу мы можем свести к предыдущей. Раз каждого пирожного по единичке мы эти пять пирожных отложим, а потом остальные пять выберем по той же самой формуле. Но можно рассуждать и в лоб, можно сразу попробовать придумать двоичный набор. Вот у нас 10 пирожных:

$$\oslash$$
. \oslash . \oslash . \bigcirc

И теперь поскольку я знаю, что хотя бы одно пирожное каждого сорта у меня должно быть в наборе, я должен поставить вот между ними на 4 выбранных мною места (где точки) палочки. То есть теперь у меня уже палочки две подряд идти не могут как раньше, они могут идти только не более чем или не менее чем через одну:

$$\emptyset. \emptyset. \emptyset | \emptyset | \emptyset. \emptyset | \emptyset. \emptyset. \emptyset | \emptyset$$

Вот четыре палочки поставил, что я купил там... Три эклера, один буше там и так далее сколькими способами тогда тут можно? Из этих точечек, которых сколько? 9. Я должен выбрать 4, на которые я должен поставить разделитель — C_9^4 . Вот, пожалуйста вам другой способ решения другой, правда, задачи, но тоже такой же прием, когда мы берем придумываем двоичный язык двоичных кодов. Ну, конечно, здесь идея похожа. Можно сказать что мы просто перешли к другой задаче.

59:33

И последнее я хочу вам тоже по рекламировать другую книгу Кнута. Она называется — "Конкретная математика". Там Грэхэм, Поташник и Кнут — три автора, и по ней тоже можно взять потом альтернативный экзамен, поэтому можете глянуть, посмотреть, если вам она интересна, то взять. Так вот там есть ещё одна идея, которую Кнут очень красиво в книжке обыгрывает. Значит вот какую

ситуацию он рассматривает, он рассматривает цепочки из домино, да, но здесь я назову как этот, как это красивая идея.

60:22

6. От комбинаторики перейти к алгебре. Значит он рассматривает такую идею: вот у нас есть домино. Я строю из них цепочку. Домино так устроено, что у него высота в два раза больше ширины:

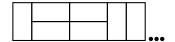


Рис. 3: Пример последовательности доминошек

Но я могу полосочки класть либо две штучки, либо одну ставить вертикально, и вот я хочу изучить множество этих цепочек. Вот Кнут упрямо даже не вводит никаких переменных, а прямо начинает с этими цепочками совершать алгебраические операции. То есть он вводит операции прямо на множестве вот таких картинок. Я могу пояснить почему. Может вы слышали о такой математическая системе верстки под называнием ТеХ. Так вот Кнут он автор не только вот таких алгоритмов и книг «Искусство программирования», но он также сделал вот эту систему TeX и он показывает, что в его системе он вместо букв может оперировать вот домино, например. И вот что говорит: «Давайте мы все эти цепочки разделим на две части». Вот у меня может цепочка начинаться с вертикальной, и тогда символ умножения означает что я ее приклеиваю. Но на самом деле этот термин будет конкатенацией у нас потом, когда мы языки будем изучать. И тогда вот эта, например, цепочка, которую я конкретно нарисовал: от нее останется вот такое продолжение, а первый элемент я отделю:

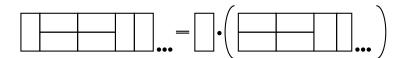


Рис. 4: Пример умножения доминошек

Мы рассматриваем множество всех цепочек. Поскольку я рассматриваю все цепочки, значит он соединяет знаком плюс первую цепочку, вторую цепочку и так далее. Следовательно, у нас какая-то цепочка могла начинаться и с двух доминошек, которые лежат горизонтально:

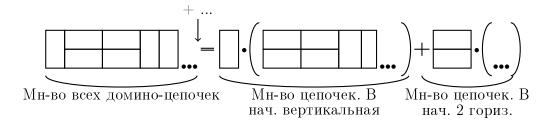


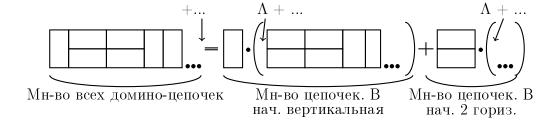
Рис. 5: Множество всех цепочек

Как вы считаете, равенство правильно написано или неправильно? Тут есть один такой фокус, может не все его заметили:

Вот это вот цепочка или не цепочка? Цепочка и цепочкой могут быть две горизонтальные. Если в левой части у меня все множество цепочек, и я из него выношу одну вертикальную доминошку, то у меня в каком-то смысле, в каком-то случае вместо неё ничего не останется. Но, если я просто ничего не буду писать, это как бы приведет потери вот этой комбинации. Значит мне нужно ввести какой-то символ, к примеру, Λ , который будет означать пустоту.

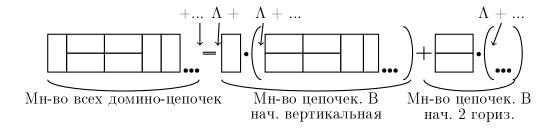
Λ — пустая цепочка

То есть это не пустое множество, а именно вот цепочка, к которой можно приклеить другое домино, и получится цепочка из одного домино, поэтому на самом деле в первую и вторую скобки нужно вставить вот такой символ, тогда у меня действительно все цепочки будут учтены:



Кроме опять же в левой части уравнения из домино у меня пустая цепочка есть, а в правой у меня пустой цепочки нет, потому что у

меня начинается сразу с вертикальной доминошки, поэтому мне и тут нужно добавить эту пустую цепочку:



Вот получается такая интересная арифметика. Значит еще раз, что в этой арифметике есть: плюс — это объединение. Я перечисляю все комбинации, и, если пишу плюсик, значит эти выражения рассматриваются вместе, умножить — это склеивание двух цепочек. Вот в следующем семестре опять же у вас будут регулярные выражения так вот это на самом деле регулярные выражения для описания множеств.

65:58

Ну а дальше, чтобы, может быть, было привычней, я уже уйду от языка домино и напишу это в более привычном алгебраическом виде. Вместо вертикальной доминошки мы будем писать x, вместо двух горизонтальных — y, а вместо Λ я буду писать 1. Понятно, почему один. Потому что при приклеивании, если бы я его назвал нулем, естественно 0 умножить на что-то будет 0, а 1 умножить на что-то будет что-то. Поэтому тут она больше похожа на единицу своим действием, чем на ноль. Тогда у меня получается, если я все множество вот это цепочек обозначу допустим за S, то у меня получится такое интересное равенство:

$$S = 1 + x...$$

А потом у меня в этой скобке опять идет то же самое множество. Похоже, когда мы с вами рассматривали периодические непрерывные дроби, то одно было частью другого, одна конструкция была под конструкцией ее же самой:

$$S = 1 + x \cdot S + y \cdot S$$

Если я перенесу все с S в одну часть, я могу написать теперь вводя минусы, хотя непонятно, что значит минус, что значит вычитать

домино. Как бы смысла в этом нет, это как мнимая единица получается:

$$S \cdot (1 - x - y) = 1 \implies S = \frac{1}{1 - x - y}$$

Посмотрите, если я это перепишу в терминах домино, получится очень забавно:

$$\Lambda + igcup + igcu$$

Вот такая какая-то совершенно абстрактная игра с символами, но оказывается она очень полезная. И вот когда я пишу $\frac{1}{1-x-y}$ это вас не удивляет и скорее всего не удивляет вот почему, потому что вы знаете, что такое сумма бесконечно убывающей геометрической прогрессии. Если вы напишете это вот так:

$$\frac{1}{1 - (x + y)} = 1 + x + y + (x + y)^{2} + (x + y)^{3} + \dots$$

Но опять же в терминах домино получается, что вот исходную последовательность, множество всех цепочек, я переписал вот в таком виде:

$$\Lambda +$$
 $+$

А что такое в квадрате здесь уже можно понять: вертикальную на вертикальную, это значит на себя умножить. Вертикальную уможить на горизонтальные будет цепочкой вот из двух частей таких:



Горизонтальные умножить на вертикальную будет то же самое, но вертикальная будет стоять сзади. Две горизонтальные в квадрате — два раза будет стоять две горизонтальные. Получается, что вот

мы множество всех цепочек оказывается переписали, переструктурировав, порядок. Так вот эта идея называется «Производящая функция». Мы с ней будем целую лекцию возиться, потому что она одна из самых употребляемых для комбинаторики в научных кругах. Вот мы с вами сейчас написали то, что потом будет называться производящими функциями. Вместо того чтобы работать с комбинаторными объектами, будем работать с такими, с алгебраическими.

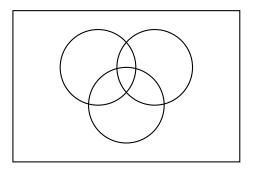
Но и последнее, что мы сегодня пройдём, это седьмое, да? Называется метод включения-исключения.

70:43

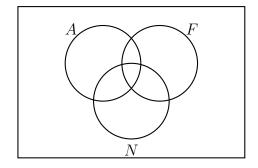
7. *Принцип включения-исключения*. Inclusion-exclusion principle, так он по-английски звучит. Мне очень удобно начать с задачки, но эту задачку мы должны сочинить вместе иначе трудно ее просто из головы извлечь.

71:51

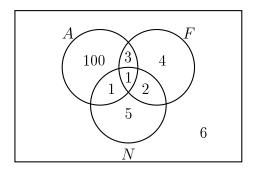
Задача. Я нарисую так называемую диаграмму Эйлера, может быть, вы уже с ней сталкивались:



Круги — это подмножество. Вот есть прямоугольник — это, как бы сказать, самое объемлющее множество, пространство и в нем есть такие подмножества: A — это будут те, кто изучает английский язык, допустим. F — это французский язык, N — это немецкий:



Тогда понятно, что множество за пределами кругов те, кто вообще этих языков не изучает, на пересечении всех множеств те, кто изучает три языка, ну, и так далее. Давайте чего-нибудь придумаем. Допустим есть всего один человек, который изучает три языка, три человека изучают немецкий и французский, два человека изучают английский и французский, четыре человека изучают английский и французский, сто человек изучают только английский, четыре человека изучают только французский, пять человек изучают только немецкий и 6 человек изучают китайский (то есть не изучают эти три языка):



Ну вот, а теперь давайте посчитаем сколько у нас всего. Вот я не хорошо N взял, потому что я обычно N означаю общее число. Ну, придется общее число обозначать за M значит сколько всего человек то получилось... 122 что ли? Правильно? M=122. Теперь задачка будет такая: вот здесь изучает английский язык, ну допустим, M_A напишу, 105 человек $M_A=105$. Здесь уже считают всех, кто изучает. Французский язык изучают 10 человек $M_F=10$, немецкий язык изучают 9 человек $M_N=9$, теперь английский и французский изучают 4 человека $M_{AF}=4$, английский и немецкий изучают 3 человека $M_{FN}=3$. Ну и наконец все три языка изучает 1 человек $M_{AFN}=1$. И вот мне нужно написать формулу, она и будет называться формула включения-исключения, называется M_0 — это сколько человек не изучает ни один из этих трех языков.

75:59

Решение. Ну вот давайте попробуем написать такую формулу. Значит я могу действовать вот так: я могу взять всех людей и вычесть оттуда тех, кто изучает английский, тех, кто изучает французский, и тех, кто изучает немецкий:

$$M(0) = M - M_A - M_F - M_N + \dots$$

Правильно я поступлю? Не очень, потому что я некоторых исключу по два раза, а одного человека который изучает 3 языка я даже трижды исключу. Но мне сразу все скомпенсировать трудно, давайте сначала скомпенсируем тех, кого я дважды вычитаю, кто знает два языка. Раз я их по два раза вычел. Допустим английский и немецкий его вычел и в M_A , и в M_N , значит я его должен прибавить:

$$M(0) = M - M_A - M_F - M_N + M_{AF} + M_{AN} + M_{FN} + \dots$$

Теперь все нормально или нет? Теперь с теми, кто знает ровно два языка, всё хорошо: я их сначала два раза вычел, а потом один раз добавил. Но человека, который знает 3 языка, я сначала три раза его убрал, а потом три раза добавил. Что нам теперь сделать? Один раз вычесть, потому что мне его считать не надо:

$$M(0) = M - M_A - M_F - M_N + M_{AF} + M_{AN} + M_{FN} - M_{AFN}$$

Вот это получилось формула включений-исключений для трех элементов. Теперь давайте ее обобщим на общий случай.

Теперь мы будем говорить о признаках. Допустим у нас есть M_1 элементов, обладающих первым признаком:

$$M_1$$
 — число элементов, обладающих первым признаком

Понятно что тут опять два языка: один язык логический, с которым мы будем с вами в следующем семестре, а другой язык теории множеств этого вот то, чем мы сейчас занимаемся. Так вот можно говорить, что элемент принадлежит к множеству А. Это я в терминах языка теории множеств говорю, а могу говорить, что он обладает признаком А, то есть знает английский язык. Это уже логическое: знает или не знает, да или нет. Поэтому я могу вот здесь и на таком языке формулировать:

 M_2 — число элементов, обладающих вторым признаком

:

 M_n — число элементов, обладающих n-ым признаком

Теперь я могу рассматривать их комбинации, но я уже напишу в общем случае. $M_{i_1i_2...i_k}$ — это как бы пересечение множеств, соответствующих M_{i_1} , M_{i_2} , и так далее до M_{i_k} , а в терминах свойств — это будет число элементов, обладающих признаками с этими номерами (признаками $i_1, i_2, ..., i_k$). Теперь давайте попробуем написать общую формулу, она и будет называться формулой включения-исключения. Значит количество элементов, не обладающих ни одним из признаков, (M(0)) будет равно все элементы вычесть все элементы, обладающие одним признаком:

$$M(0) = M - \sum_{i=1}^{n} M_i + \dots$$

Но потом по тем причинам, что я слишком много вычел, я добавлю все попарные. И вот здесь обратите внимание, когда добавляю попарные, я же не добавляю там, английский, французский и французский, английский. Поэтому мне эти признаки надо упорядочить и использовать только один раз. Английский, французский, а французский английский я уже не рассматриваю это обычно бывает, значит, берут в порядке возрастания номеров, и тогда просто пишут i < j, ну, а дальше я не пишу каждую из них от 1 до n:

$$M(0) = M - \sum_{i=1}^{n} M_i + \sum_{i < j} M_{ij} \dots$$

Но и понятно, в общем случае они чередуются. У меня будет -1 в степени k, наверное, да? Потом проверим:

$$M(0) = M - \sum_{i=1}^{n} M_i + \sum_{i < j} M_{ij} \dots (-1)^k \sum_{i_1 < \dots < i_k} M_{i_1 \dots i_k} \dots$$

Все эти признаки идут в порядке возрастания. Ну, давайте проверим, если k=1, то получается минус, если k=2, то плюс. Значит все правильно со знаками. И как выглядит последний элемент? Он на самом деле всего один, потому что там одна всего существует ситуация, когда все признаки есть от 1 до n, но в зависимости от четности, нечетности она может оказаться с минусом или с плюсом:

$$M(0) = M - \sum_{i=1}^{n} M_i + \sum_{i < j} M_{ij} \dots (-1)^k \sum_{i_1 < \dots < i_k} M_{i_1 \dots i_k} \dots (-1)^n M_{123 \dots n}$$

[Скорее всего в последнем слагаемом имелось в виду $i_1i_2...i_n$]. Поскольку у меня время заканчивается, я хочу задать вам одну задачку на метод включений-исключений дополнительно, необязательно, и кто хочет, тот решит и покажет мне, выложив на свой сайт.

82:45

Задача (Задача о шляпах). N джентльменов в клубе оставили в гардеробе N шляп, а при выходе каждый надел случайную шляпу. Какова вероятность, что никто не надел свою шляпу?

Ну, раз появилось слово вероятность, я должен связать это с термином комбинаторика. Ну все просто, если эти джентльмены могли одеть шляпу N способами и среди них $N_{\text{благоприятная}}$, вряд ли для мужчин благоприятно уйти не в свой шляпе, но с точки зрения комбинаторики это вот те варианты, когда никто из них не в своей шляпе. Вот вы делите одно на другое, получается вероятность того, что никто из них не ушел в своей шляпе:

$$P = \frac{N_{\text{благоприятная}}}{N}$$