Тема 1.1. Подсчёт комбинаторных объектов: перестановки, сочетания и размещения

На этой лекции речь пойдет о подсчёте комбинаторных объектов. Чтобы было понятно, что это такое, сразу начнём с задачи.

> **Трёхзначные числа**
> Сколько существует трёхзначных чисел, в записи которых цифры 1, 2 и 3 встречаются ровно по одному разу?
>
> www.problems.ru, №30328

В этой задаче можно просто перебрать варианты и получить ответ 6 (см. рис. 1).

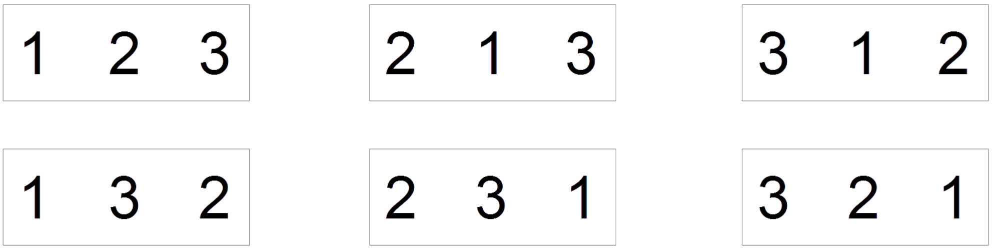


Рис. 1

Тем не менее, если мы хотим строго обосновать ответ, то нужно доказать, что мы перебрали все возможные варианты и других нет. Для этого посмотрим, как эти трёхзначные числа строятся. На первом месте в числе может стоять одна из трёх цифр: 1, 2 или 3. При фиксированной первой цифре, на втором месте может стоять одна из двух оставшихся цифр. Таким образом, числа распадаются на три группы по первой цифре, каждая группа содержит два варианта по второй цифре (см. рис. 2). Поэтому общее количество способов поставить первые две цифры получается как произведение: $3 \cdot 2 = 6$.

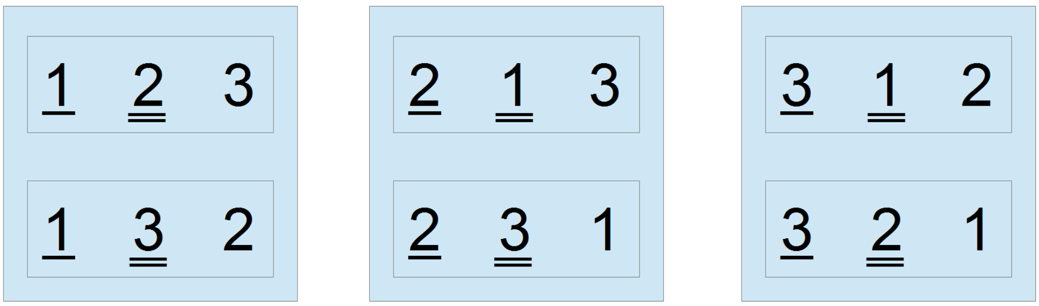


Рис. 2

Заметим, что при фиксированных двух первых цифрах последняя цифра определяется единственным способом. Поэтому ответ на задачу будет $3 \cdot 2 \cdot 1 = 6$.

Это частный случай задачи о подсчёте числа перестановок.

Перестановка из n элементов — это упорядоченный набор чисел от 1 до n.

Приведём несколько примеров перестановок для $n = 5$:

$$(1, 3, 5, 4, 2)$$
$$(2, 3, 1, 5, 4)$$
$$(3, 1, 2, 4, 5)$$

Посчитаем, сколько всего существует различных перестановок из n элементов при заданном n. Будем рассуждать так же, как в задаче про трёхзначные числа. На первом месте в перестановке может стоять один из n элементов, на втором месте — один из оставшихся, т.е. $(n-1)$ вариант, на третьем — $(n-2)$ варианта при фиксированных первых двух элементах и т.д. В результате получаем, что количество перестановок из n элементов (обозначим его P_n) вычисляется по формуле

$$P_n = n \cdot (n-1) \cdot (n-2) \cdot 2 \cdot 1 = 1 \cdot 2 \cdot \ldots \cdot n.$$

Произведение чисел от 1 до n часто встречается в комбинаторике и обозначается $n!$ (читается «n факториал»). Таким образом, число различных перестановок из n элементов равно $n!$ Например, для $n = 3$ получаем

$$P_3 = 3! = 6.$$

Для $n = 1$ произведение $1!$ состоит из одного сомножителя 1 и равно 1, то есть существует только одна перестановка из одного элемента. Также обычно считают, что $0! = 1$, то есть пустая перестановка из нуля элементов тоже одна. При решении задач всегда имеет смысл обращать внимание на крайние случаи. Мы видим, что здесь в крайних случаях общая формула работает.

Теперь задача поинтереснее.

Людоед 1

У людоеда в подвале томятся 25 пленников. Сколькими способами он может выбрать трёх из них, чтобы съесть по одному на завтрак, на обед и на ужин?

www.problems.ru, №104046

Давайте рассуждать. Выбрать человека на завтрак у нас 25 способов. После этого на обед можно выбрать одного из 24 оставшихся людей, и на ужин — одного из 23 оставшихся. Получаем ответ: $25 \cdot 24 \cdot 23 = 13800$ способов.

Немного изменим вопрос задачи.

Людоед 2
У людоеда в подвале томятся 25 пленников. Сколькими способами людоед может выбрать трёх человек, чтобы отпустить их на свободу?

www.problems.ru, №104046

Теперь порядок людей не важен, поэтому выбирать их по одному, как в прошлой задаче, будет неправильно. Посмотрим на задачу с другой стороны. Предположим, мы выбрали трёх человек, чтобы отпустить их на свободу (пусть это будут люди с номерами, например, $(7, 13, 22)$), а наш людоед передумал и решил их съесть по очереди. Подумаем, сколько существует возможных порядков съесть трёх выбранных людей. Ясно, что это наша первая задача о подсчёте числа перестановок из трёх элементов, оно равно 6. Каждому способу отпустить людей на свободу соответствует 6 способов их съесть (см. рис. 3).

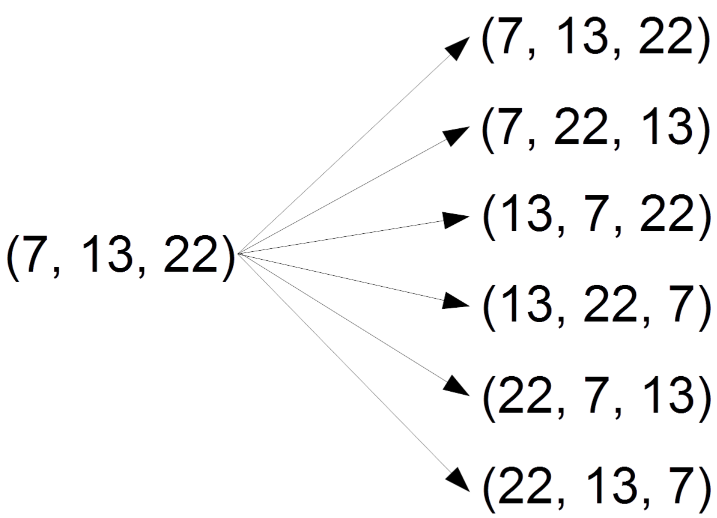


Рис. 3

Пусть x — ответ на задачу «Людоед 1», а y — ответ на задачу «Людоед 2». Тогда $x = 6y$. Мы уже знаем x, поэтому выразим y: $y = \frac{x}{6}$. Получаем ответ $\frac{13800}{6} = 2300$.

Рассмотрим теперь задачу в общем виде.

Людоед 3
У людоеда в подвале n пленников.

1. Сколькими способами он может выбрать k из них, чтобы съесть их по очереди. Порядок важен.
2. Сколькими способами он может выбрать k человек, чтобы

отпустить их на свободу?

Эта задача уже больше похожа на олимпиадную задачу по информатике. Нужен алгоритм, чтобы посчитать ответ для произвольных n и k.

Ответим сначала на первый вопрос. Первого человека можно выбрать n способами, второго — $(n-1)$ способом, и т.д. Получаем произведение

$$n \cdot (n-1) \cdot \ldots \cdot (n-k+1).$$

Всего в нем должно быть k множителей, поэтому нетрудно понять, что последний будет $(n-k+1)$. Приверим: при $k=3$ будет $n(n-1)(n-2)$. Полученное число называется числом размещений и обозначается A_n^k (читается «A из n по k»):

$$A_n^k = n \cdot (n-1) \cdot \ldots \cdot (n-k+1).$$

Размещение из n по k — это упорядоченный набор из k различных элементов n-элементного множества.

Как правило, рассматривается множество натуральных чисел от 1 до n. *Упорядоченный* набор означает, что порядок элементов в нем важен. Например, $\langle 1,3,4\rangle$ и $\langle 3,4,1\rangle$ — это разные размещения из 4 по 3, хотя они состоят из одних и тех же элементов. Вопрос задачи «сколькими способами людоед может выбрать k из n человек, чтобы съесть их по порядку» с формальной точки зрения означает «найдите число размещений из n по k».

Формулу для подсчёта размещений можно также записать следующем образом

$$A_n^k = \frac{n!}{(n-k)!}.$$

Действительно, произведение чисел от 1 до $(n-k)$ в числителе и знаменателе сокращается:

$$A_n^k = \frac{n!}{(n-k)!} = \frac{\cancel{1 \cdot 2 \cdot \ldots \cdot (n-k)} \cdot (n-k+1) \cdot \ldots \cdot n}{\cancel{1 \cdot 2 \cdot \ldots \cdot (n-k)}} = \\ = n \cdot (n-1) \cdot \cdots \cdot (n-k+1).$$

Заметим, что в этой задаче $k \leqslant n$. Если $k=n$, то получаем

$$A_n^n = \frac{n!}{0!} = n!$$

Это означает, что если людоед планирует есть всех n людей, то количество способов это сделать равно числу перестановок из n элементов.

Чтобы получить ответ на второй вопрос задачи, нужно действовать так же, как в рассмотренном частном случае (задаче «Людоед 2»). А именно, нужно полученный результат A_n^k поделить на количество перестановок из k элементов:

$$\frac{A_n^k}{P_k} = \frac{n!}{k!(n-k)!} = C_n^k.$$

Это число сочетаний, которое обозначается C_n^k (читается «C из n по k»).

Сочетание из n по k — это k-элементное подмножество n-элементного множества.

Подмножество — неупорядоченный набор. Это означает, что порядок элементов не важен и сочетания, различающиеся только порядком элементов, считаются одинаковыми. Например, сочетания $\{1, 3, 4\}$ и $\{3, 4, 1\}$ совпадают. Поэтому элементы в сочетании обычно записывают в порядке возрастания. В задаче нужно найти количество k-элементных подмножеств множества из n человек, т.е. число сочетаний из n по k.

Тема 1.2. Задачи

Упражнение 1.1
Имеется забор из n досок. Сколькими способами можно его покрасить красками k цветов? Каждую доску нужно красить целиком в один цвет. Каждый цвет разрешается использовать только один раз. Не обязательно использовать все цвета.

Например, при $n = 2$, $k = 3$ все возможные варианты покраски забора представлены на рис. 4.

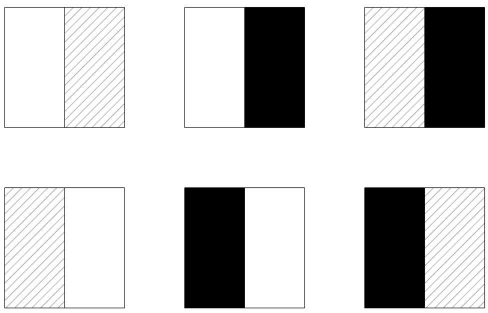


Рис. 4

По рисунку видно, что мы выбираем набор из множества цветов, а не из множества досок. Причем это упорядоченный набор, поэтому речь идет о размещениях.

Ответ на задачу будет A_k^n. Заметим, что задача имеет решение только при $k \geqslant n$. Иначе нам не хватит цветов, чтобы покрасить забор. Теперь немного изменим задачу.

Упражнение 1.2

Имеется забор из n досок. Сколькими способами можно его покрасить красками k цветов? Каждую доску нужно красить целиком в один цвет. Каждый цвет разрешается использовать любое число раз. Не обязательно использовать все цвета. Дайте ответ на задачу при $n = 4$, $k = 3$.

В этой задаче мы имеем размещения с повторениями — каждый цвет может повторяться любое число раз. Первую доску можно покрасить в k возможных цветов, вторую — тоже в k цветов, и т.д. Значит, всего будет k^n способов покрасить забор из n досок.

Упражнение 1.3

Сколькими способами можно выложить в ряд 5 красных, 5 синих и 5 зелёных шаров так, чтобы никакие два синих шара не лежали рядом?

Сначала найдём число способов выложить в ряд 5 красных и 5 зелёных шаров.

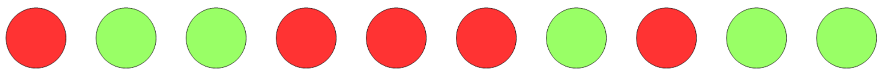


Рис. 5

Нужно выбрать среди 10 позиций 5 позиций для красных шаров, после этого позиции зелёных шаров определятся однозначно. Получится C_{10}^5 способов. Затем нужно добавить синие шары.

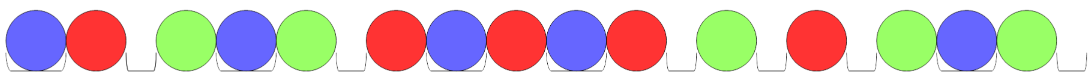


Рис. 6

Их можно размещать в позициях между красными и зелёными шарами, как показано на рис. 6, не более одного синего шара в каждую позицию. Свободных позиций 11, а синих шаров — 5, получаем C_{11}^5 способов. Итоговым ответом будет произведение $C_{10}^5 C_{11}^5$.

Упражнение 1.4

Имеется 10 юношей и 10 девушек. Сколько существует способов объединить их в пары (юноша с девушкой)?

Упражнение 1.5
Сколько существует способов разбить на пары компанию из 20 человек?

Начнём с задачи про юношей и девушек. Пронумеруем и тех, и других числами от 1 до 10. Представим, что юноши стоят в ряд по порядку. и нужно расставить перед ними девушек. Каждый способ будет соответствовать перестановке из 10 элементов. Таким образом, ответ на первую задачу — 10!

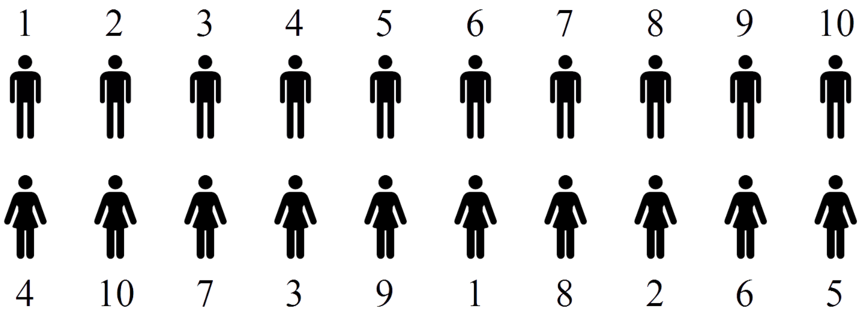


Рис. 7

Перейдём ко второй задаче. Можно попробовать свести её к первой, но есть и более простое решение. Посмотрим на человека с номером 1. Ему в пару можно определить одного из остальных 19.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Рис. 8

После вычеркивания этих двух людей, посмотрим на человека с наименьшим номером. Ему в пару можно дать одного из 17.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Рис. 9

Вычеркнем и их и продолжим процесс. У нас получится произведение нечётных чисел: $19 \cdot 17 \cdot 15 \cdot \ldots \cdot 3 \cdot 1$. Такое произведение обозначается 19!!

Упражнение 1.6
Имеется 10 юношей и 10 девушек. Сколько существует способов составить из них компанию, в которую входило бы одинаковое количество юношей и девушек? Пустое множество не считается.

www.problems.ru, №35399

В этой задаче можно перебрать число юношей в компании. Обозначим его через k. Число способов выбрать k юношей из 10 равно C_{10}^k, так же для девушек. Число способов составить компанию из k юношей и k девушек — $(C_{10}^k)^2$. Получаем ответ:

$$(C_{10}^1)^2 + (C_{10}^2)^2 + (C_{10}^3)^2 + \ldots + (C_{10}^{10})^2.$$

Вычисления по этой формуле, особенно вручную, довольно трудоёмки.

У этой задачи существует принципиально другое решение. Составим какую-нибудь компанию из одинакового числа юношей и девушек. Рассмотрим подмножество юношей, которые входят в эту компанию, и девушек, которые не входят в эту компанию. Их всего будет 10. Действительно, будет k юношей и $10 - k$ девушек. Мы установили соответствие между компаниями из одинакового числа юношей и девушек и подмножествами произвольно выбранных 10 человек из 20.

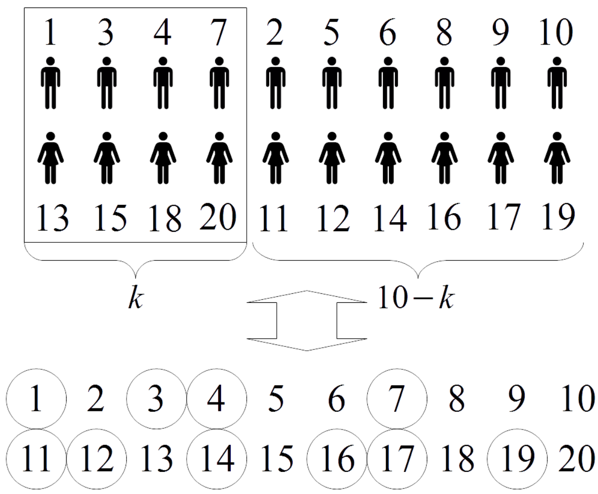


Рис. 10

Нетрудно видеть, что это соответствие — взаимно-однозначное. Каждой компании соответствует одно подмножество, и каждому подмножеству соответствует одна компания. Значит, для подсчета интересующих нас компаний можно просто найти число способов выбрать 10 человек из 20. Оно равно C_{20}^{10}. Нужно вычесть 1, чтобы не считать пустую компанию. Она будет соответствовать подмножеству из одних девушек. Таким образом, ответ: $C_{20}^{10} - 1$.

Метод построения взаимно-однозначного соответствия между комбинаторными объектами бывает очень полезен. Он позволяет свести задачу подсчёта одних комбинаторных объектов к другим, более простым.

Это были задачи на сочетания и размещения. Обсудим теперь, как посчитать сочетания с повторениями.

Сочетание с повторениями — это неупорядоченный набор из k элементов n-элементного множества, причем каждый элемент в наборе может встречаться произвольное число раз.

Примеры сочетаний с повторениями:

$$\{1, 1, 3, 4, 4, 4\}$$
$$\{1, 2, 2, 5\}$$
$$\{3, 3, 3, 3, 4\}$$
$$\{1, 2, 3, 3, 4, 5, 5\}$$

Порядок чисел в сочетании неважен, поэтому обычно используется порядок по возрастанию.

Упражнение 1.7
Найдите число сочетаний с повторениями из 3 по 2.

Выведем общую формулу для числа сочетаний с повторениями из n по k. Мы построим взаимно-однозначное соответствие между ними и обычными сочетаниями.

Рассмотрим последовательность белых и чёрных шаров длины $n + k - 1$ в которой ровно $n - 1$ шар чёрного цвета (см. рис. 11). Очевидно, число таких последовательностей C_{n+k-1}^{n-1}.

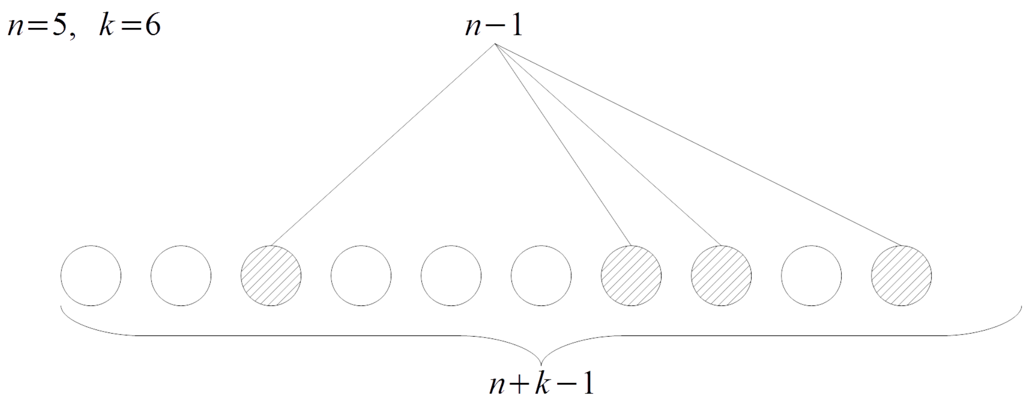


Рис. 11

Заметим, что чёрные шары разбивают белые на n отрезков. Возможно, некоторые из этих отрезков пустые: если два чёрных шара стоят подряд или если чёрный шар стоит в начале или в конце (см. рис. 12). Напишем на шарах первого отрезка число 1, на шарах второго отрезка — число 2, и т.д.

Всего белых шаров k, поэтому мы получим сочетание с повторениями из n по k: $\{1, 1, 2, 2, 2, 4\}$ для примера на рис. 12. Нетрудно показать, что это соответствие взаимно-однозначное, т.е. по последовательности шаров можно однозначно построить сочетание с повторениями (что мы уже проделали) и, наоборот, по сочетанию с повторениями можно однозначно построить последовательность шаров. Количества этих комбинаторных объектов равны между собой, и мы получили, что число сочетаний с повторениями из n по k равно C_{n+k-1}^{n-1}.

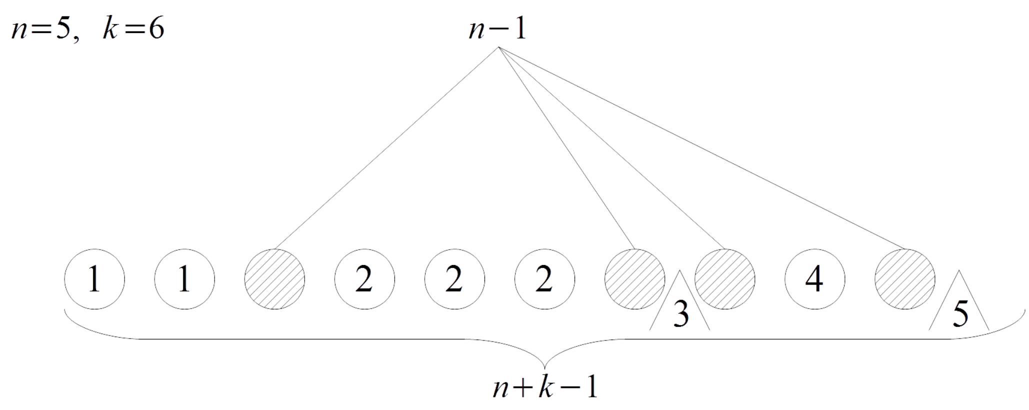


Рис. 12

Тема 1.3. Бином Ньютона

На этой лекции мы более подробно поговорим про количества сочетаний C_n^k. Эти числа еще называются **биномиальными коэффициентами**, потому что они возникают в формуле **бинома Ньютона**:

$$(a+b)^n = C_n^0 a^n b^0 + C_n^1 a^{n-1} b^1 + C_n^2 a^{n-2} b^2 + \ldots + C_n^k a^{n-k} b^k + \ldots + C_n^n a^0 b^n. \quad (1)$$

Формула (1) обобщает хорошо известные школьные формулы для квадрата суммы и куба суммы:

$$(a+b)^2 = a^2 + 2ab + b^2, \quad (a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3.$$

И для квадрата, и для куба в правой части получаются суммы, состоящие из произведений чисел a и b в некоторых степенях, причём суммарная степень a и b в каждом слагаемом равна 2 для квадрата и 3 для куба. Аналогичная формула должна быть справедлива и для n-й степени. При возведении в n-ю степень должны получаться произведения степеней a и b и перед ними будут стоять некоторые коэффициенты. В формулах квадрата и куба это числа 1, 2, 1 и 1, 3, 3, 1, соответственно. Оказывается, что для n-й степени это будут в точности количества сочетаний C_n^k. Почему так получается?

Возведение в n-ю степень означает перемножение n одинаковых скобок:

$$\underbrace{(a+b)\cdot(a+b)\cdot(a+b)\cdot\ldots\cdot(a+b)}_{n}$$

Если мы раскроем скобки по обычным правилам, то у нас получится сумма. Каждое слагаемое будет произведением n чисел a и b. Например, мы возьмём a из первой скобки, b из второй скобки, снова b из третьей скобки и т.д. Это будет одно слагаемое. Другие слагаемые строятся по тому же принципу. И так нужно перебрать все варианты. Зададим себе вопрос: сколько слагаемых будет включать k множителей b и $n-k$ множителей a? Это будет количество способов выбрать из n скобок k таких, из которых мы возьмём b. Это количество

— число сочетаний из n по k. Поэтому мы и получаем, что в формуле бинома Ньютона перед $a^{n-k}b^k$ стоит коэффициент C_n^k:

$$(a+b)^n = C_n^0 a^n b^0 + C_n^1 a^{n-1} b^1 + C_n^2 a^{n-2} b^2 + \ldots + \boxed{C_n^k a^{n-k} b^k} + \ldots + C_n^n a^0 b^n.$$

Упражнение 1.8
Выведите формулу для $(a+b)^4$. Чему равен коэффициент при ab^3?

Изучим некоторые свойства биномиальных коэффициентов. Первое свойство — это симметричность:

$$C_n^k = C_n^{n-k}. \tag{2}$$

Это свойство легко доказать, если представить биномиальные коэффициенты через факториалы:

$$C_n^k = \frac{n!}{k!(n-k)!},$$
$$C_n^{n-k} = \frac{n!}{(n-k)!(n-(n-k))!} = \frac{n!}{(n-k)!k!}.$$

Любопытно, что формула (2) имеет и чисто комбинаторную интерпретацию.

Рассмотрим задачу. Сколько существует последовательностей длины n, состоящих из нулей и единиц, содержащих ровно k единиц? Ответ на эту задачу C_n^k — число способов выбрать k позиций среди n, чтобы поставить туда единицы. С другой стороны, число последовательностей, содержащих ровно k единиц, равно числу последовательностей, содержащих ровно $n-k$ нулей. И это число равно C_n^{n-k}.

$$\underbrace{0\,1\,1\,0\,0\,0\,1\,1}_{n} \qquad k \text{ единиц, } n-k \text{ нулей}$$

Рассмотрим другое интересное свойство биномиальных коэффицентов:

$$C_n^0 + C_n^1 + C_n^2 + \ldots + C_n^k + \ldots + C_n^n = 2^n.$$

Это свойство тоже можно доказать разными способами. Первый способ — подстановка $a=1$, $b=1$ в бином Ньютона (1):

$$(1+1)^n = C_n^0 + C_n^1 + C_n^2 + \ldots + C_n^k + \ldots + C_n^n.$$

Второй способ — комбинаторный. Посчитаем число всех возможных последовательностей из 0 и 1 длины n. Это число размещений с повторениями. Как будто мы красим забор из n досок в два цвета. Получается, что всего последовательностей будет 2^n. C_n^k — это количество последовательностей, содержащих k единиц. Если мы просуммируем эти количества по всем k от 0 до n, то как раз получим общее число последовательностей 2^n.

Упражнение 1.9
Вычислите сумму

$$C_n^0 - C_n^1 + C_n^2 - C_n^3 + \cdots + (-1)^{n-1}C_n^{n-1} + (-1)^n C_n^n$$

при $n = 10$.

www.problems.ru, №30712

Докажем ещё одно свойство:

$$C_n^k = C_{n-1}^k + C_{n-1}^{k-1}. \tag{3}$$

Формулу (3) удобно использовать для вычисления биномиальных коэффициентов. Её можно доказать, представляя биномиальные коэффициенты через факториалы. Попробуйте проделать это самостоятельно. Мы сразу разберём комбинаторное доказательство.

Снова начнём с того, что C_n^k — это число последовательностей нулей и единиц длины n, содержащих k единиц. Последний элемент последовательности может быть равен либо 0, либо 1. Число последовательностей, оканчивающихся на 0, равно C_{n-1}^k, потому что фактически все k единиц стоят в первых $n-1$ позиции.

$$\underbrace{\overbrace{0\quad 1\quad 1\quad 0\quad 0\quad 0\quad 1}^{n-1}\quad \boxed{0}}_{n}$$

Рис. 13

Каждая последовательность, заканчивающаяся на 1, соответствует последовательности длины $n-1$ с $k-1$ единицей. Число таких последовательностей — C_{n-1}^{k-1}.

$$\underbrace{\overbrace{0\quad 0\quad 1\quad 0\quad 0\quad 0\quad 1}^{n-1}\quad \boxed{1}}_{n}$$

Рис. 14

Складывая полученные биномиальные коэффициенты, приходим к формуле (3).

Благодаря этому свойству, биномиальные коэффициенты удобно записывать в следующую таблицу, называемую **треугольником Паскаля** (см. рис. 15).

$$
\begin{array}{ccccccccc}
 & & & & C_0^0 & & & & \\
 & & & C_1^0 & & C_1^1 & & & \\
 & & C_2^0 & & C_2^1 & & C_2^2 & & \\
 & C_3^0 & & C_3^1 & & C_3^2 & & C_3^3 & \\
C_4^0 & & C_4^1 & & C_4^2 & & C_4^3 & & C_4^4 \\
 & & & & \dots & & & &
\end{array}
\qquad\qquad
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & 1 & & 1 & & & \\
 & & 1 & & 2 & & 1 & & \\
 & 1 & & 3 & & 3 & & 1 & \\
1 & & 4 & & 6 & & 4 & & 1 \\
 & & & & \dots & & & &
\end{array}
$$

Рис. 15

Каждый биномиальный коэффициент получается как сумма двух, расположенных над ним. Крайние элементы в строках равны единице.

Треугольник Паскаля удобно использовать для вычисления биномиальных коэффициентов в программе. Формула через факториалы

$$C_n^k = \frac{n!}{k!(n-k)!}$$

обычно не так удобна. Во-первых, факториалы очень быстро растут. Например,

$$
\begin{aligned}
10! &= 3628800,\\
100! &\approx 9,3326 \cdot 10^{157}.
\end{aligned}
$$

Биномиальные коэффициенты, конечно, тоже растут достаточно быстро, но треугольник Паскаля позволяет вычислять их, оперируя меньшими числами. Во-вторых, с помощью треугольника Паскаля можно посчитать сразу всю таблицу C_n^k для n и k, например, до 100.

Тема 2.1. Простые числа

Число называется **простым**, если оно имеет ровно два делителя: единицу и само себя. Первые простые числа — 2, 3, 5, 7, 11, 13, 17, 19, 23 и т.д. Единица имеет только один делитель, поэтому не является простым числом. Все остальные числа, кроме простых и 1, называются **составными**. Каждое составное число имеет хотя бы один простой делитель. Первый вопрос, который мы обсудим — проверка числа на простоту.

Упражнение 2.1
Являются ли простыми числа 97 и 1111111111 (10 единиц)? Решите задачу, не используя калькулятор и поиск в Интернете.

www.problems.ru, №30629

Самый очевидный способ проверки числа на простоту — перебрать все числа, меньшие него, и проверить, являются ли они его делителями. Можно ограничиться только простыми числами, но это не сильно сужает круг поиска. Заметим, что достаточно проверять делители только до квадратного корня из данного числа. Почему это так? Предположим, что наше число — составное. Тогда его можно представить в виде произведения двух множителей: $n = a \cdot b$. Покажем, что они оба одновременно не могут быть больше $\sqrt{n}$. Предположим противное:

$$a > \sqrt{n}, \quad b > \sqrt{n}.$$

Тогда

$$a \cdot b > \sqrt{n} \cdot \sqrt{n} = n.$$

Мы пришли к противоречию. Оно доказывает, что хотя бы один из сомножителей не превосходит $\sqrt{n}$:

$$a \leqslant \sqrt{n} \quad \text{или} \quad b \leqslant \sqrt{n}.$$

Они оба могут быть равны $\sqrt{n}$, как, например, для числа 4:

$$4 = 2 \cdot 2 \quad \Rightarrow \quad a = b = \sqrt{n}.$$

Таким образом, для числа 97 достаточно проверить, имеет ли оно делители до 10, т.к. $10^2 > 97$. Более того, можно проверять только простые числа. Если у числа есть составной делитель до корня, у него найдется простой делитель, который будет еще меньше.

Число 97 не делится на 2 — это число нечётное. На 3 оно тоже не делится. Для того, чтобы число делилось на 3, необходимо и достаточно, чтобы его сумма цифр делилась на 3. Для 97 это не выполняется. На 5 делятся только числа, которые оканчиваются на 5 и на 0. Число 97 к ним не относится. Для числа 7 нет простого признака делимости, но можно проверить, что 97 на него тоже не делится. Значит, 97 — простое число.

Рассмотрим теперь число из 10 единиц. Его можно разложить на множители, например, следующими двумя способами:

$$1111111111 = 11 \cdot 101010101,$$
$$1111111111 = 11111 \cdot 100001.$$

Догадаться до этих способов можно по виду числа.

Здесь может помочь также признак делимости на 11. Чтобы проверить делимость числа на 11, нужно взять сумму его цифр, стоящих на нечётных местах, и сумму цифр, стоящих на чётных местах. Если разность этих сумм делится на 11, то и само число делится на 11. Очевидно, это свойство выполняется для нашего числа:

$$\underline{1}\,\underline{\underline{1}}\,\underline{1}\,\underline{\underline{1}}\,\underline{1}\,\underline{\underline{1}}\,\underline{1}\,\underline{\underline{1}}\,\underline{1}\,\underline{\underline{1}}$$
$$(1+1+1+1+1)-(1+1+1+1+1)=0 \quad \text{— делится на 11.}$$

Вообще проверка большого числа на простоту — довольно трудная задача. Поэтому если вам предлагают решить её для такого большого числа в уме — скорее всего, число составное и нужно поискать какой-нибудь его делитель. Иногда на олимпиаде бывает полезно следовать логике жюри.

Упражнение 2.2
На сколько нулей оканчивается число 50!?

www.problems.ru, №30364

Ясно, что количество нулей в конце числа равно наибольшей степени 10, на которую делится число. Число 10, в свою очередь, представимо в виде произведения $2 \cdot 5$. Выясним, на какие наибольшие степени чисел 2 и 5 делится 50!.

Начнём с числа 2. Вспомним, что 50! — это произведение чисел от 1 до 50:

$$50! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot \ldots \cdot 49 \cdot 50.$$

Среди них на 2 делятся все чётные числа, а их — 25:

$$50! = 1 \cdot \underline{2} \cdot 3 \cdot \underline{4} \cdot 5 \cdot \underline{6} \cdot \ldots \cdot 49 \cdot \underline{50}.$$

Но это ещё не всё. Есть числа, которые делятся на 2 в степени больше первой. Это числа, делящиеся на 4. Один раз мы их уже посчитали. Чтобы посчитать их дополнительный вклад, прибавим их количество. Число 50 не делится нацело на 4, поэтому нужно брать округление вниз, чтобы точно подсчитать количество чисел, делящихся на 4 и не превосходящих 50: $\left\lfloor \frac{50}{4} \right\rfloor = 12$.

Ещё есть числа, которые делятся на 8 и дают степень двойки не меньше третьей. Мы их уже посчитали по два раза, прибавим ещё и третий. Их будет $\left\lfloor \frac{50}{8} \right\rfloor = 6$.

То же самое проделаем для следующих степеней двойки — 16 и 32, и получим ответ — 47:

$$\begin{aligned}
&\left\lfloor \frac{50}{2} \right\rfloor = 25 \text{ чисел делятся на } 2,\\
&\left\lfloor \frac{50}{4} \right\rfloor = 12 \text{ чисел делятся на } 4,\\
&\left\lfloor \frac{50}{8} \right\rfloor = 6 \text{ чисел делятся на } 8,\\
&\left\lfloor \frac{50}{16} \right\rfloor = 3 \text{ числа делятся на } 16,\\
&\left\lfloor \frac{50}{32} \right\rfloor = 1 \text{ число делится на } 32,\\
&25 + 12 + 6 + 3 + 1 = 47.
\end{aligned}$$

Аналогично поступим для числа 5:

$$\begin{aligned}
&\left\lfloor \frac{50}{5} \right\rfloor = 10 \text{ чисел делятся на } 5,\\
&\left\lfloor \frac{50}{25} \right\rfloor = 2 \text{ числа делятся на } 25,\\
&10 + 2 = 12.
\end{aligned}$$

Получаем, что $50!$ делится на 47-ю степень числа 2 и 12-ю степень числа 5, поэтому оно оканчивается на 12 нулей.

Можно обобщить эту идею и подсчитать, на какую наибольшую степень простого числа p делится $n!$:

$$\left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \ldots + \left\lfloor \frac{n}{p^k} \right\rfloor + \ldots \tag{1}$$

Формально сумма бесконечная, но фактически все слагаемые с некоторого k будут равны нулю. Для составного p эта формулу не работает. Например, один множитель в $n!$ может делиться на 2, а другой — на 5, и вместе они будут давать делимость на 10. Поэтому чтобы определить, делится ли $n!$ на составное число, нужно разложить это число на простые множители и применить к ним формулу (1) по отдельности.

Упражнение 2.3
Сколько существует троек простых чисел вида p, $p+2$, $p+4$?

www.problems.ru, №60469

Разберём решение задачи. Нетрудно подобрать подходящую тройку чисел $(3, 5, 7)$. Докажем, что она единственная. Посмотрим на остатки от деления чисел p, $p+2$ и $p+4$ на 3. Нетрудно видеть, что эти три остатка различны. Значит, одно из чисел p, $p+2$ и $p+4$ делится на 3. Но существует только одно простое число, кратное 3: это само число 3. В нашей тройке только p может быть равно 3. Получаем, что существует только одна тройка чисел, удовлетворяющих условию задачи.

Перейдём от простых чисел к составным. Для работы с составными числами очень полезна

Основная теорема арифметики. *Любое натуральное число можно представить в виде произведения простых чисел, причем такое представление единственно с точностью до порядка сомножителей.*

Например, число 60 можно представить как $2 \cdot 2 \cdot 3 \cdot 5$ или как $2 \cdot 5 \cdot 2 \cdot 3$, но фактически это один и тот же способ.

Простые числа в этом представлении обычно записывают в порядке возрастания и группируют одинаковые простые в степени:

$$n = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \cdot \ldots \cdot p_k^{\alpha_k}.$$

Здесь p_1, p_2, $\ldots$, p_k — простые числа, α_1, α_2, $\ldots$, α_k — их степени. Такое представление называется каноническим разложением числа на простые сомножители. Приведём несколько примеров канонических разложений:

$$\begin{aligned} 12 &= 2^2 \cdot 3^1, \\ 225 &= 3^2 \cdot 5^2, \\ 378 &= 2^1 \cdot 3^3 \cdot 7^1. \end{aligned}$$

По каноническому разложению числа можно составить последовательность степеней простых, которые в него входят:

$$60 = 2^2 \cdot 3^1 \cdot 5^1 \sim (2, 1, 1, 0, 0, \ldots),$$

$$21 = 3^1 \cdot 7^1 \qquad \sim (0, 1, 0, 1, 0, 0, \ldots).$$

Первое число в последовательности — степень двойки, второе — степень тройки и т.д., для всех простых по порядку. Если какое-то простое число не входит в каноническое разложение, то соответствующая ему степень будет равна нулю. Формально это бесконечные последовательности, но в них только конечное число ненулевых элементов. Простые числа будут иметь в своем представлении только одну единицу и остальные нули:

$$5 = 5^1 \sim (0, 0, 1, 0, 0, \ldots).$$

Числу 1 соответствует каноническое разложение из пустого множества простых и нулевая последовательность:

$$1 \sim (0, 0, 0, 0, 0, \ldots).$$

Числа в виде последовательностей степеней простых удобно умножать и делить. Действительно, если мы хотим перемножить два числа, то степени простых в разложениях складываются. При делении степени, наоборот, вычитаются. Можно также проверить, делится ли одно число на другое или нет. Получаем следующие правила умножения и деления:

$$\begin{aligned}
n &= p_1^{\alpha_1} \cdot p_2^{\alpha_2} \cdot \ldots \cdot p_k^{\alpha_k} && \sim (\alpha_1, \alpha_2, \alpha_3, \ldots), \\
m &= p_1^{\beta_1} \cdot p_2^{\beta_2} \cdot \ldots \cdot p_k^{\beta_k} && \sim (\beta_1, \beta_2, \beta_3, \ldots), \\
n \cdot m &= p_1^{\alpha_1+\beta_1} \cdot p_2^{\alpha_2+\beta_2} \cdot \ldots \cdot p_k^{\alpha_k+\beta_k} && \sim (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \alpha_3 + \beta_3, \ldots), \\
n : m &= p_1^{\alpha_1-\beta_1} \cdot p_2^{\alpha_2-\beta_2} \cdot \ldots \cdot p_k^{\alpha_k-\beta_k} && \sim (\alpha_1 - \beta_1, \alpha_2 - \beta_2, \alpha_3 - \beta_3, \ldots).
\end{aligned}$$

Иногда бывает удобно представлять большие числа последовательностями степеней в программе, если с ними нужно выполнять только операции умножения и деления. Простых правил сложения и вычитания в такой системе представления, к сожалению, нет.

Упражнение 2.4
Найдите количество делителей числа 720.

В этой задаче не обязательно проверять все делители до 720, и даже до $\sqrt{720}$. Можно воспользоваться каноническим разложением числа.

Рассмотрим каноническое разложение в общем виде

$$n = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \cdot \ldots \cdot p_k^{\alpha_k}.$$

Для произвольного делителя этого числа можно тоже записать каноническое разложение:

$$d = p_1^{\beta_1} \cdot p_2^{\beta_2} \cdot \ldots \cdot p_k^{\beta_k}.$$

В него входят те же самые простые, причём для степеней выполнены неравенства:

$$0 \leqslant \beta_i \leqslant \alpha_i, \quad i = 1, 2, \ldots, k.$$

Например, число 60 будет делителем числа 720, а 32 и 35 — не будут. В 32 входит слишком большая степень двойки, а 35 имеет делитель 7, который не является делителем 720:

$$
\begin{aligned}
&720 = 2^4 \cdot 3^2 \cdot 5^1,\\
&60 = 2^2 \cdot 3^1 \cdot 5^1,\\
&32 = 2^5 \cdot 3^0 \cdot 5^0,\\
&35 = 2^0 \cdot 3^0 \cdot 5^1 \cdot 7^1.
\end{aligned}
$$

Каждое β_i может принимать все целые значения от 0 до α_i включительно. Если все $\beta_i = 0$, получаем делитель 1. Если все β_i равны своим максимальным значениям α_i — получаем само число n.

Каждому набору β_i соответствует один и только один делитель.

$$(\beta_1, \beta_2, \beta_3, \ldots, \beta_k) \leftrightarrow d.$$

Это соответствие взаимно-однозначное, поэтому для получения ответа можно посчитать число вариантов составить подходящий набор из β_i. Значение β_1 можно выбрать $(\alpha_1 + 1)$ способом, β_2 — $(\alpha_2 + 1)$ способом и т.д. Ответом на вопрос задачи будет произведение

$$(\alpha_1 + 1) \cdot (\alpha_2 + 1) \cdot \ldots \cdot (\alpha_k + 1).$$

По этой формуле нетрудно посчитать ответ для числа $720 = 2^4 \cdot 3^2 \cdot 5^1$:

$$(4 + 1)(2 + 1)(1 + 1) = 30.$$

Тема 2.2. Наибольший общий делитель. Алгоритм Евклида

Используя канонические разложения чисел на простые множители, можно найти их **наибольший общий делитель** — максимальное число, на которое делятся нацело заданные числа. Например, $\text{НОД}(12, 30) = 6$.

Выпишем канонические разложения чисел 12 и 30 и составим по ним последовательности степеней простых:

$$12 = 2^2 \cdot 3^1 \qquad \sim\ (2, 1, 0, 0, \ldots),$$

$$30 = 2^1 \cdot 3^1 \cdot 5^1\ \sim\ (1, 1, 1, 0, \ldots).$$

Можно искать наибольший общий делитель отдельно по простым множителям. Максимальная степень двойки, которая входит в оба числа, равна 1, и максимальная степень тройки тоже равна 1. Пятёрка в наибольший общий делитель входить не может, т.к. 12 не делится на 5. Получаем, что наибольшему общему делителю соответствует последовательность степеней $(1, 1, 0, 0, \ldots)$, и он равен $6 = 2^1 \cdot 3^1$.

Можно обобщить этот алгоритм на произвольные числа. Предположим, два числа a и b заданы своими последовательностями степеней простых:

$$\begin{aligned} a &\sim (\alpha_1, \alpha_2, \alpha_3, \ldots), \\ b &\sim (\beta_1, \beta_2, \beta_3, \ldots). \end{aligned}$$

Тогда, чтобы составить наибольший общий делитель, мы должны выбрать минимум по каждой степени, тогда ему будет соответствовать последовательность из минимумов:

$$\begin{aligned} &\text{НОД}(a, b) \sim (\delta_1, \delta_2, \delta_3, \ldots), \\ &\delta_i = \min(\alpha_i, \beta_i), \quad i = 1, 2, \ldots \end{aligned}$$

При помощи канонических разложений можно доказать некоторые свойства наибольшего общего делителя. Например, наибольший общий делитель чисел a и b делится на все их общие делители. Действительно, пусть $d =$ НОД(a, b) и

$$\begin{aligned} a &\sim (\alpha_1, \alpha_2, \alpha_3, \ldots), \\ b &\sim (\beta_1, \beta_2, \beta_3, \ldots), \\ c &\sim (\gamma_1, \gamma_2, \gamma_3, \ldots), \\ d &\sim (\delta_1, \delta_2, \delta_3, \ldots). \end{aligned}$$

Число d делится на c тогда и только тогда, когда в их последовательностях степеней $\gamma_i \leqslant \delta_i$ для всех $i = 1, 2, 3, \ldots$. Если c — общий делитель a и b, то $\gamma_i \leqslant \alpha_i$ и $\gamma_i \leqslant \beta_i$, $i = 1, 2, 3, \ldots$. Значит, $\gamma_i \leqslant \min(\alpha_i, \beta_i) = \delta_i$, и число c является делителем d, ч.т.д.

Например, у чисел 12 и 30 наибольший общий делитель 6 делится на все их общие делители: 1, 2, 3 и 6.

Упражнение 2.5
Найдите наибольший общий делитель чисел 10, 14 и 35.

Наибольший общий делитель можно определить не только для двух чисел, но и для трёх, и вообще для любого множества. Наибольшее натуральное число, на которое делятся 10, 14 и 35 — это 1, несмотря на то, что в каждой паре из этих чисел наибольший общий делитель будет больше единицы.

Наибольший общий делитель трёх чисел можно найти по очевидной формуле через минимум всех трёх степеней простых:

$$\begin{aligned} a &\sim (\alpha_1, \alpha_2, \alpha_3, \ldots), \\ b &\sim (\beta_1, \beta_2, \beta_3, \ldots), \\ c &\sim (\gamma_1, \gamma_2, \gamma_3, \ldots), \\ \text{НОД}(a, b, c) &\sim (\min(\alpha_1, \beta_1, \gamma_1), \min(\alpha_2, \beta_2, \gamma_2), \min(\alpha_3, \beta_3, \gamma_3), \ldots). \end{aligned}$$

Эту формулу можно обобщить на случай произвольного количества чисел.

Заметим, что

$$\min(\alpha_i, \beta_i, \gamma_i) = \min(\min(\alpha_i, \beta_i), \gamma_i).$$

Благодаря этому свойству, справедлива формула

$$\text{НОД}(a, b, c) = \text{НОД}(\text{НОД}(a, b), c).$$

Пользуясь этой формулой, можно построить алгоритм нахождения наибольшего общего делителя для последовательности чисел через попарный наибольший общий делитель.

С наибольшим общим делителем тесно связано **наименьшее общее кратное** — наименьшее натуральное число, которое делится на все заданные числа. Например, $\text{НОК}(12, 30) = 60$. Наименьшее общее кратное двух чисел можно вычислить по замечательной формуле:

$$\text{НОК}(a, b) = \frac{a \cdot b}{\text{НОД}(a, b)}. \tag{2}$$

Например, для чисел 12 и 30 будет

$$\text{НОК}(12, 30) = \frac{12 \cdot 30}{6} = 60.$$

Формулу (2) легко доказать, опираясь на канонические разложения. Пусть числам a, b и их наименьшему общему кратному соответствуют следующие последовательности степеней простых:

$$\begin{aligned} a &\sim (\alpha_1, \alpha_2, \alpha_3, \ldots), \\ b &\sim (\beta_1, \beta_2, \beta_3, \ldots), \\ c = \text{НОК}(a, b) &\sim (\gamma_1, \gamma_2, \gamma_3, \ldots), \quad \gamma_i = \max(\alpha_i, \beta_i), \quad i = 1, 2, 3, \ldots \end{aligned}$$

Каждый простой множитель должен входить в наименьшее общее кратное число раз, равное максимуму из соответствующих степеней α_i и β_i, чтобы c делилось на a и b и при этом было наименьшим возможным. Для наибольшего общего делителя мы имеем аналогичную формулу с минимумом:

$$d = \text{НОД}(a, b) \sim (\delta_1, \delta_2, \delta_3, \ldots), \quad \delta_i = \min(\alpha_i, \beta_i), \quad i = 1, 2, 3, \ldots$$

а для произведения $a \cdot b$ справедливо представление

$$a \cdot b \sim (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \alpha_3 + \beta_3, \ldots).$$

Ясно, что

$$\alpha_i + \beta_i = \min(\alpha_i, \beta_i) + \max(\alpha_i, \beta_i),$$

потому что одно из чисел α_i и β_i является среди них минимумом, а другое — максимумом. Значит, и в левой, и в правой части формулы стоит одна и та же сумма. Сравнивая последовательности степеней простых для наименьшего общего кратного, наибольшего общего делителя и произведения чисел a и b, приходим к соотношению

$$a \cdot b = \text{НОД}(a, b) \cdot \text{НОК}(a, b),$$

из которого нетрудно выразить наименьшее общее кратное:

$$\text{НОК}(a, b) = \frac{a \cdot b}{\text{НОД}(a, b)}.$$

Упражнение 2.6
Верны ли следующие утверждения для любых натуральных чисел a, b, c?

1. Любое общее кратное чисел a и b делится на $\text{НОК}(a, b)$.
2. $a \cdot b \cdot c = \text{НОД}(a, b, c) \cdot \text{НОК}(a, b, c)$.

Использовать канонические разложения удобно с теоретический точки зрения, но вот практически строить их специально для нахождения наибольшего общего делителя не очень эффективно. На практике обычно используют **алгоритм Евклида**. Он состоит в следующем.

Пусть дана пара натуральных чисел (a, b). Заменим их на пару $(b \bmod a, a)$, где $b \bmod a$ — операция взятия остатка от деления числа b на a. Например, для чисел 12 и 30 получим $30 \bmod 12 = 6$. Значит, будет пара $(6, 12)$. Будем продолжать выполнять эту операцию, пока одно из чисел не обнулится. Для нашего примера это произойдет уже на следующем шаге: $12 \bmod 6 = 0$, и получится пара $(0, 6)$. Утверждается, что второе число в паре будет наибольшим общим делителем исходных чисел, потому что при преобразовании пар их наибольший общий делитель не меняется. Докажем этот факт.

Пусть d — какой-нибудь общий делитель a и b, не обязательно наибольший. Тогда число $b \bmod a$ тоже будет делиться на d. Действительно,

$$b \bmod a = b - \left\lfloor \frac{b}{a} \right\rfloor \cdot a. \tag{3}$$

Уменьшаемое и вычитаемое делятся на d, значит, и разность делится на d, т.е. оба числа в полученной паре $(b \bmod a, a)$ делятся на d. Верно и обратное: если $b \bmod a$ и a делятся на d, то и b делится на d. Таким образом, все множество общих делителей у пары при преобразовании сохраняется. Значит, сохраняется и наибольший общий делитель.

Наибольший общий делитель у пары вида $(0, a)$ равен a, потому что 0 делится на любое число. То же самое число будет наибольшим общим делителем исходной пары. Корректность алгоритма Евклида доказана.

Упражнение 2.7
Сколькими способами можно набрать 250 рублей купюрами по 20 и 50 рублей? Можно использовать купюры каждого достоинства в любом количестве, в том числе в нулевом.

Для решения задачи можно просто перебрать все способы. Ответим на более общий вопрос: как найти все решения следующего уравнения в целых числах?

$$ax + by = c. \tag{4}$$

Здесь a, b, c — известные коэффициенты, x и y — неизвестные. Уравнение (4) называется **линейным диофантовым уравнением**. Например, Упражнение 2.7 сводится к уравнению

$$20x + 50y = 250.$$

Исследуем вопрос о том, при любых ли a, b, c уравнение (4) имеет решение. Необходимо, чтобы число c делилось на $d = \text{НОД}(a, b)$. Действительно, левая часть уравнения (4) делится на d при любых целых x и y, значит, и правая часть должна делиться на это число. Если c не делится на d — уравнение (4) не имеет решений. Иначе поделим все три коэффициента a, b и c на d:

$$a' = \frac{a}{d}, \quad b' = \frac{b}{d}, \quad c' = \frac{c}{d},$$

и перейдем к уравнению

$$a'x + b'y = c',$$

в котором $\text{НОД}(a', b') = 1$. Такие числа, у которых наибольший общий делитель равен 1, называются **взаимно простыми**. Ясно, что множество решений уравнения при таком делении не изменится.

Решим сначала уравнение

$$ax + by = 1, \tag{5}$$

где a и b — взаимно простые. Найдем хотя бы одно из его решений. Оказывается, для этого можно использовать алгоритм Евклида. Вспомним, что на каждом шаге алгоритма Евклида происходит переход от пары (a, b) к паре $(b \bmod a, a)$. Представим, что мы нашли решение интересующего нас уравнения для второй пары, т.е. мы каким-то образом нашли такие числа x и y, что

$$(b \bmod a)x + ay = 1.$$

Подставим в это равенство формулу (3):

$$\left(b - \left\lfloor \frac{b}{a} \right\rfloor a\right)x + ay = 1.$$

Перегруппируем слагаемые:

$$a\left(y - \left\lfloor \frac{b}{a} \right\rfloor x\right) + bx = 1.$$

Введём новые обозначения:

$$x' = y - \left\lfloor \frac{b}{a} \right\rfloor x, \quad y' = x.$$

Получим

$$ax' + by' = 1,$$

т.е. x' и y' дают решение уравнения (5).

Таким образом, можно решить уравнение (5), переходя в алгоритме Евклида от последующей пары к предыдущей. Описанный алгоритм называется **расширенным алгоритмом Евклида**.

Заметим, что самая последняя пара чисел для взаимно простых a и b получится $(0, 1)$. Для неё решением будет, например, $x = 0, y = 1$. Действительно,

$$0 \cdot 0 + 1 \cdot 1 = 1.$$

Можно подобрать и другие решения, но нам пока достаточно одного.

Итак, мы научились искать решение уравнения (5). Чтобы перейти от его решения к решению уравнения (4) с правой частью c, достаточно умножить x и y на c:

$$\begin{aligned} ax + by &= 1, \\ a(cx) + b(cy) &= c. \end{aligned}$$

Но мы нашли пока только одно решение уравнения (4), а у него могут быть другие решения.

Предположим, у уравнения (4) есть хотя бы две пары решений (x_1, y_1) и (x_2, y_2). Составим два равенства

$$\begin{aligned} ax_1 + by_1 &= c, \\ ax_2 + by_2 &= c, \end{aligned}$$

и вычтем второе из первого. У нас получится соотношение

$$a(x_1 - x_2) + b(y_1 - y_2) = 0.$$

Поскольку a и b взаимно простые, число $(x_1 - x_2)$ должно делиться на b, а число $(y_1 - y_2)$ должно делиться на a. Более точно, необходимо выполнение условий

$$x_1 - x_2 = kb, \quad y_1 - y_2 = -ka,$$

где k — целое число. Предположим, мы нашли расширенным алгоритмом Евклида одно решение уравнения (4), пусть оно равно (x_1, y_1). Оказывается, что любая пара вида

$$x_2 = x_1 - kb,$$
$$y_2 = y_1 + ka$$

тоже будет решением. Это нетрудно проверить. Полученные формулы описывают все множество решений линейного диофантова уравнения (4).

Традиционно зададим себе вопрос: какие в этой задаче возможны крайние случаи?

1. Возможен случай $a = b = 0$, его нужно обрабатывать отдельно.

2. Также a или b могут быть отрицательными. Заметим, что наш алгоритм по умолчанию для этого случая не работает — для алгоритма Евклида числа a и b должны быть положительными. Пусть для определённости $a < 0$. Можно выполнить переход от отрицательного a к положительному, заменив также знак у x:

$$(-a)(-x) + by = c,$$

и потом для решения выполнить обратный переход.

Упражнение 2.8
Сколько целых решений имеет уравнение $20x - 50y = 250$ в диапазоне $|x| \leqslant 100$, $|y| \leqslant 100$?

Задачу, конечно, можно решать разными способами. Применим по порядку всю нашу теорию для ее решения. Имеем линейное диофантово уравнение (4) с

$$a = 20, \quad b = -50, \quad c = 250.$$

Число b отрицательное, поэтому выполним замену:

$$b := -b, \quad y := -y,$$
$$a = 20, \quad b = 50, \quad c = 250.$$

Заметим, что диапазон для y по условию симметричный, поэтому данная замена не повлияет на итоговый ответ. Далее поделим a, b и c на их наибольший общий делитель, который равен 10. Получим

$$a = 2, \quad b = 5, \quad c = 25.$$

Начнём с решения уравнения

$$2x + 5y = 1.$$

Для него нетрудно подобрать решение (можно использовать и отрицательные числа), но мы разберём на этом простом примере работу общего алгоритма. Алгоритм Евклида даст следующие пары:

$$(2, 5) \ \to \ (1, 2) \ \to \ (0, 1).$$

Для последней пары $x = 0$, $y = 1$.

a	b	x	y
2	5		
1	2		
0	1	0	1

Рис. 1

Перейдем к паре $(1, 2)$ по формулам

$$x' = y - \left\lfloor \frac{b}{a} \right\rfloor x = 1 - \frac{2}{1} \cdot 0 = 1,$$
$$y' = x = 0.$$

Получим $(1, 0)$. На всякий случай проверим, что эта пара подходит:

$$1 \cdot 1 + 2 \cdot 0 = 1.$$

a	b	x	y
2	5		
1	2	1	0
0	1	0	1

Рис. 2

Дальше выполним переход к паре $(2, 5)$:

$$x' = y - \left\lfloor \frac{b}{a} \right\rfloor x = 0 - \left\lfloor \frac{5}{2} \right\rfloor \cdot 1 = -2,$$
$$y' = x = 1.$$

Получим $(-2, 1)$. Действительно, $-2 \cdot 2 + 5 \cdot 1 = 1$.

a	b	x	y
2	5	-2	1
1	2	1	0
0	1	0	1

Рис. 3

Чтобы решить уравнение

$$2x + 5y = 25, \tag{6}$$

умножим полученные x и y на 25:

$$\begin{matrix} 2x + 5y = 1, \\ x = -2,\ y = 1 \end{matrix} \quad \Rightarrow \quad \begin{matrix} 2x + 5y = 25, \\ x = -50,\ y = 25. \end{matrix}$$

Всё множество решений уравнения (6) описывается формулами

$$x = -50 - 5k,$$
$$y = 25 + 2k.$$

Подставим эти решения в условия:

$$\begin{matrix} -100 \leqslant x \leqslant 100, \\ -100 \leqslant y \leqslant 100, \end{matrix} \quad \Rightarrow \quad \begin{matrix} -100 \leqslant -50 - 5k \leqslant 100, \\ -100 \leqslant 25 + 2k \leqslant 100, \end{matrix}$$

и получим ограничения $-30 \leqslant k \leqslant 10$. В этом диапазоне 41 целое число. Получаем 41 решение исходной задачи (Упражнения 2.8).

Описанный метод можно применить и для решения задачи с купюрами (Упражнения 2.7). Вместо диапазонов в ней будут ограничения $x \geqslant 0$ и $y \geqslant 0$.

Тема 2.3. Теория сравнений. Обратный элемент

Эта лекция посвящена теории сравнений. Как мы убедимся, она тесно связана с программированием.

Начнём с понятия сравнения. Пусть a и b — целые числа, а m — натуральное, обычно $m > 1$. Говорят, что a **сравнимо с** b **по модулю** m, если разность $a - b$ делится на m без остатка. Иначе говоря, a и b имеют один и тот же остаток при делении на m. Используется следующее обозначение:

$$a \equiv b \pmod{m}.$$

Примеры сравнений:

$$2 \equiv 12 \pmod{10},$$
$$3 \equiv -2 \pmod{5}.$$

Во многих языках программирования есть операция взятия числа по модулю. Она может обозначаться по-разному, мы продолжим использовать обозначение mod. Эта операция возвращает число от нуля до модуля без единицы, равное остатку от деления на заданный модуль. Например, $12 \bmod 10 = 2$. Операция mod может немного иначе работать для отрицательных чисел, но обычно она применяется к неотрицательным.

Отдельно отметим, что mod как операция и mod в сравнениях — это не одно и то же. Сравнение с тремя чёрточками — это отношение между парой чисел, некоторый аналог отношения равенства. Например, сравнение

$$2 \equiv 12 \pmod{10}$$

— утверждение об этих числах. В сравнении могут фигурировать любые числа, не обязательно меньшие модуля, могут быть отрицательные числа. Взятие числа по модулю (например, $12 \bmod 10$) — это операция, которая возвращает значение. Как, например, операция сложения или деления. Результат операции взятия по модулю — всегда неотрицательное число, меньшее модуля.

Между сравнениями и операцией взятия по модулю существует тесная связь. Согласно определению, верно сравнение

$$a \equiv (a \bmod m) \pmod{m}.$$

Остаток от деления числа на модуль всегда сравним с этим числом по этому модулю.

Часто в олимпиадных задачах по программированию требуется найти ответ по какому-нибудь заданному модулю, т.е. найти остаток ответа от деления на модуль. Как правило, такое требование возникает, если ответ на задачу достаточно большой, и считать его целиком нецелесообразно.

Упражнение 2.9
Найдите значение выражения

$$(172636 \cdot 35261 + 45217 \cdot 472819 + 46378 \cdot 352671 \cdot 163617)\cdot$$
$$\cdot(7662717 - 1934)$$

по модулю 10.

В этой задаче удобный модуль. Остаток от деления на 10 равен последней цифре числа. Можно заменить все числа на их последние цифры и выполнить с ними операции. При выполнении каждой операции можно тоже оставлять только последнюю цифру.

Почему при вычислениях по модулю можно вот так брать по модулю все промежуточные результаты? Хотя этот факт и кажется очевидным, он нуждается в доказательстве. Почему, например, можно сложить два слагаемых, а потом взять по модулю, а можно — сначала взять по модулю, потом сложить, потом снова взять по модулю, и получить тот же самый результат?

$$(a+b) \bmod m \underbrace{=}_{?} ((a \bmod m) + (b \bmod m)) \bmod m.$$

Строго доказать это равенство можно при помощи сравнений.

Если выполнены сравнения

$$a \equiv c \pmod m,$$
$$b \equiv d \pmod m,$$

то

$$a + b \equiv c + d \pmod m. \tag{7}$$

Действительно, рассмотрим разность $(a+b) - (c+d)$. Перегруппируем слагаемые:

$$(a+b) - (c+d) = (a-c) + (b-d).$$

Разности $(a-c)$ и $(b-d)$ делятся на m (эти пары чисел сравнимы по модулю m). Значит, и разность $(a+b)-(c+d)$ делится на m, и выполняется сравнение (7).

Аналогичное доказательство можно провести для разности и произведения:

$$\begin{array}{l} a \equiv c \pmod m, \\ b \equiv d \pmod m, \end{array} \quad \Rightarrow \quad \begin{array}{l} a + b \equiv c + d \pmod m, \\ a - b \equiv c - d \pmod m, \\ a \cdot b \equiv c \cdot d \pmod m. \end{array}$$

Полученные сравнения можно использовать для обоснования наших действий при подсчёте значения выражения в Упражнении 2.9. Мы каждый раз производили взятие по модулю, т.е. замену на сравнимое по этому модулю число.

Заметим, что для деления аналогичное утверждение неверно. Можно привести контрпример.

Упражнение 2.10
Выберите четвёрку чисел a, b, d и m, для которой $a \equiv b \pmod{m}$, но $a/d \not\equiv b/d \pmod{m}$.

1. $a = 10, b = 20, d = 2, m = 5$,
2. $a = 6, b = 9, d = 3, m = 5$,
3. $a = 5, b = 15, d = 5, m = 2$,
4. $a = 15, b = 30, d = 3, m = 15$.

Проблемы с делением возможны, если делитель и модуль не взаимно простые. Если же они взаимно простые, то делить по модулю можно. Более того, для любого числа d от 1 до $m - 1$, взаимно простого с m, можно ввести понятие обратного элемента по модулю m. **Обратный элемент** — это целое число d^{-1}, такое, что

$$1 \leqslant d^{-1} \leqslant m - 1, \quad d \cdot d^{-1} \equiv 1 \pmod{m}.$$

Можно показать, что обратный элемент определяется единственным образом.

Найти обратный элемент можно при помощи расширенного алгоритма Евклида. Рассмотрим линейное диофантово уравнение

$$dx + my = 1.$$

Если d и m взаимно простые, расширенный алгоритм Евклида позволяет найти его решение (x, y). Заметим, что

$$dx \equiv 1 \pmod{m},$$

т.е. x является обратным элементом к d, если $1 \leqslant x \leqslant m - 1$. В противном случае $x \equiv d^{-1} \pmod{m}$, и нужно привести x к отрезку $[1, m-1]$ при помощи операции взятия по модулю.

Тема 2.4. Китайская теорема об остатках

Завершим тему про теорию чисел китайской теоремой об остатках. Она дает связь между остатками от деления чисел на разные модули. Сначала проследим эту связь на примере. Построим таблицу остатков от деления на 3, на 5 и на 15 (см. рис. 4).

n	n mod 3	n mod 5	n mod 15
0	0	0	0
1	1	1	1
2	2	2	2
3	0	3	3
4	1	4	4
5	2	0	5
6	0	1	6
7	1	2	7
8	2	3	8
9	0	4	9
10	1	0	10
11	2	1	11
12	0	2	12
13	1	3	13
14	2	4	14
15	0	0	0

Рис. 4

Остатки в первом столбце циклически повторяются через 3, во втором — через 5, и во всех трёх столбцах они повторяются через 15. Действительно, $15 = 3 \cdot 5$.

Нетрудно заметить, что у чисел с одинаковыми остатками от деления на 15 будут одинаковые остатки от деления на 3 и на 5. Более того, верно обратное. Если мы возьмём произвольную пару остатков от деления на 3 и на 5 (например, $(2,1)$), эта пара найдётся в нашей таблице, и ей будет однозначно соответствовать остаток от деления на 15 (см. рис. 5). Замеченное взаимно-однозначное соответствие даёт возможность закодировать числа от 0 до $15-1$, т.е. 14, парами их остатков от деления на 3 и 5. Например, пара $(2,1)$ кодирует число 11. Аналогичный факт верен для произвольного числа модулей и называется китайской теоремой об остатках.

n	$n \bmod 3$	$n \bmod 5$	$n \bmod 15$
0	0	0	0
1	1	1	1
2	2	2	2
3	0	3	3
4	1	4	4
5	2	0	5
6	0	1	6
7	1	2	7
8	2	3	8
9	0	4	9
10	1	0	10
11	2	1	11
12	0	2	12
13	1	3	13
14	2	4	14
15	0	0	0

Рис. 5

Китайская теорема об остатках. *Если натуральные числа* $m_1, m_2, \ldots, m_n$ *попарно взаимно просты, то для любых чисел* $a_1, a_2, \ldots, a_n$, $0 \leqslant a_i < m_i$, $i = 1, 2, \ldots, n$, *найдётся число* N, $N \equiv a_i \pmod{m_i}$, $i = 1, 2, \ldots, n$. *Если найдутся два таких числа* N_1 *и* N_2, *то* $N_1 \equiv N_2 \pmod{m_1 \cdot m_2 \cdot \ldots \cdot m_n}$.

Последнее утверждение означает единственность числа N с точностью до взятия по модулю-произведению.

Посмотрим, как работает эта теорема на примере задачи.

Упражнение 2.11

На столе лежат книги, которые нужно упаковать в пачки. Если их связать в одинаковые пачки по 4, по 5 или по 6 книг, то каждый раз останется одна лишняя книга, а если связать по 7 книг в пачку, то лишних книг не останется. Какое наименьшее количество книг может быть на столе?

www.problems.ru, №60828

Обозначим ответ на задачу через x. Заметим, что к модулям 4, 5, 6, 7 нельзя применять китайскую теорему об остатках, потому что они не все попарно взаимно простые.

Будем действовать по-другому. Согласно условию задачи, число $x - 1$ делится на 4, 5 и 6, а значит, делится на наименьшее общее кратное этих чисел, равное 60. Иначе говоря,

$$x \equiv 1 \pmod{60}. \tag{8}$$

Кроме того, по условию задачи

$$x \equiv 0 \pmod{7}. \tag{9}$$

Мы свели задачу к системе двух сравнений (8) и (9). По китайской теореме об остатках система (8)–(9) имеет единственное решение, меньшее произведения $60 \cdot 7 = 420$. Можно решать эту систему по-разному. Например, из сравнения (8) находим, что

$$x = 60k + 1, \tag{10}$$

где k — некоторое целое число. Подставим (10) в сравнение (9):

$$60k + 1 \equiv 0 \pmod{7}.$$

Для упрощения возьмём 60 по модулю 7. Получим

$$4k + 1 \equiv 0 \pmod{7},$$

или

$$4k \equiv -1 \pmod{7}. \tag{11}$$

Найдём обратный элемент к числу 4 по модулю 7. Для маленьких чисел обратный элемент можно найти подбором. В данном случае $4^{-1} = 2$, потому что

$$4 \cdot 2 \equiv 1 \pmod 7.$$

Если числа были бы побольше, можно было бы воспользоваться расширенным алгоритмом Евклида.

Умножим обе части сравнения (11) на 2:

$$\begin{aligned} &2 \cdot 4k \equiv -2 \pmod 7, \\ &k \equiv -2 \pmod 7. \end{aligned}$$

Наименьшее неотрицательное k, которое нам подходит — это 5. Из соотношения (10) получаем, что

$$x = 60 \cdot 5 + 1 = 301.$$

Можно решать систему двух сравнений (8)–(9) по-другому. Из сравнения (9) найдём $x = 7m$ и подставим в сравнение (8). Но тогда нужно будет искать обратный элемент для 7 по модулю 60, что сложнее сделать при подсчётах вручную.

Разберём общий алгоритм построение числа N из китайской теоремы об остатках. Пусть мы знаем, что

$$\begin{aligned} &N \equiv a_1 \pmod{m_1}, \\ &N \equiv a_2 \pmod{m_2}, \\ &\ldots \\ &N \equiv a_n \pmod{m_n}, \end{aligned}$$

и хотим найти

$$N \bmod (m_1 \cdot m_2 \cdot \ldots \cdot m_n).$$

Представим число N в смешанной системе счисления:

$$\begin{aligned} &N = m_1 \cdot m_2 \cdot \ldots \cdot m_{n-1} \cdot x_n + m_1 \cdot m_2 \cdot \ldots \cdot m_{n-2} \cdot x_{n-1} + \\ &+ \ldots + m_1 \cdot m_2 \cdot x_3 + m_1 x_2 + x_1, \\ &0 \leqslant x_i < m_i. \end{aligned} \tag{12}$$

Здесь $x_1, x_2, \ldots, x_n$ — его цифры.

Обычно рассматриваются системы счисления с некоторым основанием b. Например, $b = 10$ для десятичной системы счисления. Цифры числа умножаются на степени этого основания:

$$N = b^n c_n + b^{n-1} c_{n-1} + \ldots + b^2 c_2 + b c_1 + c_0.$$

В смешанной системе счисления n различных оснований — $m_1, m_2, \ldots, m_n$, и вместо степеней выступают их произведения. При этом если в десятичной системе счисления все цифры лежат в одном диапазоне:

$$b = 10 \quad \Rightarrow \quad 0 \leqslant c_i \leqslant 9,$$

то здесь каждая цифра x_i может принимать значения в своем диапазоне от 0 до $m_i - 1$. Ясно, что если мы знаем все цифры x_i, то сможем найти число N по формуле (12). Обсудим, как их вычислить.

Из сравнения $N \equiv a_1 \pmod{m_1}$ находим $x_1 = a_1$. Действительно, все остальные слагаемые в разложении (12) для N кратны m_1. Из сравнения $N \equiv a_2 \pmod{m_2}$ имеем

$$m_1 x_2 + x_1 \equiv a_2 \pmod{m_2}.$$

Отсюда получаем формулу

$$x_2 = m_1^{-1}(a_2 - x_1) \bmod m_2. \tag{13}$$

Сравнение $N \equiv a_3 \pmod{m_3}$ даёт формулу для нахождения x_3 и т.д. Получаем общую формулу для нахождения x_i:

$$x_i = (m_1 \cdot m_2 \cdot m_{i-1})^{-1}(a_i - x_1 - m_1 x_2 - \ldots - m_1 \cdot m_2 \cdot \ldots m_{i-2} x_{i-1}) \bmod m_i. \tag{14}$$

Здесь нужно находить обратные элементы для одних модулей m_i по другим. Это всегда можно сделать, так как модули взаимно простые. Алгоритм, который мы разобрали, даёт возможность однозначно построить число N и конструктивно доказывает китайскую теорему об остатках.

Разберём пример. Пусть нужно найти число N, удовлетворяющее сравнениям:

$$\begin{aligned} N &\equiv 0 \pmod{2}, \\ N &\equiv 1 \pmod{3}, \\ N &\equiv 2 \pmod{5}, \\ N &\equiv 1 \pmod{7}. \end{aligned}$$

Для данного примера

$$\begin{aligned} &a_1 = 0, \quad a_2 = 1, \quad a_3 = 2, \quad a_4 = 1, \\ &m_1 = 2, \quad m_2 = 3, \quad m_3 = 5, \quad m_4 = 7. \end{aligned}$$

При $i = 1$ имеем:

$$x_1 = a_1 = 0.$$

Для $i = 2$ воспользуемся формулой (13). Нужно найти обратный элемент к $m_1 = 2$ по модулю $m_2 = 3$. Нетрудно видеть, что $m_1^{-1} = 2$, т.к. $2 \cdot 2 \equiv 1 \pmod 3$. Таким образом, формула (13) даёт

$$x_2 = 2 \cdot (1 - 0) \bmod 3 = 2.$$

Для $i = 3$ формула (14) принимает вид

$$x_3 = (m_1 \cdot m_2)^{-1}(a_3 - x_1 - m_1 x_2) \bmod m_3.$$

Произведение $m_1 \cdot m_2 = 6 > m_3$, поэтому сначала возьмём $6 \bmod 5 = 1$. Очевидно, что обратный элемент к 1 по любому модулю равен 1. Следовательно, получаем

$$x_3 = 1 \cdot (2 - 0 - 2 \cdot 2) \bmod 5 = -2 \bmod 5 = 3.$$

Здесь при взятии по модулю отрицательного числа -2 мы воспользовались сравнением $-2 \equiv 3 \pmod 5$ и получили $-2 \bmod 5 = 3$.

Формально результат операции $a \bmod m$ для $a < 0$ можно определить как целое число b из диапазона $[0, m - 1]$, такое, что $a \equiv b \pmod m$. Обратите внимание, что операция взятия по модулю в языках программирования может работать по-другому.

Для $i = 4$ формула (14) принимает вид

$$x_4 = (m_1 \cdot m_2 \cdot m_3)^{-1}(a_4 - x_1 - m_1 x_2 - m_1 m_2 x_3) \bmod m_4. \tag{15}$$

Вычисляем:

$$(m_1 \cdot m_2 \cdot m_3) \bmod m_4 = (2 \cdot 3 \cdot 5) \bmod 7 = 30 \bmod 7 = 2.$$

Нужно найти 2^{-1} по модулю 7. Для этого решим линейное диофантово уравнение

$$2x + 7y = 1$$

расширенным алгоритмом Евклида. Заполним таблицу, аналогичную приведённой на рис. 3.

a	b	x	y
2	7	-3	1
1	2	1	0
0	1	0	1

Получаем

$$\begin{array}{l} 2^{-1} \equiv -3 \pmod 7, \\ 1 \leqslant 2^{-1} \leqslant 6 \end{array} \quad \Rightarrow \quad 2^{-1} = 4.$$

Теперь вычислим x_4 по формуле (15):

$$x_4 = 4 \cdot (1 - 0 - 2 \cdot 2 - 2 \cdot 3 \cdot 3) \bmod 7 = 4 \cdot (-21) \bmod 7,$$
$$-21 \equiv 0 \pmod 7 \quad \Rightarrow \quad x_4 = 0.$$

Подставляя найденные значения x_1, x_2, x_3 и x_4 в формулу (12), находим

$$N = m_1 \cdot m_2 \cdot m_3 \cdot x_4 + m_1 \cdot m_2 \cdot x_3 + m_1 \cdot x_2 + x_1 = \\ = 2 \cdot 3 \cdot 5 \cdot 0 + 2 \cdot 3 \cdot 3 + 2 \cdot 2 + 0 = 22.$$

Поскольку

$$m_1 \cdot m_2 \cdot m_3 \cdot m_4 = 2 \cdot 3 \cdot 5 \cdot 7 = 210,$$

по китайской теореме об остатках $N = 22$ — единственное неотрицательное число, меньшее 210 и имеющее остатки $0, 1, 2, 1$ по модулям $2, 3, 5, 7$ соответственно. Всё множество решений данной системы сравнений описывается формулой

$$22 + 210k, \quad k \in \mathbb{Z}.$$

Тема 3.1. Точки и прямые на плоскости

На этой лекции мы разберём, как решать геометрические задачи, которые встречаются на олимпиадах по программированию. Для их решения обычно используется аналитический метод. Все фигуры — точки, прямые, векторы, окружности — представляются в системе координат, и геометрическая задача сводится к решению уравнений.

Введем стандартную декартову систему координат. В нашем курсе мы ограничимся изучением геометрии на плоскости. Самый простой объект на плоскости — это точка, она задается двумя координатами x и y.

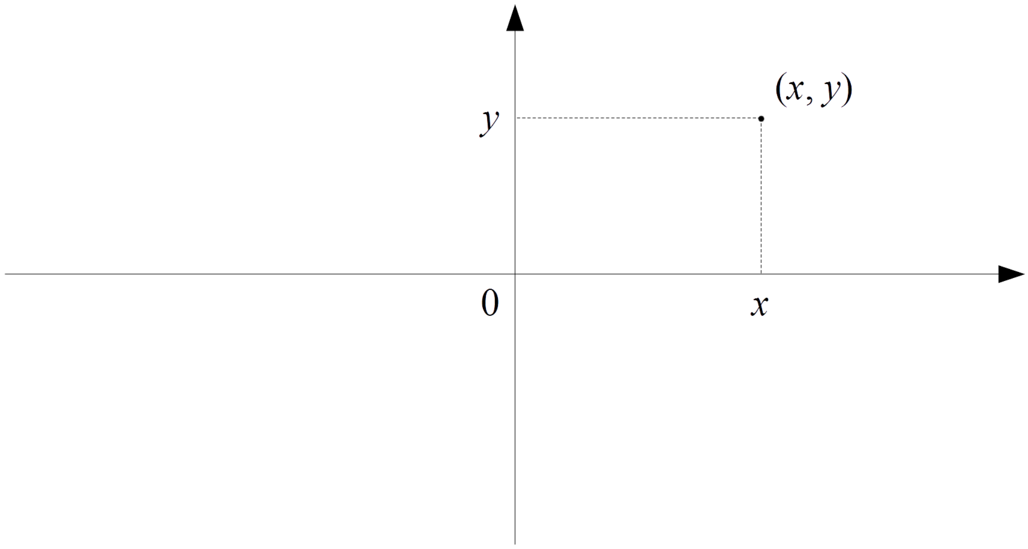


Рис. 1

Перейдём к прямым. Прямую удобнее всего задавать ее общим уравнением

$$Ax + By + C = 0. \tag{1}$$

Здесь A, B, C — коэффициенты прямой. Каждая прямая описывается тройкой чисел (A, B, C), x и y — координаты точки. Уравнение прямой обращается в равенство для тех и только для тех точек, которые лежат на этой прямой.

Например, точка $(-3, 0)$ лежит на прямой

$$x + 2y + 3 = 0,$$

поэтому подстановка даёт

$$(x, y) = (-3, 0): \quad -3 + 2 \cdot 0 + 3 = 0.$$

Точка $(0, 1)$ не лежит на этой прямой:

$$(x, y) = (0, 1): \quad 0 + 2 \cdot 1 + 3 = 5 \neq 0.$$

В школе часто используют уравнение прямой с угловым коэффициентом:

$$y = kx + b.$$

В нём k — тангенс угла между прямой и осью x (см. рис. 2).

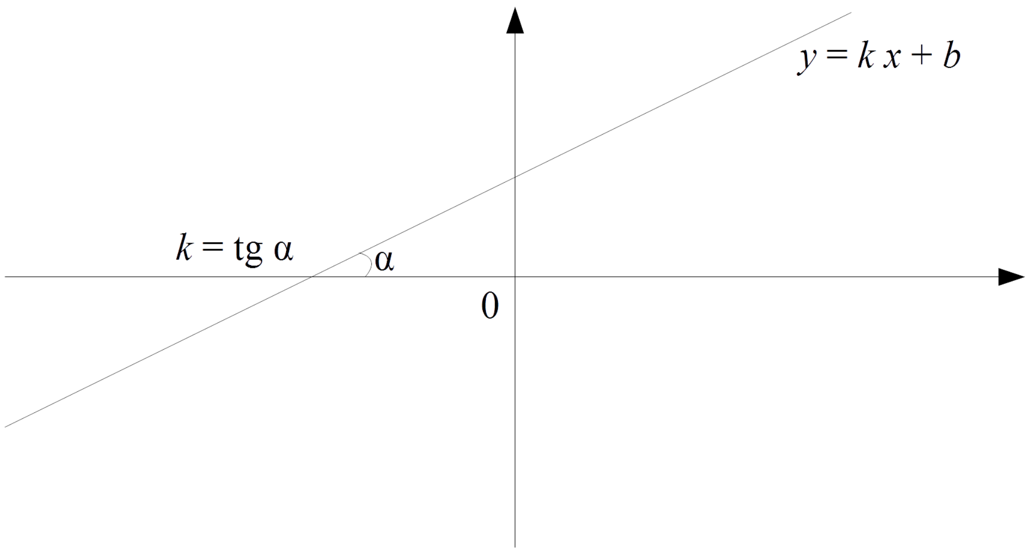


Рис. 2

При решении задач по программированию это уравнение неудобно. Во-первых, в таком виде нельзя задавать вертикальные прямые. Для них существует отдельное уравнение $x = a$. Во-вторых, даже если отдельно обрабатывать вертикальные прямые в программе, может возникнуть проблема с прямыми, которые «почти вертикальные», т.е. у которых угол с осью x близок к прямому. Тогда $\operatorname{tg}\alpha$ будет очень большим, и это может привести к проблемам с точностью. Поэтому лучше использовать общее уравнение (1), которым можно описать любую прямую.

Упражнение 3.1

Как расположены точки $(0, 0)$ и $(-1, 3)$ относительно прямой $x - y + 3 = 0$?

1. обе точки лежат на прямой,
2. одна точка лежит на прямой, другая – нет,
3. в одной полуплоскости относительно прямой,
4. в разных полуплоскостях.

Можно, конечно, построить прямую и точки графически и определить их взаимное расположение. Для построения прямой можно подобрать любые две точки, удовлетворяющие ее уравнению, и провести через них прямую. Но такой метод недоступен, если нужно написать программу, которая бы умела решать задачу для разных точек и прямых. Обсудим, как ответить на вопрос задачи аналитически, по уравнению.

Подставим координаты точек в уравнение прямой:

$$\begin{aligned}
&x - y + 3 = 0, \\
&(x, y) = (0, 0): \qquad 0 - 0 + 3 = 3 > 0, \\
&(x, y) = (-1, 3): \qquad -1 - 3 + 3 = -1 < 0.
\end{aligned}$$

Они не удовлетворяют уравнению, поэтому обе точки не лежат на прямой. Более того, левая часть уравнения для них принимает разные знаки. Оказывается, из этого следует, что точки лежат в разных полуплоскостях относительно прямой. Докажем это.

Выберем произвольную точку, не принадлежащую прямой, и будем непрерывно перемещать ее по полуплоскости, в которой она лежит. При этом левая часть уравнения $Ax + By + C$ будет тоже меняться непрерывно. Если мы не пересекаем прямую, то эта величина не будет обращаться в ноль, а значит, не поменяет знак (будет либо всегда положительной, либо всегда отрицательной). Получается, что для всех точек, лежащих в одной полуплоскости, левая часть уравнения прямой имеет одинаковый знак. Но для наших двух точек знаки разные, поэтому они не могут лежать в одной полуплоскости.

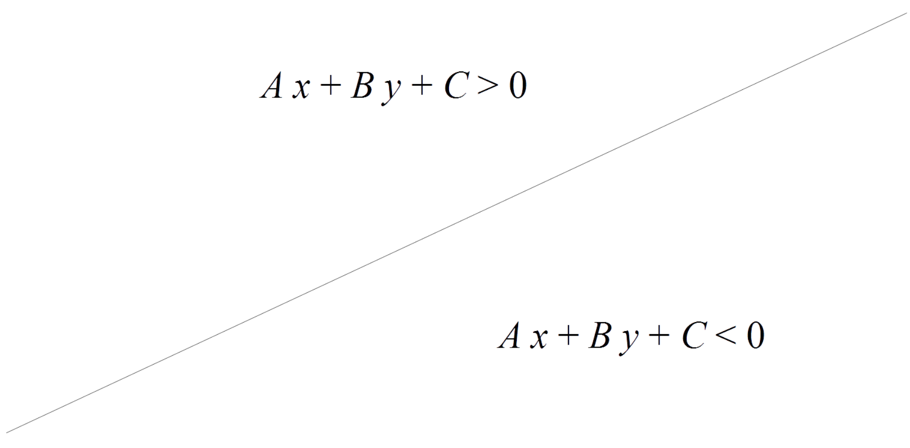


Рис. 3

Это, конечно, не совсем строгое доказательство, потому что мы понимаем непрерывное движение точки чисто интуитивно. Тем не менее, верен тот факт, что точки, расположенные в разных полуплоскостях относительно прямой, всегда дают разные знаки величины $Ax + By + C$, а точки в одной полуплоскости — одинаковый знак. В общем случае полуплоскости, на которые прямая делит плоскость, задаются неравенствами $Ax+By+C > 0$ и $Ax+By+C < 0$.

Рассмотрим теперь другую задачу: как провести прямую через две заданные точки. Пусть координаты данных точек — (x_1, y_1) и (x_2, y_2), а (A, B, C) — неизвестные нам пока коэффициенты прямой. Тогда ясно, что обе точки должны удовлетворять уравнению прямой, т.е. справедливы равенства:

$$\begin{aligned} Ax_1 + By_1 + C = 0, \\ Ax_2 + By_2 + C = 0. \end{aligned} \tag{2}$$

Отнимем второе равенство от первого. Коэффициент C при этом уничтожится, и получится соотношение

$$A(x_1 - x_2) + B(y_1 - y_2) = 0.$$

Чтобы коэффициенты A и B удовлетворяли последнему соотношению, можно взять, например,

$$A = y_1 - y_2, \quad B = x_2 - x_1. \tag{3}$$

Тогда C можно определить из уравнения (2):

$$C = -(Ax_1 + By_1). \tag{4}$$

Формулы (3) и (4) можно использовать для построения коэффициентов прямой по двум точкам в программе.

В этой задаче есть важный частный случай: когда точки совпадают. В этом случае по формулам (3) и (4) получаем $A = B = C = 0$. Тогда уравнение $Ax + By + C = 0$ не задаёт прямую, ему удовлетворяют все точки плоскости. Поэтому обычно считают, что коэффициенты A и B в уравнении прямой одновременно не равны нулю.

Упражнение 3.2
Какое множество точек описывает уравнение $Ax + By + C = 0$, если $A = 0$, $B = 0$, $C \neq 0$?

Заметим, что коэффициенты прямой определяются неоднозначно, т.е. одной прямой могут соответствовать разные уравнения. Действительно, при умножении A, B и C на любое ненулевое число k, уравнение будет описывать ту же самую прямую:

$$\begin{aligned} &Ax + By + C = 0, \\ &(kA)x + (kB)y + kC = 0. \end{aligned}$$

Это необходимо иметь в виду при сравнении прямых в программе. С этой целью прямые обычно нормируются. Если коэффициенты прямой — вещественные числа, можно добиться выполнения условия

$$A^2 + B^2 = 1,$$

разделив A, B и C на коэффициент $\sqrt{A^2 + B^2}$:

$$A' = \frac{A}{\sqrt{A^2 + B^2}}, \quad B' = \frac{B}{\sqrt{A^2 + B^2}}, \quad C' = \frac{C}{\sqrt{A^2 + B^2}}.$$

Знаменатель не равен нулю, потому что, как мы договорились, коэффициенты A и B в уравнении прямой одновременно не обращаются в ноль. Если коэффициенты прямой — целые числа, то можно разделить их на их наибольший общий делитель и таким образом нормировать прямую.

Обсудим, как найти точку пересечения двух прямых. Пусть прямые имеют коэффициенты (A_1, B_1, C_1) и (A_2, B_2, C_2). Тогда их общая точка (x, y) будет удовлетворять обоим уравнениям:

$$A_1x + B_1y + C_1 = 0,$$
$$A_2x + B_2y + C_2 = 0.$$

У нас получилась система двух уравнений с двумя неизвестными. Можно умножить первое уравнение на B_2, второе — на B_1, и вычесть второе уравнение из первого:

$$(A_1B_2 - A_2B_1)x + C_1B_2 - C_2B_1 = 0.$$

Отсюда можно выразить x:

$$x = -\frac{C_1B_2 - C_2B_1}{A_1B_2 - A_2B_1}. \tag{5}$$

Аналогично выражаем y:

$$y = -\frac{A_1C_2 - A_2C_1}{A_1B_2 - A_2B_1}. \tag{6}$$

Полученные формулы дают решение задачи о пересечении двух прямых.

Заметим, что в числителе и знаменателе формул (5) и (6) стоят похожие выражения. Эти выражения называются определителями 2×2 и обозначаются следующим образом:

$$C_1B_2 - C_2B_1 = \begin{vmatrix} C_1 & B_1 \\ C_2 & B_2 \end{vmatrix},$$

$$A_1C_2 - A_2C_2 = \begin{vmatrix} A_1 & C_1 \\ A_2 & C_2 \end{vmatrix},$$

$$A_1B_2 - A_2B_2 = \begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}.$$

Вообще **определителем** матрицы 2×2, состоящей из некоторых элементов a, b, c и d, называется следующая величина:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Вернёмся к задаче о пересечении прямых. Посмотрим, какие в этой задаче возможны крайние случаи. Прямые могут быть параллельны или совпадать. Что будет в этих случаях с нашим решением? Второй случай проще. Для равных прямых наборы коэффициентов пропорциональны, т.е. существует такой коэффициент k, что

$$A_1 = kA_2, \quad B_1 = kB_2, \quad C_1 = kC_2.$$

Если в матрице две строки пропорциональны, то её определитель равен нулю. В этом можно убедиться непосредственной проверкой:

$$\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} = \begin{vmatrix} kA_2 & kB_2 \\ A_2 & B_2 \end{vmatrix} = kA_2B_2 - A_2kB_2 = 0.$$

Значит, и числители, и знаменатели дробей в решении (5)–(6) равны нулю и возникает неопределённость:

$$x = -\frac{\begin{vmatrix} C_1 & B_1 \\ C_2 & B_2 \end{vmatrix}}{\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}} = \frac{0}{0}, \quad y = -\frac{\begin{vmatrix} A_1 & C_1 \\ A_2 & C_2 \end{vmatrix}}{\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}} = \frac{0}{0}.$$

Это неудивительно, потому что две равные прямые имеют бесконечное число общих точек.

Тема 3.2 Векторы. Скалярное и векторное произведения

Вектор — это направленный отрезок (см. рис. 4).

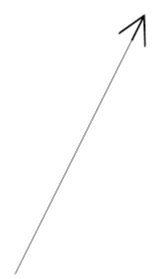


Рис. 4

Рис. 5

Векторы считаются равными, если их можно совместить параллельным переносом (см. рис. 5), т.е. неважно, от какой точки вектор отложен.

В декартовой системе координат вектор задается парой чисел (x, y) — это координаты точки, в которую попадает конец вектора, если вектор отложен от начала координат (см. рис. 6).

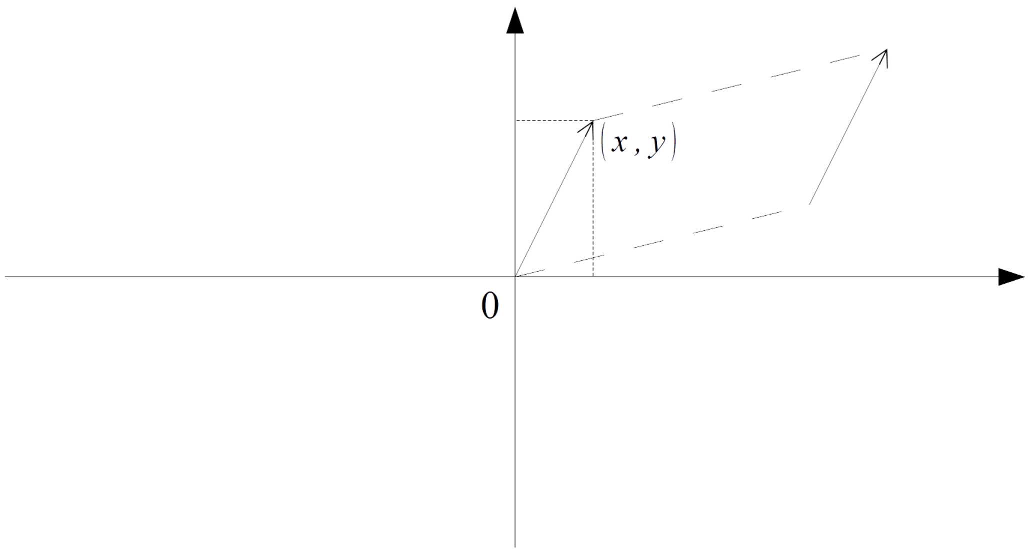


Рис. 6

Иначе координаты вектора можно найти, как разности координат его конца и начала (см. рис. 7)

Вектор можно умножать на число. При этом его длина увеличивается в данное число раз и направление меняется на противоположное, если число отрицательное. Координаты при умножении просто умножаются на это число (см. рис. 8).

Также векторы можно складывать по правилу треугольника, как показано на рис. 9. Координаты векторов при этом складываются.

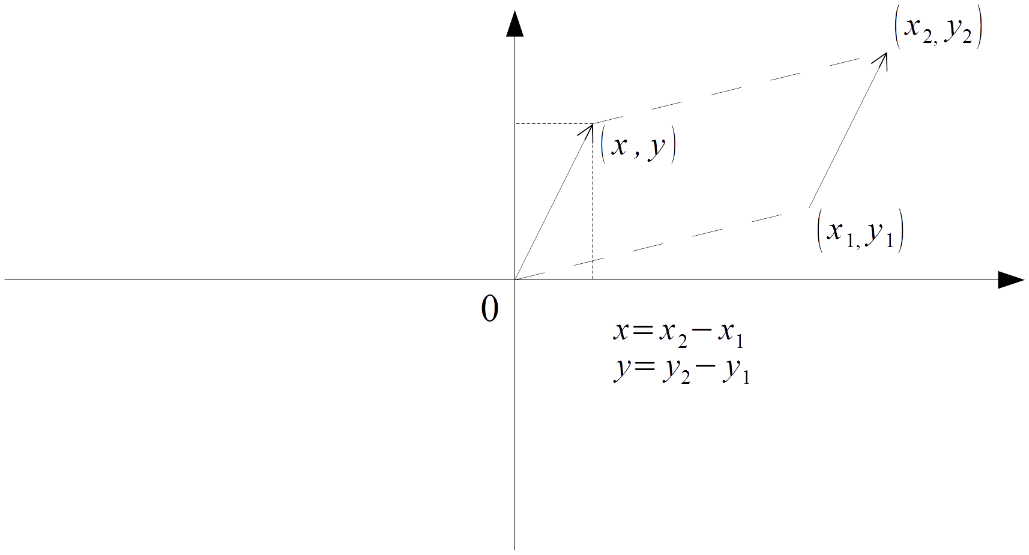


Рис. 7

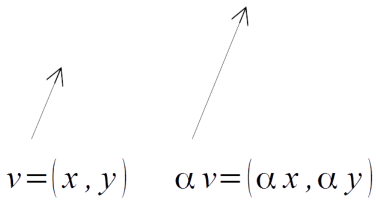


Рис. 8. Умножение вектора на число

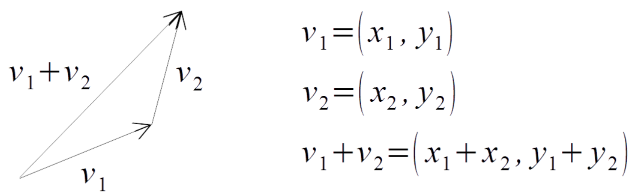


Рис. 9. Сложение векторов

В результате выполнения операций может получиться нулевой вектор, у которого начало и конец совпадают.

Длина вектора вычисляется по теореме Пифагора (см. рис. 10). Будем обозначать длину вектора $|v|$.

Другая важная операция с векторами — **скалярное произведение**. Скалярное произведение двух векторов $v_1 = (x_1, y_1)$ и $v_2 = (x_2, y_2)$ обозначается (v_1, v_2) и вычисляется по формуле

$$(v_1, v_2) = x_1 x_2 + y_1 y_2.$$

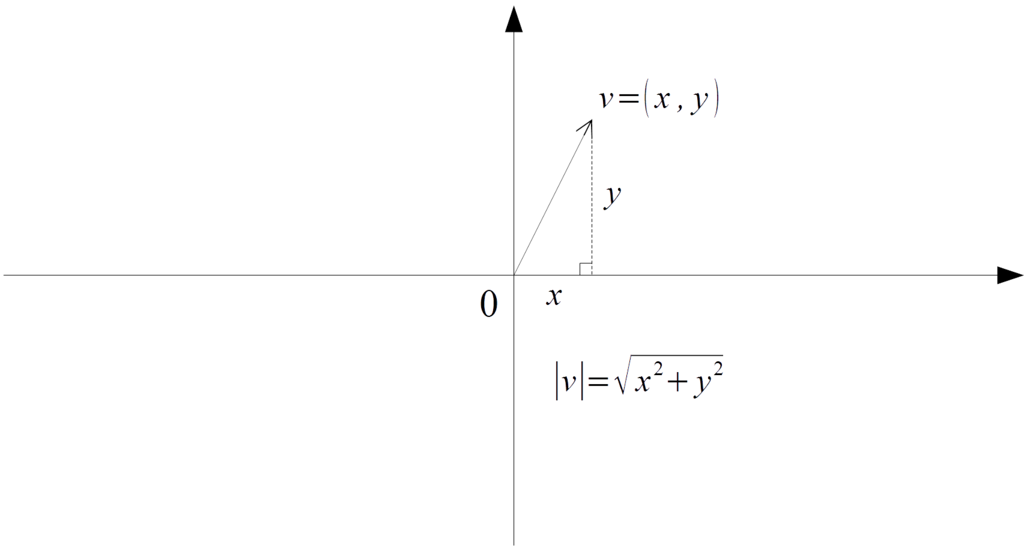


Рис. 10. Длина вектора

С другой стороны, скалярное произведение равно произведению длин векторов на косинус угла между ними:

$$(v_1, v_2) = |v_1||v_2| \cos(\widehat{v_1, v_2}).$$

Если два вектора перпендикулярны, угол между ними равен $90°$ и $\cos(\widehat{v_1, v_2}) = 0$. Значит, скалярное произведение для перпендикулярных векторов равно нулю. При помощи скалярного произведения удобно проверять векторы на перпендикулярность.

А как определить, что два вектора параллельны, или, как ещё говорят, коллинеарны (см. рис. 11)? Для этого можно использовать **векторное произведение**, которое обозначается и определяется следующим образом:

$$v_1 = (x_1, y_1), \quad v_2 = (x_2, y_2),$$
$$[v_1, v_2] = x_1 y_2 - x_2 y_1.$$

Заметим, что

$$[v_1, v_2] = \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}.$$

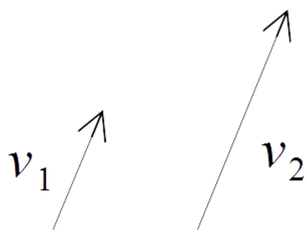


Рис. 11

Заметим, что если два вектора коллинеарны, то один можно получить из другого умножением на число: $v_1 = \alpha v_2$. Можно проверить, что определитель матрицы, у которой столбцы пропорциональны друг другу, равен нулю:

$$\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = \begin{vmatrix} \alpha x_2 & x_2 \\ \alpha y_2 & y_2 \end{vmatrix} = \alpha x_2 y_2 - \alpha y_2 x_2 = 0.$$

Существует другая формула для векторного произведения:

$$[v_1, v_2] = |v_1||v_2| \sin(\widehat{v_1, v_2}).$$

Два вектора коллинеарны тогда и только тогда, когда $\sin(\widehat{v_1, v_2}) = 0$. Поэтому для проверки векторов на коллинеарность можно просто проверять на равенство нулю их векторное произведение.

Скалярное и векторное произведения часто бывает удобно использовать для решения задач.

Упражнение 3.3
Даны четыре точки $A = (-1, 1)$, $B = (2, 3)$, $C = (3, 2)$, $D = (2, 3)$. Найдите среди них все тройки, лежащие на одной прямой.

Существуют разные способы определить, лежат ли три точки на одной прямой. Например, можно провести прямую через две точки, построить ее уравнение и проверить, принадлежит ли этой прямой третья точка. Но это решение имеет один частный случай, который надо обрабатывать отдельно. А именно, две точки могут совпадать, и тогда через них нельзя провести прямую. Поэтому здесь удобно воспользоваться векторным произведением.

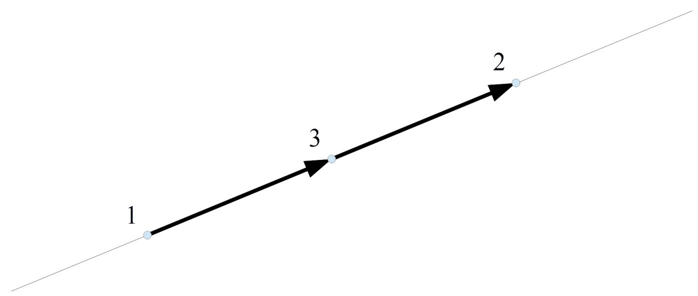


Рис. 12

Найдём координаты векторов, проведенных из первой точки во вторую и из первой точки в третью (см. рис. 12). Точки лежат на одной прямой тогда и только тогда, когда эти векторы коллинеарны. Проверим это с помощью их векторного произведения. Считается, что нулевой вектор коллинеарен с любым другим, и наше решение будет корректно работать в этом частном случае.

Вернёмся к вопросу о параллельности прямых. Как параллельность прямых связана с их уравнениями? Для ответа на этот вопрос выясним геометрический смысл коэффициентов A и B в уравнении прямой. Оказывается, это координаты вектора, перпендикулярного прямой. Если прямая нормирована, т.е.

$$A^2 + B^2 = 1,$$

вектор (A, B) имеет единичную длину и называется **вектором нормали** к прямой (см. рис. 13). Докажем, что этот вектор действительно перпендикулярен прямой.

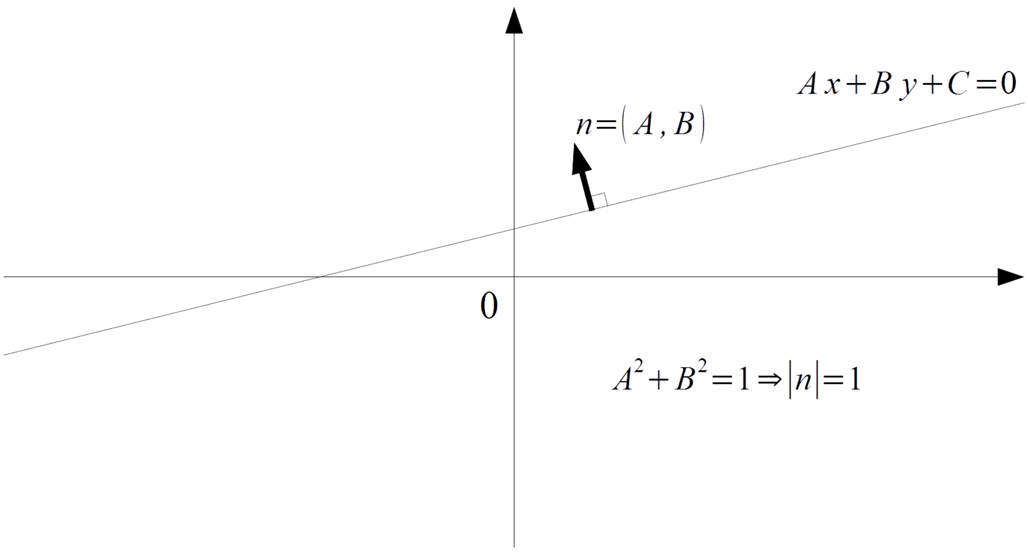


Рис. 13. Вектор нормали

Выберем на прямой две произвольные различные точки (x_1, y_1) и (x_2, y_2) и вычтем соответствующие им равенства одно из другого.

$$-\begin{cases} Ax_1 + By_1 + C = 0, \\ Ax_2 + By_2 + C = 0, \end{cases}$$

$$A(x_1 - x_2) + B(y_1 - y_2) = 0. \tag{7}$$

Тогда разности $(x_1 - x_2)$ и $(y_1 - y_2)$ дадут координаты некоторого вектора на прямой, называемого **направляющим вектором** (см. рис. 14). Заметим, что в левой части равенства (7) стоит скалярное произведение вектора с координатами (A, B) на направляющий вектор, и это скалярное произведение равно нулю. Следовательно, вектор (A, B) перпендикулярен любому направляющему вектору прямой, а значит, и самой прямой, что и требовалось доказать.

Две прямые параллельны тогда и только тогда, когда перпендикулярные им векторы коллинеарны (см. рис. 15). Это верно и для совпадающих прямых: совпадение выступает как частный случай параллельности.

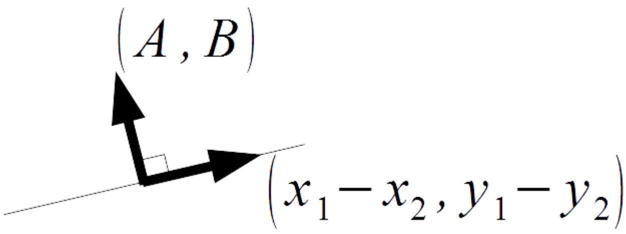


Рис. 14

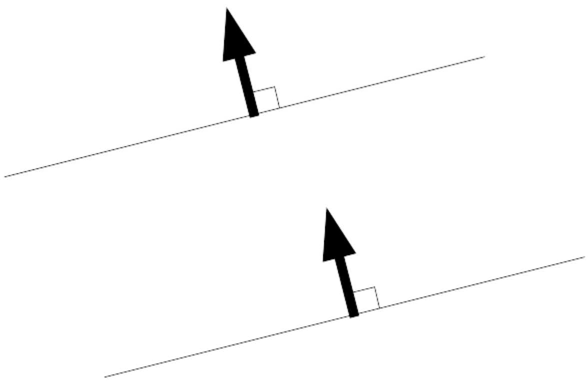


Рис. 15

Векторы нормали двух прямых имеют следующие координаты

$$n_1 = (A_1, B_1), \quad n_2 = (A_2, B_2).$$

При коллинеарности их векторное произведение будет равно нулю:

$$[n_1, n_2] = A_1B_2 - B_1A_2 = \begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} = 0.$$

Этот же определитель стоит в знаменателе формул для нахождения точки пересечения:

$$x = -\frac{\begin{vmatrix} C_1 & B_1 \\ C_2 & B_2 \end{vmatrix}}{\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}}, \quad y = -\frac{\begin{vmatrix} A_1 & C_1 \\ A_2 & C_2 \end{vmatrix}}{\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}}.$$

Получается, что при параллельных прямых наши формулы для точки пересечения не работают, чего и следовало ожидать. При решении задачи о пересечении прямых на компьютере этот случай нужно обрабатывать отдельно.

Упражнение 3.4
Рыболов Вася каждое утро ходит на берег реки ловить рыбу. Найдите расстояние, которое проходит Вася от его домика до реки по кратчайшему пути, если домик имеет координаты $(1, 3)$, а реку можно считать прямой с уравнением $x - 2y = 0$.

Выведем общую формулу для решения этой задачи. Ясно, что чтобы найти путь, по которому ходит рыболов Вася, нужно опустить перпендикуляр из точки на прямую. Вектор, соединяющий точку-домик с ближайшей точкой на реке, перпендикулярен прямой. А значит, он коллинеарен ее вектору нормали.

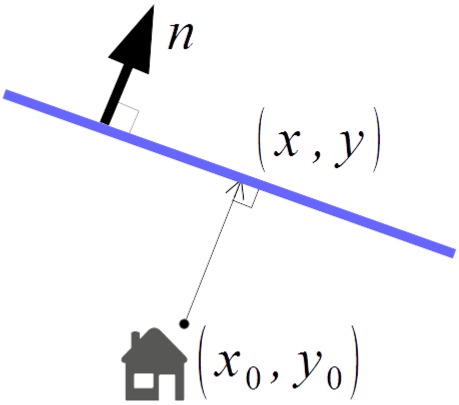


Рис. 16

Вектор нормали можно найти по уравнению прямой:

$$Ax + By + C = 0, \tag{8}$$
$$n = \left(\frac{A}{\sqrt{A^2+B^2}}, \frac{B}{\sqrt{A^2+B^2}}\right).$$

Обозначим координаты домика (x_0, y_0), а координаты неизвестной пока точки на прямой — (x, y) (см. рис. 16). Тогда выполнены следующие соотношения:

$$x = x_0 + \frac{Ad}{\sqrt{A^2+B^2}}, \quad y = y_0 + \frac{Bd}{\sqrt{A^2+B^2}}, \tag{9}$$

где число d по абсолютной величине равно длине перпендикуляра, которую мы хотим найти. Это число может быть как положительным, так и отрицательным в зависимости от того, направлен ли вектор нормали к точке-домику или от неё.

Точка (x, y) должна удовлетворять уравнению прямой (8). Подставим формулы (9) в это уравнение:

$$A\left(x_0 + \frac{Ad}{\sqrt{A^2+B^2}}\right) + B\left(y_0 + \frac{Bd}{\sqrt{A^2+B^2}}\right) + C = 0,$$

и выразим отсюда d:

$$\frac{A^2 + B^2}{\sqrt{A^2 + B^2}} d = -(Ax_0 + By_0 + C).$$

Получаем общую формулу для расстояния от точки до прямой:

$$|d| = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}}. \tag{10}$$

Упражнение 3.5
Выйдя утром на рыбалку, рыболов Вася увидел НЛО, которое приземлилось на другом берегу реки в точке, симметричной его домику относительно реки. Помогите Васе определить координаты НЛО, чтобы он успел предупредить спецслужбы. Напомним, что домик Васи находится в точке с координатами $(1, 3)$, а река представляет собой прямую, заданную уравнением $x - 2y = 0$.

Для решения этой задачи нужно взять вектор нормали к прямой, умножить его на удвоенное расстояние до прямой $2|d|$ (см. формулу (10)) и отложить этот вектор от точки-домика (см. рис. 17). Необходимо правильно учесть направление, чтобы точка-НЛО оказалась на другом берегу реки. Удобнее всего в данном случае проверить оба варианта — взять $+d$ и $-d$ и определить, в каком из случаев точки, соответствующие домику и НЛО, будут лежать в разных полуплоскостях относительно прямой-реки. Подобная задача уже разбиралась (см. Упражнение 3.1).

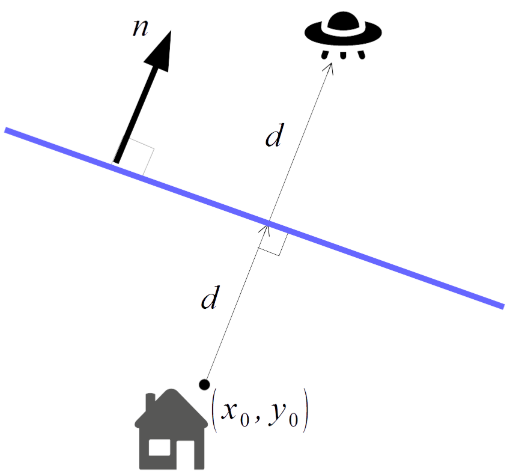


Рис. 17

Тема 3.3. Окружность

Окружность описывается уравнением

$$(x-a)^2+(y-b)^2=R^2, \tag{11}$$

где (a, b) — координаты её центра, R — радиус. Этому уравнению удовлетворяют те и только те точки (x, y), которые лежат на окружности. Равенство (11) означает, что квадрат расстояния от центра окружности (a, b) до точки (x, y) равен квадрату радиуса (см. рис. 18).

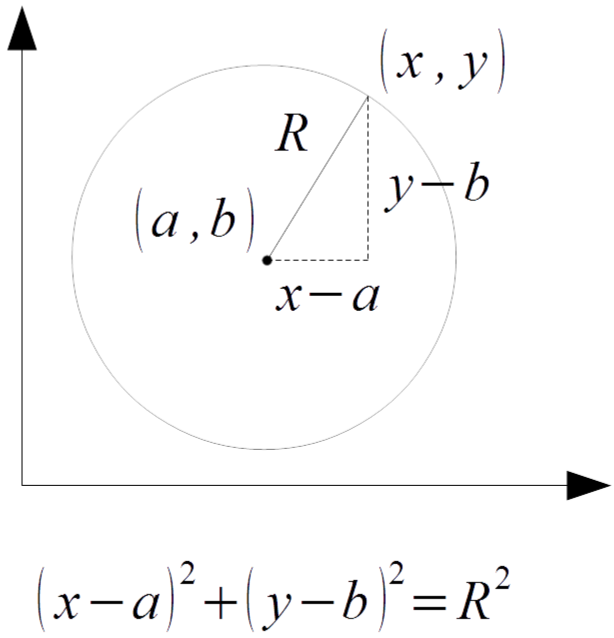


Рис. 18

Если точка лежит внутри окружности, то расстояние от неё до центра меньше радиуса, поэтому будет выполняться неравенство

$$(x-a)^2+(y-b)^2<R^2.$$

Для точки вне окружности будет выполняться неравенство

$$(x-a)^2+(y-b)^2>R^2.$$

Задача, с которой мы начнём — найти точки пересечения прямой и окружности. Пусть окружность и прямая заданы своими уравнениями:

$$\begin{aligned}&(x-a)^2+(y-b)^2=R^2,\\&Ax+By+C=0.\end{aligned}$$

Немного упростим задачу. Выполним параллельный перенос и переместим центр окружности в начало координат. Тогда её уравнение примет совсем простой вид:

$$x^2 + y^2 = R^2.$$

Выясним, что при этом будет с уравнением прямой. При параллельном переносе фактически вводятся новые координаты:

$$x' = x - a, \quad y' = y - b.$$

Отсюда выражаем

$$x = x' + a, \quad y = y' + b,$$

и подставляем в уравнение прямой:

$$A(x' + a) + B(y' + b) + C = 0.$$

Раскроем скобки и получим новое уравнение:

$$Ax' + By' + Aa + Bb + C = 0.$$

Это тоже уравнение прямой с другим коэффициентом

$$C' = Aa + Bb + C$$

вместо C. Коэффициенты A и B не изменятся. Действительно, вектор нормали при параллельном переносе не меняется. Для простоты дальнейших действий уберём штрихи и будем считать, что центр окружности изначально лежит в начале координат, и уравнение прямой задано с учётом этого факта.

Итак, нам нужно решить систему двух уравнений

$$\begin{cases} x^2 + y^2 = R^2, \\ Ax + By + C = 0. \end{cases} \tag{12}$$

с двумя неизвестными x и y. Выразим x из второго уравнения:

$$x = -\frac{By + C}{A}. \tag{13}$$

Так можно делать, только если $A \neq 0$. Но нетрудно проверить, что формулы (15) и (16), к которым мы придём в результате, работают и для $A = 0$.

Подставим выражение (13) в первое уравнение системы и раскроем скобки:

$$\left(\frac{By + C}{A}\right)^2 + y^2 = R^2,$$
$$\frac{B^2y^2 + 2BCy + C^2}{A^2} + y^2 = R^2.$$

Умножим полученное уравнение на A^2 и сгруппируем слагаемые с одинаковыми степенями y:

$$(A^2 + B^2)y^2 + 2BCy + C^2 - A^2R^2 = 0.$$

Мы получили квадратное уравнение. Найдём его дискриминант и два корня:

$$D = A^2(R^2(A^2 + B^2) - C^2), \tag{14}$$

$$y_{1,2} = \frac{-BC \pm A\sqrt{R^2(A^2 + B^2) - C^2}}{A^2 + B^2}. \tag{15}$$

Подставим (15) в (13) и найдём соответствующие значения x:

$$x_{1,2} = \frac{-AC \mp B\sqrt{R^2(A^2 + B^2) - C^2}}{A^2 + B^2}. \tag{16}$$

Полученные формулы (15) и (16) дают решение задачи. По ним можно найти координаты точек пересечения окружности и прямой (x_1, y_1) и (x_2, y_2).

Для взаимного расположения прямой и окружности возможны три варианта: они пересекаются в двух точках, касаются или не имеют общих точек (см. рис. 19).

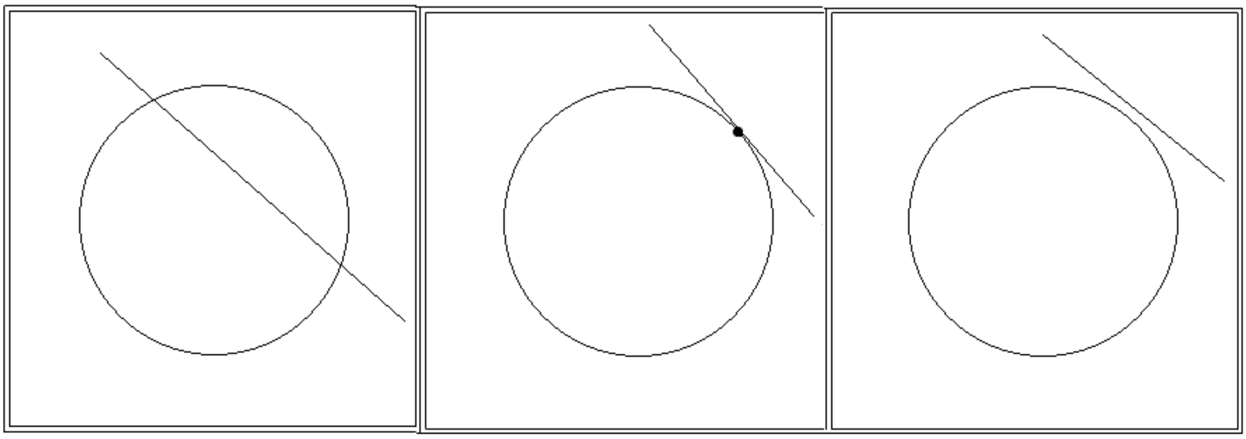


Рис. 19

Те же случаи возможны и для квадратного уравнения: оно либо имеет два корня, когда дискриминант положителен, либо имеет один корень, когда дискриминант равен нулю, либо не имеет корней, когда дискриминант отрицателен. Оказывается, эти три алгебраических случая в точности соответствуют трём геометрическим случаям (см. рис. 20).

Посмотрим на выражение под корнем в формулах (15)–(16) для точки пересечения окружности и прямой:

$$R^2(A^2 + B^2) - C^2.$$

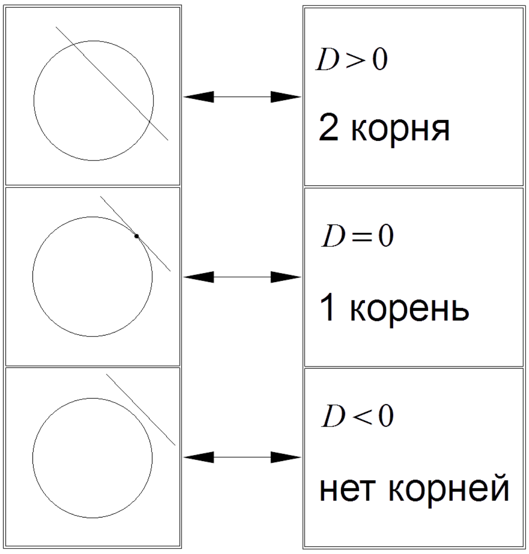


Рис. 20

Поделим его на $(A^2 + B^2)$:

$$R^2 - \frac{C^2}{A^2 + B^2}.$$

Заметим, что вычитаемое равно расстоянию от центра окружности до прямой, возведенному в квадрат. Это нетрудно проверить подстановкой координат центра окружности $(x_0, y_0) = (0, 0)$ в формулу (10) для расстояния от точки до прямой:

$$d^2 = \frac{C^2}{A^2 + B^2}.$$

Значит, дискриминант квадратного уравнения D (см. (14)) будет больше нуля, когда расстояние $d < R$, $D = 0$, если $d = R$, и $D < 0$, если $d > R$. С геометрической точки зрения, это как раз соответствуют трём случаям для числа точек пересечения.

Упражнение 3.6
Красная Шапочка идет в гости к бабушке по прямолинейному отрезку, соединяющему их домики. В лесу есть поляна, на которой Красную Шапочку подстерегает Серый Волк. Бабушка очень волнуется за внучку и хочет определить, пересекает ли ее путь поляну. Дайте ответ на этот вопрос, если известно, что домик Красной Шапочки имеет координаты $(1, 4)$, домик бабушки — координаты $(5, 1)$, а поляна представляет собой круг с центром $(8, 0)$ и радиусом 2. Считайте, что путь Красной Шапочки пересекает поляну, если они имеют хотя бы одну общую точку.

Для решения этой задачи совсем не обязательно пересекать прямую с окружностью. Можно найти кратчайшее расстояние от отрезка до центра окружности и сравнить его с радиусом.

Найдём наиболее близкую к центру окружности точку на прямой. Это можно сделать, как в решении задачи про рыболова Васю (Упражнения 3.4). Затем нужно проверить, попадает ли эта точка на путь Красной Шапочки, т.е. будет ли она лежать именно на заданном отрезке (см. рис. 21).

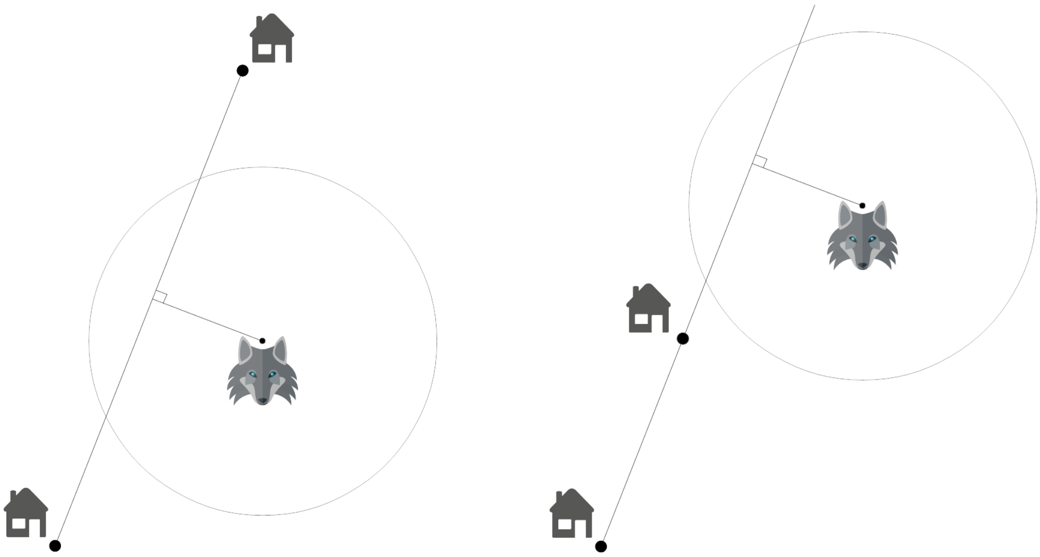


Рис. 21

Как проверить, лежит ли точка на прямой между двумя другими? Это можно сделать при помощи скалярного произведения. Возьмём два вектора с началом в проверяемой точке и концами в концах отрезка (см. рис. 22).

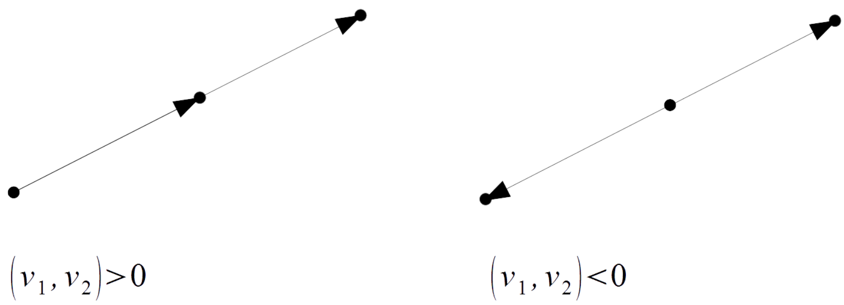


Рис. 22

Скалярное произведение векторов будет положительным, если они «смотрят в одну сторону», и отрицательным, если они «смотрят в разные стороны». Действительно, для скалярного произведения справедлива формула

$$(v_1, v_2) = |v_1||v_2| \cos(\widehat{v_1, v_2}),$$

причём

$$\cos(\widehat{v_1, v_2}) = 1, \qquad \text{если} \quad \widehat{v_1, v_2} = 0^\circ,$$
$$\cos(\widehat{v_1, v_2}) = -1, \qquad \text{если} \quad \widehat{v_1, v_2} = 180^\circ.$$

Если $(v_1, v_2) = 0$, то точка совпадает с одним из концов отрезка, а значит, попадает на отрезок. Если ближайшая точка к центру окружности попала на отрезок, нужно сравнить расстояние до неё с радиусом окружности. Если оно меньше либо равно радиусу — Красная Шапочка попадёт на поляну, иначе не попадёт.

При решении Упражнения 3.6 получается, что ближайшая к центру поляны точка не лежит на отрезке. Это означает, что минимум расстояния от центра окружности до отрезка достигается в одном из концов отрезка. Проверим, что ни один из домиков не лежит на поляне. Для этого посчитаем расстояния d_1 и d_2 и сравним их с радиусом R (см. рис. 23). Поскольку $d_1 > R$ и $d_2 > R$, кратчайшее расстояние от центра окружности до отрезка больше радиуса, а значит, опасность Красной Шапочке не угрожает.

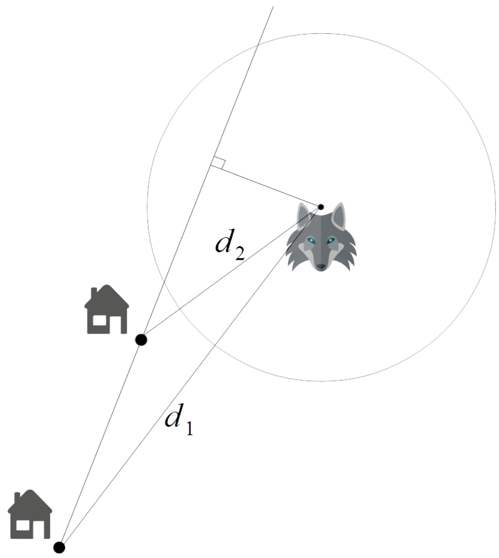


Рис. 23

Наконец, мы подошли к самой сложной из наших геометрических задач — пересечению двух окружностей. Начнём с того, что существует шесть способов взаимного расположения двух окружностей (см. рис. 24):

- одна окружность лежит внутри другой;
- касание внутренним образом;
- касание внешним образом;
- пересечение в двух точках;
- окружности не имеют общих точек и не лежат одна внутри другой;

- окружности совпадают.

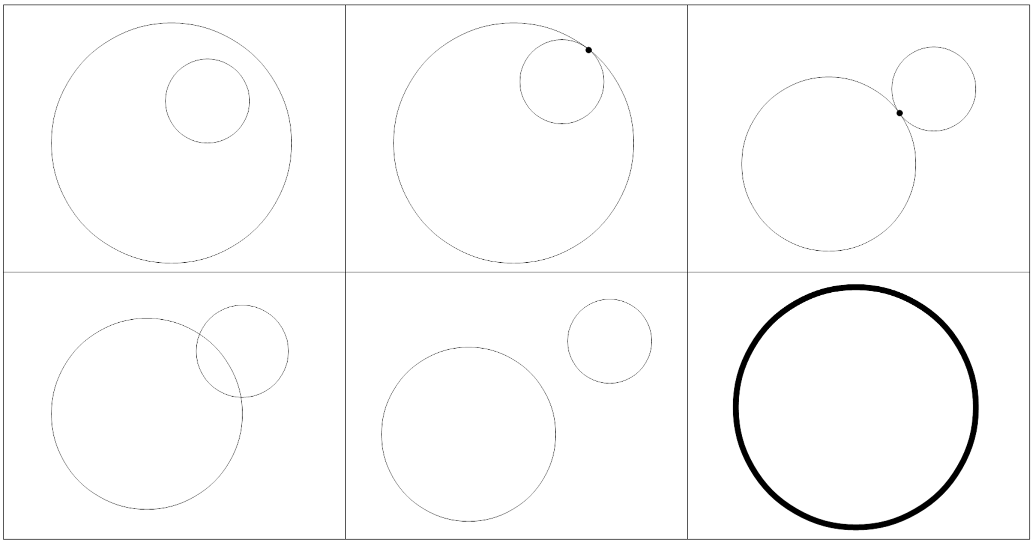


Рис. 24

Сначала нужно определить, какой из перечисленных случаев имеет место. Это можно сделать, используя расстояние между центрами окружностей и радиусы.

Упражнение 3.7
Определите тип взаимного расположения окружностей $(x-1)^2+y^2=25$, $(x+1)^2+(y+1)^2=4$.

Предположим, у нас есть две окружности, которые пересекаются в двух точках. Найдём точки пересечения. Параллельным переносом переместим центр первой окружности в начало координат (см. рис. 25).

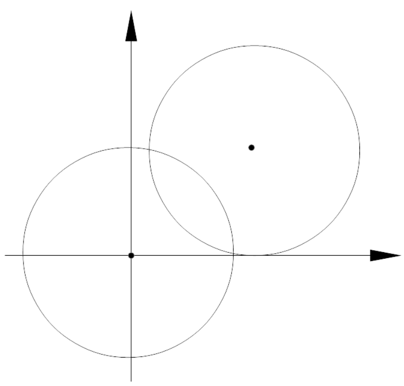


Рис. 25

Тогда окружности заданы уравнениями:

$$x^2 + y^2 = R_1^2, \tag{17}$$
$$(x-a)^2 + (y-b)^2 = R_2^2. \tag{18}$$

Вычтем (18) из (17). У нас получится соотношение

$$2ax - a^2 + 2by - b^2 = R_1^2 - R_2^2.$$

Это уравнение прямой:

$$\underbrace{2a}_{A}x + \underbrace{2b}_{B}y \underbrace{-a^2 - b^2 + R_2^2 - R_1^2}_{C} = 0. \tag{19}$$

С геометрической точки зрения, это прямая, проходящая через точки пересечения окружностей (см. рис. 26).

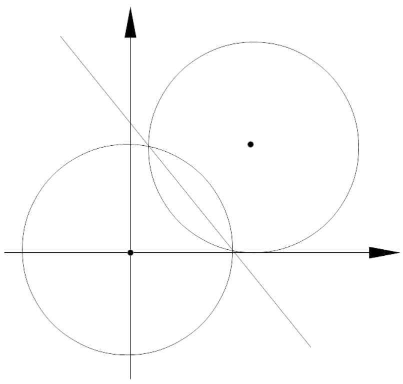


Рис. 26

Составим систему из уравнений (17) и (19). Мы свели задачу к пересечению прямой и окружности, и теперь ее можно решить по формулам (15)–(16).

Тема 4.1. Инварианты

Эта лекция посвящена инвариантам и полуинвариантам, а также их использованию для доказательства правильности программ.

Стаканы
На столе стоят 16 стаканов. При этом 15 стаканов стоят правильно, а один перевёрнут донышком вверх. Разрешается одновременно переворачивать любые 4 стакана. Можно ли при помощи таких операций добиться того, чтобы все стаканы стояли правильно?

www.problems.ru, №32042

Разберём решение этой задачи. Первым ходом можно перевернуть либо один перевёрнутый стакан и три правильных, либо четыре правильных (см. рис. 1).

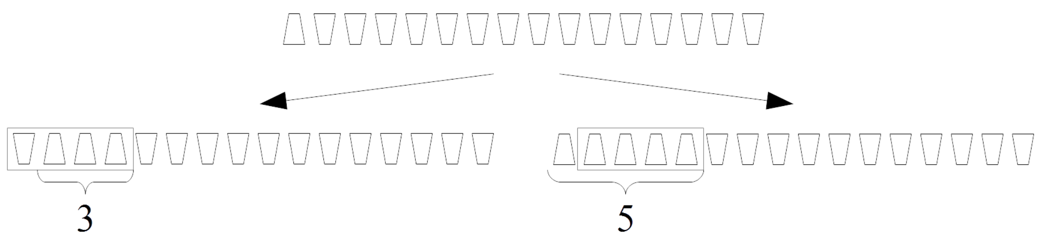


Рис. 1

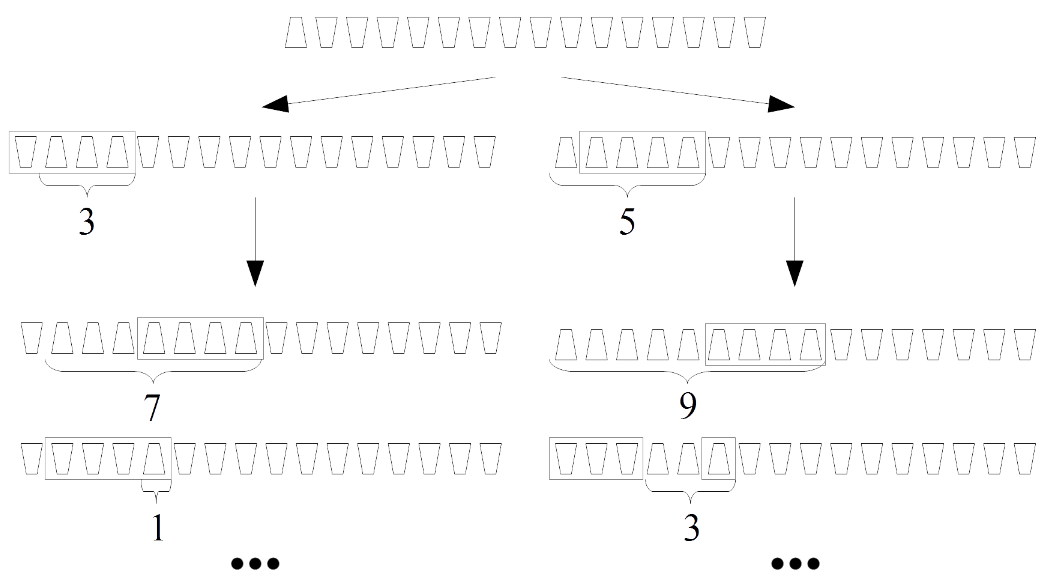


Рис. 2

Для следующего хода у нас уже больше способов (см. рис. 2), но можно заметить, что число стаканов, стоящих донышком вверх, всегда будет нечётным. Это **инвариант**, то есть условие, которое всегда выполняется. Действительно, изначально это число нечётное. При каждом ходе среди переворачиваемых стаканов количества стоящих донышком вверх и донышком вниз имеют одинаковую чётность, т.е. у нас либо чётное число перевёрнутых стаканов и чётное число правильных, либо нечётное число и тех и других. Поэтому при перевороте чётность числа стаканов, стоящих донышком вниз, не меняется. Вопрос задачи состоит в том, можем ли мы добиться нулевого количества перевёрнутых стаканов. Ответ: не можем, потому что выполняется инвариант — число перевёрнутых стаканов всегда нечётное, а нуль — чётное число.

Упражнение 4.1
100 фишек выставлены в ряд. Разрешено менять местами две фишки, стоящие через одну фишку. Можно ли с помощью таких операций переставить все фишки в обратном порядке?

www.problems.ru, №35583

Здесь можно найти много разных инвариантов. Например, проследим за движением первой фишки. Пронумеруем позиции числами от 1 до 100. После каждого хода наша фишка будет находиться на нечётной позиции. Это инвариант. Действительно, позиции через одну имеют одинаковую чётность, поэтому при всех ходах инвариант будет сохраняться. Но при обратном порядке первая фишка должна оказаться на чётной 100-й позиции. Значит, так переставить фишки невозможно. Заметим, что в этой задаче вообще не важно, что происходит с остальными фишками. Можно получить противоречие уже на одной.

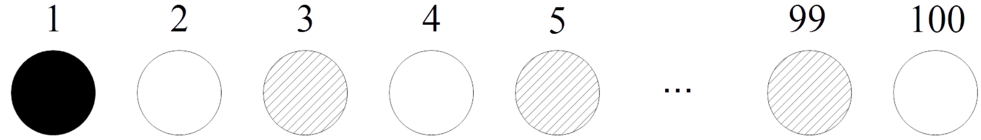


Рис. 3

Какова общая стратегия решения задач на инварианты? Во-первых, инвариант нужно заметить. Для этого можно смоделировать разные последовательности применения операций и выделить какое-то характерное условие, которое всегда выполняется. Во-вторых, нужно доказать, что инвариант будет выполняться при любом применении операций. В разобранных задачах это довольно очевидно. Строго говоря, для доказательства используется **метод математической индукции**.

Пусть нужно доказать, что некоторый инвариант будет выполняться после любого числа допустимых операций. Обозначим число операций через n и начнём доказательство по индукции. Сначала нужно доказать **базу индукции** — инвариант выполняется в самом начале, т.е. при $n = 0$. Затем доказывается **переход индукции**: показывается, что из выполнения инварианта после n операций следует его выполнение после $n + 1$ операции, т.е. что допустимые операции не влияют на инвариант. По индукции отсюда следует, что инвариант будет выполняться после любых n операций для любого целого $n \geqslant 0$.

Упражнение 4.2

На шести ёлках сидят шесть чижей, на каждой ёлке — по чижу. Ёлки растут в ряд с интервалами 10 метров. Если какой-то чиж перелетает с одной ёлки на другую, то какой-то другой чиж обязательно перелетает на столько же метров, только в обратном направлении. Могут ли все чижи собраться на одной ёлке? А если чижей и ёлок – семь?

www.problems.ru, №30754

Инвариантом здесь будет сумма номеров ёлок, на которых сидят чижи. Для шести ёлок эта сумма изначально равна $1+2+3+4+5+6 = 21$ (см. рис. 4).

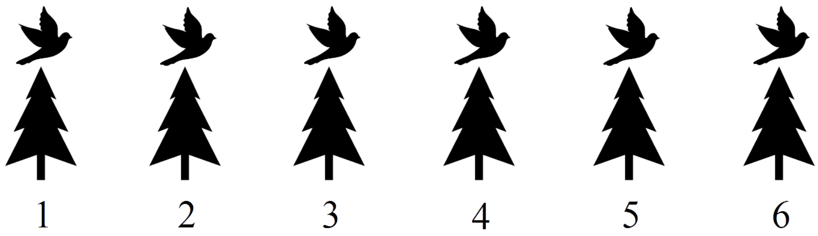


Рис. 4

При перелёте чижей слагаемые для двух чижей увеличиваются и уменьшаются на одно и то же число, поэтому сумма не меняется. Если все чижи окажутся на одной ёлке, то сумма будет делиться на 6, но для числа 21 это не выполняется, поэтому такого быть не может.

Во втором случае сумма равна $1+2+3+4+5+6+7 = 28$ и она делится на 7 — на количество ёлок, т.е. нахождение чижей на одной ёлке не препятствует выполнению инварианта. Можно даже посчитать, на какой ёлке могут находиться все чижи: на 4-й. Сам по себе этот факт ещё ничего не доказывает. Может существовать какой-то другой инвариант, который приводит к противоречию. Для доказательства положительного ответа нужен не инвариант, а подтверждающий пример. Соответствующий пример представлен на рис. 5.

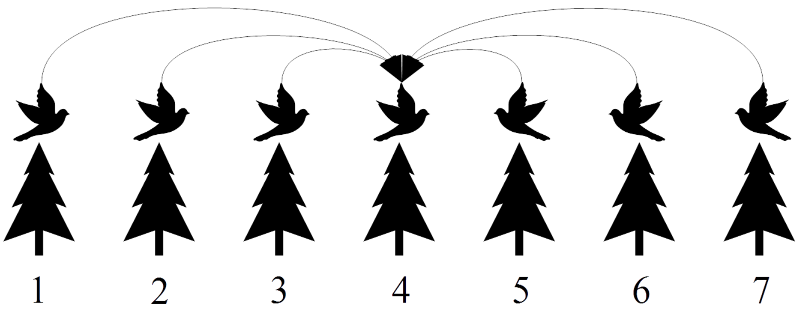


Рис. 5

Упражнение 4.3

На столе рубашкой вниз лежит игральная карта. Можно ли, перекатывая ее по столу через ребро, добиться того, чтобы она оказалась на прежнем месте, но:

1. рубашкой вверх,
2. рубашкой вниз и вверх ногами?

www.problems.ru, №108402

Раскрасим возможные позиции карты как шахматную доску (см. рис. 6).

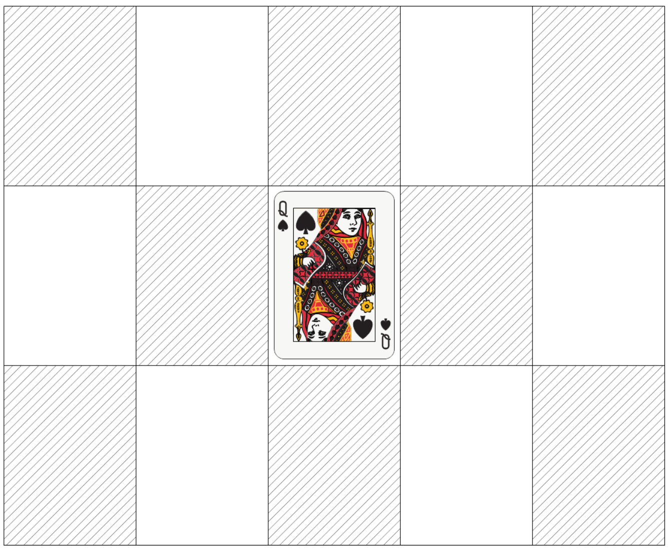


Рис. 6

Заметим, что при перемещении на соседнюю клетку карта переворачивается и оказывается рубашкой вниз на белых клетках и рубашкой вверх на чёрных. Здесь, в отличие от предыдущих задач, инвариант будет зависеть от позиции, в которой оказалась карта, или от чётности числа операций. Значит, ответ на первый вопрос — нет.

Для ответа на второй вопрос нужно проследить, как перемещаются верх и низ карты. Они не меняются при перемещении по горизонтали и каждый раз меняются местами при перемещении по вертикали. Т.е. положение карты зависит от чётности числа перемещений по вертикали. Для возврата в исходную позицию это число — чётное, поэтому карта не может лежать там вверх ногами.

Теперь задача потруднее.

Упражнение 4.4
На столе лежит куча из n ракушек. Из нее убирают одну ракушку и кучу делят на две (не обязательно поровну). Затем из какой-нибудь кучи, содержащей более одной ракушки, снова убирают одну ракушку и снова кучу делят на две. И так далее. Можно ли через несколько ходов оставить на столе только кучи, состоящие из трёх ракушек при $n = 635$ и $n = 637$?

www.problems.ru, №98630

Заметим, что после раздвоения кучки число ракушек на одну уменьшается, а число кучек на одну увеличивается. Выполняется следующий инвариант: после k ходов будет $n - k$ ракушек и $k + 1$ кучка. Нужно проверить, возможно ли равенство $n - k = 3(k + 1)$, или $n = 4k + 3$. Иначе говоря, число n должно иметь остаток 3 при делении на 4. Это верно для числа 635 и неверно для числа 637. Поэтому для второго числа сразу получаем противоречие и ответ «нет». Для первого числа недостаточно инварианта. Чтобы доказать положительный ответ, нужно привести пример. Пример легко получить, если каждый раз отделять от большой кучи кучку из трёх ракушек.

Тема 4.2. Полуинварианты

Разберём понятие полуинварианта на примере задачи.

Сержант и новобранцы
Шеренга новобранцев стояла лицом к сержанту. По команде сержанта повернуться налево некоторые повернулись налево, некоторые — направо, а остальные — кругом. Верно ли, что сержант всегда сможет встать в какое-то место в шеренге так, чтобы с обеих сторон от него было одинаковое число новобранцев, стоящих к нему лицом?

www.problems.ru, №105126

С первого взгляда в этой задаче нет никаких операций, но мы сейчас сами их введём. Пусть сержант сначала встанет слева от шеренги и дальше за одну операцию каждый раз перемещается вправо на одного человека. Обозначим через a число человек, стоящих слева лицом к сержанту, и через b — число человек, стоящих к нему лицом справа. Вначале $a = 0$, $b \geqslant 0$ (см. рис. 7).

Рис. 7. $a = 0, b = 2$

Если человек, которого проходит сержант, стоит лицом влево, то a не меняется, b уменьшается на 1 (см. рис. 8).

Рис. 8. $a = 0, b = 1$

Если очередной человек стоит лицом вправо, то a увеличивается на 1, а b не меняется (см. рис. 9).

Рис. 9. $a = 1, b = 1$

Если человек стоит спиной, то и b не меняются (см. рис. 10).

Рис. 10. $a = 1, b = 1$

Рис. 11. $a = 2, b = 1$

Рис. 12. $a = 3, b = 1$

Рис. 13. $a = 3, b = 0$

Когда сержант дойдёт до правого края, будет $b = 0, a \geqslant 0$. Если оба числа равны нулю в начале или в конце — задача решена. В противном случае разность $a - b$ в начале отрицательна, а в конце — положительна. На каждом шаге это число увеличивается на единицу или не изменяется. Значит, в какой-то промежуточный момент $a - b = 0$, т.е. $a = b$. Именно в это место в шеренге и нужно встать сержанту (см. рис. 14).

$a-b=-1$ $a-b=0$ $a-b=3$

Рис. 14

Число $a-b$ является **полуинвариантом**. В отличие от инварианта, который никогда не меняется, полуинвариант меняется в одну сторону. В данном случае всегда увеличивается. Точнее не уменьшается, и это помогает нам решить задачу.

Упражнение 4.5
На плоскости даны 10 точек: несколько из них — белые, а остальные — чёрные. Некоторые точки соединены отрезками. Назовем точку особой, если более половины соединённых с ней точек

имеют цвет, отличный от ее цвета. Каждым ходом выбирается одна из особых точек (если такие есть) и перекрашивается в противоположный цвет. Докажите, что через несколько ходов не останется ни одной особой точки.

www.problems.ru, №35056

Здесь полуинвариантом является количество отрезков, у которых концы разного цвета. Назовем такие отрезки плохими. Рассмотрим особую точку. По условию из отрезков, соединяющих ее с соседями, больше плохих, чем хороших. После перекрашивания плохие отрезки становятся хорошими, а хорошие — плохими. Поэтому число плохих отрезков каждый раз уменьшается хотя бы на единицу. Н о оно не может уменьшаться бесконечно, значит, процесс закончится.

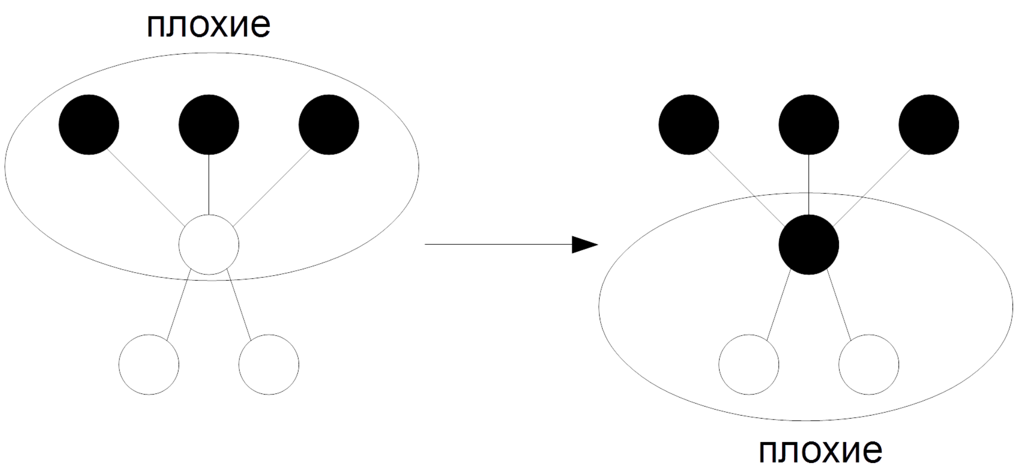


Рис. 15

Тема 4.3. Анализ алгоритмов: сортировка пузырьком

Инварианты и полуинварианты можно применять для доказательства корректности алгоритмов. Рассмотрим алгоритм сортировки пузырьком. Задача сортировки состоит в том, чтобы упорядочить числа в заданной последовательности (массиве) по возрастанию. Один из наиболее простых алгоритмов сортировки — сортировка пузырьком. Этот алгоритм работает следующим образом.

Будем выполнять проходы по массиву слева направо. При каждом проходе по порядку просматриваются пары соседних элементов. Если элементы в паре стоят по убыванию — они меняются местами. Такие проходы выполняются до тех пор, пока массив не будет отсортирован.

Рассмотрим пример. Пусть в массиве записаны числа $[3, 2, 1, 5, 4]$. При первом проходе поменяются элементы $(3, 2)$, затем $(3, 1)$ и $(5, 4)$ (см. рис. 16). При втором проходе поменяются местами элементы 1 и 2 и массив будет отсортирован.

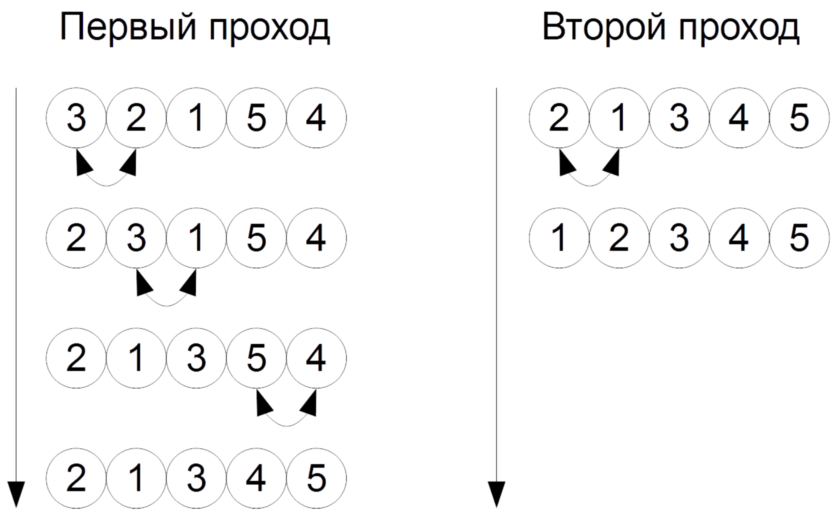


Рис. 16. Сортировка пузырьком

Этот алгоритм можно записать на псевдокоде:

пока массив не отсортирован
  для i **от** 1 **до** $n - 1$
    если $(a[i] > a[i + 1])$
      поменять местами $a[i]$ и $a[i + 1]$

Здесь $a[i]$ — i-й элемент массива, n — число элементов в массиве, i — переменная цикла, которая последовательно принимает значения 1, 2, 3, и т.д. до $n-1$. На каждом шаге цикла сравниваются соответствующие элементы — сначала $a[1]$ и $a[2]$, затем $a[2]$ и $a[3]$ и т.д. и при необходимости меняются местами. Это в точности алгоритм, который мы разобрали на примере. Для уверенности в том, что вы правильно его поняли, выполните следующее упражнение.

Упражнение 4.6
Как будет выглядеть массив $[5, 2, 1, 4, 3]$ после одного прохода по нему алгоритма сортировки пузырьком?

Докажем, что алгоритм сортировки пузырьком корректно работает, т.е. он всегда завершится и в результате получится отсортированный массив. Для простоты рассмотрим случай, когда все элементы массива — различные числа от 1 до n, как было в наших примерах. Значит, в массиве содержится перестановка из n элементов. **Инверсией** в перестановке называется пара элементов, стоящих в неправильном порядке относительно друг друга (больший элемент стоит раньше меньшего). Это не обязательно соседние элементы.

Например, в данной перестановке $(5, 2, 1, 4, 3)$ инверсии образуют, например, пары элементы $(4, 3)$, $(5, 1)$, $(5, 3)$ и т.д.

> **Упражнение 4.7**
> Сколько всего инверсий в перестановке $(5, 2, 1, 4, 3)$?

Перестановка, в которой все элементы стоят в порядке возрастания, называется **тождественной**. Количество инверсий в ней равно нулю. Например, перестановка $(1, 2, 3, 4, 5)$ — тождественная.

Заметим, что при каждой операции обмена число инверсий уменьшается ровно на одну. Действительно, инверсия между двумя соседними элементами исчезает, а на остальные инверсии обмен никак не влияет. Число инверсий будет полуинвариантом — оно каждый раз уменьшается.

пока массив не отсортирован
  для i **от** 1 **до** $n-1$
    если $(a[i] > a[i+1])$
      поменять местами $a[i]$ и $a[i+1]$
      {полуинвариант: число инверсий уменьшается}

Пока массив не отсортирован, в нем будет хотя бы одна инверсия, и можно показать, что найдётся хотя бы одна пара соседних элементов, стоящих не по порядку. Алгоритм будет продолжать работать и совершать обмены. Это будет продолжаться до тех пор, пока число инверсий не уменьшится до нуля, т.е. пока не получится тождественная перестановка. Таким образом, используя полуинвариант, мы показали, что алгоритм выполнится за конечное время и отсортирует перестановку.

Подумаем над вопросом: сколько проходов по массиву будет выполнено? Для этого нам понадобится инвариант. Проследим за наибольшим элементом массива. При первом проходе он будет меняться со всеми элементами, стоящими после него, потому что они меньше, и в результате встанет на последнее место (см. рис. 17).

При следующем проходе то же самое произойдёт с элементом $(n-1)$ и он встанет на $(n-1)$-е место (см. рис. 18), и т.д. Выполняется инвариант: после j проходов последние j элементов будут стоять на своих местах. После $(n-1)$ прохода $(n-1)$ последний элемент будет стоять на своём месте, и оставшийся первый элемент тоже окажется на своём месте.

Значит, внешний цикл «пока» выполнится не более $(n-1)$ раза. Учитывая это обстоятельство, можно внести изменения в псевдокод:

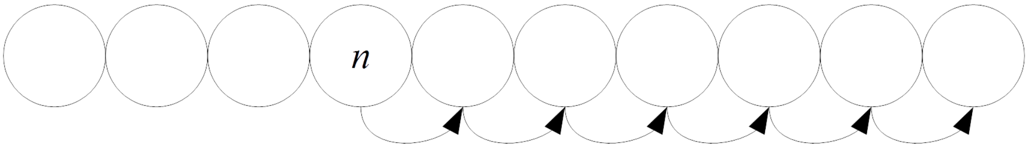


Рис. 17

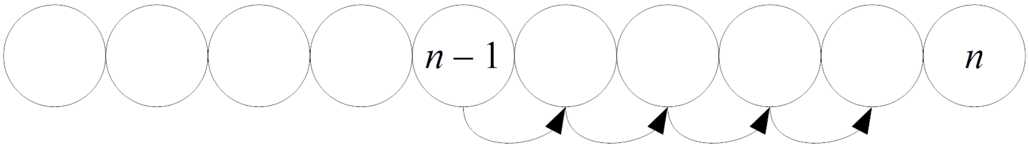


Рис. 18

для j **от** 1 **до** $n-1$
  для i **от** 1 **до** $n-j$
    если $(a[i] > a[i+1])$
      поменять местами $a[i]$ и $a[i+1]$
  {инвариант: последние j элементов стоят на своих местах}

Теперь выполняется фиксированное количество проходов по внешнему циклу. Переменная i изменяется не до $(n-1)$, а до $(n-j)$, потому что нам теперь не нужно проходить по последним элементам, которые отсортированы. Отступы от начала строки в псевдокоде характеризуют вложенность циклов, т.е. инвариант будет выполняться внутри цикла по j.

По данному псевдокоду уже можно оценить число операций сравнения, которые выполняет алгоритм. Внешний цикл выполнится не более n раз (а именно, $(n-1)$ раз). При каждом выполнении внешнего цикла внутренний тоже выполнится не более n раз. Всего будет не более n^2 действий. Эта оценка позволяет сравнивать алгоритм сортировки пузырьком с другими алгоритмами.

Тема 4.4. Анализ алгоритмов: бинарный поиск

Рассмотрим другой хорошо известный алгоритм. Пусть дан отсортированный массив и число x. Нужно определить, встречается ли число x в массиве, и если встречается, то в какой позиции. Можно, конечно, пройти все элементы в массиве и сравнить их с x, но это будет неэффективно. Хочется использовать то, что массив отсортирован.

Алгоритм бинарного поиска описывается следующим псевдокодом:

$[L, R] := [1, n]$
пока $L <= R$
    $m := (L + R)/2$
    если $(a[m] = x)$
        найден элемент x
        выход из цикла
    если $(a[m] < x)$
        $L := m + 1$
    если $(a[m] > x)$
        $R := m - 1$

Итак, имеется отрезок $[L, R]$. Изначально он равен отрезку $[1, n]$, где n — число элементов в массиве. Идея алгоритма состоит в том, чтобы на каждом шаге цикла «пока» делить отрезок пополам. При этом элемент x либо совпадает с элементом в середине отрезка, либо попадает в одну из половинок. Во втором случае отрезок $[L, R]$ заменяется на ту половинку, в которую попадает x.

На каждом шаге цикла «пока» находится середина отрезка $m = (L + R)/2$. Символы $:=$ обозначают оператор присваивания, он записывает в переменную новое значение. Деление здесь целочисленное, с округлением в меньшую сторону. Если в отрезок попадает чётное число элементов, в качестве середины будет взят левый из средних. Элемент $a[m]$ сравнивается с x. Если он равен x, то мы решили задачу. Если $a[m] < x$, то значение L присваивается $m + 1$, иначе R присваивается $m - 1$, т.е. происходит присваивание отрезка $[L, R]$ нужной половине. Отступы в начале строк отвечают за уровень вложенности. Например, все три оператора «если» вложены в цикл, т.е. проверяются все три условия, выполняются соответствующие действия, и только потом происходит переход на следующий шаг цикла. Эти операции выполняются до тех пор, пока ответ не найдётся, либо пока отрезок не станет пустым.

Почему этот алгоритм корректно работает? Если элемент x встречается в массиве, то на каждом шаге цикла выполняется следующий инвариант:

$[L, R] := [1, n]$
пока $L <= R$
    {инвариант: $a[L] <= x <= a[R]$}
    $m := (L + R)/2$
    если $(a[m] = x)$
        найден элемент x
        выход из цикла
    если $(a[m] < x)$
        $L := m + 1$

если $(a[m] > x)$
$R := m - 1$

Соответственно, индекс элемента x будет находиться в отрезке $[L, R]$.

Разберём работу алгоритма на примере. Пусть $a = [1, 3, 5, 7, 9]$, $x = 7$. Сначала $L = 1$, $R = 5$.

На первом шаге находим $m = (1 + 5)/2 = 3$. Имеем $a[3] = 5 < x$, поэтому присваиваем $L := 4$.

На втором шаге $m = (4 + 5)/2$. С учётом округления вниз, $m = 4$. Так как $a[4] = 7$, мы нашли элемент x.

Благодаря тому, что индекс элемента x всегда находится между L и R, мы его никогда не потеряем. Но как гарантировать, что программа не уйдет в бесконечный цикл? Этого не случится, потому что выполняется полуинвариант: длина отрезка $[L, R]$ уменьшается. В итоге мы или найдём элемент x, или длина станет нулевой.

$[L, R] := [1, n]$
пока $L <= R$
{инвариант: $a[L] <= x <= a[R]$}
$m := (L + R)/2$
если $(a[m] = x)$
найден элемент x
выход из цикла
если $(a[m] < x)$
$L := m + 1$
если $(a[m] > x)$
$R := m - 1$
{полуинвариант: длина отрезка $[L, R]$ уменьшается}

Алгоритм бинарного поиска напоминает то, как мы ищем нужную страницу в книге. Открываем какую-то страницу. Если её номер меньше нужного — продолжаем поиск в правой половине, иначе — в левой. Различие состоит в том, что мы не всегда делим книгу точно пополам. За счёт деления пополам в алгоритме бинарного поиска мы можем улучшить наш полуинвариант: длина отрезка $[L, R]$ на каждом шаге уменьшается не менее чем в два раза. Действительно, мы заменяем отрезок на одну из его половинок, отбрасывая при этом ещё и средний элемент.

Упражнение 4.8
Сколько раз выполнится цикл в алгоритме бинарного поиска при $n = 10$ в худшем случае?

В худшем случае каждый раз будет осуществляться переход к отрезку наибольшей длины. Для $n = 10$ изначальный отрезок равен $[1, 10]$. Его середина будет $m = 5$, поэтому возможен переход к отрезкам $[1, 4]$ и $[6, 10]$. Отрезок наибольшей длины — $[6, 10]$, поэтому предположим, что мы перешли к нему. Продолжая процесс и выбирая каждый раз наибольший отрезок, получаем варианты, представленные на рис. 19.

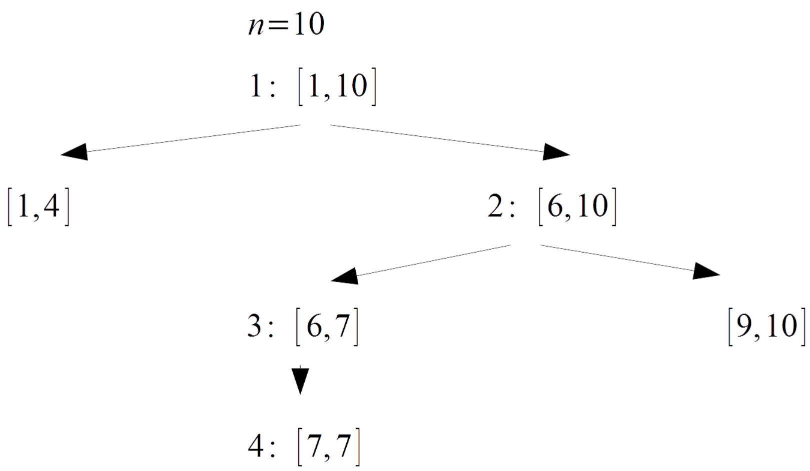


Рис. 19

Ответ в худшем случае — 4. Мы сравним с x только 4 элемента вместо 10. Можно посчитать, что для массива из 100 элементов будет только 7 сравнений, что значительно меньше, чем 100.

В общем случае количество шагов цикла будет близко к **двоичному логарифму** n, то есть такому числу $\log_2 n = k$, что $2^k = n$. Например, $\log_2 8 = 3$, потому что $2^3 = 8$; $\log_2 100 \approx 6.64385619$. Последнее число не целое, потому что 100 не является степенью двойки, но округлённо оно как раз равно 7 (числу сравнений в алгоритме бинарного поиска).

Логарифм растет достаточно медленно с ростом n. Например, при $n = 1000000$, $\log_2 n$ — число порядка 20, и бинарный поиск быстро решит задачу для массива из 1000000 элементов.

Тема 5.1. Симметричная стратегия

Эта лекция посвящена задачам на игры, похожим на те, которые встречаются на олимпиадах по программированию. Например, таким.

> **Игра со спичками**
> На столе лежит кучка спичек. Двое играют в игру. Ходят по очереди. За ход нужно забрать одну, две или три спички. Проигрывает тот, кто не может сделать ход. Кто выиграет при правильной игре?
>
> www.problems.ru, №102818

Ответ на задачу будет зависеть от начального количества спичек в кучке.

Сначала нужно пояснить вопрос задачи: что означает «выиграет при правильной игре». Вообще говоря, события в игре могут развиваться по-разному в зависимости от действий игроков. Даже при одинаковом начальном количестве спичек может так получиться, что выиграет первый игрок, а может так, что второй. Но возможна ситуация, когда один из игроков имеет выигрышную стратегию, т.е. он может так отвечать на любые ходы противника, чтобы выиграть. Мы будем считать, что если у игрока есть выигрышная стратегия, то он действует в соответствии с ней.

В данном случае выигрышной будет стратегия оставлять после своего хода на столе число спичек, делящееся на 4. Например, если начальное число спичек равно 10, то согласно этой стратегии, первый игрок должен своим ходом взять 2 спички, чтобы осталось 8 (см. рис. 1).

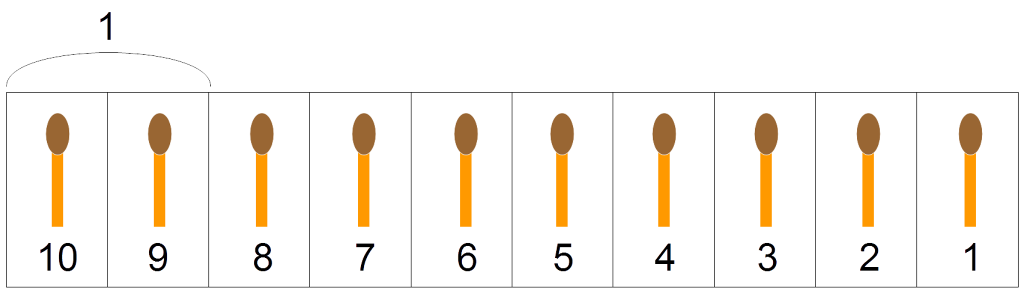


Рис. 1

После этого второй игрок возьмёт 1, 2 или 3 спички, но затем первый может всегда добрать столько спичек, чтобы осталось 4 (см. рис. 2). Теперь как бы ни ходил второй, первому останется меньше четырёх спичек. Он их возьмёт за один ход и второму больше ничего не достанется. Значит, первый выиграет (см. рис. 3).

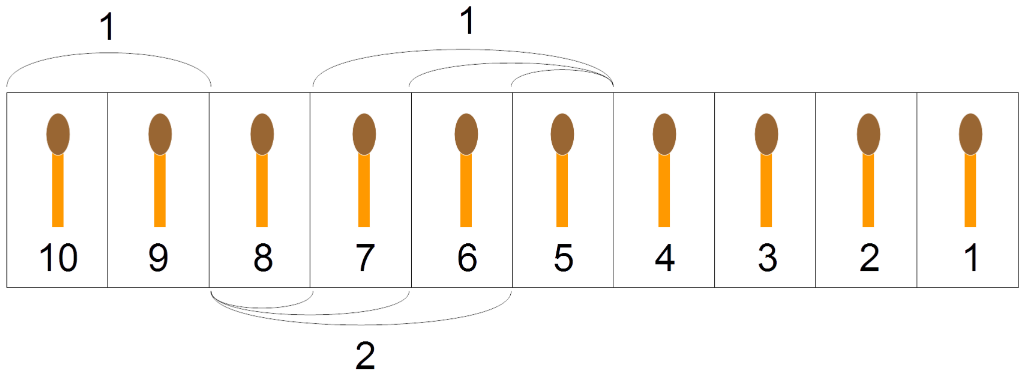


Рис. 2

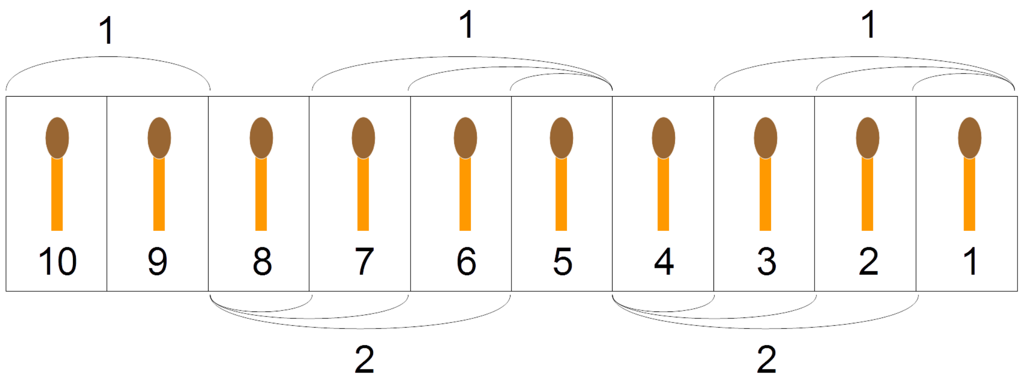


Рис. 3

Если начальное количество спичек делится на 4, то та же самая выигрышная стратегия будет работать для второго игрока. Если, например, спичек 8, то первый игрок оказывается в роли второго из примера для 10. Он уже не может следовать стратегии в свой первый ход и вынужден взять 1, 2 или 3 спички, а второй будет добирать до числа, кратного 4 (см. рис. 4).

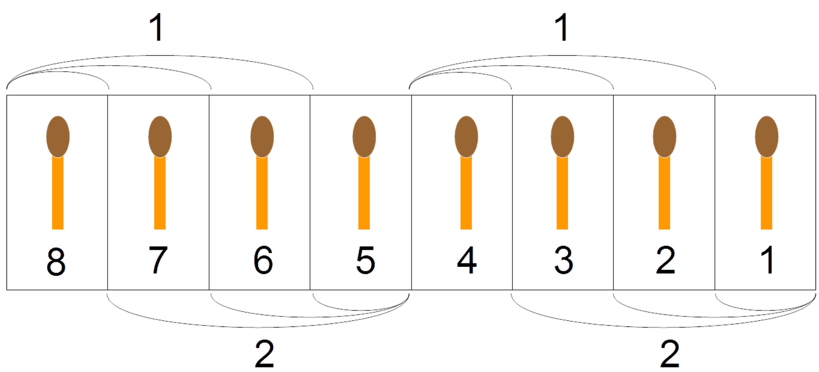


Рис. 4

Таким образом, ответ на задачу: если начальное число спичек не делится на 4, выиграет первый игрок, иначе выиграет второй.

Как догадаться до этой выигрышной стратегии? Как и во многих других задачах, здесь полезно рассмотреть случаи маленького начального количества спичек: 1, 2, 3 и т.д. Для них задачу можно решить перебором вариантов и заметить, что ответ зависит от делимости на 4, и затем разобраться, какая стратегия приводит к выигрышу.

Упражнение 5.1

Дана клетчатая доска. Двое играют в игру. За ход разрешается покрыть любые две соседние клетки доминошкой (прямоугольником 1×2) так, чтобы доминошки не перекрывались. Проигрывает тот, кто не сможет сделать ход. Кто выиграет при правильной игре, если размеры доски

1. 10×10,
2. 10×9?

www.problems.ru, №30446

Выигрышная стратегия в данной игре состоит в том, чтобы симметрично отражать ходы противника относительно центра доски (см. рис. 5).

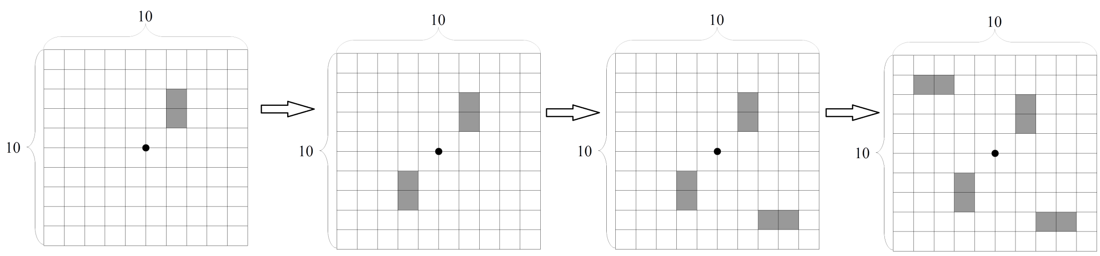


Рис. 5

В первом случае выигрывает второй игрок. Ему нужно своим ходом добиваться выполнения инварианта: в любой паре клеток, расположенных симметрично относительно центра доски, обе клетки либо свободны, либо заняты (см. рис. 6). Поддерживая этот инвариант, второй всегда имеет симметричный ход и выигрывает.

На доске 10×9 выигрывает первый. Он должен первым ходом занять две центральные клетки, а дальше придерживаться симметричной стратегии (см. рис. 7).

Симметричная стратегия очень часто встречается в задачах на игры. Попробуйте найти симметричную стратегию в следующей задаче.

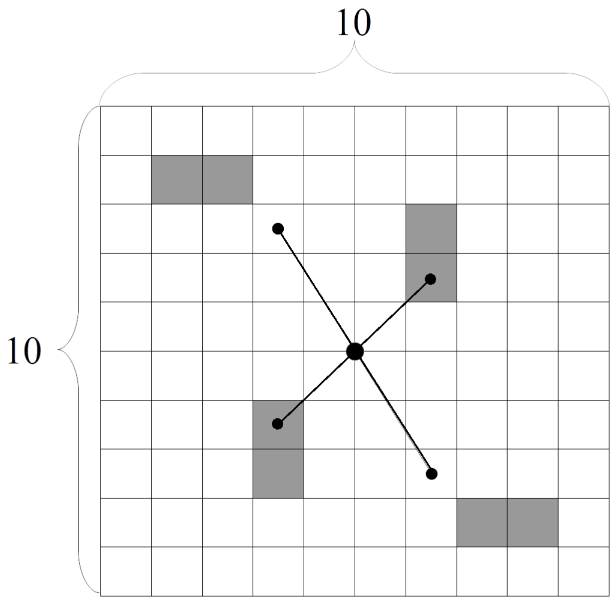


Рис. 6

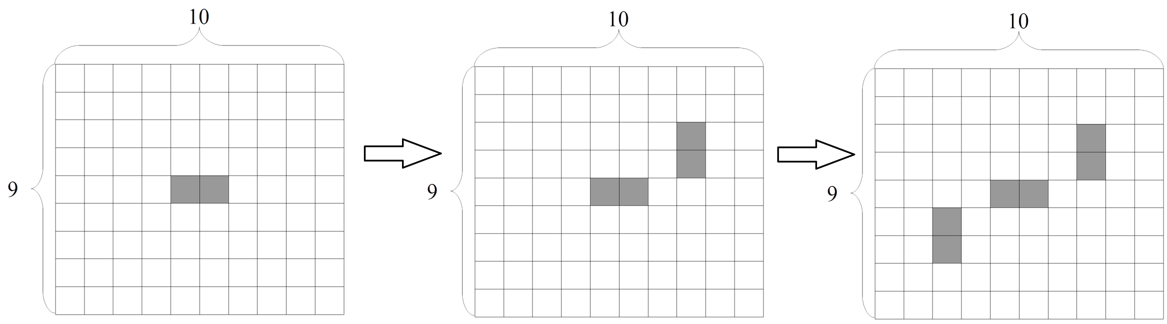


Рис. 7

Упражнение 5.2
Двое по очереди ставят слонов в клетки шахматной доски так, чтобы слоны не били друг друга. (Цвет слонов значения не имеет). Проигрывает тот, кто не может сделать ход. Кто выиграет при правильной игре? Кто выиграет в аналогичной игре с конями?

www.problems.ru, №30441
www.problems.ru, №30443

В этой задаче удобно использовать осевую симметрию. Проведем вертикальную прямую, разделяющую доску на две половины. Второй игрок сможет отражать ходы первого симметрично относительно этой прямой, т.е. всегда ставить своего слона или коня в симметричное поле другой половины. Разберём сначала задачу со слонами (см. рис. 8).

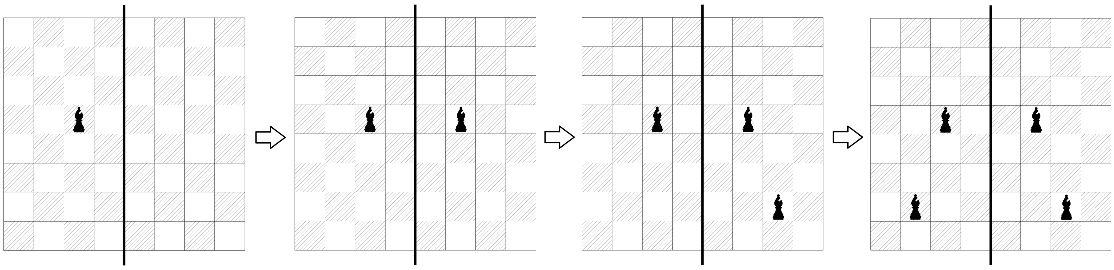


Рис. 8

При симметричной стратегии второго будет поддерживаться инвариант: битые поля расположены симметрично относительно вертикальной прямой (см. рис. 9). Если первый поставил слона на некоторое поле, то симметричное поле имеет другой цвет и новый слон не бьёт его. Значит, второй игрок всегда имеет ход и выигрывает. То же самое верно и в задаче с конями. В ней можно использовать или осевую, или центральную симметрию.

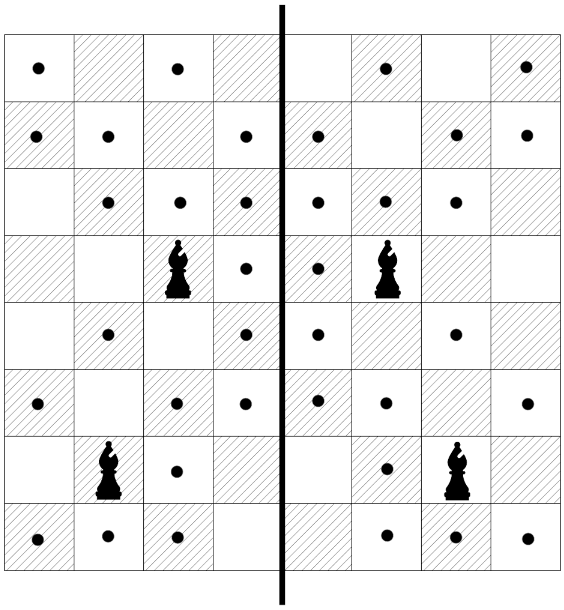


Рис. 9

Упражнение 5.3
На окружности расставлено 20 точек. Двое играют в игру. За ход разрешается соединить любые две из них отрезком, не пересекающим отрезков, проведенных ранее. Проигрывает тот, кто не может сделать ход. Кто выиграет при правильной игре?

www.problems.ru, №30449

Первым ходом нужно провести отрезок, разделяющий остальные точки на две группы по 9. Теперь каждый ход второго игрока будет происходить в одной из двух половин, а первый может отражать ходы второго в другой половине (см. рис. 10). Выиграет первый.

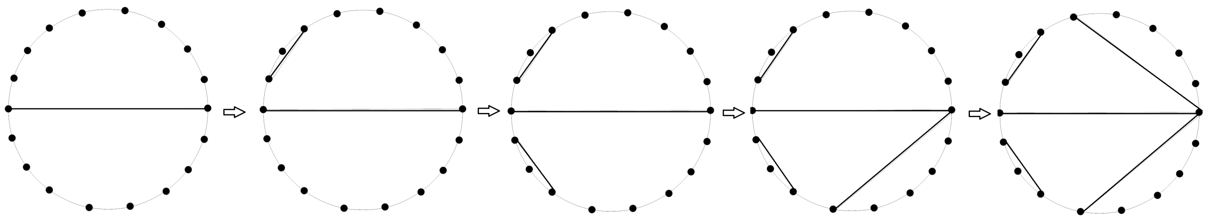


Рис. 10

Тема 5.2. Выигрышные и проигрышные позиции

Упражнение 5.4
На столе лежат n спичек. Двое играют в игру. За один ход нужно взять положительное число спичек, не превосходящее $\sqrt{k}$ спичек, где k — число спичек на столе в данный момент. Проигрывает тот, кто возьмёт последнюю спичку. Кто выиграет при правильной игре для $n = 10$ и $n = 15$?

В этой задаче не получается применить симметричную стратегию. Остаётся только один известный нам метод, хорошо себя зарекомендовавший: посмотреть, что будет для маленьких n. Заполним табличку (см. рис. 11).

n	1	2	3	4	5
	–	+	–	+	+

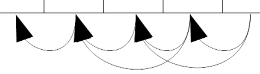


Рис. 11

- При $n = 1$ первый игрок возьмёт последнюю спичку, и эта игра будет для него проигрышной (в табличке стоит «-»).
- При $n = 2$ первый игрок возьмёт одну спичку, и после этого второй останется с одной спичкой и проиграет. Значит, игра выигрышная для первого игрока. В табличке стоит «+».
- Для $n = 3$ первый игрок берёт одну спичку, после чего второй оказывается с двумя спичками и, согласно предыдущим рассуждениям, выигрывает, а первый проигрывает.

- Из состояния с 4 спичками имеется два перехода: можно оставить второму игроку 2 или 3 спички. Если ему останется 2 спички, он выиграет, а если 3 — проиграет. Играя за первого игрока, мы хотим, чтобы второй проиграл, и оставляем ему 3 спички. Значит, 4 — выигрышное состояние.
- При $n = 5$ первый игрок тоже может оставить после своего хода 3 спички и обеспечить себе выигрыш.

Сформулируем правила, по которым заполняется таблица. **Позиция** в игре характеризуется одним числом — количеством спичек на столе. При ходах игроков происходят переходы из одной позиции в другую, т.е. получается исходная игра с меньшим числом спичек. Позиция называется **выигрышной**, если ходящий из неё игрок имеет выигрышную стратегию, и **проигрышной**, если выигрышную стратегию имеет его противник.

Начинающие олимпиадники часто привязывают к позиции также номер ходящего игрока. У нас позиция характеризуется только одним числом — количеством спичек, и неважно, чей сейчас ход. При этом мы оцениваем, выигрышная позиция или проигрышная, относительно того игрока, который из этой позиции будет ходить. Если первый игрок оказался в выигрышной позиции — выиграет первый, если второй — то выиграет второй. Так удобнее, если игроки равноправные, т.е. они ходят одинаково. Если игроки неравноправные, то, конечно, нужно привязывать к позиции номер ходящего игрока.

По каким правилам можно определить, будет ли позиция выигрышной или проигрышной?

1. *Позиция выигрышная, если из неё есть хотя бы один переход в проигрышную позицию.* Действительно, при переходе игроки меняются местами. Если мы из выигрышной позиции перейдём в проигрышную, то она будет проигрышной для нашего противника, а мы выиграем.
2. *Позиция проигрышная, если все переходы из нее ведут в выигрышные позиции.* Если мы оказались в такой позиции, то любой ход переведёт нас в позицию, выигрышную для противника, и мы в любом случае проиграем.

Из этих правил следует выигрышная стратегия для игрока, который оказался в выигрышной позиции — всегда ходить в проигрышную позицию.

Продолжим заполнять таблицу по этим правилам (см. рис. 12).

- Из позиции 6 есть переходы в позиции 4 и 5, которые обе выигрышные. Значит, 6 — проигрышная позиция.
- Из 7, 8 и 9 есть переход в проигрышную позицию 6, значит, они выигрышные.

n	1	2	3	4	5	6	7	8	9	10
	–	+	–	+	+	–	+	+	+	–

Рис. 12

- Из 10 есть переходы только в выигрышные позиции 7, 8, 9, и она проигрышная.

Получаем ответ на первый вопрос задачи: для $n = 10$ выиграет второй игрок. Продолжая заполнять таблицу дальше, определяем, что 15 — выигрышная позиция, значит, выиграет первый игрок (см. рис. 13).

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	–	+	–	+	+	–	+	+	+	–	+	+	+	–	+

Рис. 13

Отметим один важный момент. В этой игре у нас все позиции либо выигрышные, либо проигрышные, и никаких других не бывает, т.е. один из игроков всегда имеет выигрышную стратегию. Это можно строго доказать методом математической индукции. При $n = 1$ мы знаем, что позиция проигрышная по условию задачи. База индукции выполняется. Пусть мы доказали этот факт до некоторого n и хотим доказать для $(n+1)$. Из позиции $(n+1)$ переходы ведут только в позиции с меньшим количеством спичек, которые по предположению индукции либо выигрышные, либо проигрышные, и по правилам мы можем определить, что позиция $(n + 1)$ тоже либо выигрышная, либо проигрышная.

Упражнение 5.5

Имеется три кучки камней: в первой — 10, во второй — 15 и в третьей — 20 камней. Двое играют в игру. За ход разрешается разделить любую кучку на две меньшие части. Проигрывает тот, кто не сможет сделать ход. Кто выиграет при правильной игре?

www.problems.ru, №86556

В этой задаче не нужно анализировать выигрышные и проигрышные позиции. Можно заметить, что исход игры не зависит от действий игроков. Здесь работает инвариант: за ход число кучек увеличивается на одну. В начале игры 3 кучки, в конце — 45. Значит, будет выполнено 42 хода и не сможет сделать хода первый игрок, т.е. выиграет второй.

В следующей задаче уже пригодится анализ выигрышных и проигрышных позиций.

Упражнение 5.6

Игра начинается с числа 60. За ход разрешается уменьшить число на любой из его делителей. Проигрывает тот, кто получит ноль. Кто выиграет при правильной игре?

www.problems.ru, №30460

В этой игре позиция соответствует числу на доске. Заполнять таблицу до 60 долго, но можно заметить закономерность уже на маленьких начальных числах (см. рис. 14). Все нечётные числа — проигрышные, а чётные — выигрышные. Значит, при начальном числе 60 выигрывает первый игрок.

n	1	2	3	4	5	6	7	8
	–	+	–	+	–	+	–	+

Рис. 14

Но заметить закономерность — это только полдела. Нужно её доказать.

Итак, мы докажем, что все нечётные числа соответствуют проигрышным позициям, а все чётные — выигрышным. Воспользуемся методом математической индукции. База индукции — число 0. Это выигрышное состояние. Тот игрок, который получит 0 после своего хода, проигрывает, но другой игрок, который окажется в состоянии 0, при этом выигрывает.

Пусть утверждение доказано для чисел от 0 до n. Перейдём к следующему числу $(n+1)$. Если оно чётное, от него можно всегда отнять 1 и перейти в проигрышное нечётное число. В этом случае новое число будет выигрышным. Пусть теперь очередное число $(n+1)$ нечётное. Тогда все его делители тоже нечётные, и из него возможны переходы только в чётные числа, т.е. выигрышные позиции. Значит, позиция $(n+1)$ — проигрышная, и по индукции утверждение верно для любого n.

Упражнение 5.7

В коробке лежит 300 спичек. За ход разрешается взять из коробка не более половины имеющихся в нём спичек. Проигрывает тот, кто не может сделать ход. Докажите, что выиграет первый игрок, и укажите, какое количество спичек он должен взять первым ходом.

www.problems.ru, №30458

Так же как и в предыдущей задаче, начнём перебирать позиции, чтобы найти закономерность. Заметим, что 1 — проигрышная позиция, в ней одну спичку брать нельзя. Следующими проигрышными позициями будут 3, 7, 15, т.е. числа вида $2^k - 1$. Докажем методом математической индукции, что все такие позиции будут проигрышными, а остальные — выигрышными. База индукции ($n = 1$) уже доказана.

Рассмотрим некоторое количество спичек n, и пусть для всех меньших количеств мы уже доказали утверждение. Если $n = 2^k - 1$, то из этой позиции есть переходы в позиции с 2^{k-1} до $2^k - 2$. По предположению индукции, все эти позиции — выигрышные, значит, позиция n — проигрышная. Теперь пусть n не имеет вид $2^k - 1$. Тогда n можно представить в виде $2^k + m$, где m от 0 до $2^k - 2$. Заметим, что $2^k + m \leqslant 2(2^k - 1)$, поэтому, взяв не более половины спичек, из позиции n можно перейти в позицию $2^k - 1$, которая по предположению индукции является проигрышной. Следовательно, n — выигрышная позиция. По индукции получаем, что замеченная закономерность выполняется для всех позиций.

Позиция 300 выигрышная, и чтобы выиграть, первый игрок должен сделать ход в доступную проигрышную позицию $2^k - 1 = 255$. Значит, он должен взять 45 спичек.

Упражнение 5.8

Шашка стоит в левом нижнем углу шахматной доски. Двое играют в игру. За ход можно передвинуть шашку на одну клетку вправо, на одну клетку вверх или на одну клетку вправо-вверх по диагонали. Проигрывает тот, кто не может сделать ход. Кто выиграет при правильной игре на доске размера

1. 8×8,
2. 1000000×1000000?

www.problems.ru, №35426

Позиция в игре определяется клеткой, в которой стоит шашка. Таблица выигрышных и проигрышных позиций для доски 8×8 представлена на рис. 15. Левый нижней угол соответствует выигрышной позиции, поэтому выигрывает первый игрок. Для решения задачи на доске 1000000×1000000 можно заметить закономерность. Если мы пронумеруем строки и столбцы доски от правого верхнего угла, как показано на рис. 15, то проигрышными будут те и только те позиции, у которых номера строки и столбца — нечётные числа. Попробуйте доказать эту закономерность самостоятельно методом математической индукции. Таким образом, позиция $(1000000, 1000000)$ будет выигрышной.

8	7	6	5	4	3	2	1	
+	–	+	–	+	–	+	–	1
+	+	+	+	+	+	+	+	2
+	–	+	–	+	–	+	–	3
+	+	+	+	+	+	+	+	4
+	–	+	–	+	–	+	–	5
+	+	+	+	+	+	+	+	6
+	–	+	–	+	–	+	–	7
+	+	+	+	+	+	+	+	8

Рис. 15

Тема 5.3. Игры на ациклических графах

Все рассмотренные игры можно представить как игры на ациклических графах.

Граф представляет собой множество точек, называемых **вершинами**, некоторые пары из которых соединены отрезками, называемыми **рёбрами**. Пример графа представлен на рис. 16. Мы будем рассматривать **ориентированные** графы, у которых рёбра являются направленными отрезками.

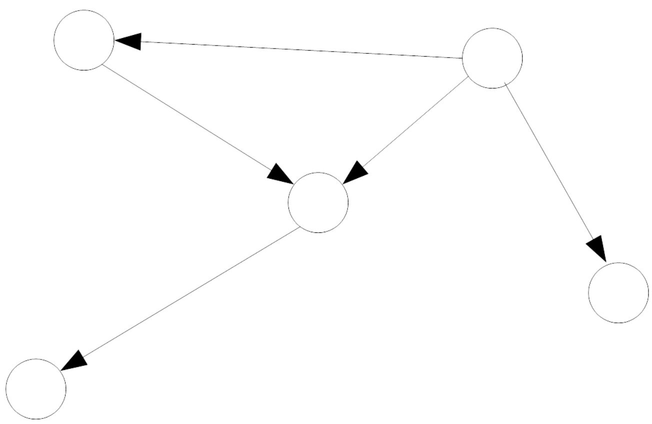
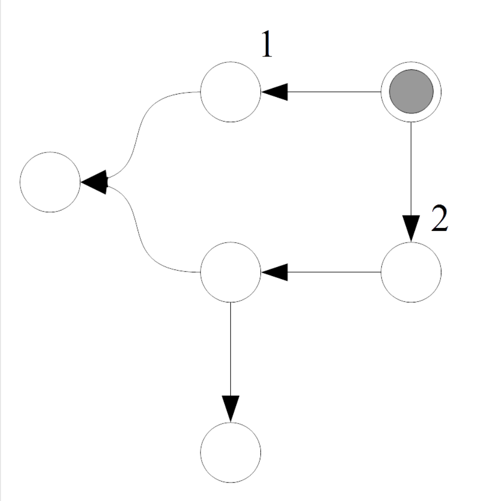


Рис. 16. Ориентированный граф

Игра на графе
Дан ориентированный граф, и в одной из его вершин стоит фишка. Два игрока ходят по очереди. За ход нужно передвинуть фишку в соседнюю вершину по одному из рёбер в соответствии с направлением ребра. Проигрывает тот, кто не сможет сделать ход.

Упражнение 5.9

Кто выиграет при правильной игре на графе, изображенном на рисунке? Если это первый игрок, то в какую вершину он должен пойти первым ходом?



Позиция в этой игре — вершина, в которой находится фишка. Те вершины, из которых нет исходящих ребер, являются проигрышными. Дальше, пользуясь обычными правилами, нетрудно определить, что исходная позиция — выигрышная, и ходить из нее нужно в проигрышную (см. рис. 17).

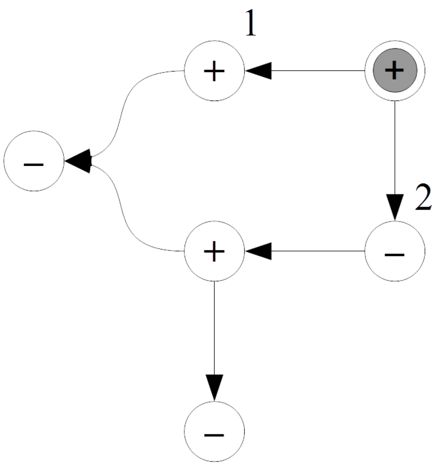
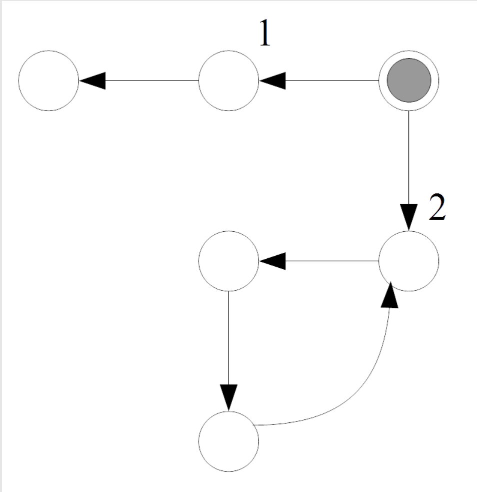


Рис. 17

Упражнение 5.10

Кто выиграет при правильной игре на графе, изображенном на рисунке? Если это первый игрок, то в какую вершину он должен пойти первым ходом?



В этой игре ни один из игроков не имеет выигрышной стратегии. В вершину 1 первый игрок ходить не будет, потому что это выигрышная позиция для второго (см. рис. 18). Если же первый пойдёт в вершину 2, то фишка попадёт в цикл и дальше будет бесконечно ходить по циклу. При этом никто не проиграет — ведь игроки всегда смогут сделать ход. Такая ситуация интерпретируется как ничья. Игры на графах с циклами сложнее, чем на графах без циклов.

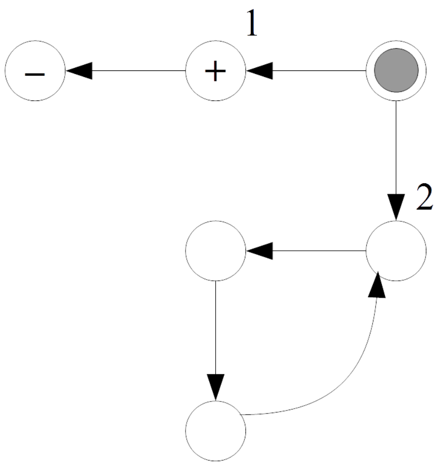


Рис. 18

Дадим формальные определения пути и цикла в графе. **Путём** в графе называется последовательность вершин v_1, v_2, $\ldots$, v_m и соединяющих их ориентированных рёбер (v_1, v_2), (v_2, v_3), $\ldots$, (v_{m-1}, v_m). т.е. из первой вершины есть ребро во вторую, из второй в третью и т.д. (см. рис. 19).

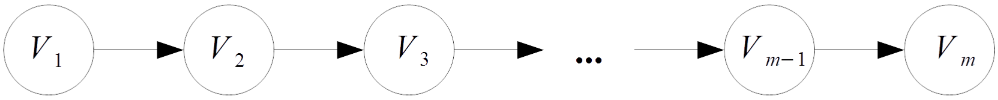


Рис. 19. Путь

Цикл — это путь, у которого первая и последняя вершины совпадают (см. рис. 20).

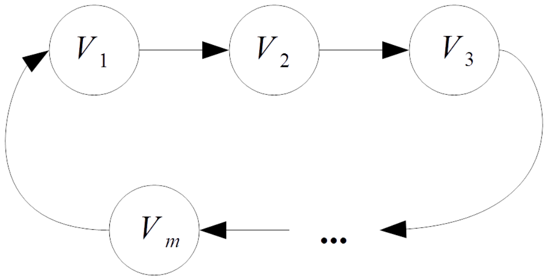


Рис. 20. Цикл

Ориентация рёбер в цикле важна. Например, граф, представленный на рис. 21, не является циклом.

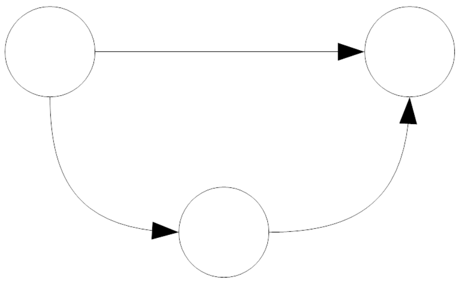


Рис. 21

Граф называется **ациклическим**, если в нём нет циклов.

В нашем курсе мы ограничимся изучением игр на ациклических графах. Игры на графах с циклами — более сложная тема. Как мы убедились на примере, в них возможны ничьи, хотя также могут быть выигрышные и проигрышные позиции.

Покажем, что если граф ациклический, то игра с фишкой на нём закончится за конечное число ходов, и все вершины ациклического графа соответствуют выигрышным и проигрышным позициям. Для этого нам понадобится топологическая сортировка вершин ациклического графа.

Задача **топологической сортировки** состоит в том, чтобы пронумеровать вершины графа так, чтобы каждое ориентированное ребро вело из вершины с бóльшим номером в вершину с меньшим номером. На рис. 22 представлен пример нумерации графа из Упражнения 5.9.

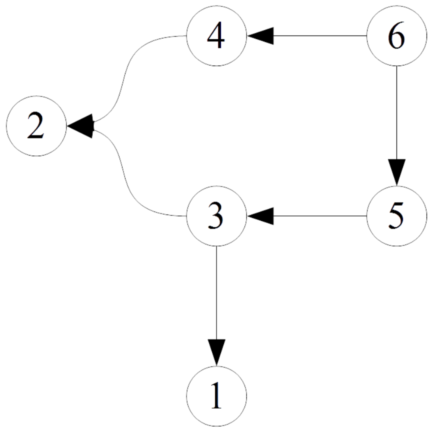


Рис. 22. Топологическая сортировка

Заметим, что если в графе есть цикл, то топологическая сортировка для него невозможна. Действительно, если мы присвоили вершине некоторый номер, то все вершины, из которых существует путь в неё, имеют бóльшие номера. Но если наша вершина находится в цикле, из неё существует путь ненулевой длины в саму себя, и это приводит к противоречию (см. рис. 23).

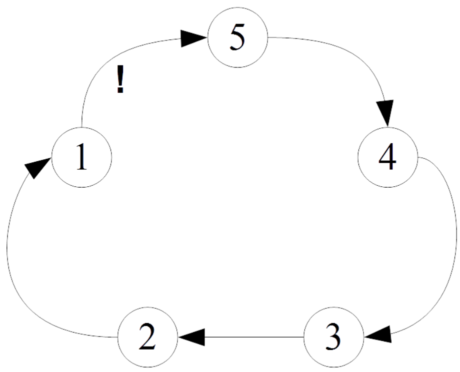


Рис. 23

С другой стороны, для ациклических графов верна следующая

Теорема. *Для любого ациклического графа возможна топологическая сортировка.*

Доказательство. Покажем сначала, что в любом ациклическом графе найдётся хотя бы одна вершина, из которой нет исходящих ребер. Докажем этот факт от противного. Предположим, что, наоборот, из каждой вершины есть хотя бы одно исходящее ребро. Тогда, начиная с любой вершины, мы можем построить бесконечный путь. Но число вершин в графе — конечное, поэтому наш путь рано или поздно придет в уже посещённую вершину. Эта вершина — не обязательно первая, возможно и ситуация, представленная на рис. 24. В любом случае мы показали, что в графе будет цикл. Но по условию теоремы граф ациклический. Мы пришли к противоречию и доказали, что найдётся хотя бы одна вершина, из которой не исходит ни одно ребро.

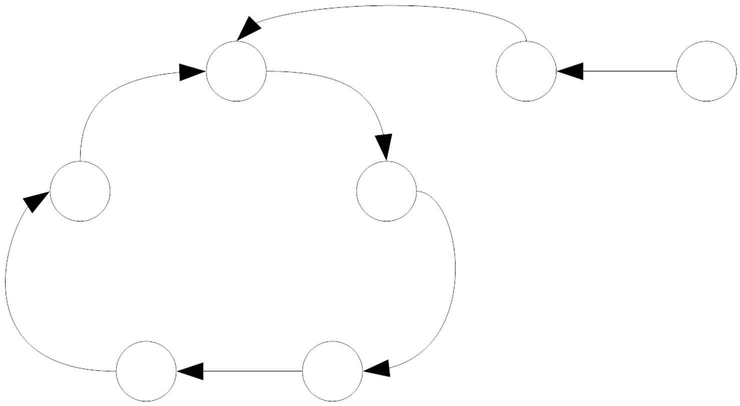


Рис. 24

Выберем любую такую вершину. Присвоим ей номер 1 и отбросим её вместе со всеми входящими в неё рёбрами (см. рис. 25).

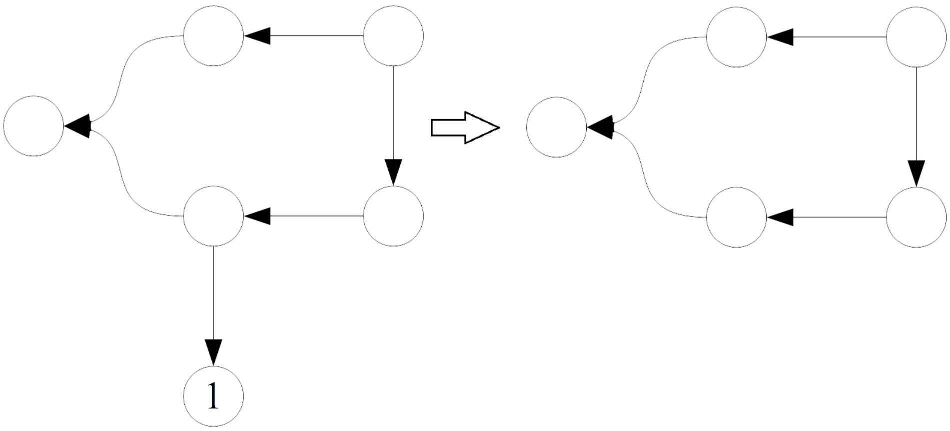


Рис. 25

Граф после этого останется ациклическим, т.к. новые циклы в нём после этой операции не образуются. Значит, в нём снова можно найти вершину, из которой не исходит ни одно ребро, присвоить ей номер 2 и отбросить (см. рис. 26), и т.д.

Нетрудно видеть, что порядок вершин, полученный по этому алгоритму, будет порядком топологической сортировки, т.е. все ребра исходного графа будут идти из вершин с бо́льшими номерами в вершины с меньшими номерами (см. рис. 22). Теорема доказана. □

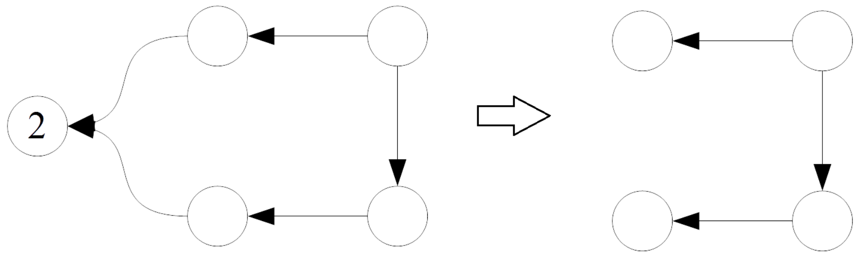


Рис. 26

Рассматривая вершины в порядке топологической сортировки, можно определить, какие из них соответствуют выигрышным состояниям, а какие — проигрышным (см. рис. 27). Действительно, если из вершины нет исходящих рёбер, то согласно правилам игры, это проигрышное состояние. Если исходящие ребра есть, то они ведут в вершины с меньшими номерами, для которых тип позиции уже определён.

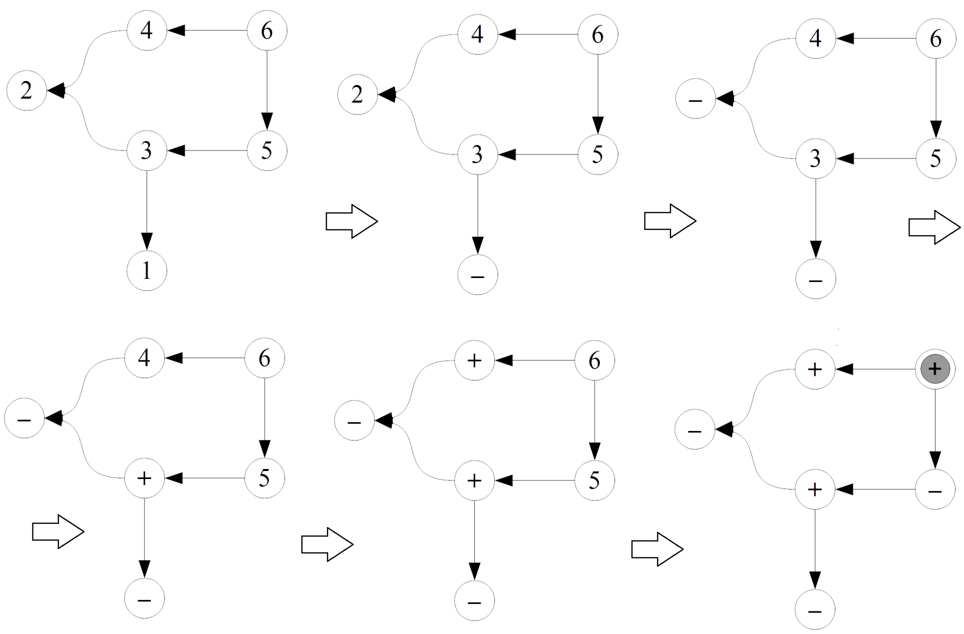


Рис. 27

Таким образом, мы полностью исследовали игру с фишкой на произвольном ациклическом графе. Эта игра — обобщение всех предыдущих игр. Действительно, по любой из них можно построить ориентированный граф, в котором вершинам соответствуют позиции, а рёбрам — переходы. Важно, что все эти графы будут ациклическими. Это можно показать, выбрав в каждой игре полуинвариант. Например, у нас были три игры со спичками. В них во всех полуинвариантом будет число спичек на столе: оно с каждым ходом уменьшается, поэтому в графе позиций циклов не будет. В задаче с покрытием доски доминошками полуинвариантом будет количество свободных клеток, в задаче с движением шашки из левого нижнего угла в правый верхний — расстояние до правого верхнего угла.

В задаче по программированию часто оказывается возможным построить граф явно и дальше просто решать задачу с фишкой. Хотя это делать не обязательно, можно просто моделировать ходы в игре и заполнять таблицу выигрышных и проигрышных позиций. Тем не менее, при этом полезно иметь в виду граф, потому что могут понадобиться какие-то его свойства, например, для оценки быстродействия программы.

Упражнение 5.11

На столе лежит кучка спичек. Двое играют в игру. Ходят по очереди. За ход нужно забрать одну, две или три спички. Проигрывает тот, кто не может сделать ход.

Постройте по этой игре граф с 6 вершинами, соответствующими исходным количествам спичек от 0 до 5. Сколько рёбер будет в этом графе?