# Что такое комбинаторика

**А.ЛЕВИН**

КОМБИНАТОРИКА – ЭТО раздел математики, связанный с методами подсчета числа объектов определенной природы. По смыслу задачи обычно очевидно, что существует лишь конечное число интересующих нас объектов; все дело в том, чтобы найти это число.

Рассматриваемые объекты, как правило, являются определенными комбинациями других объектов (чисел, букв и т.д.). Отсюда и название – комбинаторика. В более широком понимании комбинаторика – это теория конечных множеств; мы здесь будем рассматривать только задачи подсчета, так что такое расширенное толкование нам не потребуется.

Из сказанного ясно, что комбинаторика имеет дело лишь с натуральными числами. Может показаться, что она поэтому более «элементарна», чем другие разделы математики, оперирующие с богатым числовым материалом (отрицательные числа, дробные, иррациональные, комплексные...). Но такое суждение было бы поспешным. Практика показывает, что многие, впервые сталкиваясь с комбинаторикой, с трудом привыкают к комбинаторным рассуждениям (более близким к программированию, чем, скажем, к геометрии). Наша цель состоит в том, чтобы помочь преодолеть эти трудности. Лучший способ освоения комбинаторики – решение задач. Начинать, естественно, надо с простейших. Именно о простых, типовых (и в то же время важнейших) задачах и пойдет речь ниже.

Интересная сама по себе, комбинаторика важна и для многих других разделов математики. Ее связи с алгеброй и теорией вероятностей будут вкратце затронуты позднее.


Иллюстрация В.Власова

## Нельзя ли просто пересчитать?

Этот вопрос напрашивается. Если интересующих нас объектов конечное число, почему бы не составить полный их перечень и попросту пересчитать безо всяких там «комбинаторных рассуждений»?

Рассмотрим пример.

**Задача 1.** *Сколько существует различных двоичных (т.е. состоящих только из нулей и единиц) последовательностей длины $m = 2$?*

Перечень составляется без всякого труда:

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

Задача, как видим, оказалась проста, как дважды два – четыре (в данном случае буквально).

Но если тем же методом решать задачу при $m = 10$, уйдет куда больше времени (в перечень войдет свыше тысячи последовательностей). При $m = 20$ осилить такой перечень сможет лишь компьютер. Ну а при $m = 1000$ окажутся бессильными все компьютеры мира, вместе взятые. Правда, в принципе такой перебор все же возможен (если отвлечься от таких «мелочей», как миллиарды лет машинного времени). А как быть, если надо найти число двоичных последовательностей произвольной длины $m$?

Здесь перебор невозможен в принципе. А между тем очень простые соображения общего характера позволяют мгновенно дать ответ: искомое число есть $2^m$. Это сразу вытекает из принципа умножения (см. ниже). Читатель, знакомый с двоичной системой счисления, может здесь обойтись и без этого принципа: все наши последовательности – это записи в двоичной системе чисел $0, 1, 2, \ldots, 2^m - 1$ (записи, содержащие менее $m$ разрядов, дополняются нулями впереди).

Данный пример ясно показывает мощь общих комбинаторных соображений по сравнению с примитивным перебором. Не следует, конечно, думать, что все комбинаторные задачи можно решить так же мгновенно. Рассмотренная задача относилась к числу простейших. В комбинаторике много трудных задач, а есть и такие, решения которых еще никому не удалось найти.

## Сколько вариантов у понятия «вариант»

Теперь мы хотим предупредить об одной особенности комбинаторики: в ней исключительно большую роль играет точная формулировка (и точное понимание) задачи. Именно с этим связано большинство ошибок начинающих, да и не только начинающих; в некоторых задачниках можно встретить задачи, некорректные ввиду неопределенности формулировки.

Разумеется, никакой неопределенности не возникает, когда речь идет о том, чтобы подсчитать число учеников в классе или число окон в комнате. Но когда речь идет о числе различных вариантов (или способов), ситуация бывает куда менее определенной. Приведем пример.

**Задача 2.** *Сколькими способами можно распределить три конфеты между тремя лицами?*

Тут впору искать ответ на другой вопрос: сколькими способами можно понимать эту задачу. В зависимости от ответа на него возможны шесть разных ответов на поставленный вопрос – 1, 3, 5, 6, 10, 27!

Первый источник неопределенности – термин «распределить». Можно ли сказать, что конфеты распределены между тремя лицами, если, скажем, все они отданы одному? Примем, так сказать, социально-справедливый вариант распределения. Если же понимать распределение в широком смысле слова, по-прежнему считая конфеты тождественными, получаем 10 вариантов – 3 варианта типа (3,0,0) (когда все конфеты отдаются кому-либо из трех) + 6 вариантов типа (2,1,0) + 1 вариант типа (1,1,1). Если же вдобавок и все конфеты различны, получаем максимальное число $3^3 = 27$ вариантов. Другие ответы предоставляем читателю получить самостоятельно. Отметим лишь, что цифры 3 и 5 получаются, если считать неразличимыми людей, что не особенно естественно (но становится вполне естественным при замене людей тремя одинаковыми коробками).

Некоторая расплывчатость понятия о числе вариантов связана с тем, что варианты – умозрительные понятия, их нельзя увидеть непосредственно (если нет перечня). Полезно поэтому мысленно представить себе полный перечень различных вариантов (закодированных каким-либо образом). Число вариантов при этом превращается в число записей. А это уже нечто материальное и может интерпретироваться, например, как количество ячеек машинной памяти. Такая мысленная «материализация» не имеет, конечно, ничего общего с решением задачи прямым подсчетом, требующим отнюдь не мысленного, а реального составления перечня. Да и речь идет не о решении, а лишь о лучшем понимании постановки задачи.

Итак, Главное Правило Комбинаторики:

*прежде чем подсчитывать число различных вариантов, необходимо точно выяснить смысл слов «различные варианты».*

Само по себе это правило не позволит решить ни одной задачи, зато поможет избежать путаницы и недоразумений при решении множества задач.

## Принцип сложения

Мы обсудили некоторые общие особенности комбинаторики. Теперь можно переходить к конкретным способам подсчета. Начнем с простейшего.

*Если все варианты делятся на $k$ взаимоисключающих типов, причем имеется $n_1$ вариантов 1-го типа, $n_2$ вариантов 2-го типа, ..., $n_k$ вариантов $k$-го типа, то общее число вариантов есть $n = n_1 + n_2 + \ldots + n_k$.*

Мы фактически уже применяли этот «принцип сложения». Он совершенно очевиден и выделен, главным образом, для того, чтобы читатель мог сопоставить его с «принципом умножения» (см. ниже).

## Кортежи, или упорядоченные наборы

Кортеж длины $m$

$$a_1, a_2, \ldots, a_m \qquad (1)$$

– это просто *упорядоченный набор* (или, что то же, конечная последовательность) *$m$ элементов произвольной природы, при этом $a_k$ называется $k$-й компонентой кортежа* (1). Если все компоненты – числа, то кортеж длины $m$ называется также $m$-мерным вектором. Если все компоненты – буквы некоторого алфавита, то кортеж длины $m$ называется также $m$-буквенным словом в этом алфавите. Лингвистическая сторона дела здесь не учитывается; так, ЙЙЬЪ – 4-буквенное слово в русском алфавите (хотя с его произношением и истолкованием не все ясно).

Обозначение шахматных полей (например, e4) или номеров автомобилей (например, МЯУ 1999) – примеры кортежей, где одни компоненты – буквы, другие – цифры.

Как видим, кортеж — весьма общее понятие. Главным характерным признаком кортежей является их упорядоченность. Кортежи $a_1, \ldots, a_m$ и $a'_1, \ldots, a'_{m'}$ считаются совпадающими (равными) в том и только в том случае, если $m = m'$ и $a_1 = a'_1$, $a_2 = a'_2, \ldots, a_m = a'_m$; в противном случае кортежи, естественно, считаются различными. В частности, $a,b$ и $b,a$ – различные кортежи (при $a \neq b$). Итак, равенство кортежей – такое же, как у векторов или слов (никто ведь не считает слова БАР и БРА совпадающими). Иначе обстоит дело с неупорядоченными наборами, о которых речь будет идти позднее.

## Принцип умножения

Большое число задач комбинаторики (среди них есть и трудные) связано с подсчетом количества кортежей определенного вида. Однако удивительно часто приходится иметь дело с множествами кортежей, имеющими особо простую структуру, позволяющую определить число кортежей в множестве без труда. Суть дела такова.

Пусть множество $S$ кортежей длины $m$ порождается следующим образом: компонента $a_1$ пробегает $n_1$ различных значений, при любой фиксированной $a_1$ компонента $a_2$ пробегает $n_2$ различных значений и т.д., вплоть до $a_m$, которая при любых фиксированных $a_1, \ldots, a_{m-1}$ пробегает $n_m$ различных значений. Тогда общее число кортежей в множестве $S$ есть

$$n = n_1 n_2 \ldots n_m. \qquad (2)$$

Отметим, что значения $a_1$ (но не их число) могут зависеть от $a_2$, то же относится и к другим компонентам.

Например, для множества $m$-мерных двоичных векторов, о которых шла речь в начале, имеем $n_1 = n_2 = \ldots = n_m = 2$. Отсюда по принципу умножения получаем, что их число равно $2^m$.

Другой ранее рассмотренный пример – задача о распределении трех конфет между тремя лицами. Пусть все конфеты различны (как и лица), распределения допускаются любые. Обозначая лиц через $A$, $B$, $C$ и нумеруя конфеты, можем каждый способ распределения кодировать кортежем $a_1a_2a_3$, где $a_k$ – лицо, получившее $k$-ю конфету (например, код $CAA$ отвечает варианту, когда первую конфету получил $C$, а обе остальные $A$). Так как $n_1 = n_2 = n_3 = 3$, то всего получаем $3^3$ вариантов.

Далее, число всех $m$-буквенных слов в английском алфавите есть $26^m$, а в русском – $33^m$. Число всех семизначных телефонных номеров есть $10^7$ (что, очевидно, вполне достаточно для Москвы; для небольших городов можно обойтись и четырехзначными номерами, количество которых равно $10^4$). Число всех автомобильных номеров, образованных тремя русскими буквами и следующими за ними четырьмя цифрами, равно $33^3 \cdot 10^4$.

Дальнейшие примеры, связанные с принципом умножения, будут приведены позднее.

Принцип умножения достаточно очевиден, хотя и не в такой степени, как принцип сложения. Для тех, кому он не кажется очевидным, приведем доказательство с помощью индукции по $m$. При $m = 1$ утверждение верно. Пусть оно доказано для кортежей длины $m - 1$, т.е. последовательность $a_1, a_2, \ldots, a_{m-1}$ может быть выбрана $n_1n_2 \ldots n_{m-1}$ способами. Каждому из них соответствует $n_m$ кортежей длины $m$, получающихся добавлением того или иного значения $a_m$. Все полученные кортежи различны (любые два, у которых совпадают первые $m - 1$ компонент, отличаются значением $m$-й компоненты). Поэтому общее число кортежей в множестве $S$ есть $(n_1n_2 \ldots n_{m-1})n_m = n_1n_2 \ldots n_m$.

Если читатель незнаком с математической индукцией, можно обойтись и без педантичного доказательства. По правде говоря, принцип умножения мало отличается от арифметических задач типа: «сколько всего листов в 10 стопках тетрадей, если в каждой стопке по 20 тетрадей, а в каждой тетради по 12 листов?». Всякий школьник даст ответ $10 \cdot 20 \cdot 12 = 2400$, без упоминаний комбинаторики или индукции. Но ведь листов столько, сколько кортежей $a_1, a_2, a_3$, где $a_1$ – значения от 1 до 10 (номер стопки), $a_2$ – значения от 1 до 20 (номер тетради в стопке), $a_3$ – значения от 1 до 12 (номер листа в тетради). Таким образом, решая эту простенькую задачу, мы неявно пользуемся принципом умножения.

## Складывать или умножать – вот в чем вопрос

В такой гамлетовской ситуации иногда оказываются начинающие. В условиях принципа умножения проводится примерно следующее рассуждение: «Для первой компоненты имеется $n_1$ вариантов, да для второй компоненты $n_2$ вариантов, ..., да для последней $n_m$ вариантов. Итого получаем $n_1$, да $n_2$, ..., да $n_m$, т.е. $n_1 + n_2 + \ldots + n_m$ вариантов». Обращаем внимание на коварное слово «да» (с этим словом и вообще, как известно, надо обращаться осторожно).

Ошибка в том, что $n_1$ вариантов для $a_1$ – это отнюдь не те варианты, общее число которых надо найти. Нужные нам варианты даются ведь кортежами длины $m$ (а не 1, 2...). Варьируя $a_1$, затем $a_2$ и т.д., мы получаем (пока не дойдем до $a_m$) как бы «заготовки», каждая из которых в дальнейшем, после разветвлений, породит много кортежей длины $m$ (т.е. нужных нам вариантов). В «предметных» задачах такая ошибка практически исключена, никто ведь не станет подсчитывать число листов так: «10 стоп, да в каждой по 20 тетрадей, да в каждой по 12 листов – итого 10 + 20 + + 12 листов». Но с абстрактными вариантами такое, увы, бывает.

## Число подмножеств

**Задача 3.** *Сколько различных подмножеств имеется у $n$-элементного множества?*

Может показаться, что принцип умножения не имеет отношения к этой задаче. Подмножества ведь неупорядоченные наборы (например, $\{a,b\}$ и $\{b,a\}$ – одно и то же), к тому же они могут содержать разное число элементов, а в принципе умножения речь идет о кортежах одинаковой длины. Все дело, оказывается, в том, чтобы удачно закодировать подмножества. Именно, нумеруем элементы исходного множества $M = \{S_1, S_2, \ldots, S_n\}$ и ставим в соответствие каждому его подмножеству $M'$ двоичный $n$-мерный вектор $(x_1, x_2, \ldots, x_n)$, где $x_i = 1$, если $S_i \in M'$, и $x_i = 0$ в противоположном случае. (Например, если $M = \{S_1, S_2, S_3, S_4, S_5\}$, то подмножеству $\{S_2, S_5\}$ отвечает вектор (0,1,0,0,1); пустому подмножеству будет отвечать нулевой вектор.) Так как соответствие между подмножествами и векторами взаимно однозначно, то подмножеств столько же, сколько $n$-мерных двоичных векторов, т.е. $2^n$.

## О кодировании

Рассмотренная задача показывает, какую роль играет удачное кодирование объектов; часто оно является основным звеном решения вопроса. Наоборот, неудачное кодирование обычно заводит в тупик. Вернемся к нашей простенькой задаче о распределениях (произвольных) трех различных конфет между тремя лицами. Выше мы кодировали распределение кортежами $(a_1, a_2, a_3)$, где $a_k$ – лицо, получившее $k$-ю конфету. Нельзя ли кодировать не «по конфетам», а «по лицам», обозначая через $a_k$ то, что получает $k$-е лицо ($k = 1, 2, 3$)?

Каждая компонента при этом будет подмножеством (множества из 3 конфет) и может сама по себе принимать $2^3 = 8$ значений. Как уже говорилось, компоненты кортежей могут иметь любую природу, им не возбраняется быть и подмножествами. Хуже другое: при таком кодировании нельзя применить принцип умножения, так как не выполняется его основное условие. Например, если $a_1 = \varnothing$ (пустое множество), т.е. первому не досталось ничего, то для $a_2$ остаются те же 8 вариантов, если же первый получает все, то для $a_2$ возможен лишь один вариант ($\varnothing$). Так что кодирование по лицам является неудачным; это различие между двумя способами кодирования исчезает при «справедливых» распределениях (каждому по конфете), когда лица и конфеты входят в задачу равноправно.

Разумеется, всегда надо следить за взаимно однозначным соответствием между объектами и кодами – каждому объекту (из нашего множества) должен отвечать ровно один код, и обратно. Только в этом случае подсчет числа объектов можно заменить подсчетом числа кодов.

## Размещения

Далее речь будет часто идти о кортежах, все компоненты которых принимают значения из одного и того же непустого конечного множества. Иначе говоря, мы будем иметь дело со словами в некотором алфавите. Разумеется, в качестве «букв» алфавита могут выступать и числа – например, $1, 2, \ldots, n$. Так как элементы конечного множества всегда можно занумеровать, то удобная алфавитно-буквенная терминология не меняет сути дела.

Как уже говорилось, число $m$-буквенных слов в $n$-буквенном алфавите равно $n^m$. Иногда это число называют числом размещений из $n$ по $m$ с повторениями. Обычно же термин «размещение» связывается с требованием, чтобы все буквы слова были различны (не повторялись).

**Задача 4.** *Сколько существует в $n$-буквенном алфавите $m$-буквенных слов, состоящих из различных букв?*

Задача представляет интерес лишь при $m \le n$ (иначе ответ, очевидно, нуль). Можно представить себе, что имеется $n$ карточек с изображением всех букв алфавита (по одной карточке на букву) и мы формируем слова, размещая друг за другом те или иные $m$ карточек в том или ином порядке.

Задача легко решается с помощью принципа умножения, первая буква $a_1$ может пробегать $n$ значений, при любой фиксированной $a_1$ для $a_2$ остается $n-1$ значений (все буквы алфавита, кроме $a_1$) и т.д., вплоть до последней буквы, для которой остается $n - m + 1$ значений. Итак, искомое число равно

$$n(n-1)\ldots(n-m+1) = \frac{n!}{(n-m)!} \quad (m \le n). \qquad (3)$$

Здесь, как и всегда в математических выражениях, $k!$ (читается «ка факториал») есть сокращенная запись произведения $1 \cdot 2 \cdot \ldots \cdot k$, где $k$ – любое натуральное число (например, $4! = 1 \cdot 2 \cdot 3 \cdot 4 = 24$). По определению полагается также $0! = 1$.

Величина (3) называется числом размещений из $n$ элементов по $m$ и часто обозначается через $A_n^m$. Если нет специальной оговорки о повторениях, то под числом размещений из $n$ по $m$ всегда подразумевается величина (3).

Пример: если в турнире участвуют 20 команд и дележ мест исключен, то первые 3 места (т.е. золотая, серебряная, бронзовая медали) могут распределяться

$$A_{20}^3 = 20 \cdot 19 \cdot 18 = 6840$$

способами. Повторения здесь, очевидно, невозможны, так как одна команда не может занять сразу два места.

## Перестановки

Это просто частный случай размещений при $m = n$. Итак, число перестановок из $n$ элементов есть $n!$. Например, в алфавите $\{a, b, c\}$ имеется $3! = 6$ перестановок: *abc, acb, bac, bca, cab, cba*. С ростом $n$ величина $n!$ быстро растет. В отличие от размещений (при $m < n$), перестановки отличаются друг от друга только порядком букв, но не «буквенным составом». Мы просто переставляем те же самые $n$ карточек в том или ином порядке, отсюда и название.

Вернемся к турниру 20 команд из предыдущего пункта. Если интересоваться распределением всех мест (а не только трех первых), то число возможных вариантов становится равным $20!$ ($> 2 \cdot 10^{18}$). Любопытно, что перебор всех этих вариантов мощным компьютером, обрабатывающим миллиард вариантов в секунду, займет больше времени, чем реальное проведение турнира – даже если проводить по одному туру в год!

В теории дискретной оптимизации хорошо известна так называемая *задача коммивояжера*. Речь в ней идет о городах $T_1, T_2, \ldots, T_n$, расстояние между которыми заданы; требуется выбрать такой порядок объезда этих городов, начиная с $T_1$, чтобы суммарная длина маршрута была минимальной. Поскольку число маршрутов конечно, казалось бы, тут нет проблемы – просто перебрать все варианты и выбрать наилучший, тем более если под рукой хороший компьютер. После всего сказанного ранее читатель, несомненно, догадывается, в чем тут загвоздка. Перебирать ведь придется $(n-1)!$ маршрутов – число, хотя и конечное, но с ростом $n$ быстро становящееся «практически бесконечным». При $n = 5$ компьютер выдаст кратчайший маршрут мгновенно, при $n = 15$ провозится куда больше, чем хотелось бы, а при $n = 20$ до получения ответа можно и не дожить…

Мы говорим о гигантских, мучительно долгих вычислениях, а сами, между прочим, подсчитываем все очень легко, можно сказать, на пальцах. Нет ли здесь противоречия? Нет, конечно. Мы ведь только находим число вариантов, применяя при этом «сверхскоростные» (по сравнению с примитивным перебором) приемы комбинаторики. Вместо самого перебора мы лишь прикидываем, сколько времени он бы занял. Чтобы пройти пешком 1000 километров, нужно примерно 200 часов. А чтобы это подсчитать, достаточно секунды.

А как же все-таки решать задачу коммивояжера? Ну, во-первых, это уже относится к оптимизации, а не к элементарной комбинаторике. А во-вторых, на сегодня никто не знает по-настоящему быстрого алгоритма решения этой знаменитой и важной (к ней сводятся многие другие) задачи. Вполне возможно, что такого алгоритма вообще не существует.

Помимо прочего, комбинаторика обычно позволяет быстро оценить, что реально, а что нет, – особенно, когда речь идет об алгоритмах переборного типа. Жизнь, как известно, дается только раз, и не стоит тратить ее, сидя перед компьютером, перебирающим 19! маршрутов.

## Когда нужно заниматься ненужным

Пусть имеется известное число $N$ объектов и требуется выяснить, сколько из них обладает определенным свойством. Искомое число «нужных» объектов обозначим через $N_1 (\le N)$. Как видно из предыдущего, множество нужных объектов часто имеет простую структуру, позволяющую применить принцип умножения. А если это не так – как в таком случае искать $N_1$?

Как правило, полезно проверить множество «ненужных» (т.е. не обладающих требуемым свойством) объектов – не имеет ли оно простой структуры. Если для него удастся применить принцип умножения, то будет найдено число $N_2$ ненужных объектов, после чего легко определяется и $N_1 = N - N_2$. Приведем пример.

**Задача 5.** *Сколько существует $m$-буквенных слов в русском алфавите, содержащих букву $A$?*

Всего $m$-буквенных слов, как мы знаем, $N = 33^m$. Попробуем подсчитать, сколько среди них нужных (содержащих $A$), принципом умножения. Сначала все идет хорошо: на 1-м, 2-м, ..., $(m-1)$-м месте могут стоять по 33 буквы, независимо от предыдущих. Увы, на последнем шаге рушится все: если среди предыдущих была буква $A$, то для $m$-й буквы имеем опять-таки 33 варианта, если нет – всего 1 вариант (буква $A$).

Зато простой структурой обладает множество ненужных (не содержащих $A$) слов – ведь это всевозможные $m$-буквенные слова в 32-буквенном алфавите! Всего их $N_2 = 32^m$, стало быть, искомое число $N_1 = 33^m - 32^m$.

Изложенный прием перехода к дополнительному множеству (к множеству «ненужных» объектов) по идее очень прост и не понять его невозможно. А вот забыть про него очень легко, все внимание часто сосредотачивается на «нужных» объектах. В рассмотренной задаче, например, напрашивается такой путь: разбить все нужные слова на $m$ типов – содержащие ровно одну букву $A$, ровно две и т.д., попытаться подсчитать число слов каждого типа, а затем сложить результаты. Можно и на этом пути получить решение, правда в громоздком виде. Но это значило бы ломиться в открытую дверь – ведь переход к дополнительному множеству дает ответ сразу и в простейшем виде.

К сожалению, довольно часто оказывается, что ни множество нужных объектов, ни дополнительное к нему не обладают простой структурой. Это значит, что задача не сводится к простому применению принципа умножения и нужны новые идеи. Одна из них излагается в следующем пункте.

## «Растождествление» и перестановки с повторениями

Слово «растождествление» вряд ли есть в словарях, но оно точно передает суть дела. Воспользуемся тем, что комбинаторика позволяет считать словом любую комбинацию букв...

Пусть $A_1, A_2, \ldots, A_k$ – различные буквы некоторого алфавита. Рассматривается следующая задача.

**Задача 6.** *Пусть $n_1, n_2, \ldots, n_k$ – заданные натуральные числа, сумма которых равна $n$. Сколько существует $n$-буквенных слов, содержащих $n_1$ букв $A_1$, $n_2$ букв $A_2$, ..., $n_k$ букв $A_k$?*

В частном случае $n_1 = \ldots = n_k = 1$, $k = n$ это знакомая нам задача о числе перестановок. Но если хотя бы одно из чисел $n_k$ больше 1, возникает нечто новое – перестановки с повторениями. Как и в обычных перестановках, «буквенный состав» слов здесь заранее определен, так что варьируется лишь порядок. Но теперь среди букв есть тождественные (совпадающие, идентичные, неразличимые – все это одно и то же). Можно считать, что у нас есть $n_i$ карточек с буквой $A_i$, $i = 1, \ldots, k$, и надо выяснить, сколько различных $n$-буквенных слов можно из них составить.

Задача эта несколько труднее предыдущих, решить ее уже знакомыми приемами не удастся. Основную идею решения лучше всего пояснить на конкретном примере.

**Задача 7.** *Сколько 7-буквенных слов можно составить из букв К,К,Л,Л,О,О,О?*

В данном случае число вариантов не слишком велико, и ответ можно получить прямым перебором. Все это потребовало бы времени и внимания – дабы ни одно слово не пропустить и ни одно не засчитать дважды (Сцилла и Харибда всех комбинаторных подсчетов!). Главное же – для решения общей задачи такой подсчет ничего не даст и будет, как сказал бы К.Прутков, «пустою забавою».

Вообще-то от перечня мы не отказываемся; более того, у нас их будет даже два – основной (куда входят нужные слова) и вспомогательный. Вот только реально составлять их мы не собираемся, «работать» с ними будем с помощью воображения. Что ж, не впервой.

Говорят, математики больше всего на свете любят сводить новые задачи к уже решенным. Здесь, к примеру – как хорошо бы применить знакомые способы подсчета числа перестановок! Жаль, не получается – тождественные буквы мешают, путаются друг с другом... А нельзя ли их «растождествить» – хотя бы временно – и посмотреть, что получится? В случае с карточками это особенно естественно – ведь наши 7 карточек действительно физически различны (мы отождествили их, так сказать, искусственно).

Сказано – сделано. После индексации получаем 7 различных букв $К_1$, $К_2$, $Л_1$, $Л_2$, $О_1$, $О_2$, $О_3$. Из них, как мы знаем, можно составить 7! различных 7-буквенных слов. Они и образуют вспомогательный перечень.

Он, очевидно, «раздут» по сравнению с основным: если взять какое-нибудь слово основного перечня – например *КОЛОКОЛ*, то ему будет соответствовать много слов вспомогательного перечня:

$$К_1О_1Л_1О_2К_2О_3Л_2,\ К_2О_1Л_2О_3К_1О_2Л_1$$

и т.д.

Оказывается, нетрудно вычислить, во сколько раз вспомогательный перечень больше основного. Каждое слово вспомогательного перечня закодируем кортежем длины 4, где первая компонента – соответствующее слово основного перечня (число таких слов обозначим через $x$), а остальные определяют порядок следования индексов. Так, слово $O_2K_2K_1O_1Л_1Л_2O_3$ получит код

$$ОККОЛЛО,\ K_2K_1,\ Л_1Л_2,\ O_2O_1O_3$$

(для определенности выбираем алфавитный порядок букв). Такие коды образуют множество простой структуры. Действительно, компоненты пробегают возможные значения независимо одна от другой, причем первая пробегает $x$ значений, а остальные 2!, 2!, 3! значений соответственно (это обычные числа перестановок). Итак, по принципу умножения число слов вспомогательного перечня (т.е. 7!) равно в то же время $x \cdot 2! \cdot 2! \cdot 3!$, откуда

$$x = \frac{7!}{2!2!3!} = 210.$$

Данные рассуждения переносятся на общий случай практически без изменений. Новый индекс можно ставить, например, вверху: получаем $n$ различных букв $A_1^1, \ldots, A_1^{n_1}, \ldots, A_k^1, \ldots$ $\ldots, A_k^{n_k}$.

Из них можно составить $n!$ слов вспомогательного перечня. Кодируя их, как выше, и пересчитывая вторым способом с помощью принципа умножения, получаем равенство ($x$ – искомое число)

$$x \cdot n_1! \cdot n_2! \cdot \ldots \cdot n_k! = n!,$$

поэтому число слов основного перечня равно

$$\frac{n!}{n_1!n_2!\ldots n_k!}. \qquad (4)$$

Из доказанного следует, между прочим, что эта дробь всегда является целым числом. Полученное выражение для числа перестановок с повторениями – центральный результат элементарной комбинаторики. Как частные случаи, оно содержит выражения для числа обычных перестановок (при $k = n$) и для числа сочетаний (при $k = 2$, см. ниже). Стоит еще упомянуть, что в знаменитой работе физика Л.Больцмана, где впервые был выяснен статистический смысл второго начала термодинамики, также не обошлось без выражения (4). Речь там шла не о буквах и словах, а о молекулах, энергетических уровнях и состояниях физической системы, но формула от этого, конечно, не меняется.

Помимо принципа умножения – этого излюбленного орудия элементарной комбинаторики – выше был применен новый для нас прием «растождествления». Он показывает, как с пользой для дела можно переходить от одного понятия тождества (или различия) к другому. Если при нечеткости формулировок в связи с этим возникали путаница и недоразумения, то при точном понимании терминов – т.е. при соблюдении Главного Правила Комбинаторики – здесь открываются новые возможности решения задач.

## Сочетания

В задаче 3 мы подсчитывали число всех подмножеств у $n$-множества. Поставим теперь вопрос иначе.

**Задача 8.** *Сколько существует $m$-подмножеств у $n$-множества?*

Естественно, имеются в виду различные $m$-подмножества (эта оговорка ради краткости обычно опускается). Подразумевается, конечно, что $m \le n$.

Неискушенного человека могло бы удивить, что ответ на этот вопрос по существу дан в предыдущем пункте. Там ведь речь шла об упорядоченных наборах с повторениями, а подмножества – неупорядоченные наборы без повторений. Но наш читатель, надеемся, не забыл, как удачно кодируются подмножества двоичными векторами. А двоичный вектор – это ведь слово в алфавите из двух «букв» 0 и 1. При таком кодировании каждому $m$-подмножеству $n$-множества отвечает слово, содержащее $m$ «букв» 1 и $n - m$ «букв» 0. Итак, искомое число равно

$$\frac{n!}{m!(n-m)!} = \frac{n(n-1)\ldots(n-m+1)}{1 \cdot 2 \cdot \ldots \cdot m}. \qquad (5)$$

Величина (5) обозначается обычно $C_n^m$ или $\binom{n}{m}$ и называется *числом сочетаний* из $n$ по $m$; часто эти числа называются также *биноминальными коэффициентами* (причина такого названия станет ясной чуть позже). Встречаются эти величины очень часто. Вот несколько примеров.

**Задача 9.** *Сколько встреч будет проведено в турнире 20 команд по круговой системе (в 1 круг)?*

Очевидно, столько же, сколько 2-подмножеств у 20-элементного множества, т.е.

$$\binom{20}{2} = \frac{20 \cdot 19}{1 \cdot 2} = 190.$$

К тому же результату легко прийти и непосредственно: каждая из 20 команд проводит по 19 встреч; всего вроде бы получается $20 \cdot 19$ встреч, при этом каждая встреча засчитывается дважды, так что надо еще поделить на 2. Фактически это повторение общего рассуждения для очень простого частного случая.

Столько же рукопожатий при встрече 20 человек – если все здороваются друг с другом.

**Задача 10.** *Сколькими способами можно выбрать 3 дежурных в классе из 25 человек?*

Ответ очевиден:

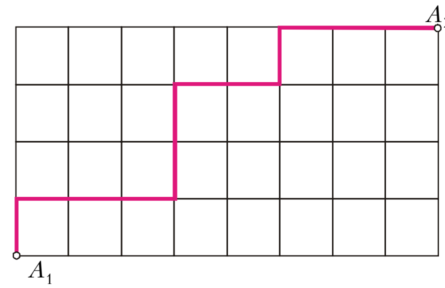
$$\binom{25}{3} = \frac{25 \cdot 24 \cdot 23}{1 \cdot 2 \cdot 3} = 2300.$$

Тут важно, что все трое дежурных равноправны, и, значит, от порядка, в котором названы 3 фамилии, выбор не зависит. Иное дело, если бы речь шла о трех разных поручениях – допустим, выбираются староста, кассир и физорг. Это были бы уже не сочетания, а размещения, и способов было бы в 3! = 6 раз больше (предполагается, что школьнику дается не более одного поручения).

Рассмотрим теперь прямоугольную сетку, образованную равноотстоящими прямыми (см. рисунок).

Можно считать, например, что это уличная сеть (чересчур правильная для реального города). Сколько существует кратчайших путей из $A_1$ в $A_2$, проходящих по этой сетке?

Если расстояния между соседними прямыми равны 1, то все кратчайшие пути имеют, очевидно, длину 12. Один из них выделен на рисунке.




Основной момент в решении – кодирование путей. Каждый из 12 единичных этапов кратчайшего пути направлен либо вверх, либо вправо. Обозначая эти возможности буквами В и П соответственно, получаем код в виде 12-буквенного слова, содержащего 4 В и 8 П. Например, для выделенного пути он имеет вид

В П П П В В П П В П П П.

Остается воспользоваться числом перестановок с повторениями, либо, поскольку букв всего две, числом сочетаний. Так как для букв В надо выбрать 4 места из 12, то ответ равен $\binom{12}{4} = 495$.

Вместо выбора 4 мест для букв В можно, конечно, выбирать 8 мест для букв П. На ответ это не повлияет, так как

$$\binom{n}{m} = \binom{n}{n-m}.$$

Данное равенство можно доказывать двумя методами – комбинаторным и аналитическим. Комбинаторный таков: каждому $m$-подмножеству отвечает дополнительное к нему $(n - m)$-подмножество, поэтому число тех и других совпадает. Аналитический же метод сводится к тому, чтобы посмотреть на выражение (5).

Существует целый ряд других, не столь очевидных, соотношений для чисел $\binom{n}{m}$. Некоторые из них будут упомянуты позднее. Расчет, проведенный выше для прямоугольника $4 \times 8$, читателю предлагается самостоятельно провести для прямоугольника общего вида $m_1 \times m_2$. А что будет, если перейти от плоского случая к пространственному?

## Диофантово уравнение $x_1 + \ldots + x_n = m$

Давно уже мы не занимались распределением материальных благ – с тех пор, как делили три конфеты между тремя лицами. Теперь, вооружившись формулой для числа сочетаний, вернемся к этому увлекательному занятию. Для разнообразия пусть пираты делят между собой золотые монеты (добытые, естественно, специфическими методами).

**Задача 11.** *Сколькими способами могут 5 пиратов разделить между собой 10 монет?*

Блеск золота не должен ослеплять нас настолько, чтобы мы забыли о Главном Правиле. Итак, уточняем: пираты все различны (ярко выраженные индивидуальности!); монеты все одинаковы; допускается любой способ дележа, при котором каждый получает хотя бы одну монету. Теперь задачу можно решать.

На первый взгляд неясно, при чем тут число сочетаний. Но мы уже убедились, что первый взгляд часто обманчив. Перейдем сразу ко второму и направим его на 10 монет

○○○|○|○○○|○○|○,

разделенных «перегородками», указывающими конкретный способ дележа. В данном случае, например, расположение перегородок означает, что первый пират получает 3 монеты, второй – 1, третий – 3, четвертый – 2, пятый – 1.

Теперь ситуация проясняется. Ввиду правил дележа перегородки должны находиться в промежутках между монетами (а не слева или справа от монет) и в каждом промежутке должно быть не более одной перегородки. Итак, речь идет о выборе 4 мест из 9. Стало быть, способов дележа столько, сколько 4-подмножеств у 9-множества, т.е. $\binom{9}{4} = 126$.

Заметим, что фактически найдено число решений в натуральных числах «уравнения дележа»

$$x_1 + x_2 + x_3 + x_4 + x_5 = 10.$$

Здесь $x_i$ – число монет, получаемых $i$-м пиратом ($i = 1, \ldots, 5$).

Переходя к общему случаю $n$ пиратов и $m (\geq n)$ монет,

$$x_1 + x_2 + \ldots + x_n = m, \qquad (6)$$

имеем $\binom{m-1}{n-1}$ различных решений в натуральных числах. (Этот ответ остается в силе для любых натуральных $m$ и $n$, если принять, что $\binom{k}{l} = 0$ при $k < l$.)

Рассмотрим теперь случай совершенно аморальных пиратов, которые допускают любые способы дележа (вплоть до того, что все монеты достанутся одному). Итак, речь теперь идет о числе решений уравнения (6) в целых неотрицательных числах (для $x_i$ допускается и значение 0). Неравенство $m > n$ при такой постановке излишне, т.е. $m$, $n$ – произвольные натуральные числа.

Если снова вернуться к геометрической иллюстрации, то теперь возможны любые расположения $m$ кружков и $n - 1$ перегородок. Например, для 5 пиратов и 10 монет расположение

|○○○○○|||○○○○○

отвечает случаю, когда первый, третий и четвертый пираты не получают ничего, а второй и пятый – по пять монет каждый. Итак, для перегородок надо выбрать 4 места из 14, а это можно сделать $\binom{14}{4} = 1001$ способом.

Для общего случая находим, что уравнение (6) имеет

$$\binom{m+n-1}{n-1} = \binom{m+n-1}{m} \qquad (7)$$

решений в целых неотрицательных числах.

Пожалуй, более простое доказательство состоит в замене переменных $y_i = x_i + 1$, $i = 1, \ldots, n$. Каждому решению уравнения (6) в целых неотрицательных числах взаимно однозначно соответствует решение уравнения

$$y_1 + y_2 + \ldots + y_n = m + n$$

в натуральных числах. Таких решений, как показано ранее, $\binom{m+n-1}{n-1}$, т.е. опять приходим к (7).

Уравнения в целых числах обычно называют *диофантовыми* по имени много занимавшегося ими древнегреческого математика Диофанта. Итак, мы нашли число решений диофантова уравнения (6) при любом из предположений: а) все $x_i > 0$; б) все $x_i \geq 0$. А что будет с числом решений, если на целые $x_i$ не накладывается никаких дополнительных ограничений?

Величину (7) часто называют числом сочетаний с повторениями из $n$ по $m$. Такое название связано с классом задач, типичным представителем которых является следующая.

**Задача 12.** *Имеется n сортов пирожных, требуется купить m пирожных. Сколькими способами это можно сделать?*

Пирожные одного сорта, естественно, считаются неразличимыми. Данная совсем простенькая на вид задача сама по себе не так уж проста. Но после того как найдено число решений уравнения (6) в целых $x_i \geq 0$, тут и в самом деле почти нечего делать – разве что обозначить через $x_i$ количество покупаемых пирожных $i$-го сорта ($i = 1, \ldots, n$). Искомое число равно (7).

Конечно, речь могла идти не о пирожных, а о чем-либо другом. Важно лишь, что ищется число неупорядоченных наборов из $m$ элементов, причем каждый элемент принадлежит какому-либо из $n$ типов. Так как элементы одного типа считаются неразличимыми, то допускаются, таким образом, «повторения» элементов в наборе.

## «Повторения с повторениями»

Размещения, перестановки, сочетания – и сразу, как грибы после дождя, размещения с повторениями, перестановки с повторениями, сочетания с повторениями… С непривычки можно растеряться, тем более, что повторения не всегда понимаются одинаково. Стоит, пожалуй, вернуться к этим терминам и еще немного о них поговорить. Так сказать, повторение на тему повторений.

Вот, к примеру, такой вопрос. Перестановки – частный случай размещений, и формула для числа перестановок – частный случай формулы для числа размещений. С другой стороны, перестановки с повторениями – вроде бы частный случай размещений с повторениями, но формула для числа перестановок с повторениями отнюдь не является частным случаем формулы для числа размещений с повторениями. В чем тут дело?

Когда речь идет о повторениях элементов в наборе (упорядоченном или нет), возможны две противоположные ситуации:

а) нет никаких ограничений на число повторений элементов (кроме того, что общее число элементов набора равно заданному числу);

б) каждый элемент должен повторяться в наборе заданное число раз.

Встречаются и варианты, промежуточные между этими двумя, но сейчас они нам не нужны.

Размещения с повторениями и сочетания с повторениями объединяет то, что имеется в виду а), тогда как число перестановок с повторениями определяется в ситуации б). Конечно, размещения и перестановки (с повторениями или без) близки в другом отношении – как упорядоченные наборы, тогда как сочетания (с повторениями или без) – наборы неупорядоченные. Кстати, для неупорядоченных наборов постановка б) бессодержательна (ответ, очевидно, равен 1).

Насколько полезна и удобна вся эта «повторительная» терминология?

Для очень простого (по виду и смыслу) выражения $n^m$ название «число размещений с повторениями из $n$ по $m$» является, пожалуй, излишне длинным и торжественным. Но в ясности ему не откажешь.

Перестановки с повторениями – термин ясный и удачный.

В случае сочетаний с повторениями с ясностью не все благополучно. Вернемся к задачам из предыдущего пункта. При покупке пирожных проблем не возникает (кроме финансовой); ясно, что здесь повторяются пирожные одного сорта. Задача дележа рассматривалась в двух постановках – I (пираты с признаками морали) и II (аморальные пираты). В случае I возникали сочетания, в случае II – сочетания с повторениями. А что, собственно, повторяется в варианте II?

Монеты? Они, естественно, «повторяются», поскольку идентичны. Но они ведь были идентичными и в варианте I.

Пираты? Но каждый из них у нас в единственном экземпляре. Правда, можно кодировать дележ, сопоставляя каждой монете имя владельца; тогда пираты (вернее, их имена) будут повторяться. Но, опять-таки, в этом смысле они «повторялись» и в I постановке – однако там были сочетания без повторений.

Если исходить из метода решения, следует признать, что повторяются промежутки между монетами – как места для перегородок. При II постановке, в отличие от I, один промежуток может повториться для нескольких перегородок.

Прямо скажем, промежуток между монетами – вещь куда менее ощутимая, чем монета или пирожное. Два пирата, которым при дележе достались лишь «промежутки между монетами», вряд ли станут выяснять, один и тот же у них промежуток или разные – скорее, они станут выяснять отношения с другими членами шайки.

Для тех, кто еще не устал, можно добавить следующее. Если забыть о способе решения и думать лишь об аналогии с задачей о покупке пирожных, то придется, пожалуй, принять – несмотря на все сказанное ранее, – что «повторяются» все-таки пираты!

В общем, неразбериха что надо.

Мораль проста: начинающим не рекомендуется связываться с термином «сочетания с повторениями» без особой надобности. Можно ведь говорить о числе решений уравнения (6) в целых $x_i \geq 0$. Это кристально ясное выражение, не дающее повода для «кривотолков». Задачи о дележе (II постановка) и о покупке пирожных внешне различны. Но стоит написать уравнение (6), как сразу становится очевидным, что это одна и та же задача.

Иногда вместо количества решений уравнения (6) в целых $x_i \geq 0$ говорят о количестве разбиений числа $m$ на $n$ неотрицательных слагаемых. Это хуже. При наличии нулевых слагаемых не очень уместно говорить о разбиениях (можно ли считать, что $3 = 3 + 0 + 0 + 0$ означает «разбиение» числа 3 на 4 слагаемых?), но главное в другом. Разбиения $8 = 5 + 3$ и $8 = 3 + 5$ – это два разных разбиения или это по существу одно и то же разбиение? Можно понимать и так и этак, стало быть, нарушается Главное Правило Комбинаторики.

В то же время вряд ли кто-нибудь усомнится в том, что $x_1 = 5$, $x_2 = 3$ и $x_1 = 3$, $x_2 = 5$ – это разные решения уравнения $x_1 + x_2 = 8$. На сей счет в математике устойчивые традиции.

Стоит ли уделять столько внимания терминологии – ведь в конечном счете она несущественна? В конечном, пожалуй, и впрямь несущественна. Но для начинающих «начальный счет» куда важнее «конечного».

*(Окончание следует)*