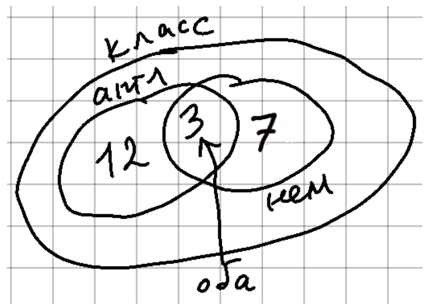
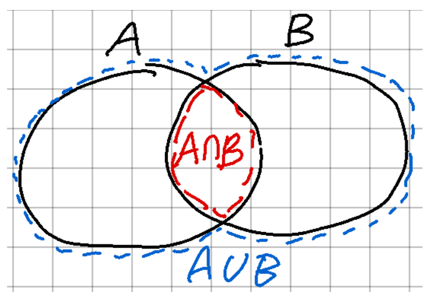
6. Включения и исключения

$\boxed{\textbf{6.1}}$ В классе 10 учеников знают немецкий, 15 знают английский, и среди них есть трое, которые знают оба языка (они входят в 10 и 15, о которых шла речь). Сколько учеников знает хотя бы один из этих двух языков?





На более математическом языке (мы уже упоминали это в разделе 1) говорят так. Есть два конечных множества A и B (тех, кто знает немецкий, и тех, кто знает английский). Тогда можно рассмотреть их *пересечение* $A \cap B$ (тех, кто знает оба языка) и *объединение* $A \cup B$ (тех, кто знает хотя бы один язык). Тогда

$$|A \cup B| = |A| + |B| - |A \cap B|,$$

где через $|X|$ обозначено число элементов в (конечном) множестве X. В нашем примере было $22 = 15 + 10 - 3$.

• Аналогичное утверждение можно сформулировать для площадей: площадь объединения фигур A и B на рисунке равна сумме площадей фигур минус площадь их пересечения.

Можно задать такой же вопрос для большего числа множеств — известны размеры множеств и размеры всех пересечений, надо найти размер объединения. Можно было бы задать такой вопрос про английский, немецкий и французский, но чтобы не запутывать дело, будем сразу говорить о множествах.

6.2 Докажите, что

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$$

для любых трёх конечных множеств A, B, C.

Словами: чтобы найти число людей, знающий хотя бы один язык из трёх, надо сложить числа людей, знающих немецкий, знающих французский и знающих английский, затем вычесть числа людей, знающих одновременно английский и немецкий, английский и французский, немецкий и французский, и, наконец, прибавить число людей, знающих все три языка.

6.3 Как изменятся слагаемые в правой части, если из множества A удалить элемент, который не входил ни в B, ни в C? Если удалить элемент, который входил в A и B, но не входил в C? Если удалить элемент, который входил во все три множества?

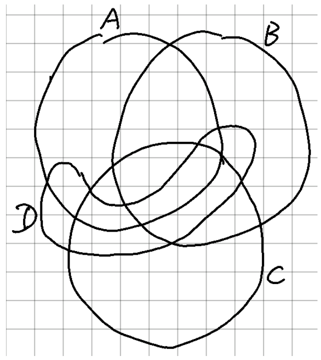
• Эта задача показывает, что во всех трёх случаях правая часть уменьшается на единицу — синхронно с левой. Если по очереди удалить все элементы, то она станет равной нулю — значит, изначально она была равна числу элементов в $A \cup B \cup C$, как и утверждает формула включений и исключений.

Теперь уже можно догадаться, как будет выглядеть формула для четырёх множеств: чтобы найти число элементов в их объединении, надо

- сложить числа элементов в этих множествах;
- вычесть все шесть попарных пересечений;
- добавить четыре тройных пересечения;
- наконец, вычесть пересечение всех четырёх множеств.

(Понятно, почему попарных пересечений шесть? потому что $C_4^2 = 6$.)

▷ Чтобы доказать эту формулу, тоже можно нарисовать «представительную» картинку для четырёх множеств, где были бы все 16 вариантов принадлежности и непринадлежности. Это уже не так просто, но если не настаивать на симметрии, то можно:



По этой картинке тоже можно посчитать, сколько раз вошла и вышла каждая часть. Но нам хотелось бы доказать аналогичную формулу для любого числа множеств, так что надо искать другое доказательство. ◁

6.4* Докажите общую *формулу включений и исключений*:

$$|A_1 \cup \ldots \cup A_n| = \sum_i |A_i| - \sum_{i<j} |A_i \cap A_j| + \sum_{i<j<k} |A_i \cap A_j \cap A_k| - \ldots - (-1)^n |A_1 \cap \ldots \cap A_n|.$$

Здесь написано, что для подсчёта размера объединения надо сложить размеры множеств, вычесть размеры попарных пересечений, прибавить размеры тройных пересечений и так далее (последний знак будет плюс при нечётном n и минус при чётном n, поэтому там и стоит коэффициент $-(-1)^n$).

6.5* Докажите, что выражение

$$\sum_i |A_i| - 2\sum_{i<j} |A_i \cap A_j| + 4 \sum_{i<j<k} |A_i \cap A_j \cap A_k| - \ldots - (-2)^{n-1} |A_1 \cap \ldots \cap A_n|.$$

(добавлены степени двойки) тоже имеет комбинаторный смысл: оно равно числу элементов, которые входят в *нечётное* число множеств $A_1, \ldots, A_n$. (Если множества A_i не пересекаются, то по-прежнему получается сумма.)

Приведём два примера, где можно использовать формулу включений и исключений: (1) перестановки без неподвижных точек и (2) слова, использующие все буквы (сюръекции).

6.6* Слово состоит из n различных букв. Сколько разных слов можно составить, переставляя эти буквы, если дополнительно требуется, чтобы ни одна буква не осталась на своём прежнем месте?

▷ Можно было бы спросить, сколькими способами n писем для n адресатов можно разложить по n надписанным конвертам, если нужно, чтобы каждый адресат получил бы неправильное письмо. (А математики сказали бы коротко «найти число перестановок без неподвижных точек».) ◁

▷ Знатоки математического анализа сразу скажут, что выражение в скобках с ростом n быстро приближается к значению $1/e$, где $e = 2{,}7182818284\ldots$ — так называемое «основание натуральных логарифмов», так что в примерно 37% случаев ни одна буква не стоит на своём месте.

(А знатоки теории вероятностей добавили бы, что в среднем в перестановке n букв есть одна неподвижная точка, так что не очень и удивительно, что около трети слов не имеют её — зато другие имеют больше одной.) ◁

А вот второй пример, где полезна формула включений и исключений.

Мы знаем, что слов длины n в алфавите из k букв (то есть последовательностей из n символов, выбираемых из k-элементного списка) всего k^n (для каждой буквы есть k вариантов). Наложим дополнительное ограничение: *в слове должны встречаться все буквы*. Сколько тогда останется слов?

Вопрос имеет смысл при $n \geqslant k$ (иначе мест для всех букв не хватит). Скажем, можно спросить (и мы сделаем это в следующей задаче), сколько существует десятизначных чисел с нечётными цифрами, в которых каждая из пяти нечётных цифр встречается не менее одного раза. Это соответствует $n = 10$, $k = 5$. Но нет смысла спрашивать, сколько бывает пятизначных чисел, в которых встречаются все 10 цифр — их не бывает.

В случае $k = n$ ответ будет $n!$ (или $k!$), потому что мест столько же, сколько букв, и если каждая буква встречается, то ровно один раз (иначе мест для других не хватит). Но как найти ответ при $n > k$?

$\boxed{\mathbf{6.7}^*}$ Сколько существует десятизначных чисел, составленных из нечётных цифр, в которых каждая нечётная цифра использована (встречается хоть раз)? (Скажем, таковы числа 1975311111 или 5533997711, но не 1111335777.)

▷ Для вычисления по рекуррентной формуле можно написать простую программу типа такой (это `python`):

```
N=10
K=5
T = [[0 for i in range(K+1)] for j in range(N+1)]
```

```
T[1][1]=1
for n in range(2,N+1):
    for k in range(1,K+1):
        T[n][k]=k*(T[n-1][k-1]+T[n-1][k])
print (T[10][5])
```

В этой программе мы окружаем интересующую нас таблицу нулями, чтобы не разбирать случаи $k = 1$ и $k = n$ отдельно. Благодаря этому единственное ненулевое значение, которое нам надо явно указать, это $T(1, 1) = 1$. ◁

• Посмотрим ещё раз на десятизначные числа из нечётных цифр, в которых встречаются все нечётные цифры. В таком числе можно переставить цифры по какому-то правилу. Скажем, если применить перестановку (1→5, 3→3, 5→7, 7→1, 9→9) к числу 5533997711, то получится другое число того же типа, а именно 7733991155. (Как сказали бы математики, «на нашем множестве действует группа перестановок пяти цифр».)

Сколько чисел можно получить из одного такими перестановками? («Сколько элементов в орбите одного числа при действии группы перестановок?») Разные перестановки дадут разные числа (поскольку все цифры входят, отличие проявится), поэтому получится $120 = 5!$ чисел в каждой группе. Такая группа соответствует разбиению всех разрядов числа (позиций в его записи) на пять классов (групп) — в каждом классе стоит одна и та же цифра. Можно сказать, что мы сначала разбиваем разряды на пять групп, а потом решаем, какую цифру назначить для каждой группы. Способов такого разбиения (на пять групп) будет $5\,103\,000/120 = 42\,525$ (из нашего рассуждения было сразу ясно, что поделится нацело).

Тот же вывод в общем случае: *число разбиений множества из n элементов на k различных классов равно $T(n, k)/k!$*. Если обозначить его через $S(n, k)$, то получается рекуррентная формула

$$S(n, k) = S(n - 1, k - 1) + kS(n - 1, k).$$

Надо только отчётливо понимать, чтó именно мы подсчитываем, говоря о числе разбиений: разбиение определяется тем, какие позиции входят в один класс, а какие в разные (а, скажем, порядок перечисления классов не учитывается). Математики сказали бы, что мы подсчитываем «отношения эквивалентности на n-элементном множестве с k классами эквивалентности».

▷ Для чисел $S(n, k)$ есть традиционное название: *числа Стирлинга второго рода*. (Это тот же самый английский математик Стирлинг (1692–1770), который нашёл приближённую формулу для факториалов: $n! \approx \sqrt{2\pi n}(n/e)^n$.) ◁