Set theory and combinatorics

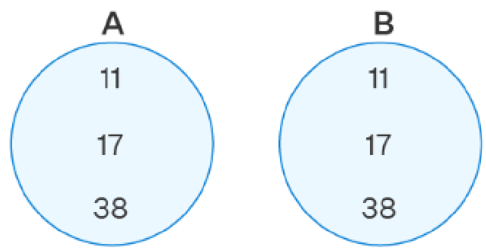
Olli-Pekka Hämäläinen

Set theory

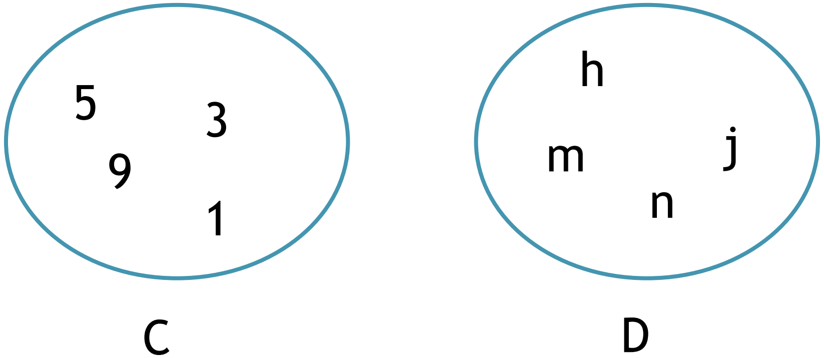
- Set = a group of certain elements
- Elements are in random order unless otherwise specified (ordered set)
- A set can be defined by
  - Listing the elements (for example $\{1,3,5,7,9\}$)
  - Specifying the property that defines the set (for example $A = \{x \in \mathbb{Z} \mid p(x)\}$, where set *A* consists of all elements that have property *p*)
- Example: $G = x \in \mathbb{N} \middle| \sqrt{3x+1} < 15$
  - Solving the inequality gives us the elements, so we could write the set G as a list: $G = \{1,2,3,\ldots,74\}$

Equal vs. equivalent sets

- Two sets A and B are *equal*, if they have exactly the same elements and same *cardinality* (# of elements)
  - Notation: A = B



- Two sets C and D are said to be *equivalent* if they have the same cardinality
  - Notation: C ~ D or C ≡ D



Subsets

- A set can be divided to subsets
- Set A is a subset of B, if every element of A is a member of B
  - In this case it can be said that set A is included in B (or: B includes A)
  - Notation: $A \subseteq B$ (or $B \supseteq A$)
- If B includes also other elements on top of the elements in A, set A is called a proper subset or strict subset of B
  - Mathematically speaking: $A \subseteq B$, but $A \neq B$
- If set A is not included in B, then it’s not a subset of B
  - Notation : $A \not\subset B$

Properties of sets

- Reflexivity: each set includes itself

$$A \subseteq A$$

- Antisymmetry: if A is included in B and B is included in A, then A and B are equal

$$A \subseteq B \ and \ B \subseteq A \iff A = B$$

- Transitivity: If A is included in B and B is included in C, then naturally A is also included in C

$$A \subseteq B \ and \ B \subseteq C \Rightarrow A \subseteq C$$

Basic set

- If a set is defined using a predicate p(x), then set theory analysis is limited only to elements and subsets of a certain basic set X
- The domain of predicates is basic set X
- It is good practice to define the basic set in order to avoid misinterpretations
  - For example, $\{x \mid 3 \leq x \leq 6\}$ is not a good definition for a set, because the interpretation varies with the domain of x (real numbers or integers?)
  - Also, remember to define the basic set if the predicate has limitations!
  - For example, $\left\{x \in \mathbb{Z} \,\middle|\, \sqrt{x+2} > 3\right\}$ is not ok, because the predicate is not defined for all elements of the basic set!

Empty set and universal set

- A set which contains no elements is called an empty set
  - Notation: ∅ tai { }
- Empty set is a subset of every set
- Universal set, on the other hand, is the set that is formed by all possible sets
  - Notation: U
  - Useful mostly in definitions and theorems

Power set

- The power set of A is the set of all subsets of A
  - Notation: $\mathcal{P}(A)$
  - For example, if $A = \{0\}$ and $B = \{0,1,2\}$, then

$$\begin{aligned}
\mathcal{P}(A) &= \{\emptyset, \{0\}\}, \\
\mathcal{P}(\mathcal{P}(A)) &= \mathcal{P}(\{\emptyset, \{0\}\}) = \{\emptyset, \{\emptyset\}, \{\{0\}\}, \{\emptyset, \{0\}\}\}, \\
\mathcal{P}(B) &= \{\emptyset, \{0\}, \{1\}, \{2\}, \{0,1\}, \{0,2\}, \{1,2\}, \{0,1,2\}\}.
\end{aligned}$$

- The number of A's subsets (= number of elements in power set $\mathcal{P}(A)$) is very interesting in multiple applications
- Calculation of this is surprisingly easy: if there are n elements in set A, the number of A's subsets is 2^n

Calculations with sets

- Many set calculations are already familiar to us from statistics, where they play an important role
- For example, union ($A \cup B$), intersection ($A \cap B$) and difference ($A \setminus B$) are probably well understood:

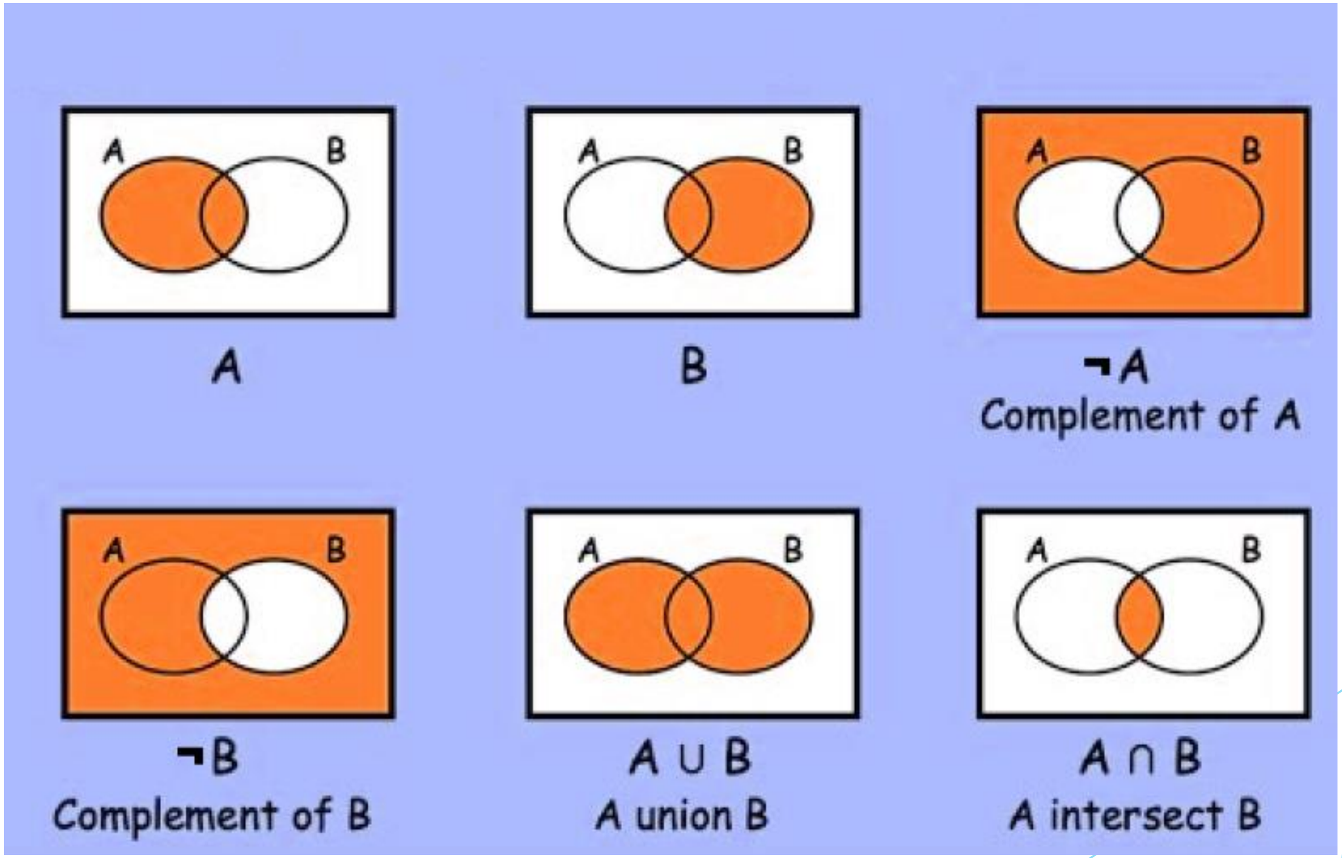
$$A \cup B = \{ x \in X \mid x \in A \vee x \in B \},$$
$$A \cap B = \{ x \in X \mid x \in A \wedge x \in B \},$$
$$A \setminus B = \{ x \in X \mid x \in A \wedge x \notin B \}.$$

- Also complement is probably a familiar concept – especially because it is analogous to the negation in logic:
  - Complement of A = difference of basic set and A

$$\bar{A} = X \setminus A$$

Venn diagram

- Venn diagrams provide us pictorial views of sets and enhance understanding:



Associative laws

- In earlier courses we've used Venn diagrams only for such cases where there are two sets: A and B
- Using the same calculation principles we can also handle cases where we have three or more sets
- First we just need to define the associative laws that we need for calculation of unions and intersections:
  - Quite logical, don't need proof (even though it could be done very simply via Venn diagrams)
  - Analogous to associative laws in logic

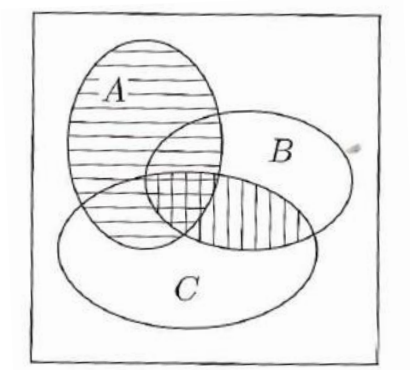
$$A \cap (B \cap C) = (A \cap B) \cap C$$
$$A \cup (B \cup C) = (A \cup B) \cup C$$

Distributive laws

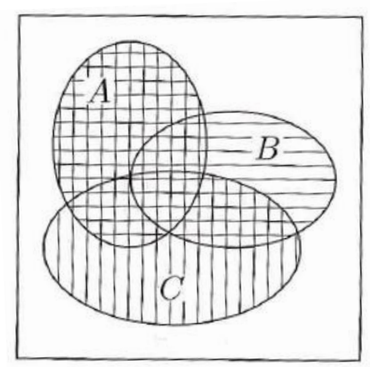
- If we have unions and intersections mixed, the situation is not so obvious
- Distributive laws help:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

  - We can "prove" these via Venn diagrams, too



A: *horizontal lines*
$B \cap C$: *vertical lines*
$A \cup (B \cap C)$: *lines or squares*



$A \cup B$: *horizontal lines*
$A \cup C$: *vertical lines*
$(A \cup B) \cap (A \cup C)$: *squares*

Generalizations of union and intersection

- From associative laws we can make a conclusion that if we have only unions or only intersections in our expression, parentheses don’t matter
- Therefore, the definitions of union and intersection can be generalized to n sets:

$$\bigcup_{k=1}^{n} A_k = A_1 \cup A_2 \cup \cdots \cup A_n = \{\, x \in X \mid \exists k \in \{1, 2, \ldots, n\} \colon x \in A_k \,\},$$

$$\bigcap_{k=1}^{n} A_k = A_1 \cap A_2 \cap \cdots \cap A_n = \{\, x \in X \mid \forall k \in \{1, 2, \ldots, n\} \colon x \in A_k \,\},$$

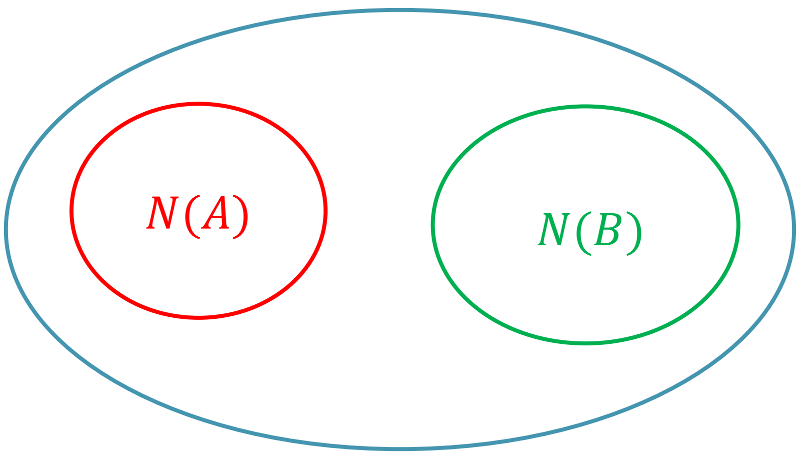
Combinatorics

- In combinatorics, we examine methods for calculating the number of possible elements or subsets of finite sets
- Basics of combinatorics have been studied in previous mathematics courses
- Earlier these have been connected to probability calculations
- This connection is not mandatory: we can be interested in the number of possible combinations without thinking about probabilities
  - …for example, when solving counting problems!
- Let’s revise the earlier topics a bit and then go deeper in the subject

Addition principle

- Let's denote the number of elements in a finite set as N(A)
- Addition principle: if A and B are finite sets that have no common elements, then the number of elements in their union is

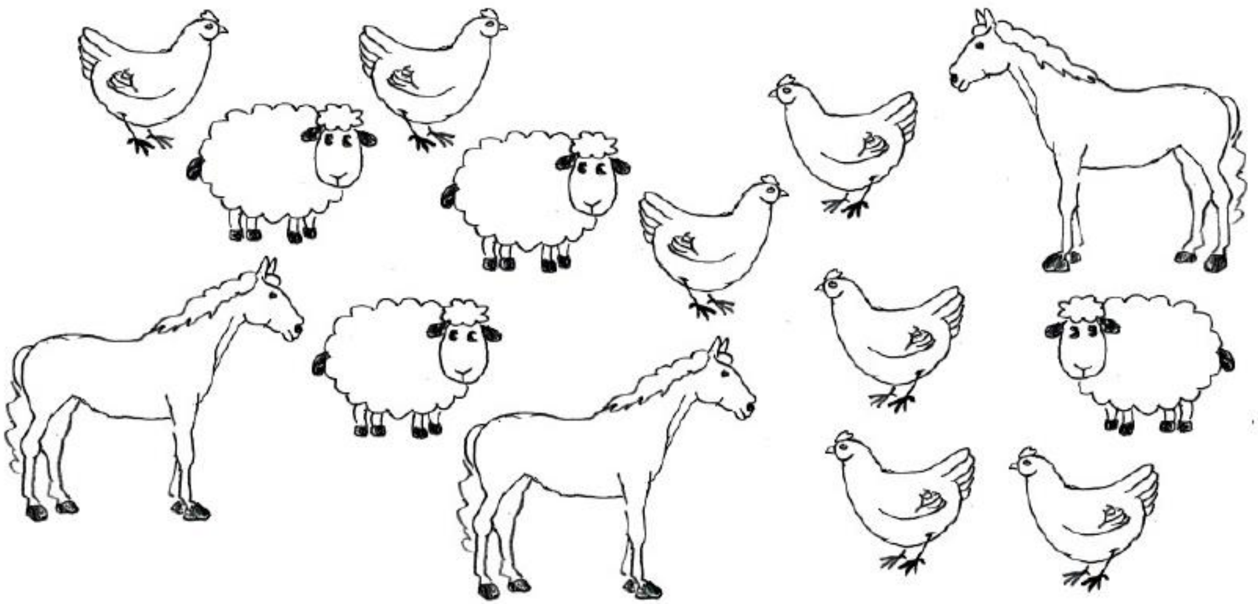
$$N(A \cup B) = N(A) + N(B)$$



Addition principle

- The addition principle can be extended to several sets
- Pictorial example:

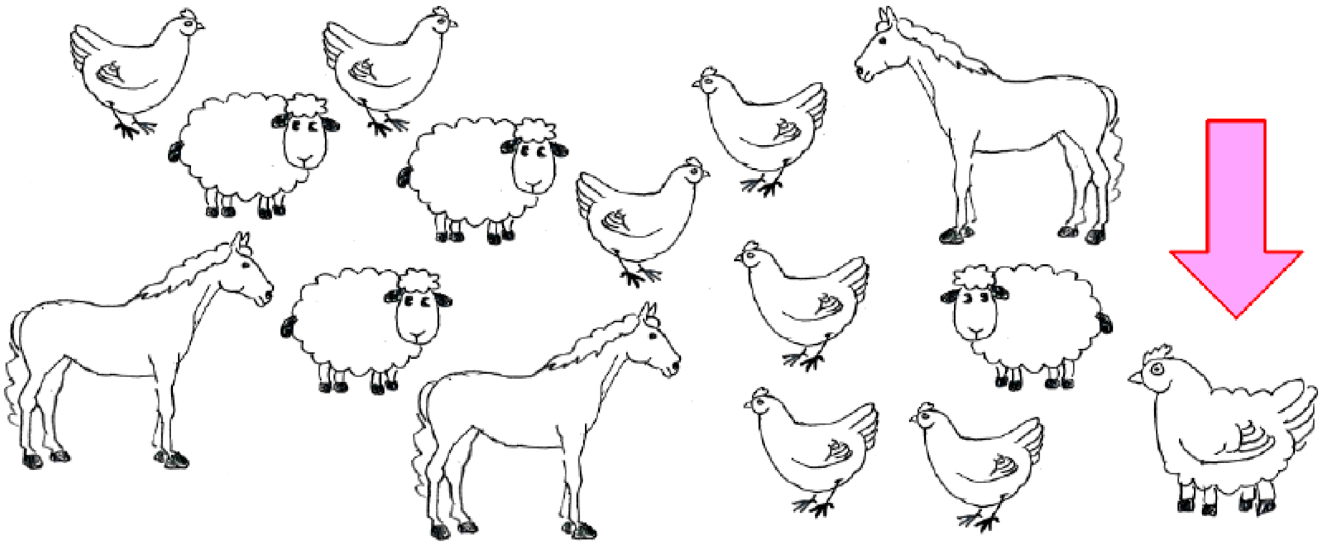
  N(sheep) + N(chickens) + N(horses) = 4 + 7 + 3 = 14



Addition principle

- Remember that the sets may have no overlap!
- If we have an element ("chickensheep") which belongs to two sets at the same time, this principle can't be used

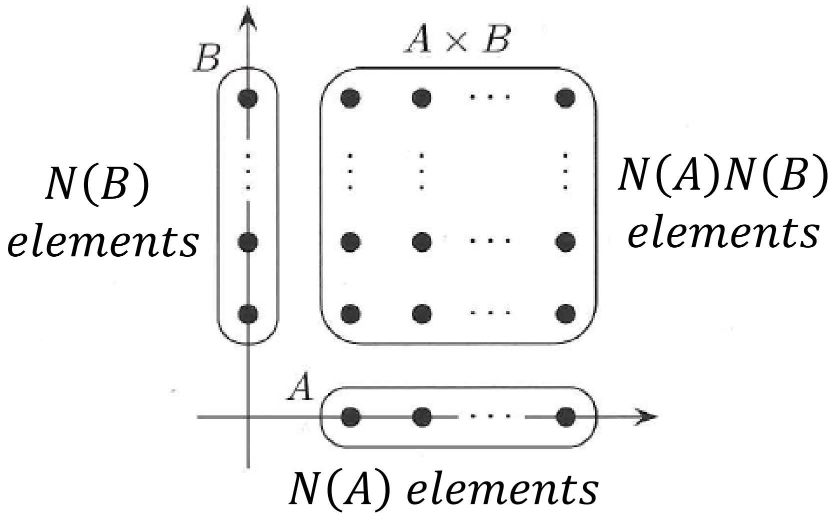
N(sheep) + N(chickens) + N(horses) = 5 + 8 + 3 = 16 $\neq$ 15



Multiplication principle

- If we formulate a new set by selecting one element from A and one element from B, the number of possible elements (ordered pairs) is $N(A \times B)$
- This can be calculated by using the multiplication principle – directly by multiplying the number of elements in each set:

$$N(A \times B) = N(A) \cdot N(B)$$



Addition and multiplication principle: Example 1

- a) How many different 5-bit strings exist?

- b) How many different 5-bit strings, whose 1st bit is 0, exist?

- c) How many different n-bit strings, where at least bits of even order number are 1, exist?

Addition and multiplication principle: Example 1

- a) How many different 5-bit strings exist?

$$A = \{0,1\} \rightarrow N(A) = 2$$

$$N(A) \cdot N(A) \cdot N(A) \cdot N(A) \cdot N(A) = 2^5 = 32$$

- b) How many different 5-bit strings, whose 1st bit is 0, exist?

- c) How many different n-bit strings, where at least bits of even order number are 1, exist?

Addition and multiplication principle: Example 1

- a) How many different 5-bit strings exist?

$$A = \{0,1\} \rightarrow N(A) = 2$$

$$N(A) \cdot N(A) \cdot N(A) \cdot N(A) \cdot N(A) = 2^5 = 32$$

- b) How many different 5-bit strings, whose 1st bit is 0, exist?

$$\{0\} \cdot N(A) \cdot N(A) \cdot N(A) \cdot N(A) = 1 \cdot 2^4 = 16$$

- c) How many different n-bit strings, where at least bits of even order number are 1, exist?

Addition and multiplication principle: Example 1

- a) How many different 5-bit strings exist?

$$A = \{0,1\} \rightarrow N(A) = 2$$

$$N(A) \cdot N(A) \cdot N(A) \cdot N(A) \cdot N(A) = 2^5 = 32$$

- b) How many different 5-bit strings, whose 1st bit is 0, exist?

$$\{0\} \cdot N(A) \cdot N(A) \cdot N(A) \cdot N(A) = 1 \cdot 2^4 = 16$$

- c) How many different n-bit strings, where at least bits of even order number are 1, exist?

$$N(A) \cdot \{1\} \cdot N(A) \cdot \{1\} \cdot \dots \cdot N(A) \cdot \{1\} = 2^{\frac{n}{2}} \text{ (if n even)}$$

$$N(A) \cdot \{1\} \cdot N(A) \cdot \{1\} \cdot \dots \cdot N(A) = 2^{\frac{n+1}{2}} \text{ (if n odd)}$$

Addition and multiplication principle: Example 2

- In an ancient version of BASIC programming language, the variable name could be formed using the letters of the English alphabet A,B,…,Z (26 pcs) and numbers 0,1,…,9 in such a way that the variable name could include one or two characters – with the first one being a letter. Also, the language had five two-letter reserved words (IF, OR etc). Variables are not case sensitive.
- What is the total number of possible variable names?

Addition and multiplication principle: Example 2

- In an ancient version of BASIC programming language, the variable name could be formed using the letters of the English alphabet A,B,…,Z (26 pcs) and numbers 0,1,…,9 in such a way that the variable name could include one or two characters – with the first one being a letter. Also, the language had five two-letter reserved words (IF, OR etc). Variables are not case sensitive.
- What is the total number of possible variable names?
  - Nominate: A = single character (=letter) variable names and B = two-character variable names

$$N(A) = 26$$

Addition and multiplication principle: Example 2

- In an ancient version of BASIC programming language, the variable name could be formed using the letters of the English alphabet A,B,…,Z (26 pcs) and numbers 0,1,…,9 in such a way that the variable name could include one or two characters – with the first one being a letter. Also, the language had five two-letter reserved words (IF, OR etc). Variables are not case sensitive.
- What is the total number of possible variable names?
  - Nominate: A = single character (=letter) variable names and B = two-character variable names

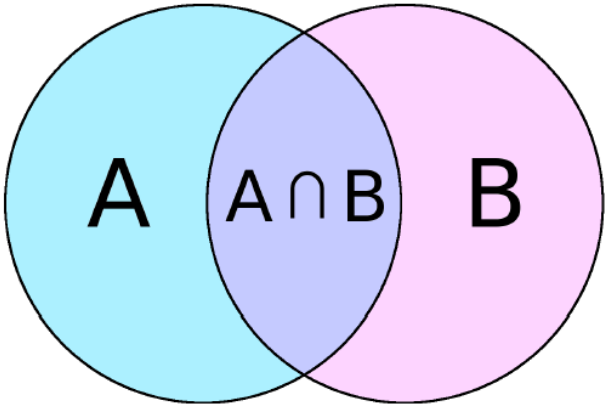
$N(A) = 26 \qquad N(B) = 26(26+10) - 5 = 931$

$$N(C) = N(A) + N(B) = 26 + 931 = \mathbf{957}$$

Inclusion-exclusion principle

- What if sets A and B are not separate but have common elements?
- We can't use addition principle now
- The number of elements in the union of A and B can be solved by using the inclusion-exclusion principle: sum the number of elements of A and B and then remove the number of common elements

$$N(A \cup B) = N(A) + N(B) - N(A \cap B)$$



Inclusion-exclusion principle

- This calculation principle happens to be familiar; we just didn’t know it had a name
- What we (possibly) didn’t know is that this principle can be applied to multiple sets – for example for three:

$$N(A \cup B \cup C) = N(A \cup (B \cup C)) = N(A) + N(B \cup C) - N\big(A \cap (B \cup C)\big) \quad (1)$$

Inclusion-exclusion principle

- This calculation principle happens to be familiar; we just didn't know it had a name
- What we (possibly) didn't know is that this principle can be applied to multiple sets – for example for three:

$$N(A \cup B \cup C) = N(A \cup (B \cup C)) = N(A) + N(B \cup C) - N\big(A \cap (B \cup C)\big) \quad (1)$$

  - The last term can be expressed using the distributive law as

$$N(A \cap (B \cup C)) = N\big((A \cap B) \cup (A \cap C)\big)$$

Inclusion-exclusion principle

- This calculation principle happens to be familiar; we just didn't know it had a name
- What we (possibly) didn't know is that this principle can be applied to multiple sets – for example for three:

$$N(A \cup B \cup C) = N(A \cup (B \cup C)) = N(A) + N(B \cup C) - N\big(A \cap (B \cup C)\big) \quad (1)$$

  - The last term can be expressed using the distributive law as

$$N(A \cap (B \cup C)) = N\big((A \cap B) \cup (A \cap C)\big)$$

  - Apply inclusion-exclusion principle again:

$$N\big((A \cap B) \cup (A \cap C)\big) = N(A \cap B) + N(A \cap C) - N((A \cap B) \cap (A \cap C))$$

Inclusion-exclusion principle

- This calculation principle happens to be familiar; we just didn't know it had a name
- What we (possibly) didn't know is that this principle can be applied to multiple sets – for example for three:

$$N(A \cup B \cup C) = N(A \cup (B \cup C)) = N(A) + N(B \cup C) - N\big(A \cap (B \cup C)\big) \quad (1)$$

  - The last term can be expressed using the distributive law as

$$N(A \cap (B \cup C)) = N\big((A \cap B) \cup (A \cap C)\big)$$

  - Apply inclusion-exclusion principle again:

$$N\big((A \cap B) \cup (A \cap C)\big) = N(A \cap B) + N(A \cap C) - N((A \cap B) \cap (A \cap C))$$

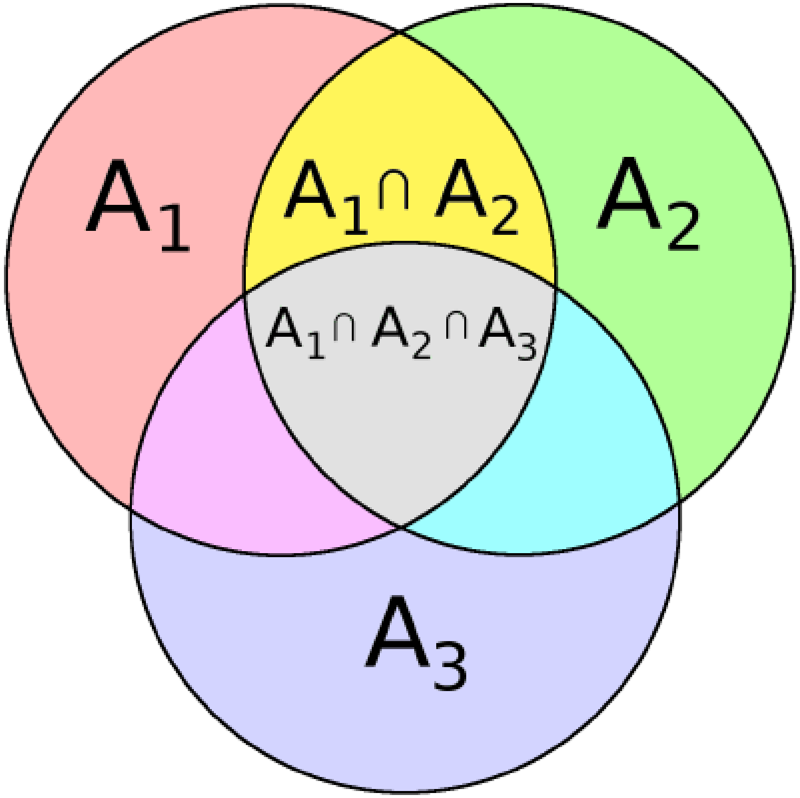
  - Remove unnecessary parentheses by the name of associative law and substitute to (1):

$$N(A \cup B \cup C) = N(A) + N(B \cup C) - \big(N(A \cap B) + N(A \cap C) - N(A \cap B \cap C)\big)$$
$$= \boldsymbol{N(A) + N(B) + N(C) - N(A \cap B) - N(B \cap C) - N(A \cap C) + N(A \cap B \cap C)}$$

Inclusion-exclusion principle

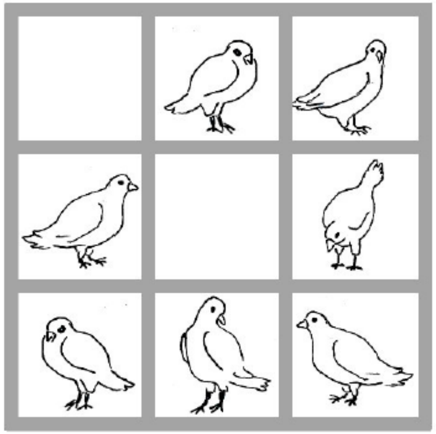
- Graphical representation looks like this:

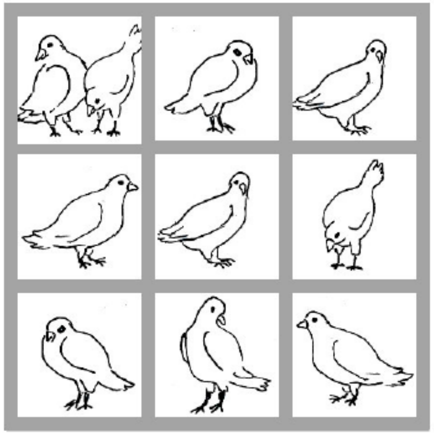
When three sets are summed up and then their intersections are deducted, the elements in gray area (intersection of all 3 sets) get deducted three times; that's why the last term in the expression is positive.



Pigeonhole principle

- If elements are divided to subsets, the number of elements in these subsets is surprisingly often of interest
- Pigeonhole principle presents a rule for defining a minimum for the number of elements
- Represent subsets as boxes and elements as pigeons
- If at least k+1 pigeons must be placed in k boxes, then at least one box will contain at least 2 pigeons





Pigeonhole principle

- …wait, THAT'S math?
- This principle, which at first glance may seem trivial and idiotic, can be generalized followingly:
- If n elements must be placed in k subsets, at least one subset will contain at least [n/k] elements
- Here, brackets represent a ceiling function: [x] means the smallest integer that is ≥ x
  - In Matlab: ceil(x)
  - Rounding rule which rounds upwards – compare to floor(x)

- Example: we have a crowd of 16 people, who are asked to form seven groups. Based on the pigeonhole principle, at least one group will have [16/7] = [2,2857…] = 3 people.

Permutations

- Next a little bit revision: how many ordered queues can we form by selecting elements from an n-element set?
- If the length of the queue will be k elements, then
  - 1^{st} element can be selected from a population of n elements
  - 2^{nd} element can be selected from a population of (n-1) elements
  - 3^{rd} element can be selected from a population of (n-2) elements
  - … k^{th} element can be selected from a pop. of (n-(k-1)) elements
- Based on the multiplication principle, number of queues is

$$n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot (n-k+1)$$

- Using factorial notation we can formulate this a bit shorter:
  - This is the number of k-permutations of n
  - Found in most calculators as nPr (syntax: n nPr k or nPr(n,k))

$$\frac{n!}{(n-k)!}$$

Combinations

- Next we investigate how many ways there are for us to select a subset of k elements from a set of n elements (NOTE! A set has no order!)
- Let's start from the previous result: amount of ordered queues of length k is

$$\frac{n!}{(n-k)!}$$

- These queues contain k elements each, so by the multiplication principle, these elements can be arranged in k! different orders
- So, if we remove the effect of different orders by dividing by k!, we get the number of subsets:
  - This is the number of k-combinations of n

$$\frac{n!}{k!\,(n-k)!}$$

Binomial coefficient

- Previously derived result for number of different subsets is called binomial coefficient

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

- This operation can be found in most calculators, usually by the name nCr (syntax: n nCr k or nCr(n,k))
- Should be well-known for the majority of students
- Also this is possible to generalize to a situation, where the original set of n elements is divided in multiple subsets
  - In this case it will be a multinomial coefficient

Multinomial coefficient

- In how many ways can we divide a set of n elements to m separate subsets in such a way that the numbers of elements in subsets are k_1, k_2, …, k_m?
- If we start from set 1, its elements k_1 can be selected from the whole population n of the original set
- Next, the set 2 elements k_2 can be selected from a population of n - k_1 elements
- Next, the set 3 elements k_3 can be selected from a population of n - k_1 - k_2 elements etc.
- Following this logic, we get the number of all possible subsets by multiplication:

$$\binom{n}{k_1}\binom{n-k_1}{k_2}\binom{n-k_1-k_2}{k_3}\dots\binom{n-k_1-k_2-\cdots-k_{m-1}}{k_m}$$

Multinomial coefficient

- The previously derived product of binomial coefficients can be simplified to a multinomial coefficient:

$$\binom{n}{k_1, k_2, \ldots, k_m} = \frac{n!}{k_1!\, k_2! \ldots k_m!}$$

- Example: how many different strings of characters can we form from letter tokens of word MIMMIMAMMA?
  - 10 letters; 6 pcs of M, 2 pcs of I and 2 pcs of A
  - Different combinations for order numbers of letters:

$$\binom{10}{6,2,2} = \frac{10!}{6! \cdot 2! \cdot 2!} = \frac{3\,628\,800}{720 \cdot 2 \cdot 2} = 1260$$

Multicombination

- In previous examples the elements have been non-replaceable – that is, if some element has been picked from set n to subset k, it is removed from n and can't be selected again
- If the elements are replaceable (so, we can select the same element again), our theory requires some changes
- In this case we're talking about multicombinations
- If there are n elements in the original set and k get selected, the number of possible multicombinations is

$$\binom{n+k-1}{k} = \frac{(n+k-1)!}{(n-1)!\,k!}$$

  - Formula: BETA 17.1 -> Sampling

Multiset

- In previous case our base set n included only separate elements, and all elements were in the base set only once
- If we were to have such a situation where our sets can include same elements multiple times, we talk about multisets and their combinations
- Calculation of these is quite complicated
- For example, the possible 4-combinations of a multiset {0,1,1,2,2,2,3} are:

```
[0, 1, 1, 2]     [1, 1, 2, 2]
[0, 1, 1, 3]     [1, 1, 2, 3]
[0, 1, 2, 2]     [1, 2, 2, 2]
[0, 1, 2, 3]     [1, 2, 2, 3]
[0, 2, 2, 2]     [2, 2, 2, 3]
[0, 2, 2, 3]
```

Example: Pizza!

- Pizza place has an advertisement:
  - 21 different toppings, pick four!
  - Over 10 000 different pizzas!
- Is the advertisement correct?

Example: Pizza!

- Pizza place has an advertisement:
  - 21 different toppings, pick four!
  - Over 10 000 different pizzas!
- Is the advertisement correct?
- Calculate the amount of possibilities for each number of different toppings:
  - 4 diff. toppings: $\binom{21}{4} = 5985$ pcs (select four out of 21)

Example: Pizza!

- Pizza place has an advertisement:
  - 21 different toppings, pick four!
  - Over 10 000 different pizzas!
- Is the advertisement correct?
- Calculate the amount of possibilities for each number of different toppings:
  - 4 diff. toppings: $\binom{21}{4} = 5985$ pcs
  - 3 diff. toppings: $3 \cdot \binom{21}{3} = 3 \cdot 1330 = 3990$ pcs (three toppings, of which one is put on top of pizza as double)

Example: Pizza!

- Pizza place has an advertisement:
  - 21 different toppings, pick four!
  - Over 10 000 different pizzas!
- Is the advertisement correct?
- Calculate the amount of possibilities for each number of different toppings:
  - 4 diff. toppings: $\binom{21}{4} = 5985$ pcs
  - 3 diff. toppings: $3 \cdot \binom{21}{3} = 3 \cdot 1330 = 3990$ pcs
  - 2 diff. toppings: $3 \cdot \binom{21}{2} = 3 \cdot 210 = 630$ pcs (two toppings, which can have divisions 1+3, 2+2 or 3+1)

Example: Pizza!

- Pizza place has an advertisement:
  - 21 different toppings, pick four!
  - Over 10 000 different pizzas!
- Is the advertisement correct?
- Calculate the amount of possibilities for each number of different toppings:
  - 4 diff. toppings: $\binom{21}{4} = 5985$ pcs
  - 3 diff. toppings: $3 \cdot \binom{21}{3} = 3 \cdot 1330 = 3990$ pcs
  - 2 diff. toppings: $3 \cdot \binom{21}{2} = 3 \cdot 210 = 630$ pcs
  - 1 topping: 21 pcs (one quadruple topping)

Example: Pizza!

- Pizza place has an advertisement:
  - 21 different toppings, pick four!
  - Over 10 000 different pizzas!
- Is the advertisement correct?
- Calculate the amount of possibilities for each number of different toppings:
  - 4 diff. toppings: $\binom{21}{4} = 5985$ pcs
  - 3 diff. toppings: $3 \cdot \binom{21}{3} = 3 \cdot 1330 = 3990$ pcs
  - 2 diff. toppings: $3 \cdot \binom{21}{2} = 3 \cdot 210 = 630$ pcs
  - 1 topping: 21 pcs
- Combined: 5985 + 3990 + 630 + 21 = 10 626 pcs
- Conclusion: advertisement is correct!

Thank you!