1 Inclusion-Exclusion for two sets

Suppose we have two sets $A, B \subseteq S$, and want to find $|A \cup B|$. We know that if $A \cap B = \emptyset$, then this is exactly the job for the sum rule:

$$|A \cup B| = |A| + |B|.$$  

But this might not always be the case and $A$ and $B$ might have a nonzero intersection. What do we do then? You have seen in 205 that, more generally,

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

This is called the inclusion-exclusion formula for two sets. It is called so because to find $|A \cup B|$, we start by “including” too much in the our count by including all of $A$ and $B$ (that’s how we get $|A| + |B|$). But clearly, this is an overcount, because the part that is common to $A$ and $B$, i.e. $A \cap B$, is counted twice, once as a part of $A$ and then the second time when included as part of $B$, and so we need “exclude” $|A \cap B|$ to fix our calculation, and this is why we subtract $|A \cap B|$. The result is the above formula. Because we did an “inclusion” followed by an “exclusion”, the principle/process is called inclusion-exclusion. This will become even clearer as we move to more sets.

**Question.** Find the number of passwords of length 10 that can be made using lower case letters, digits, and the symbols $+$ and $\ast$, and contain at least one lower case letter and at least one digit.

**Proof.** Let $U$ be the set of all passwords of length 10 that can be made using lower case letters, digits, and the given symbols (this is the universe), let $A$ be the set of passwords in $U$ that contain at least one lower case letter, and let $B$ the set of passwords in $U$ that contain at least one digit. Our goal is to find $|A \cap B|$ (do you see why?).

We want to use inclusion-exclusion but we are working with an intersection of two sets and not union. How can we change an intersection into a union? De Morgan’s rule!

$$\overline{A \cap B} = \overline{A} \cup \overline{B},$$

here the complement is done inside the universe $U$.

Thus, we can find $|A \cap B|$ by finding $|\overline{A} \cup \overline{B}|$. But how do use $|A \cap B|$ to then find what we really want, which is $|A \cap B|$. We can use the difference method:

$$|A \cap B| = |U| - |\overline{A} \cap \overline{B}| = |U| - |\overline{A} \cup \overline{B}|.$$
Note that \(|U| = 38^{10}\) using the product rule, since we have 10 slots to fill (in order to construct a string), and there are 38 choices for each slot that are independent across slots.

Let us now find \(|\bar{A} \cup \bar{B}|\) using the inclusion-exclusion formula for two sets. But first, note that \(\bar{A}\) is the set of all passwords in the universe \(U\) that do not contain a lower case letter and \(\bar{B}\) is the set of all passwords that do not contain any digits. Thus,

\[
|\bar{A}| = (10 + 2)^{10} = 12^{10},
\]

and

\[
|\bar{B}| = (26 + 2)^{10} = 28^{10}.
\]

Using inclusion-exclusion for two sets, we know that

\[
|\bar{A} \cup \bar{B}| = |\bar{A}| + |\bar{B}| - |\bar{A} \cap \bar{B}|.
\]

All we need to do now is find \(|\bar{A} \cap \bar{B}|\). Observe that \(\bar{A} \cap \bar{B}\) is the set of all passwords in \(U\) that neither contain lower case letters nor contain digits, and so they only contain the symbols + and *. Thus,

\[
|\bar{A} \cup \bar{B}| = 2^{10}.
\]

Putting in the values, we see that

\[
|\bar{A} \cup \bar{B}| = 12^{10} + 28^{10} - 2^{10},
\]

and so

\[
|A \cap B| = 38^{10} - (12^{10} + 28^{10} - 2^{10}).
\]

\[
2 \quad \text{Inclusion-exclusion for three sets}
\]

What do we do to find the size of the union of three sets, i.e. find \(|A \cup B \cup C|\)? Let us try and derive a formula for the cardinality in this case.

**Theorem 1.** Let \(A, B, C\) be subsets of \(S\), then

\[
|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|.
\]

**Proof.** Let \(X = B \cup C\). Then we have using the formula for union of two sets that

\[
|A \cup B \cup C| = |A \cup X| = |A| + |X| - |A \cap X| = |A| + |B \cup C| - |A \cap (B \cup C)| = |A| + |B \cup C| - |(A \cap B) \cup (A \cap C)|.
\]

We now again appeal to the formula for union of two sets to compute \(|B \cup C|\) and \(|(A \cap B) \cup (A \cap C)|\). We see that

\[
|B \cup C| = |B| + |C| - |B \cap C|,
\]

and

\[
|(A \cap B) \cup (A \cap C)| = |A| - |A \cap (B \cap C)|.
\]

We thus have

\[
|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|,
\]

as claimed.

\[\square\]

If we substitute these back into Equation 1, we get that

\[ |A ∪ B ∪ C| = |A| + |B| + |C| - |A ∩ B| - |A ∩ C| - |B ∩ C| + |A ∩ B ∩ C|. \]

This is the inclusion-exclusion formula for three sets: we first overestimate \(|A ∪ B ∪ C|\) by including too much (this is the \(|A| + |B| + |C|\) part), then we underestimate it by excluding too much (this is the \(-|A ∩ B| - |A ∩ C| - |B ∩ C|\) part), and then we finally include again to get the right count for \(|A ∪ B ∪ C|\) (this is the \(+|A ∩ B ∩ C|\) part).

For the sake of clarity, and to help with recalling the formula, we call will group terms in the formula as follows:

\[ |A ∪ B ∪ C| = \text{level 1 terms} - \text{level 2 terms} + \text{level 3 terms}. \]

Level 1 terms are always positive (or “added” or “included”) and involve cardinalities of the sets participating in the union. As we go from level \(i\) to level \(i+1\) the always sign changes (from + to − or − to +, and so level 2 terms are all negative (or “subtracted” or “excluded”). Level 2 terms involve the cardinalities of all possible intersections of pairs of sets from the participating sets. In this case we have 3 participating sets, and so we can form \(\binom{3}{2} = 3\) pairs, which is why we have 3 level 2 terms. Finally, when we go from level 2 to level 3, the sign changes from − to + and so the terms in level 3 are positive (i.e., “included”). Level 3 terms involve cardinalities of intersections of all 3-tuples of participating sets. There are only 3 participating sets in our case, and so there is only one 3-tuple that can be formed, namely \(A ∩ B ∩ C\).

In the next lecture we will see how to generalize this to the union of an arbitrary number of sets, but for now let’s see some applications of the formula for union of three sets.

### 3 General type of problems that can be solved by inclusion-exclusion

Typically, an inclusion-exclusion problem will be disguised in different ways but the following are some good cues/steps to realize if it indeed is an inclusion-exclusion problem:

- The problem will always involve some sort of “objects”: could be strings, numbers, or something else. Call this set \(U\), the universe of all objects.
- You will be able to realize that there are some \(k\) “conditions” on objects, and not all objects will satisfy all \(k\) conditions: it could vary from object to object. For example, if the universe \(U\) is the set of all integers between 1 and 1000. Then one condition could be “divisible by 2”, another could “is a perfect square”, etc.
- Let \(A_i\) be the set of all “objects” that satisfy the \(i^{th}\) condition. Of course, it could happen that \(A_i ∩ A_j ≠ \emptyset\), where \(i\) and \(j\) represent two distinct conditions.
• The problem will usually ask something like “find the number of objects that satisfy none/all/at least one of \( k \) conditions”.

• Depending on what the problem asks you should write down what is being asked as the cardinality of a set that is a union/intersection of the \( A_i \)'s. For example, if the problem says find the number of objects that satisfy at least one condition, then basically you are being asked to find \( |A_1 \cup A_2 \cup \ldots \cup A_k| \). Or, if the problem asks find all objects that satisfy none of the \( k \) conditions, then you are being asked to find \( |\overline{A_1} \cup \overline{A_2} \cup \ldots \cup \overline{A_k}| \), where the complementation is done inside the universe \( U \).

• Once you have converted the problem into finding the cardinality of unions/intersections. You may have to use De Morgan’s law and/or the difference method to convert the problem into one where we have to find the cardinality of union of sets. Why? Because inclusion-exclusion can only help us find the cardinality of unions of sets. For example, if we want to find \( |A_1 \cup A_2 \cup \ldots \cup A_k| \), we can use the difference method:

\[
|A_1 \cup A_2 \cup \ldots \cup A_k| = |U| - |A_1 \cup A_2 \cup \ldots \cup A_k|.
\]

Typically, it will be easy to find \( |U| \), and to find \( |A_1 \cup A_2 \cup \ldots \cup A_k| \) we can use inclusion-exclusion.

Or, if the problem says find the number of objects that satisfy all \( k \) conditions, then that is equivalent to finding \( |A_1 \cap \ldots \cap A_k| \). This is however an intersection of sets, and not a union. To go to union, we can apply the difference method followed by De Morgan’s law as follows:

\[
|A_1 \cap \ldots \cap A_k| = |U| - |\overline{A_1} \cup \ldots \cup \overline{A_k}|,
\]

which, using De Morgan’s rule, can be written as

\[
|A_1 \cap \ldots \cap A_k| = |U| - |\overline{A_1} \cup \ldots \cup \overline{A_k}|.
\]

Note that \( \overline{A_i} \) is the set of all objects in \( U \) that do not satisfy condition \( i \). Once we have the above expression, all we need to do is find \( |U| \) and use inclusion-exclusion to find \( |\overline{A_1} \cup \ldots \cup \overline{A_k}| \).

**Question**. Let \( S = \{1, \ldots, 100\} \). How many numbers are there in \( S \) that are either multiples of 2 or 3 or 5?

**Proof**. Like we just discussed, inclusion-exclusion comes in handy when you are dealing with a set of objects and want to know how many of those objects satisfy at least one out of two or more given conditions. In this case, the “objects” are the numbers from 1 to 100, and there are three given conditions:

1. divisible by 2
2. divisible by 3
3. divisible by 5

and we want to find the number of “objects”, i.e. numbers, that satisfy at least one of the three above conditions.

The next step in solving such problems is to define a set for every condition: the set of the objects that satisfy that condition. So in this case we will define three sets for the three conditions:
1. Let $A$ be the set of numbers in $S$ that are divisible by 2
2. Let $B$ be the set of numbers in $S$ that are divisible by 3
3. Let $C$ be the set of numbers in $S$ that are divisible by 5

Notice that we want to find all the numbers that are either divisible by 2 or by 3 or by 5, and this translates to $|A \cup B \cup C|$ in the language of set theory ("or" is $\cup$, i.e. union, and "and" is $\cap$, i.e., intersection). Now we can apply the inclusion-exclusion formula for the union of three sets:

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

To find the cardinality of the union of the three sets, we need to find the cardinality of all the sets that occur on the right hand side of the above formula:

- $|A|$ is just the number of numbers in $S$ that are divisible by 2, and this is just 50.
- $|B|$ is the number of numbers divisible by 3, and this is just 33.
- $|C|$ is the number of numbers divisible by 5, and this is just 20.
- $|A \cap B|$ is the number of numbers divisible by both 2 and 3 (i.e. by 6), and this is just 16.
- $|A \cap C|$ is the number of numbers divisible by both 2 and 5 (i.e. by 10), and this is just 10.
- $|B \cap C|$ is the number of numbers divisible by both 3 and 5 (i.e. by 15), and this is just 6.
- $|A \cap B \cap C|$ is the number of numbers divisible by 2, 3, and 5 (i.e. by 30), and this is just 3.

We can now substitute in all these values in the inclusion-exclusion formula we stated above, and we see that

$$|A \cup B \cup C| = 50 + 33 + 20 - 16 - 10 - 6 + 3 = 74.$$ 

In the next lecture we will see many more examples of inclusion-exclusion and will state a general formula for the union of an arbitrary number of sets.