Section 9.3 – Counting Elements of Disjoint Sets: the Addition Rule

The ***Addition Rule***

If set $A = A_1 \cup A_2 \cup \ldots \cup A_k$,where the sets A_i are mutually disjoint (sounds like a partition, by the way…),

Then $N(A) = N(A_1) + N(A_2) + \ldots + N(A_k)$

The difference between ***mutually exclusive*** and ***independent*** events is: a mutually exclusive event can simply be defined as a situation when two events cannot occur at same time whereas independent event occurs when one event remains unaffected by the occurrence of the other event.

https://byjus.com/maths/difference-between-mutually-exclusive-and-independent-events/#:~:text=The%20difference%20between%20mutually%20exclusive,occurrence%20of%20the%20other%20event.

Mutually exclusive → addition rule

*"Can't have an identifier that is both 2 characters **AND** 3 characters long."*

Independent → multiplication rule

"Picking the next character in an identifier doesn't depend on the previous character."*

*** -** Note: identifiers with no repeated characters still are independent, there are just fewer choices.

Application of the Addition Rule

How many Python Identifiers are there of 3 or fewer characters?

First character can be	A, B, C, … , Z,	a, b, c, … , z,	or the _	
	26	26	1	= 53 symbols

Additional characters after the first can be those 53 above, plus the 10 digits (0 .. 9) = 63 symbols

So, there are:
53 identifiers of length 1
53*63 identifiers of length 2 (note the use of the Multiplication Rule)
$53*63^2$ identifiers of length 3 (note the use of the Multiplication Rule)

We can apply the Addition Rule, since the sets of these identifiers of length 1, 2, and 3 are mutually disjoint. For example, there isn't an identifier of both length 2 AND 3.

$53 + 53*63 + 53*63^2$

Addition Rule Multiplication Rule

If we factor out the 53 we have 53 * (1 + 63 + 63^2) = 53 * (63^0 + 63^1 + 63^2), and then we can apply the formula for a geometric series: $53 * \left(\frac{63^3-1}{63-1}\right) = 213{,}749$.

By the way, for completeness, we should acknowledge that we've counted thing(s) that can't be Python identifiers? What are they?

The Python *keywords*, such as ***and***, ***if***, ***print***, and so on. How many are there? It depends on the implementation. The textbook uses 31, while here is a list of 32, https://www.programiz.com/python-programming/keyword-list . Obviously, in the example, we'd need to subtract the number of keywords of length 3 or fewer.

The Difference Rule

Sometimes, it's easier to calculate the complement of what you actually want, rather than calculating it directly. For example, we've discussed 4 digit PIN numbers. Note that here we are not varying the length of the PIN, it is always 4 digits.

We know that there are 10 choices for the first digit, 10 for the second, 10 for the third, and 10 for the fourth. If we apply the Multiplication Rule, that would give us 10^4 = 10,000 different 4 digit PINs.

How many 4 digit PIN numbers have repeated digits? Calculating that directly would be difficult, but it is much easier to calculate how many 4 digit PIN numbers have no repeated digits. We know that there are 10 choices for the first digit, 9 for the second, 8 for the third, and 7 for the fourth. If we apply the Multiplication Rule, that would give us 10*9*8*7, which we could also formulate as an r-Permutation, P(10,4). This means there are 5,040, 4 digit PINs which have no repeated digits.

The Difference Rule states that if A is a finite set, and B is a subset of A, then:

$$N(A - B) = N(A) - N(B)$$

In this example, the larger set A would be all the 4 digit PINs, and the subset B would be the set of 4 digit PINs which have no repeated digits. A – B is then the set of 4 digit PINs with repeated digits.

10,000 – 5,040 = 4,960

So there are 4,960 4 digit PINs with repeated digits.

Residence Hall example

The Inclusion-Exclusion Rule for Two or Three Sets

If A, B, and C are any finite sets, then:

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

and

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

In general, if you know any 7 of the 8 terms (or 3 of 4 terms) in the formula, you can solve for the eighth (or fourth) term.

<u>Application of the Inclusion-Exclusion Rule</u>

A professor surveys the 50 students in his class, asking them to check which of the three Mathematics and Computer Sciences courses, MA 100, MA 130, and CS 110, they have completed. The results are:

30 students completed MA 100
18 students completed MA 130
26 students completed CS 110
16 students completed both MA 100 and CS 110
8 students completed both MA 130 and CS 110
9 students completed both MA 100 and MA 130
47 students completed at least one of the three courses

How many students didn't complete any of the three courses?
Using the Difference Rule, the number of students who did not complete any of the three courses is the total number of students minus the number who completed at least one course.
Here, 50 – 47 = 3 students did not complete any of the three courses.

How many students took all three courses?
If we consider the following application of the Inclusion-Exclusion Rule:

$$N(MA100 \cup MA130 \cup CS110) = N(MA100) + N(MA130) + N(CS110)$$
$$- N(MA100 \cap MA130) - N(MA100 \cap CS110) - N(MA130 \cap CS110)$$
$$+ N(MA100 \cap MA130 \cap CS110)$$
$$47 = 30 + 18 + 26$$
$$-9 \quad -16 \quad -8$$
$$+ N(MA100 \cap MA130 \cap CS110)$$

$$47 = 30 + 18 + 26 - 9 - 16 - 8 + N(MA100 \cap MA130 \cap CS110)$$

$$47 = 41 + N(MA100 \cap MA130 \cap CS110)$$

$$6 = N(MA100 \cap MA130 \cap CS110)$$

So, 6 students completed all three of MA 100, MA 130, and CS 110.