C1. Counting.

Notes originally prepared by Judy-anne Osborn.
Editing, expansion and additions by Malcolm Brooks.
Tinkering by Adam Piggott.

Text Reference (Epp)	3ed: Sections	6.1-7, 7.3
	4ed: Sections	9.1-7
	5ed: Sections	9.1-7

Cardinality

This section is mostly about calculating the number of objects of some specified type; for example counting all five digit numbers with no repeated digits. Counting like this can be viewed as finding the number of members of some set, also known as finding the *size* of the set.

For a finite set A, this 'size' or 'cardinality' is just the number of members of A. However it can be defined formally as follows:

Let A be a set. Suppose there exists a bijection (one-to-one correspondence) from A to a subset of the natural numbers of the form $\{1, 2, ..., n\}$ for some $n \in \mathbb{N}$. Then the **cardinality**, or **size** of the set A, written $|A|$, is n. Thus $|A| = n$.

Earlier in the course we saw that this definition is suitable for generalisation to infinite sets (not all infinite sets have the same cardinality.) It also points to some practical counting techniques (see next slide).

Example: numbers in an interval

- What is the cardinality of the set of natural numbers in an interval?
- Example: $S = \{150, 151, 152, ..., 160\}$
- Subtract 149 from each:

$$\begin{array}{ccccc} 150 & 151 & 152 & \dots & 160 \\ \downarrow & \downarrow & \downarrow & \dots & \downarrow \\ 1 & 2 & 3 & \dots & 11 \end{array}$$

- We have made a bijection to the set $\{1, 2, 3, ..., 11\}$, so $|S| = 11$.

Example: numbers in an interval, generalized

- Let $S = \{a, a+1, a+2, ..., b\} \subseteq \mathbb{N}$
- A nice bijection subtracts '$a-1'$ from each element of S. We have

$$\begin{array}{ccccc} a & a+1 & a+2 & \cdots & b \\ \downarrow & \downarrow & \downarrow & \cdots & \downarrow \\ 1 & 2 & 3 & \cdots & b-a+1 \end{array}$$

- Therefore $|S| = b - a + 1$.

A slightly harder but similar example

How many numbers from 150 to 330 inclusive are congruent to 5 mod 7?

150 mod $7 = 3$ so the lowest number is $150 + 2 = 152$.

330 mod $7 = 1$ so the highest number is $330 - 3 = 327$.

Now use the following composition of bijections:

$$
\begin{array}{rccccc}
 & 152 & 159 & 166 & \cdots & 327 \\
\text{subtract 5:} & \downarrow & \downarrow & \downarrow & \cdots & \downarrow \\
 & 147 & 154 & 161 & \cdots & 322 \\
\text{divide by 7:} & \downarrow & \downarrow & \downarrow & \cdots & \downarrow \\
 & 21 & 22 & 23 & \cdots & 46 \\
\text{subtract 20:} & \downarrow & \downarrow & \downarrow & \cdots & \downarrow \\
 & 1 & 2 & 3 & \cdots & 26
\end{array}
$$

Since the composition of bijections is a bijection, the answer is 26.

Finite and infinite sets

The **cardinality of the empty set** is defined to be 0. Thus $|\emptyset| = 0$.

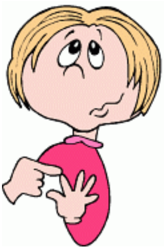
Any set S with cardinality $|S| = n \in \mathbb{N}^\star$ it is said to be **finite**.

Any set S that is not finite is said to be **infinite**. We write $|S| = \infty$.

Examples:

Finite Sets	Infinite Sets
$\{1, 2, 3\}$	$\mathbb{N}$ ... natural numbers
{red, orange, yellow, green, blue, purple}	$\mathbb{Z}$ ... integers
$\{b$: b is a book in the Hancock library$\}$	$\mathbb{Q}$ ... rational numbers
$\{s$: s is a star in the Milky Way Galaxy$\}$	$\mathbb{R}$ ... real numbers
$\{\}$	$\mathcal{P}(\mathbb{R})$ ... power set of $\mathbb{R}$

Countability

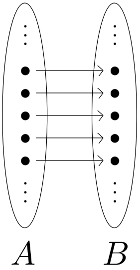


A set S is called **countable** if (and only if) there is a bijection (one-to-one-correspondence) from S to a subset of the set $\mathbb{N}$ of natural numbers.

Examples:

- Any *finite* set is countable.
- The *empty set* is countable (because $\emptyset \subseteq \mathbb{N}$).
- The set $\mathbb{P}$ of all primes is countable (because $\mathbb{P} \subseteq \mathbb{N}$).
- $\mathbb{N}$ itself is countable (because $\mathbb{N} \subseteq \mathbb{N}$).
- The sets $\mathbb{N}$ and $\mathbb{P}$ are each both countable and infinite. Such sets are called **countably infinite**.

Comparing cardinalities



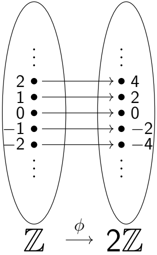
Generalising from the case of finite sets, we say that two sets A and B have **the same cardinality**, written $|A| = |B|$, provided that there exists a bijection (one-to-one correspondence) from A to B.

Remember that for $\phi : A \to B$ to be a bijection:

- no two arrows point to the same element of B, *i.e.*
$$\forall x, y \in A \;\; x \neq y \implies \phi(x) \neq \phi(y).$$
- each element of B must have an arrow pointing to it: *i.e.*
$$\forall b \in B \; \exists a \in A \;\; \phi(a) = b.$$

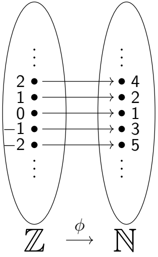
Also remember that $\phi : A \to B$ is a bijection if and only if it has an inverse $\phi^{-1} : B \to A$.
Since the inverse is also a bijection, we say that A and B have the same cardinality if and only if there is a bijection **between** them. (i.e. we don't have to specify the direction of the isomorphism).

Examples of sets with the same cardinality



1 You might expect the set $\mathbb{Z}$ of integers to have a 'bigger' cardinality than the set $2\mathbb{Z} = \{2z : z \in \mathbb{Z}\}$ of even integers, since there are 'twice as many' integers as even integers.

But actually the two sets have the **same** cardinality, because the function $\phi : \mathbb{Z} \to 2\mathbb{Z} : \quad \phi(z) = 2z$ is (clearly) a bijection.



2 By tweaking the function ϕ a little we can establish the perhaps more surprising fact that $\mathbb{Z}$ also has the same cardinality as $\mathbb{N}$.

The modified version of ϕ is

$$\phi : \mathbb{Z} \to \mathbb{N} : \quad \phi(z) = \begin{cases} 2z & \text{if } z > 0 \\ 1 - 2z & \text{if } z \leq 0 \end{cases}$$

It not difficult to see that this is a bijection.

All countably infinite sets have the same cardinality

Recall that we call a set S countably infinite if (and only if) it is infinite and there is a bijection from S to a subset of $\mathbb{N}$.
In other words, S has the same cardinality as some infinite subset of $\mathbb{N}$. However

Any infinite subset S of $\mathbb{N}$ has the same cardinality as $\mathbb{N}$ itself.

Proof: Since every bijection has an inverse, it is enough to establish a bijection $\phi : \mathbb{N} \to S$. This can be done recursively as follows:

$$\begin{aligned} \phi(1) &= \text{least member of } S \\ \phi(2) &= \text{least member of } S \setminus \{\phi(1)\} \quad (2^{\text{nd}} \text{ least member}) \\ \forall n \in \mathbb{N} \;\; \phi(n+1) &= \text{least member of } S \setminus \{\phi(1), \ldots, \phi(n)\}. \end{aligned}$$

Note: It follows from the result above (and its proof) that proving that an arbitrary infinite set S (not necessarily a subset of $\mathbb{N}$) is countably infinite amounts to showing that it can be '**well-ordered**'. This means that it is possible to order the elements of S in some (perhaps ingenious) way so that S and every subset of S has a 'least' member.

The Pigeonhole Principle

The Pigeonhole Principle

> If $k+1$ or more pigeons occupy k pigeonholes, then at least one hole must contain two or more pigeons!

Examples:

1. If there are 11 players in a soccer team that wins $12-0$, there must be at least one player in the team who scored more than once (assuming no own-goals).

2. If a molecule can exist in 2 different configurations, and you have 10^9 such molecules, at least two of them must be in the same configuration.

Generalized Pigeonhole Principle

In the second example, of 10^9 molecules each exhibiting one or the other of two distinct configurations, we concluded that there must be a pair in the same configuration.

But of course something stronger can be said. **What?**

Yes, at least half of the molecules must be in the same configuration.

More generally:

> If N objects are classified in k categories, then at least one category must contain $\lceil \frac{N}{k} \rceil$ objects.

($\lceil \frac{N}{k} \rceil$ denotes the **ceiling** of $\frac{N}{k}$, the least integer not less than $\frac{N}{k}$.)

Example:
In any set of a thousand words, there must be at least 39 words that start with the same letter, because $\lceil \frac{1000}{26} \rceil = \lceil 38.46 \rceil = 39$.

Harder pigeon hole example I

(Epp(4ed) Q9.4.33)

Let A be a set of six [distinct] positive integers each of which is less than 15. Show that there must be two distinct subsets of A whose elements when added up give the same sum.

The phrase '*there must be two distinct subsets*' in the second sentence of the question suggests that the subsets should play the rôle of the pigeons.

Then the phrase '*give the same sum*' suggests that the possible sums should play the rôle of the pigeon holes. So

$$A = \{a, b, c, d, e, f \in \mathbb{N} \,:\, a < b < c < d < e < f < 15\}$$

Pigeons: subsets of A

Pigeon holes: possible element sums of subsets of A

Now we have to count the pigeons and pigeon holes.

Harder pigeon hole example I (cont.)

How many pigeons?

The set of subsets of A is the power set $\mathcal{P}(A)$, which has cardinality $2^{|A|} = 2^6 = 64$. However the empty subset $\emptyset$ and the entire set A cannot have the same sum as any other subset, so we can ignore these and concentrate on the 62 *proper non-empty* subsets.

How many pigeon holes?

The proper non-empty subset with with the least possible element sum is $\{1\}$, with element sum 1.
The proper non-empty subset with the greatest possible element sum is $\{10, 11, 12, 13, 14\}$, with element sum $5(10 + 14)/2 = 60$.

So there are more pigeons than pigeon holes.

Hence at least two pigeons share the same pigeon hole; *i.e* at least two subsets have the same element sum.

Harder pigeon hole example II

(Epp(4ed) Q9.4.35)

Given a set of 52 distinct integers, show that there must be two whose sum or difference is divisible by 100.

The phrase '*there must be two [integers]*' suggests that the 52 integers should play the rôle of the pigeons.

But note that, assuming the integers are written in ordinary base 10 form, it is only their last two digits that matter.

If a and b are two of the integers and if

$$a' = a \bmod 100, \qquad b' = b \bmod 100,$$

then

(1) $a - b$ is divisible by 100 if and only if $a' = b'$;

(2) $a + b$ is divisible by 100 if and only if $a' + b' = 100$.

So we can take the pigeons to be two-digit numbers, 00 – 99.
In view of (1) we may assume they are distinct and by (2) try to prove that two of them have sum 100.

Harder pigeon hole example II (cont.)

For the pigeon hole principle to work, three things are required:

1. each pigeon's number must direct it to one of the holes;
2. there must be less than 52 holes;
3. two pigeons in the same hole must have number sum 100.

Requirement 3 eventually leads to the idea of numbering each pigeon hole with **two** two-digit numbers whose sum is 100:

01	02	03		49
$\square$	$\square$	$\square$	....	$\square$
99	98	97		51

Requirement 1 means that we will need two extra pigeon holes numbered 00 and 50 respectively. These will be occupied by at most one pigeon each (by assumption, pigeons are distinct).

There are now exactly 51 holes, so Requirement 2 is also met.

Thus at least two of the 52 pigeons share a hole and so have sum 100.

Permutations

Andy, Beth and Cai are standing in line.
In how many different orders could this queue be arranged?

Answer: **6**.

To see this, list all possible orderings:

A, B, C	B, A, C	C, A, B
A, C, B	B, C, A	C, B, A

Alternatively, notice that:

there are	3	choices for who is first in the queue,
leaving only	2	choices for who is second,
and then only	1	choice for last place.

So there are $3 \times 2 \times 1 = 3!$ possibilities. By extension we get:

$\boxed{\text{There are } n! \text{ ways to arrange } n \text{ distinct objects in a list.}}$

r-Permutations

A pet show awards 1st, 2nd and 3rd prizes. There are 5 entrants:

Rachel the Rabbit Charles the Chicken Tilly the Terrier
Bob the Bilby Karen the Kangaroo

In how many ways can the prizes be handed out? Answer: **60**

To see this it is not really practicable to list all the possibilities.

Instead, we can proceed by the alternative 'choices' method:

There are	5	choices for which pet gets 1st prize,
leaving	4	choices for which pet gets 2nd prize,
and finally	3	choices for which pet gets 3rd prize.

So there are $5 \times 4 \times 3 = \frac{5!}{2!}$ possibilities. By extension we get:

> There are $\boldsymbol{P(n,r)} = n\times(n-1)\times\cdots\times(n-r+1) = \dfrac{n!}{(n-r)!}$
> ways to select and order r out of n distinct objects.

These ordered selections (lists) are called ***r*-permutations**.

Combinations

Suppose now you just want to know which pets were prizewinners, and you don't care about the order.

How many possible sets of prizewinners? Answer: **10**

The 10 possibilities are:

$$\begin{array}{llll} \{R,C,T\} & \{R,T,B\} & \{C,T,B\} & \{T,B,K\} \\ \{R,C,B\} & \{R,T,K\} & \{C,T,K\} & \\ \{R,C,K\} & \{R,B,K\} & \{C,B,K\} & \end{array}$$

How do we know that we listed all possibilities?

Well, each set corresponds to 6 possible orderings.

For example the set $\{R, C, T\}$ corresponds to orderings

R,C,T R,T,C C,R,T C,T,R T,R,C T,C,R.

So take the $P(5, 3) = 60$ ordered lists of 3 out of 5 objects, and then divide by the number of orderings of 3 objects, namely $3! = 6$, to arrive at the answer $60/6{=}10$.

Combinations and binomial coefficients

Generalising the previous example we get

> There are $\boldsymbol{C(n,r)} = \dfrac{P(n,r)}{r!} = \dfrac{n!}{r!(n-r)!}$ ways to choose a set of r objects from a set of n candidates.
> *I.e.* A set of cardinality n has $\frac{n!}{r!(n-r)!}$ subsets of cardinality r.

The subsets are called ***r*-combinations**.

We say 'n choose r' for $C(n,r)$ and often write it $\dbinom{n}{r}$.

These numbers $\binom{n}{r}$ arise as coefficients in the algebraic expansion of the n-th power of the 'binomial' $(x+y)$ and are consequently also known as **binomial coefficients**. The expansion is

$$(x+y)^n = \sum_{r=0}^{n} \binom{n}{r} x^{n-r} y^r.$$

Two important properties of $\binom{n}{r}$

1. $\forall r, n \in \mathbb{N}^\star \quad 0 \le r \le n \implies$ $\boxed{\binom{n}{r} = \binom{n}{n-r}.}$ *e.g.* $\binom{5}{3} = \binom{5}{2}$.

   Proof: Choosing the r elements of a subset S of U, with $|U| = n$, is exactly equivalent to choosing the $n-r$ elements of U to be left out.

2. $\forall r, n \in \mathbb{N}^\star \quad 0 < r \le n \implies$ $\boxed{\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}.}$
   e.g. $\binom{5}{3} = \binom{4}{3} + \binom{4}{2}$.
   (Pascal's Triangle Identity)

   Proof: Let u be a fixed member of U, with $|U| = n$.
   Subsets S of U with $|S| = r$ are of two types; those that don't contain u and those that do.
   There are $\binom{n-1}{r}$ of the first kind, since the r members of S are chosen from the $n-1$ members of $U \backslash u$.
   There are $\binom{n-1}{r-1}$ of the second kind, since the $r-1$ members of $S \backslash u$ are also chosen from the $n-1$ members of $U \backslash u$.

A mixed example

How many distinguishable ways can the letters of the word

MILLIMICRON

be arranged?

If we were to distinguish between like letters using labels, as in

$$\mathrm{M_1 I_1 L_1 L_2 I_2 M_2 I_3 C\,R\,O\,N}$$

there would be $\quad 11! = 39\,916\,800 \quad$ different arrangements.

Now we must compensate for the over-counting induced by this distinguishing between indistinguishable arrangements.

Since MILLIMICRON has 2 M's, 3 I's and 2 L's the true answer is :

$$\frac{11!}{2!\,3!\,2!} = 1\,663\,200.$$

This example generalises both permutations and combinations.
Can you see how?

'Stars and Bars' example

(Epp(4ed) Q9.6.15)

For how many integers from 1 through 99 999 is the sum of their digits equal to 10?

By inserting leading zeros if necessary, all these integers can be considered as 5-digit strings $abcde$ with $a+b+c+d+e=10$.

Each of these 5-digit strings can be represented as a length-14 pattern of 10 **stars** and 4 **bars**. For example:

$$\begin{array}{lll} \star\star|\star\star\star\star|\star|\star\star|\star & \text{represents} & 24\,121. \\ \star\star\star||\star\star\star\star|\star\star\star| & \text{represents} & 30\,430. \\ |||\star\star\star\star\star\star\star\star\star|\star & \text{represents} & 00\,091. \end{array}$$

There are $\binom{14}{4}$ of these 10-star-4-bar patterns (choose 4 of the 14 possible positions for the bars). But five of the patterns have ten stars in a row and so don't count (ten is not a digit).

So the number of integers is $\displaystyle\binom{14}{4}-5=\frac{14\times13\times12\times11}{4\times3\times2\times1}-5=996.$

Counting 'Multisets'

What is a 'multiset'?

It's a 'set' with multiple copies of elements allowed and acknowledged.
An example is $\{c, b, a, c, a\}$, which has 2 a's, 1 b and 2 c's.
As for ordinary sets, order is irrelevant: $\{c, b, a, c, a\} = \{a, a, b, c, c\}$.
But the multiplicities **do** matter.

Formally, a **size-r multiset** is a set S together with a 'multiplicity function' $m : S \to \mathbb{N}$, where,

$$\forall s \in S \quad m(s) = \text{number of copies of } s \quad \text{and} \quad r = \textstyle\sum_{s \in S} m(s).$$

So, for example, $\{c, b, a, c, a\}$ has size $r = 2 + 1 + 2 = 5$.

How many different size-r multisets can be formed from members of a set S of cardinality n?

Represent each multiset as a pattern of r stars and $n-1$ bars.
For example if $S = \{a, b, c, d\}$ and $r = 5$ then $\{c, b, a, c, a\}$ is represented by $\star\star\,|\,\star\,|\,\star\star|$ $(m(a)=2, m(b)=1, m(c)=2, m(d)=0)$.

$$\boxed{\text{There are } \tbinom{r+n-1}{r} \text{ size-}r \text{ multisets with members from a set of size } n.}$$

Multisets example

(Epp(4ed) Q9.6.6)

If n is a positive integer, how many 5-tuples of integers from 1 through n can be formed in which the elements of the 5-tuple are written in non-increasing order?

For $n = 9$ some 5-tuples are $(8, 6, 4, 2, 1)$, $(9, 3, 3, 2, 2)$, $(6, 6, 6, 6, 6)$.

There is a bijection (one-to-one correspondence) between the set of all these 5-tuples and the set of all size-5 multisets chosen from $\{1, \ldots, n\}$, because the r 'members' of the multiset can only be arranged in one way in non-increasing order.

So by stars-and-bars, there are $\binom{5+n-1}{5} = \binom{n+4}{5}$ of these 5-tuples.

For example for $n = 3$ there are $\binom{7}{5} = \binom{7}{2} = 21$ such 5-tuples:

33333 33332 33331 33322 33321 33311 33222 33221 33211 33111
32222 32221 32211 32111 31111 22222 22221 22211 22111 21111 11111

New counts from old

- The Sum Rule
- The Product Rule
- Inclusion-Exclusion

The Sum Rule

If sets A and B are finite and *disjoint* then the cardinality of their union $A \cup B$ is the sum of the their individual cardinalities, i.e.

$$\boxed{A \cap B = \emptyset \implies |A \cup B| = |A| + |B|.}$$

More generally, if $\{A_1, A_2, ..., A_m\}$, $m \in \mathbb{N}$, is a *partition* of the finite set A then

$$\boxed{|A| = |A_1| + |A_2| + ... + |A_m|.}$$

Example:

For $U = \{-10, \ldots, 10\} \subseteq \mathbb{Z}$ and $S = \{n \in U \ : \ |20 - n^2| > 10\}$, find $|S|$.

Observe that $|20 - n^2| > 10 \iff n^2 < 10 \ \vee \ n^2 > 30$.

So $S = \{-3, -2, \ldots, 2, 3\} \cup \{6, 7, \ldots, 10\} \cup \{-10, -9, \ldots, -6\}$.

Hence $|S| = 7 + 5 + 5 = 17$.

The Product Rule

For finite sets A and B the cardinality of their cartesian product $A \times B$ is the product of the their individual cardinalities, i.e.

$$\boxed{|A \times B| = |A| \times |B|.}$$

More generally, for finite sets $A_1, A_2, ..., A_m$, $m \in \mathbb{N}$,

$$\boxed{|A_1 \times A_2 \times \cdots \times A_m| = |A_1| \times |A_2| \times \cdots \times |A_m|.}$$

Example:

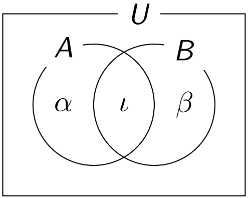
A regular **cubic die, D6,** has face set $C = \{1, 2, 3, 4, 5, 6\}$.

An **octahedral die, D8,**
has face set $O = \{1, 2, \ldots, 8\}$.
A **dodecahedral die, D12,**
has face set $D = \{1, 2, \ldots, 12\}$.

The number of possible outcomes from throwing the three dice together is $|C \times O \times D| = |C| \times |O| \times |D| = 6 \times 8 \times 12 = 576.$

Inclusion-Exclusion

If A and B are finite sets which *may not be disjoint* the sum rule has to be modified to the **inclusion-exclusion rule:**



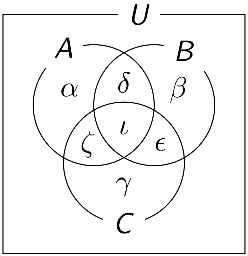
$$\boxed{|A \cup B| = |A| + |B| - |A \cap B|}$$

This is because the plain sum rule $|A \cup B| = |A| + |B|$ includes the intersection $A \cap B$ twice, so it has to be excluded once:

$$|A \cup B| = \alpha + \iota + \beta = (\alpha + \iota) + (\iota + \beta) - \iota = |A| + |B| - |A \cap B|.$$

The inclusion-exclusion rule can be generalised to deal with more than two sets, but it quickly gets very messy.

Can you figure out how to extend the rule to deal with just three sets A, B and C?



Bit-String Example of Inclusion-Exclusion

How many bit-strings of length 8 can be constructed that start with ‘1’ OR end with ‘00’?

Task 1: Construct a string of length 8 that starts with ‘1’.

- There is one way to choose the first bit (1)
- There are two ways to choose the second bit (0 or 1)
- There are two ways to choose the third bit (0 or 1)

$\vdots$

- There are two ways to choose the eighth bit (0 or 1)

Product Rule: Task 1 can be done in $1 \times 2^7 = 128$ ways.

Bit-String Example of Inclusion-Exclusion

Task 2

Task 2: Construct a string of length 8 that ends with '00'.

- There are two ways to choose the first bit (0 or 1)
- There are two ways to choose the second bit (0 or 1)

  ⋮

- There are two ways to choose the sixth bit (0 or 1)
- There is one way to choose the seventh bit (0)
- There is one way to choose the eighth bit (0)

Product Rule: Task 2 can be done in $2^6 \times 1^2 = 64$ ways.

Bit-String Example of Inclusion-Exclusion

Task 3

Is the answer $128+64 = 196$? **NO!**

That would be overcounting.
We have to subtract off the cases we counted twice.

Task 3: Construct a string of length 8 that both starts with '1' and ends with '00'.

- There is one way to choose the first bit (1)
- There are two ways to choose the second bit (0 or 1)

$\vdots$

- There are two ways to choose the sixth bit (0 or 1)
- There is one way to choose the seventh bit (0)
- There is one way to choose the eighth bit (0)

Product Rule: Task 3 can be done in $1 \times 2^5 \times 1^2 = 32$ ways.

Bit-String Example of Inclusion-Exclusion

Conclusion

Finally, the number of ways to construct a bit string of length 8 that starts with '1' or ends with '00' is equal to:

- the number of ways to do task 1, **plus**
- the number of ways to do task 2, **minus**
- the number of ways to do both at the same time (task 3), *i.e.*

$$128 + 64 - 32 = 160.$$

Note: An alternative, and quite different, way to solve this problem is to use **complementary counting**; *i.e.* calculate $|S^c|$ where S is the set of strings we are interested in and U is the set of all (8-bit) strings. Then $|S| = |U| - |S^c| = 2^8 - 1 \times 2^5 \times 3 = 256 - 96 = 160.$

Can you see how to get the $1 \times 2^5 \times 3$?

END OF SECTION C1