Example 9.1 Let $A, B, C$ be finite non-empty sets. Assume that there exist injections $f : A \to C$ and $g : B \to C$ such that

1. $\forall x \in A \cap B \ f(x) = g(x)$, AND
2. $\forall x \in A \ \forall y \in B \ f(x) = g(y) \Rightarrow x, y \in A \cap B$.

Prove that $|A \cup B| \leq |C|$.

Solution: A solution using just the injection rule will be done in recitation.

Here we show another solution that is more roundabout but gives us some useful additional facts.

Proposition 9.2 Let $\iota : X \to Y$ be an injection. Then $|X| = |\text{Ran}(\iota)|$.

Proof: Let $\iota' : X \to \text{Ran}(\iota)$ defined by $\iota'(x) = \iota(x)$. From the definition of range it follows that this is a surjection. 

However, $\iota$ was an injection so it immediately follows that $\iota'$ is an injection too. Hence $\iota'$ is a bijection and by the bijection rule applied to $\iota'$ we have $|X| = |\text{Ran}(\iota)|$. ■

Back to our solution. Applying Proposition 9.2 to $f$ and $g$ it follows that $|A| = |\text{Ran}(f)|$ and $|B| = |\text{Ran}(g)|$.

However, our problem is about $|A \cup B|$. Can we relate it to $|A|$ and $|B|$? We might want to use the sum (addition) rule, however, our sets $A$ and $B$ are not disjoint.

Still, there is a way to reason about counting $|A \cup B|$. (At this point you should draw the potato diagram.) If we add $|A|$ and $|B|$ then we have, of course, counted all the elements in $A \cup B$ but we have counted some of them twice, namely those in $|A \cap B|$. Therefore, we must have

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

This is called the Inclusion-Exclusion rule (principle) for two sets.

This is still not solving our problem because we have to deal with $|A \cap B|$. Consider $h : A \cap B \to C$ defined by $h(z) = f(z)$, or by $h(z) = g(z)$. It follows immediately that $h$ is injective, because $f$ is (or because $g$ is, take your pick).

\footnote{The moral of the story here is that any function gives rise to a surjection when we “restrict” the codomain to the range of the original function.}
Lemma 9.3  \( \text{Ran}(h) = \text{Ran}(f) \cap \text{Ran}(g) \).  

Proof:  (of Lemma)  Let \( u \in \text{Ran}(h) \). Then, \( u = h(z) \) for some \( z \in A \cap B \). Hence \( u = f(x) = g(y) \). Therefore \( u \in \text{Ran}(f) \) and \( u \in \text{Ran}(g) \) so it must be in their intersection.

Conversely, let \( u \in \text{Ran}(f) \cap \text{Ran}(g) \). then \( u = f(x) \) for some \( x \in A \) and \( u = g(y) \) for some \( y \in B \). Hence \( f(x) = g(y) \) and we can apply property (2) deriving \( x, y \in A \cap B \). It follows that \( u \in \text{Ran}(h) \).

So we can finally derive the desired inequality using the inclusion-exclusion rule for both \( A \cup B \) and \( \text{Ran}(f) \cup \text{Ran}(g) \):

\[
\left| A \cup B \right| = \left| A \right| + \left| B \right| - \left| A \cap B \right| = \left| \text{Ran}(f) \right| + \left| \text{Ran}(g) \right| - \left| \text{Ran}(f) \cap \text{Ran}(g) \right| = \left| \text{Ran}(f) \cup \text{Ran}(g) \right| \leq \left| C \right|
\]

(Because \( \text{Ran}(f) \cup \text{Ran}(g) \subseteq C \).)

Inclusion-Exclusion  In fact, the inclusion-exclusion rule for two sets is a consequence of the sum rule. Indeed, using it on several pairs of disjoint sets\(^2\) we have

\[
\begin{align*}
\left| A \cup B \right| &= \left| A \right| + \left| B \setminus A \right| \\
\left| A \cup B \right| &= \left| B \right| + \left| A \setminus B \right| \\
\left| A \right| &= \left| A \setminus B \right| + \left| A \cap B \right| \\
\left| B \right| &= \left| B \setminus A \right| + \left| A \cap B \right|
\end{align*}
\]

Now with a bit of algebra (do it!) we can derive the inclusion-exclusion rule.

Recall in full generality the sum rule refers to two or more sets. Namely if the sets \( A_1, \ldots A_n \) are pairwise disjoint then

\[
\left| \bigcup_{i=1}^{n} A_i \right| = \sum_{i=1}^{n} \left| A_i \right|
\]

How about inclusion-exclusion for more than two sets? For three sets we have

Proposition 9.4

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

\(^2\)Since we have understood how to prove properties of sets if we need to, we will, from now on, just state these properties when needed, relying just on the potato diagrams to get them right. Yay.
Proof: If we add $|A| + |B| + |C|$ (inclusion) then we count the elements of $A \cap B$, $B \cap C$, $C \cap A$ twice (actually, those of $A \cap B \cap C$ thrice!), so we subtract their counts (exclusion), but now we have subtracted the elements of $A \cap B \cap C$ three times so we add back their count (inclusion again). You can also prove using just the sum rule and lots of disjoint sets.

Example 9.5  How many integers between 1 and 150 (inclusive) are divisible by 3, or by 5, or by 7?

Solution: Let’s introduce some notation

$A = \{ n \mid n \in [1..150] \text{ and } 3 \mid n \}$  $B = \{ n \mid n \in [1..150] \text{ and } 5 \mid n \}$  $C = \{ n \mid n \in [1..150] \text{ and } 7 \mid n \}$

We want to compute $|A \cup B \cup C|$. A bit of thought gives $|A| = 50$ and $|B| = 30$. As for $C$, note that 147 is a multiple of 7 and 147/7 = 21 therefore $|C| = 21$. Unfortunately we cannot conclude $|A \cup B \cup C| = 50 + 30 + 21$ because this counts twice the numbers that are divisible by both 3 and 5 (such as 15 or 150), etc. This is where inclusion-exclusion helps. Note that $n \in A \cap B$ iff $15 \mid n$ so $|A \cap B| = 150/15 = 10$. Similarly $|A \cap C| = 147/21 = 7$ and $|B \cap C| = 140/35 = 4$. Finally, $|A \cap B \cap C| = 105/105 = 1$. By inclusion-exclusion we have

$|A \cup B \cup C| = 50 + 30 + 21 - 10 - 7 - 4 + 1 = 81$

More than three sets...  For four sets $A, B, C, D$ we may still have the patience to write down the inclusion-exclusion rule but for five sets it takes utmost dedication. It’s easier to figure out a way to write it in general:

Proposition 9.6

$|\bigcup_{i=1}^{n} A_i| = \sum_{k=1}^{n} (-1)^{k-1} \sum_{J \subseteq [1..n]} \sum_{|J| = k} |\bigcap_{j \in J} A_j|$

We accept this without proof.

Example 9.7  Let $Y$ be a non-empty set. A sequence of elements from $Y$ is called “sursequence” \footnote{I just made up this terminology for this problem.} if each element of $Y$ occurs in the sequence at least once (one or more times). Note that the sursequences of length $r$ correspond one-to-one to the surjections from $[1..r]$ to $Y$. How many sursequences of length $r$ are there? This is the same as counting the surjections from a domain of size $r$ to a codomain of size $n$. 

$\text{Example 9.7}$  Let $Y$ be a non-empty set. A sequence of elements from $Y$ is called “sursequence” \footnote{I just made up this terminology for this problem.} if each element of $Y$ occurs in the sequence at least once (one or more times). Note that the sursequences of length $r$ correspond one-to-one to the surjections from $[1..r]$ to $Y$. How many sursequences of length $r$ are there? This is the same as counting the surjections from a domain of size $r$ to a codomain of size $n$. 

$3$
Solution: Warm-up: If \( r < n \) then we don’t have enough length to use all the elements of \( Y \) so the answer is 0. If \( r = n \) then using each element of \( Y \) one or more times means using it exactly once. Such sequences are permutations of \( n \) and their number is \( n! \).

In general, we will count complementarily, as follows. The total number of sequences, whether sursequences or not, is \( n^r \). We will compute the number of sequences which are not sursequences and then subtract it from \( n^r \).

Let \( Y = \{b_1, \ldots, b_n\} \). For each \( i \in [1..n] \) define \( B_i \) to be the set of sequences of length \( r \) that do not contain \( b_i \). Each sequence that is not a sursequence belongs to at least one of the sets \( B_1, \ldots, B_n \) so their total number is \( |B_1 \cup \cdots \cup B_n| \). These sets are not pairwise disjoint so we use inclusion-exclusion to count the elements of the union.

- For each \( i \in [1..n] \) we have \( |B_i| = (n - 1)^r \). Their sum over all \( i \) is \( n(n - 1)^r = \binom{n}{1}(n - 1)^r \).
- For each \( \{i, j\} \subseteq [1..n] \) \( i \neq j \) we have \( |B_i \cap B_j| = (n - 2)^r \). There are \( \binom{n}{2} \) sets of two distinct indices from \( [1..n] \). The sum over all such is \( \binom{n}{2}(n - 2)^r \).
- Now we generalize this. Let \( J \subseteq [1..n] \) and let \( |J| = k \). Then

\[
\sum_{J \subseteq [1..n] \atop |J| = k} \left| \bigcap_{j \in J} B_j \right| = \binom{n}{k} (n - k)^r
\]

Therefore

\[
|\bigcup_{i=1}^{n} B_i| = \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} (n - k)^r
\]

It follows that the number of sursequences is

\[
n^r - |\bigcup_{i=1}^{n} B_i| = \sum_{k=0}^{n} (-1)^{k} \binom{n}{k} (n - k)^r
\]

This also counts the number of surjections. (Recall that we stated this formula in Lecture Notes 8; if you are interested in more such combinatorics, google “Stirling numbers of the second kind”.)

By the way, using the formula we just obtained for \( r = n \) we have two different ways of counting the permutations of \( n \) and hence a combinatorial proof for the following identity:

\[
\sum_{k=0}^{n} (-1)^{k} \binom{n}{k} (n - k)^n = n!
\]

Maybe you like it more like this?

\[
n^n - \binom{n}{1}(n - 1)^n + \binom{n}{2}(n - 2)^n - \cdots (-1)^{n-1} \binom{n}{n-1} = n!
\]
We also have the following intriguing identity:

\[ \sum_{k=0}^{n} (-1)^k \binom{n}{k} (n-k)^r = 0 \quad (r < n) \]

For \( r = 0 \) we get an identity that can also be immediately derived from the Binomial Theorem applied to \((1 - 1)^n\):

\[ \sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0 \]

For \( r = 1 \) we get

\[ 0 = \sum_{k=0}^{n} (-1)^k \binom{n}{k} (n-k) = n \sum_{k=0}^{n} (-1)^k \binom{n}{k} - \sum_{k=0}^{n} (-1)^k \binom{n}{k} k \]

From this another useful identity follows:

\[ \sum_{k=1}^{n} (-1)^k \binom{n}{k} k = 0 \]

(You can get this one also from the Binomial Theorem: use \((1 + x)^n\), take the derivative of both side, then set \( x = -1 \).)

Example 9.8 \( n \) associates of Al Capone leave their hats with a cloakroom attendant. (Everything is distinguishable.) The attendant gives them back in such a way that none of the gangsters gets their own hat. The returned hats form what is called a “derangement” or “deranged permutation” in memory of the attendant who was clearly deranged. How many deranged permutations are possible?

Solution: We want to compute the number of permutations of \( H = \{h_1, \ldots, h_n\} \) in which \( h_i \) does not occur in position \( i \) for any \( i \in [1..n] \).

We define \( B_i \) to be the set of permutations in which \( h_i \) does occur in position \( i \) and we subtract from the total number of permutations, \( n! \), the count \( |B_1 \cup \cdots \cup B_n| \).

For any \( J \subseteq [1..n] \) let \( |J| = k \) and observe that

\[ |\bigcap_{j \in J} B_j| = (n-x)^k \]

because we already know which elements go in the positions from \( J \). There are \( \binom{n}{k} \) such sets \( J \). Hence

\[ \sum_{\substack{J \subseteq [1..n] \\mid |J|=k}} |\bigcap_{j \in J} B_j| = \binom{n}{k} (n-k)! \]
After applying the inclusion-exclusion formula, subtracting from \( n! \), and doing some algebra (do it!), the number of deranged permutations turns out to be

\[ n! \sum_{k=0}^{n} \frac{(-1)^k}{k!} \]

\[ \blacksquare \]