CSCI/MATH 2113 Discrete Structures

Chapter 8 The Principle of Inclusion and Exclusion

Alyssa Motas

April 7, 2021

Contents

1 8.1 The Principle of Inclusion and Exclusion **3**

1.1 Theorem 5

1.2 Corollary: At least one condition 6

1.3 Notation 6

1.4 Examples 7

2 8.3 Derangements: Nothing Is in Its Right Place **10**

2.1 Definition 10

2.2 Euler's totient function 12

1 8.1 The Principle of Inclusion and Exclusion

Let S represent the set of 100 students. Now let c_1, c_2 denote the following conditions satisfied by some of the elements of S:

c_1 : a student is enrolled in Writing

c_2 : a student is enrolled in Economics

Suppose that 35 students are enrolled in Writing and 30 of them are enrolled in Economics. We shall denote this by

$$N(c_1) = 35 \quad \text{and} \quad N(c_2) = 30.$$

If nine of the students are enrolled in both Writing and Economics, we write $N(c_1c_2) = 9$. Furthermore, there are $100 - 35 = 65$ who are *not* taking Writing and we denote this by writing

$$N(\overline{c_1}) = N - N(c_1) = 65.$$

Similarly,

$$N(\overline{c_2}) = N - N(c_2) = 100 - 30 = 70.$$

The number who are taking Writing and who are *not* taking Economics is

$$N(c_1\overline{c_2}) = N(c_1) - N(c_1c_2) = 35 - 9 = 26.$$

Conversely, we also have

$$N(\overline{c_1}c_2) = N(c_2) - N(c_1c_2) = 30 - 9 = 21.$$

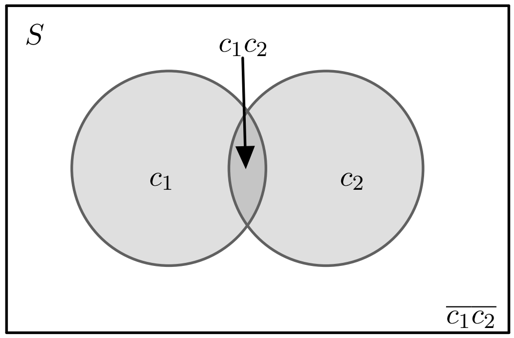
For students who are not taking Writing and Economics, we have

$$N(\overline{c_1}\overline{c_2}) = N(\overline{c_1}) - N(\overline{c_1}c_2) = 65 - 21 = 44$$

because we have $N(\overline{c_1}) = N(\overline{c_1}c_2) + N(\overline{c_1}\overline{c_2})$. Note that

$$\begin{aligned}
N(\overline{c_1}\overline{c_2}) &= N(\overline{c_1}) - N(\overline{c_1}c_2) = [N - N(c_1)] - [N(c_2) - N(c_1c_2)] \\
&= N - N(c_1) - N(c_2) + N(c_1c_2) = N - [N(c_1) + N(c_2)] + N(c_1c_2) \\
&= 100 - [35 + 30] + 9 = 44, \text{ as we saw above.}
\end{aligned}$$

In diagrams, we have



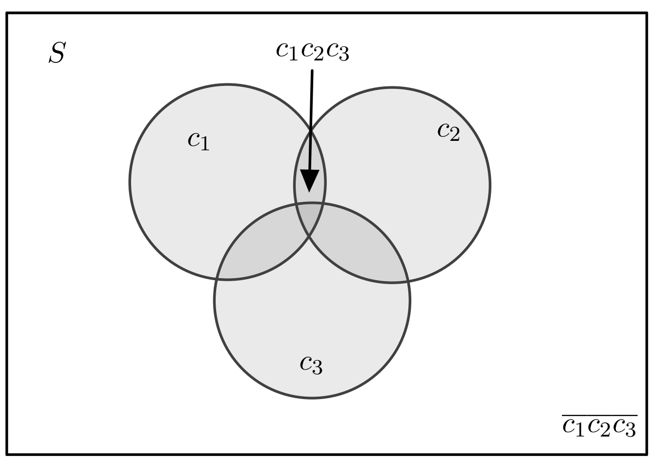
Suppose that we introduce a new condition:

c_3 : a student enrolled in Programming

We then have

$$N(\overline{c_1c_2c_3}) = N - [N(c_1) + N(c_2) + N(c_3)] + [N(c_1c_2) + N(c_1c_3) + N(c_2c_3)] \\ - N(c_1c_2c_3).$$

In diagrams, we have



1.1 Theorem

Let S be a set with $|S| = N$, and conditions $c_i, 1 \leq i \leq t$, each of which may be satisfied by some of the elements of S. The number of elements of S that satisfy *none* of the conditions $c_i, 1 \leq i \leq t$, is denoted by $\overline{N} = N(\overline{c_1}\overline{c_2}\overline{c_3}\ldots\overline{c_t})$ where

$$\begin{aligned}\overline{N} = N &- [N(c_1) + N(c_2) + N(c_3) + \cdots + N(c_t)] \\ &+ [N(c_1c_2) + N(c_1c_3) + \cdots + N(c_1c_t) + N(c_2c_3) + \cdots + N(c_{t-1}c_t)] \\ &- [N(c_1c_2c_3) + N(c_1c_2c_4) + \cdots + N(c_1c_2c_t) + N(c_1c_3c_4) + \ldots \\ &+ N(c_1c_3c_t) + \cdots + N(c_{t-2}, c_{t-1}, c_t)] + \cdots + (-1)^t N(c_1c_2c_3\ldots c_t),\end{aligned}$$

or

$$\begin{aligned}\overline{N} = N - \sum_{1\leq i\leq t} N(c_i) + \sum_{1\leq i<j\leq t} N(c_ic_j) - \sum_{1\leq i<j<k\leq t} N(c_ic_jc_k) + \ldots \\ + (-1)^t N(c_1c_2c_3\ldots c_t).\end{aligned}$$

Proof. We give a combinatorial proof. For $x \in S$, we have:

- If x satisfies none of the conditions then x contributes 1 to each side of the equality.
- If x satisfies exactly r of the conditions ($1 \leq r \leq t$) then x contributes 0 to the LHS of the equality.

Considering the RHS, we get that x adds to

- 1 time in N
- r times in $\sum N(c_i)$
- $\binom{r}{2}$ times in $\sum N(c_ic_j)$
- $\binom{r}{3}$ times in $\sum(c_ic_jc_k)$

  $\vdots$
- $\binom{r}{r}$ times in $\sum N(c_{i_1}c_{i_2}\ldots c_{i_r})$.

In total, the contribution of x is

$$1 - r + \binom{r}{2} - \binom{r}{3} + \cdots + (-1)^r\binom{r}{r} \Rightarrow (1 + (-1))^r = 0.$$

This completes the proof. □

1.2 Corollary: At least one condition

Under the hypotheses of the previous theorem, the number of elements in S that satisfy at least one of the conditions c_i, where $1 \leq i \leq t$, is given by

$$N(c_1 \text{ or } c_2 \text{ or } \ldots \text{ or } c_t) = N - \overline{N}.$$

1.3 Notation

To simplify the theorem above, we write

$$\begin{aligned}
&S_0 = N, \\
&S_1 = [N(c_1) + N(c_2) + \cdots + N(c_t)], \\
&S_2 = [N(c_1c_2) + N(c_1c_3) + \cdots + N(c_1c_t) + N(c_2c_3) + \cdots + N(c_{t-1}c_t)],
\end{aligned}$$

and, in general,

$$S_k = \sum N(c_{i_1}c_{i_2}\ldots c_{i_k}), 1 \leq k \leq t,$$

where the summation is taken over all selections of size k from the collection of t conditions. Hence S_k has $\binom{t}{k}$ summands in it. We can also rewrite the equation above as

$$\overline{N} = S_0 - S_1 + S_2 - S_3 + \cdots + (-1)^t S_t.$$

Remark: Let $U_1, \ldots, U_t$ be sets. Then we have

$$\left|\bigcup_{i=1}^{t} U_i\right|$$

is the number of elements that belong to at least one of the U_i. Thinking of $x \in U_i$ as "x satisfies condition c_i" then we have

$$\begin{aligned}
\left|\bigcup_{i=1}^{t} U_i\right| &= N - \overline{N} = N - (S_0 - S_1 + S_2 - \cdots + (-1)^t S_t) \\
&= S_1 - S_2 + \cdots + (-1)^{t+1} S_t.
\end{aligned}$$

Alternatively, we could write

$$\left|\bigcup_{i=1}^{t} U_i\right| = \sum_{1\leq i\leq t} |U_i| - \sum_{1\leq i<j\leq t} |U_i \cap U_j| + \cdots + (-1)^{t+1}|U_1 \cap U_2 \cap \cdots \cap U_t|,$$

or more concisely as

$$\sum_{\varnothing \neq J \subseteq 1,\ldots,t} (-1)^{|J|+1} \left| \bigcap_{j \in J} U_j \right|.$$

Remark: When $t = 2$, we get the "usual" equality:

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

1.4 Examples

(a) How many integers n $(1 \leq n \leq 100)$ are not divisble by $2, 3$, or 5? Let

$$c_1 : x \text{ is divisible by } 2$$
$$c_2 : x \text{ is divisible by } 3$$
$$c_3 : x \text{ is divisible by } 5$$

We are looking for $N(\overline{c_1} \overline{c_2} \overline{c_3})$. We have the following:

- $N(c_1) = \lfloor \frac{100}{2} \rfloor = 50;$
- $N(c_2) = \lfloor \frac{100}{3} \rfloor = 33;$
- $N(c_3) = \lfloor \frac{100}{5} \rfloor = 20;$
- $N(c_1 c_2) = \lfloor \frac{100}{6} \rfloor = 16;$
- $N(c_1 c_3) = \lfloor \frac{100}{10} \rfloor = 10;$
- $N(c_2 c_3) = \lfloor \frac{100}{15} \rfloor = 6;$
- $N(c_1 c_2 c_3) = \lfloor \frac{100}{30} \rfloor = 3.$

Hence, we have

$$N(\overline{c_1} \overline{c_2} \overline{c_3}) = S_0 - S_1 + S_2 - S_3 = 100 - (50+33+20) + (16+10+6) - 3 = 26.$$

(b) Let A and B be sets with $|A| = m \geq n = |B|$. Let $A = \{a_1, a_2, \ldots, a_m\}$ and $B = \{b_1, b_2, \ldots, b_n\}$. Let us count the number of surjective functions from A to B. For $1 \leq i \leq n$, we define the condition c_i by

$$b_i \text{ is not in the range of } f.$$

Note that c_i is a condition on functions A to B. If f satisfies *none* of the conditions, then f is onto. So we are looking for $N(\overline{c_1} \overline{c_2} \ldots \overline{c_n})$. We have

- $N(c_i) = (n-1)^m \Rightarrow S_1 = n \cdot (n-1)^m$
- $N(c_i c_j) = (n-2)^m \Rightarrow S_2 = \binom{n}{2}(n-2)^m$
- In general, for $1 \leq k \leq n$ we have $S_k = \binom{n}{k}(n-k)^m$.

Therefore, we have

$$\begin{aligned} N(\overline{c_1}\overline{c_2}\ldots\overline{c_n}) &= S_0 - S_1 + S_2 + S_3 + \cdots + (-1)^n S_n \\ &= \sum_{i=0}^{n}(-1)^i S_i \\ &= \sum_{i=0}^{n}(-1)^i \binom{n}{i}(n-i)^m \\ &= \sum_{i=0}^{n}(-1)^i \binom{n}{n-i}(n-i)^m, \end{aligned}$$

which is the number of surjections from A to B.

(c) In how many ways can the 26 letters of the alphabet be arranged so that none of the patterns "car," "dog," "pun," or "byte" appear? Let c_1 be the condition "the arrangement does contain the pattern car." Similarly c_2, c_3, and c_4 are defined for dog, pun, or byte, respectively. Then, we are looking for $\overline{N} = N(\overline{c_1}\overline{c_2}\overline{c_3}\overline{c_4})$. By the Inclusion-Exclusion Principle, we have

$$\begin{aligned} \overline{N} = N &- (N(c_1) + N(c_2) + N(c_3) + N(4)) \\ &+ (N(c_1c_2) + N(c_1c_3) + N(c_1c_4) + N(c_2c_3) + N(c_2c_4) + N(c_3c_4)) \\ &- (N(c_1c_2c_3) + N(c_1c_2c_4) + N(c_1c_3c_4) + N(c_2c_3c_4)) \\ &+ N(c_1c_2c_3c_4). \end{aligned}$$

Then we have the following:

- $N = 26!$
- $N(c_1) = 24! = N(c_2) = N(c_3)$ which is the number of ways we can arrange the "letter" CAR and the 23 remaining letters
- $N(c_4) = 23!$ since "byte" has 4 letters
- $N(c_1c_2) = 22! = N(c_1c_3) = N(c_2c_3)$
- $N(c_ic_4) = 21!$
- $N(c_1c_2c_3) = 20!$

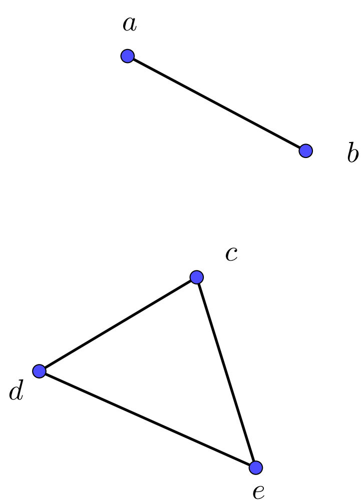
- $N(c_i c_j c_4) = 19!$
- $N(c_1 c_2 c_3 c_4) = 17!$

Hence, we have

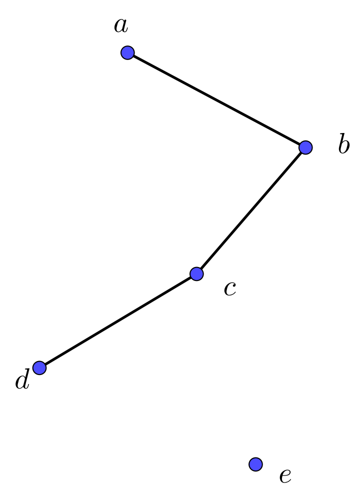
$$\overline{N} = 26! - (3(24!) + 23!) + (3(22!) + 3(21!)) + (20! + 3(19!)) + 17!$$

(d) How many arrangements contain the words "bald" and "blad"? Aside, fun problem to look at: superpermutations.

(e) There are 5 villages. You want to devise a system of roads connecting the villages such that no village is completely isolated. In how many ways can you do this?



No village is isolated.



Village e is isolated.

Let S be the set of all (undirected, loop-free) graphs on the vertices $\{a, b, c, d, e\}$. We know that

$$|S| = 2^{\binom{5}{2}} = 2^{10}.$$

Now, for $i \in \{1, \ldots, 5\}$, the condition c_i is "the system of roads isolates the i-th village." Then for $N(c_1)$, we have the roads $\{b, c\}, \{b, d\}, \{b, e\}, \{c, d\}, \{c, e\}, \{d, e\}$.

Thus

$$N(c_1) = 2^6.$$

Similarly, $N(c_i) = 2^6$ for $1 \le i \le 5$. Reasoning in the same way, we find that $N(c_1c_2) = 2^3$ and similarly $N(c_ic_j) = 2^3$. Finally, $N(c_ic_jc_k) = 2^1$ and $N(c_ic_jc_kc_l) = 2^0 = N(c_1c_2c_3c_4c_5)$. In total, we have

$$N(\overline{c_1c_2c_3c_4c_5}) = 2^{10} - \binom{5}{1}2^6 + \binom{5}{2}2^3 - \binom{5}{3}2^1 + \binom{5}{4}2^0 - \binom{5}{5}2^1.$$

2 8.3 Derangements: Nothing Is in Its Right Place

A professor wants the students to grade the assignments. In this case, we would like every student to receive a single assignment and not their own. This is a derangement.

Write $[n] = \{1, 2, \ldots, n\}$. A *derangement* of $[n]$ is a permutation of $[n]$ such that no element is left in place.

$$\begin{aligned}
[2] \Rightarrow & 21 \text{ since the ``standard'' order is } 12 \\
[3] \Rightarrow & 312, 231 \\
[4] \Rightarrow & 4123 \quad 3142 \quad 2143 \\
& 3412 \quad 4312 \quad 2413 \\
& 2341 \quad 3421 \quad 4321
\end{aligned}$$

2.1 Definition

We write $d(n)$ for the number of derangements of $[n]$.

Proposition: We have

$$d(n) = \sum_{k=0}^{n} (-1)^k \cdot \frac{n!}{k!}$$

Proof. Write T for the set of all permutations of $[n]$. Then we have $|T| = n!$. Now, let $T_i, 1 \le i \le n$, be the collection of permutations that fix i. For example, $[3] \Rightarrow 213 \in T_3$. Then

$$d(n) = n! - \underbrace{\left|\bigcup_{i=1}^{n} T_i\right|}_{\text{all the perm. that fix } i}$$

By Inclusion-Exclusion, we have

$$\left|\bigcup_{i=1}^{n} T_i\right| = \sum_i |T_i| - \sum_{i<j} |T_i \cap T_j| + \sum_{i<j<k} |T_i \cap T_j \cap T_k| - \cdots + (-1)^{n+1}\left|\bigcap_i T_i\right|.$$

- For $\sum |T_i|$, we have $|T_i| = (n-1)!$ hence

$$\sum_i |T_i| = n \cdot (n-1)!$$

- For $\sum_{i<j}$, we have $|T_i \cap T_j| = (n-2)!$ hence

$$\sum_{i<j} |T_i \cap T_j| = \binom{n}{2}(n-2)!$$

Hence we have

$$\begin{aligned}\left|\bigcup_{i=1}^{n} T_i\right| &= n \cdot (n-1)! - \binom{n}{2}(n-2)! + \binom{n}{3}(n-3)! - \dots \\ &= \sum_{k=1}^{n} (-1)^{k+1}\binom{n}{k}(n-k)!\end{aligned}$$

Therefore, we have

$$\begin{aligned} d(n) &= n! - \left|\bigcup_i T_i\right| \\ &= n! - \sum_{k=1}^{n}(-1)^{k+1}\binom{n}{k}(n-k)! \\ &= n! + \sum_{k=1}^{n}(-1)^{k}\binom{n}{k}(n-k)! \\ &= \sum_{k=0}^{n}(-1)^{k}\binom{n}{k}(n-k)! \\ &= \sum_{k=0}^{n}(-1)^{k} \cdot \frac{n!}{k!} \end{aligned}$$

where $\binom{n}{k} = \dfrac{n!}{k!(n-k)!}$. $\square$

2.2 Euler's totient function

Let $n \geq 2$, then we have

$$\phi(n) = |\{x \in [n] \mid \gcd(x, n) = 1\}|$$

is *Euler's totient function.* For example, we have

- For $n = 2$, we have $\phi(2) = 1$ which is 1.
- For $n = 3$, we have $\phi(3) = 2$ which is 1 and 2.
- For $n = 4$, we have $\phi(4) = 2$ which is 1 and 3.
- For $n = 5$, we have $\phi(5) = 4$ which is 1,2,3 and 4.

Remark: If n is prime, then $\phi(n) = n - 1$.

Proposition: We have

$$\phi(n) = n \cdot \prod_{p|n} (1 - \frac{1}{p})$$

where the product is taken over the primes that divide n.

Proof. Given n, we write it as

$$n = P_1^{e_1} P_2^{e_2} \dots P_t^{e_t}$$

where P_i are primes and $e_i \geq 1$. Suppose, for simplicity, that $t = 4$. Let $S = [n]$ and let c_i be the condition "x is divisible by P_i" for $1 \leq i \leq 4$. Then

$$\phi(n) = N(\overline{c_1 c_2 c_3 c_4}).$$

Now,

- $N(c_i) = \dfrac{n}{P_i}$
- $N(c_i c_j) = \dfrac{n}{P_i P_j}$
- $N(c_i c_j c_k) = \dfrac{n}{P_i P_j P_k}$
- $N(c_1 c_2 c_3 c_4) = \dfrac{n}{P_1 P_2 P_3 P_4}$

Then, we have

$$
\begin{aligned}
\phi(n) &= S_0 - S_1 + S_2 - S_3 + S_4 \\
&= n - \left(\frac{n}{P_1} + \cdots + \frac{n}{P_4}\right) + \left(\frac{n}{P_1P_2} + \cdots + \frac{n}{P_3P_4}\right) \\
&\quad - \left(\frac{n}{P_1P_2P_3} + \cdots + \frac{n}{P_2P_3P_4}\right) + \frac{n}{P_1P_2P_3P_4} \\
&= \frac{n}{P_1P_2P_3P_4}(P_1P_2P_3P_4 - (P_2P_3P_4 + \cdots + P_1P_2P_3) \\
&\quad - (P_4 + \cdots + P_1) + 1) \\
&= \frac{n}{P_1P_2P_3P_4}((P_1-1)(P_2-1)(P_3-1)(P_4-1)) \\
&= n\left(\frac{P_1-1}{P_1} \cdot \frac{P_2-1}{P_2} \cdot \frac{P_3-1}{P_3} \cdot \frac{P_4-1}{P_4}\right) \\
&= n \cdot \prod_{i=1}^{4}\left(1 - \frac{1}{P_i}\right),
\end{aligned}
$$

which finishes the proof. ∎