1) Let G be the undirected weighted multigraph having vertices $V = \{0, 1, \ldots, 7\}$ and edges $E = \{0, 1, \ldots, 8\}$ with end-points r and weights b given by:

e	$r(e)$	$b(e)$
0	{0, 1}	3
1	{0, 2}	2
2	{0, 3}	1
3	{1, 3}	2
4	{2, 3}	1
5	{4, 5}	2
6	{4, 6}	3
7	{5, 7}	1
8	{6, 7}	3

Compute a minimal spanning forest of G using Kruskal's algorithm. Is it unique? How many possible spanning forests (not necessarily minimal) are there?

Solution:
When we look at the right component of Figure **??** we have four possibilities to remove one edge. Whereas in the left component we have 5 choose 2 possibilities to remove two edges. However, two of the ten possibilities are not allowed (remove both edges of node 1 or both edges of node 2). Therefore there are eight different spanning trees of the left component. So, in total there are $8 \cdot 4 = 32$ spanning forests.

A minimal spanning forest of G using Kruskal's algorithm is not unique since we have the choice between the edges 6 and 8 to be in our spanning tree. Kruskal's algorithm starts with $F = \varnothing; P = \{\{0\}, \{1\}, \ldots, \{8\}\}$ and terminates with $P = \{\{0, 1, 2, 3\}, \{4, 5, 6, 7\}\}$ and either $F = \{\{0, 3\}, \{2, 3\}, \{1, 3\}, \{4, 5\}, \{5, 7\}, \{4, 6\}\}$ or $F = \{\{0, 3\}, \{2, 3\}, \{1, 3\}, \{4, 5\}, \{5, 7\}, \{6, 7\}\}$.

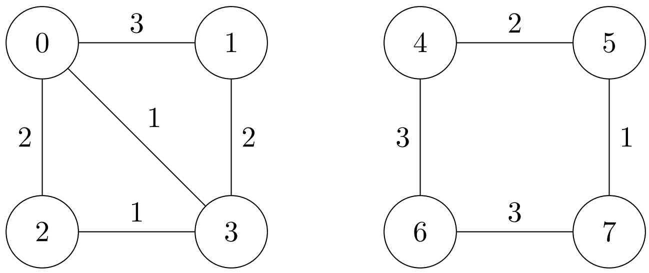


Figure 1: The undirected weighted multigraph G.

2) Consider a set $S = \{1, 2, \ldots, 16\}$ where we pick nine different numbers. Prove that there are always two picked numbers whose sum is 17.

Can you find a straightforward generalisation of your proof when $S = \{1, \ldots, 2n\}$?

Solution:
There are eight possible pairs with sum 17: $\{1, 16\}$, $\{2, 15\}$, $\{3, 14\}$, $\{4, 13\}$, $\{5, 12\}$, $\{6, 11\}$,

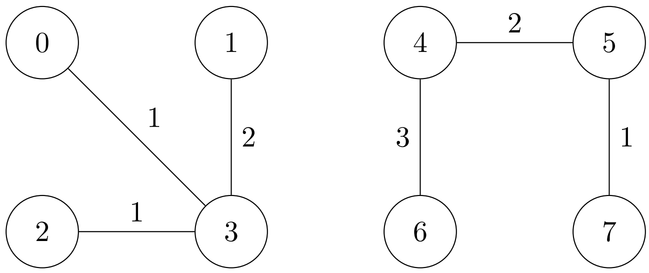


Figure 2: A minimal spanning forest of G.

$\{7,10\}$, $\{8,9\}$. So we pick nine numbers out of eight pairs. From the pigeonhole principle we get that there must be a pair where we picked both numbers. Therefore we picked two numbers whose sum is 17.

More formally: Let $P = \{(a,b) \mid a,b \in S,\ a+b=17\}$ be the set of pairs whose sum is 17. Let $S' \subseteq S$ be the set of numbers we picked. Now consider the function $f : S' \to P,\ f(x) = \{x,\ 17-x\}$ that maps every number in S' to its pair in P. Since $|S'| = 9$ and $|P| = 8$, the Pigeonhole principle tells us that there exist $a, b \in S'$ with $a \neq b$ and $f(a) = f(b)$, i.e. two of the numbers we picked are in the same pair and thus sum to 17.

The argument generalises in the following way: If we pick $n+1$ different numbers between 1 and $2n$, there are always two picked numbers whose sum is $2n+1$.

3) We roll 5 dice after another, each having the numbers 1 to 6 on its faces. The dice are distinguishable (i.e. there is a first, second, third, etc. die) and we look at the results that they show and denote tham as e.g. 21135 (the first die shows a 2, the second one a 1, etc.) Note that e.g. the outcomes 12345 and 21345 are considered different.

   There are $6^5 = 7776$ possible outcomes. Find out how many of these satisfy the following:

   a) Yahtzee: all five dice show the same number
   b) Large straight: The five dice, when ordered in a certain way, show consecutive numbers, e.g. 42153 or 23654
   c) Four of a kind: At least four dice show the same number
   d) Full House: Three dice show one number a and the remaining two show another number b. Note that this means that a Yahtzee is *not* a full house.
   e) Three of a kind: At least three dice show the same number
   f) **Optional exercise (no points):** Small straight: We can pick four dice from our five and arrange them in a way so that they show consecutive numbers, e.g. 42613 or 54352. Note that this means that a large straight is also a small straight.
      Caution: This one is a bit more difficult.

   Solution:
   Let [6] denote the set $\{1,2,3,4,5,6\}$.

   a) The function $f : [6] \to [6]^5,\ x \mapsto xxxxx$ is a bijection between the set [6] and the set of Yahtzees. Consequently, there are 6 Yahtzee outcomes. (namely: 11111, 22222, 33333, 44444, 55555, 66666)
   b) If we sort the outcomes of a large straight, the only possibilities are 12345 and 23456. Each of these corresponds to exactly 5! permutations of dice, so we get $2 \cdot 5! = 240$ outcomes.

c) First, let's find the number of outcomes where *exactly* four dice show the same number. That means four dice show the a number $a \in [6]$ and the remaining die shows a number $b \in [6] \setminus \{a\}$. Given a and b and the index of that one non-conforming die (between 1 and 5), we can construct an *exactly four of a kind* outcome, and of course we can also compute a, b, and the index of the non-conforming die given an "exactly four of a kind" outcome, so this is a bijection between the set of "exactly four of a kind" outcomes and the set $\{(a, b, i) \mid a \in [6], b \in [6] \setminus \{b\}, i \in [5]\}$. Since we have 6 choices for a, 5 choices for b, and 5 for i, we get $6 \cdot 5 \cdot 5 = 150$ outcomes.

   Since we want to know that number of *at least four of a kind* outcomes, we still need to add the outcomes where all five dice show the same number (i.e. the Yahtzees). We showed above that there are 6 of them, so the number of four-of-a-kinds is 156.

d) A full house consists of a dice pair showing the same number and a dice triple showing another number. Any full house is therefore uniquely determined by the number $a \in [6]$ shown by the pair, the number $b \in [6] \setminus \{a\}$ shown by the triple, and the set of indices I of the two dice that show b. There are clearly 6 possible choices for a and 5 for b. For the set I, we are picking a 2-element subset of a 5-element set, so there are $\binom{5}{2} = 10$ possibilities for I, i.e. $6 \cdot 5 \cdot 10 = 300$ possibilities altogether.

e) Again, we first find the number of outcomes where *exactly* three dice show the same number. We have $\binom{5}{3} = 10$ ways to pick those three dice and 6 numbers to choose from for them. For each of the remaining two dice, we can choose from 5 numbers. That gives us a total of $10 \cdot 6 \cdot 5 \cdot 5 = 1500$ outcomes. Together with the 156 four-of-a-kind outcomes, we then have 1656 outcomes.

f) Let us first look at small straights that are not large straights. There are two possibilities for this:

   - the dice show 5 different numbers, spanning the set $\{1, 2, 3, 4, 6\}$ or the set $\{1, 3, 4, 5, 6\}$
   - the dice show 4 different numbers, spanning one of the sets $\{1, 2, 3, 4\}$ or $\{2, 3, 4, 5\}$ or $\{3, 4, 5, 6\}$, with two dice showing the same number

   The first two cases are easy to count: there are five numbers to distribute over the 5 dice, giving us 5! possibilities each time. For each of the remaining three cases, we have $\binom{5}{2} = 10$ possibilities to pick the pair of dice showing the same number, 4 possibilities to pick the number that they show, and 3! possibilities to distribute the remaining 3 numbers onto the remaining 3 dice.

   Thus, in total, we have $2 \cdot 5! + 3 \cdot 10 \cdot 4 \cdot 3! = 960$ "proper small straight" outcomes. Together with the 240 large straight outcomes we computed earlier, there are 1200 small straight outcomes.

4*) Consider Kruskal's algorithm from the lecture.

a) Propose a modification of the algorithm to compute a spanning forest with a maximum possible weight instead of the minimal. Prove its correctness.

b) Prove the following. If the weight function is injective, then the minimal spanning forest is unique (that is, all other spanning forests are not minimal).

c) Is the following true? If the minimal spanning forest of a weighted multigraph is unique, then the weight function must be injective.

d) Consider the fully connected undirected graph G_5 on 5 nodes where every node is connected to every other node except itself. How many edges does G_5 have? How many spanning trees of G_5 are there?

Solution:

a) Observe that Kruskal's algorithm does not need the weights to be positive. Hence, given the weight function w, we can define a new weight function $w'(e) = -w(e)$ and run the standard Kruskal's algorithm. The output is the minimal spanning forest with respect to w' which is the maximal spanning forest with respect to w.

b) If the weight function is injective, then the order of edges considered by Kruskal's algorithm is uniquely determined. The proof of the claim follows.

Let $G = (V, E)$ is a graph with n nodes and weights w. Let $F = \{e_1, e_2, \dots, e_m\}$ be the output of Kruskal's algorithm executed on G. For the contradiction, let $F' = \{e'_1, e'_2, \dots, e'_m\}$ be a different minimal spanning forest. Let us assume the edges are sorted by their increasing weights. Hence the edges of F were selected in this order by the run of Kruskal's algorithm. Since the weight function w is injective and because $w(F) = w(F')$, there must be an index k such that $w(e_k) > w(e'_k)$. Let us consider the minimal such index k. Clearly $k > 1$, because e_1 has the least weight. Let us consider the following sets of edges.

$$L = \{e_1, e_2, \dots, e_{k-1}\}$$
$$L' = \{e'_1, e'_2, \dots, e'_{k-1}, e'_k\}$$

Since the edges are sorted by weights, we have $w(e') < w(e_k)$ for all $e' \in L'$.

Now, L was the intermediate result of the run of Kruskal's algorithm on G after selecting edges from L, and before selecting the edge e_k. Let us consider the partitioning into connected components (V_1, E_1),...,(V_s, E_s) at this point. This is a forest with s trees and thus it has exactly $\sum_{i=1}^{s}(|V_i| - 1) = n - s$ edges. These edges are L and hence $|L| = n - s$. Now we observe that there must be an edge $e' \in L'$ which connects two different trees of this forest. This is because L' does not contain a cycle. If we try to place the edges of L' into the components V_1,...,V_s, we can place at most $|V_i| - 1$ edges in the component V_i, otherwise we create a cycle. In this way can place at most $\sum_{i=1}^{s}(|V_i| - 1) = n - s$ edges but there are $|L'| = |L| + 1 = n - s + 1$ edges in L' and hence some edge $e' \in L'$ must connect two different components. Clearly $e' \notin L$ and we know $w(e') < w(e_k)$ from above. But then e' should have been selected by Kruskal's algorithm after selecting the edges from L, and not e_k. Hence contradiction.

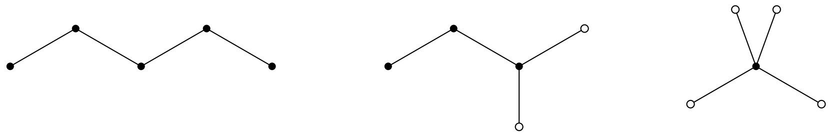
c) No. For an example, consider a multigraph on 2 nodes $\{a, b\}$ with three (multi-)edges between a and b with weights $1, 2, 2$.

In fact, *any* forest clearly has a unique spanning forest, so any forest where two edges have the same weight is a counterexample.

d) The number of edges is simply the number of 2-element subsets of $\{1, 2, 3, 4, 5\}$, which is simply $\binom{5}{2} = 10$. This generalises to $\binom{n}{2} = \frac{1}{2}n(n-1)$ for n nodes.

For the spanning trees, we can first think about what general shapes of trees with 5 elements are possible. There are only these three:[1]

[1] For those of you who took organic chemistry in high school, these are the three different structural isomers of pentane: Pentane, Isopentane, and Neopentane



We must now count how many ways we have to distributes the labels $\{1, 2, 3, 4, 5\}$ onto these.

- For the first type of tree (where the nodes are arranged in a single line), we can distribute our 5 nodes onto the 5 positions in that line in any way we like. However, clearly e.g. 12345 and 54321 result in the same tree, so we end up counting each tree twice when we do this. Since there are 5! ways to distribute the nodes onto the 5 positions and count every tree twice that way, we have $\frac{5!}{2} = 60$ such trees.
- For the second type of tree, we can pick the nodes marked by a black dot freely, and the remaining two (marked by a black dot) are then already fixed. Thus, we have $5 \cdot 4 \cdot 3 = 60$ such trees.
- For the last type of tree, we can pick the centre node freely and the rest are then already fixed. Such we have 5 such trees.

All in all, this is 125 trees. In general, there are n^{n-2} labelled trees with n nodes and thus also n^{n-2} spanning trees of a complete graph with n nodes. This is known as *Cayleys's formula* (which we will not prove here).