0.1 Graphs and their corresponding Matroid Theory

In this section we introduce the graph theoretic deifnitions and results needed to understand the algorithms and subsequent optimisation in later sections. These definitions and results come mainly from Jungnickel's[**?**] text. We will also see how the subsequent properties of graphs naturally align with the properties of matroids and in particular to bases of matroids which are the cornerstone of the results in the later sections.

Definition 0.1 (Connected)**.** A graph is connected when there is a path between each pair of vertices.

Definition 0.2 (Acyclic)**.** An acyclic graph is a graph which contains no closed walks.

Definition 0.3 (Walk)**.** If there are vertices $v_{i-1}v_i$ for $i = 1, ..., n$ the sequence is called a walk. If $v_0 = v_n$ this sequence is called a closed walk.

Definition 0.4 (Tree)**.** A connected graph containing no circuits. In other words, acyclic graphs.
A forest is a disconnected graph containing no circuits

We know from *theorem 1.7.* If we let E be the edge set of our graph G. And we define $\mathscr{I}$ as the subsets of E not containing all of the edges of any cycle in G we have $M(G)$ the cycle matroid of G. We know that this is a matroid, and it is easy to see that in this case the elements of $\mathscr{I}$ correspond exactly to the trees of G.

Definition 0.5 (Spanning Tree)**.** A spanning tree T of an undirected graph G is a subgraph that is a *tree* which includes all of the vertices of G.
A disconnected graph cannot contain a spanning tree as we cannot find a walk which brings us to all of the disconnected vertices.

Proposition 0.6. *A theorem of Cayley(1889) states that the number of distinct laballed trees which can be drawn using* n *labelled points is* n^{n-2}.

Corollary. *The number of distinct labelled spanning trees which can be drawn using* n *labelled points is* n^2.

Remark. Determining the number of spanning trees of a graph in polynomial time is NP-hard.

Lemma 0.7. *[**?**] Any acyclic graph on* n *vertices has at most* $n - 1$ *edges.*

Proof: Let G be an acyclic graph with n vertices.
if n = 1 then we have no edges hence nothing to prove.
Assumse $n > 1$, let e be an edge in G connecting two vertices ab in the vertex set of G.
Let $H = G\backslash\{e\}, H$ has one more connected component than G. H has two maximal acyclic connected componenets, and thus can be decomposedinto acyclic

connected graphs $H_1, H_2, ..., H_k where k \geq 2$.
By induction, we can assume each graph H_i contains at most $n_i - 1$ edges where n_i is the number of vertices of H_i.
Then G has at most $n - 1$ edges.
$(n_1 - 1) + ... + (n_k - 1) + 1 = (n_1 + ... + n_k) - (k - 1) \leq n - 1$ edges.

$\square$

Definition 0.8 (Bridge)**.** A bridge(or cut edge) is an edge of a graph whose deletion increases the numbe rof connected components. Equivalently, an edge is a bridge is and only if it is not contained in any cycle.

The above definition is mainly useful for us in proving the next lemma, which comes from Jungnickel's text and will help us visualise the characteristics of a spanning tree of a graph. In particular condition (1) and (2), will allow us to see that the spanning trees of a graph directly correspond to the bases of a matroid.

Lemma 0.9. *[**?**] Let G be a graph. Then the following conditions are equivalent:*
1) G is a tree.
2) G does not contain any cycles, but adding any further edge yields a cycle.
3) Any two vertices of G are connected by a unique path.
4) G is connected, and any edge of G is a bridge.

Proof: (1) $\implies$ (2)
Suppose that G is a tree,then G is a connected graph with no circuits. Let e be a new edge in G with $e = g_i g_k$ where g_i, g_k are in the vertex set of G. Then as $G \cup \{e\}$ must be connected, there exists a walk between any pair of vertices of G. So there is a walk K from $g_j \to g_k$ and there is also a walk L from $g_k \to g_j$ where K does not traverse e and L does traverse e and so we have a cycle.

Proof: (2) $\implies$ (3)
Let u, v be vertices of G. If there was not path joining uv in G then $e = uv$ does not create a cycle in G. Thus G must be connected.
Suppose G contained two different paths W_1, W_2 from u to v.
Then $u \longrightarrow v \longrightarrow u$ would be a closed walk in G.
$\implies$ G contains a cycle. Which is a contradiction.

Proof: (3) $\implies$ (4)
G is connected by hypothesis. Let $e = uv$ be an edge in G.
Suppose e is not a bridge, then $G \setminus \{e\}$ is still connected. But then we have two distinct path from u to v in G.

Proof: (4) $\implies$ (1)
G is connected by hypotheis.
Suppose G contains a cycle K. Then any edge of K could be ommited from G, and the resulting graph would still be connected. In other words, no edge of K would be a bridge, a contradiction.

$\square$

The vertices of a graph/network can be labelled and referred to as nodes. Information may then be recorded in them along with a cost, penalty or probability associated with each edge. For example, the problem of joining all nodes in a graph by the minimum length using our respective metric to a tree known as a *minimum spanning tree*.
Weights can be assigned using a process as detailed below. We will later redefine this process in a more suitable way to take advantage of matroid properties in order to determine these minimum spanning trees in graphs.

Definition 0.10. let (G, ω) be a network. For any subset T of the edge set of G, ω is called the weight of T.

$$\omega(T) = \sum_{e \in T} \omega(T) \tag{1}$$

Definition 0.11 (Minimal Spanning Tree)**.** A spanning tree is a *minimal* spanning tree if its weight is minimal of all the weights of spanning trees. A forest can be considered by finding a minimal spanning tree for each connected component of G.

Remark. If the weight ω is constant, any spanning tree is minimal.
In this case, determining a minimal spanning tree could be done using a breadth-first search.