Data Structures and Algorithms Notes

paraphrased by Tyler Wright

An important note, these notes are absolutely ***NOT*** *guaranteed to be correct, representative of the course, or rigorous. Any result of this is not the author's fault.*

1 Graph Theory

1.1 Definition of a Graph

A graph is a pair of sets $G = (V, E)$, where V is a set of vertices (or nodes) and E is a set of edges (or arcs).

1.2 Definition of an Edge

An edge of a graph $G = (V, E)$ is $e = \{u, v\}$ in E where u, v are vertices in V.

1.3 Definition of a Neighbourhood

For a graph $G = (V, E)$ with v in V, the neighbourhood of v is the set $V' \subseteq V$ of vertices connected to v by an edge in E.

The neighbourhood of v is denoted by $N(v)$.

The neighbourhood of a set of vertices is the union of the neighbourhoods of each vertex.

1.4 Definition of Degree

For a graph $G = (V, E)$ with v in V, the degree of v is the size of its neighbourhood.

The degree of v is denoted by $d(v)$.

1.5 The Handshake Lemma

For a graph $G = (V, E)$, we have that:

$$|E| = \frac{\sum_{v \in V} d(v)}{2}.$$

This is because each edge visits two vertices, so by counting the degree of each vertex we count each edge exactly twice.

1.6 k-regular Graphs

For a graph $G = (V, E)$, we have that G is k-regular for some k in $\mathbb{Z}_{>0}$ if for all v in V, we have:

$$d(v) = k.$$

We cannot have a k-regular graph where k is odd and $|V|$ is odd by the Handshake Lemma.

1.7 Isomorphic Graphs

Graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are called isomorphic if there exists a bijection $f : V_1 \to V_2$ such that:

$$\{u, v\} \in E_1 \iff \{f(u), f(v)\} \in E_2.$$

This relationship is denoted by $G_1 \cong G_2$.

1.8 Definition of a Subgraph

A graph $G' = (V', E')$ is a subgraph of $G = (V, E)$ if $V' \subseteq V$ and $E' \subseteq E$.

1.9 Definition of an Induced Subgraph

An induced subgraph generated from $G = (V, E)$ by $V' \subseteq V$ is the graph $G' = (V', E')$ where:

$$E' = \{\{u, v\} \in E \text{ such that } u, v \in V'\}.$$

Essentially, you generate an induced subgraph from a subset of the vertices of a graph by selecting edges that join vertices in the subset.

1.10 Walks

1.10.1 Definition of a walk

A walk in a graph $G = (V, E)$ is a set of vertices in V connected by edges in E. The length of the walk is the number of edges traversed in the walk.

1.10.2 Definition of a path

A path is a walk where no vertices are repeated.

1.10.3 Definition of an Euler walk

An Euler walk is a walk such that every edge is traversed exactly once. Thus, for a graph $G = (V, E)$, the length is $|E|$.

1.10.4 Conditions for an Euler walk

For an Euler walk to be possible on a given graph, all vertices must have an even degree **or** exactly two vertices have odd degree.

If all vertices have even degree we have that the Euler walk is a cycle, if exactly two vertices have odd degree then we have that these vertices are the start and end points of our Euler walk.

1.11 Definition of a Connected Graph

A connected graph is a graph where for each pair of vertices, there is a path connecting them.

1.12 Definition of a Component

A component of a graph $G = (V, E)$ is a maximal connected induced subgraph of G. This means an induced subgraph of G that is connected but is not longer connected if a vertex is removed.

Connected graphs have a single component, the entire graph.

1.13 Digraphs

1.13.1 Definition of a digraph

A digraph (or directed graph) is a graph where each of the edges has a direction. This direction means the edge can only be traversed in a single direction.

1.13.2 The Directed Handshake Lemma

For a digraph $G = (V, E)$, we have that:

$$\sum_{v \in V} d^-(v) = \sum_{v \in V} d^+(v) = |E|.$$

This is because if we consider the 'tail' of an edge (the vertex it leaves), each edge has exactly one tail.

1.13.3 Definition of a strongly connected digraph

A digraph $G = (V, E)$ is strongly connected if for each u, v in E, there exists a path from u to v **and** from v to u.

1.13.4 Definition of a weakly connected digraph

A digraph $G = (V, E)$ is weakly connected if for each u, v in E, there exists a path from u to v **or** from v to u.

1.13.5 Definition of components of digraphs

A strong component of a digraph is the maximal *strongly* connected induced subgraph.

A weak component of a digraph is the maximal *weakly* connected induced subgraph.

So, these are induced subgraphs that are strongly/weakly connected but are no longer strongly/weakly connected once a vertex is removed.

1.13.6 Definition of neighbourhoods in digraphs

The neighbourhood of a vertex in a digraph can be considered by looking at the edges *from* the vertex and the edges *to* the vertex.

The in-neighbourhood of a vertex v are the edges that enter v. The out-neighbourhood of a vertex v are the edges that exit v. These are denoted by $N^{-}(v)$ and $N^{+}(v)$ respectively.

1.13.7 Definition of degrees in digraphs

For a vertex v, the in-degree of the vertex $d^{-}(v)$ is the size of the in-neighbourhood and the out-degree of the vertex $d^{+}(v)$ is the size of the out-neighbourhood.

It can be seen that the degree of a given vertex is the sum of its in and out degree (in a digraph).

1.13.8 Conditions for an Euler walk in a digraph

For an Euler walk to be possible on a given digraph, we have two cases, either:

- the digraph is strongly connected and every vertex has equal in and out degrees, or

- one vertex has an in-degree one greater than its out-degree, another has an out-degree one greater than its in-degree, and all remaining vertices have equal in and out degrees.

In the first case we have that the Euler walk is a cycle, in the second we have that the special vertices are the start and end points of our Euler walk.

1.13.9 Cycles

1.13.10 Definition of a cycle

A cycle is a walk where the first and last vertices are the same and each vertex appears at most once (barring the first and last vertex).

1.13.11 Definition of a Hamiltonian cycle

A Hamiltonian cycle is a cycle where each vertex is visited.

1.13.12 Conditions for a Hamiltonian cycle

Whilst the conditions necessary for a Hamiltonian cycle in general are unknown, by Dirac's theorem, we know that for a graph with n vertices, if every vertex has degree $\frac{n}{2}$ or greater then a Hamiltonian cycle exists.

1.14 Trees

1.14.1 Definition of a forest

A forest is a graph with no cycles.

1.14.2 Definition of a tree

A tree is a connected forest (or a connected graph with no cycles).

1.14.3 Path uniqueness of trees

For a tree $T = (V, E)$, we have that for any u, v in V, there exists a unique path from u to v.

To prove this, suppose there are two unique paths between u *and* v*. These paths must diverge and if we connect them, they form a cycle which contradicts the definition of a tree.*

1.14.4 The magnitude of edges in trees

For a tree $T = (V, E)$, we have that $|E| = |V| - 1$.

1.14.5 Rooted trees

For a tree $T = (V, E)$, we can root T with some r in V. For v in $V \setminus r$, we define P_v to be the path from r to v, we then direct the edges from r to v for each P_v.

For u, v in $V \setminus \{r\}$, we say that:

- u is an **ancestor** of v if u lies on P_v
- u is the **parent** of v if u is in the in-neighbourhood of v
- v is a **leaf** if it has degree 1
- $L_0 = \{r\}$ and $L_n = \{v : |P_v| = n\}$ are the **levels** of T
- The **depth** of a tree is the greatest n where L_n is non-empty.

1.14.6 Lower bound on the amount of leaves in a tree

For a tree with $T = (V, E)$, if $V > 1$, there must be at least 2 leaves.

1.14.7 Equivalent statements to the tree definition

For a graph $T = (V, E)$, we have that the following are equivalent:

- T is a tree
- T is connected and has no cycles
- $|E| = n - 1$ and T is connected
- $|E| = n - 1$ and T has no cycles
- T has a unique path between any two vertices

1.15 Bipartitions

1.15.1 Definition of a bipartite graph

For $G = (V, E)$, we have that G is bipartite if there exists $A \subset V$, $B \subset V$ such that A and B are disjoint and the induced subgraphs of A and B have no edges. A and B are bipartitions of G.

Saying G is bipartite is equivalent to saying G has no cycles of odd length.

1.15.2 Definition of a matching

A matching in a graph is a set of disjoint edges.

A matching is **perfect** if each vertex is contained in some matching edge.

1.15.3 Definition of a semi-matching

For k in $\mathbb{Z}_{>0}$, a k to 1 semi-matching in a bipartite graph G with a bipartition $\{A, B\}$ is a subgraph of G where each vertex in A has degree at most k and each vertex in B has degree at most 1.

1.15.4 Definition of an augmenting path

Given a matching M in a bipartite graph $G = (V, E)$, an augmenting path is a set of vertices in V connected by edges e_i in E such that:

$$e_i \text{ is } \begin{cases} \text{in } M & \text{for } i \text{ odd} \\ \text{not in } M & \text{for } i \text{ even.} \end{cases}$$

With the condition that the first and last vertices in the path are not in the matching.

1.15.5 Hall's Theorem

For a bipartite graph $G = (V, E)$ with the bipartition (A, B) has a perfect matching if and only if $|A| = |B|$ and for all $X \subseteq A$, $|N(X)| \geq |X|$.

2 Types of Algorithms

2.1 Greedy Algorithms

These types of algorithms start with a trivial solution and iteratively optimise their solution based on the information available at the time. They do not retroactively change the solution based on new data, only add to it.