# A Directed Acyclic Graph Map in Java

# CITS2200 (Data Structures and Algorithms)

## By

## Caleb Fetzer (21384976) Reece Notargiacomo (21108155)

# *Abstract*

We implemented a Directed Acyclic Graph Map, which invovles the functionality of a directed graph with no cycles, and the Key-Value pairing of a map. The implementation was used with an Adjacency List data structure as an excercise for recursive data structures, and Java API's Util LinkedHashSet functions for the storing of Keys and Values. The time complexity of the standard lookups are reflected by the general time complexities (in Big O) of the respective implementation – adjacency list.

We found the theoretical growth rates of method functionality matched practicality, using JUnit testing as a basis for a time vs operations comparison.The main methods, getLongestPath and getMax for finding, the longest path of the DAG and the Max-flow of the DAG respectively, were not near the theoretical rates. This is due to the lack of test data in the testing. Due to inconclusive test data, a very small graph of of 6-10 vertices, would calculate the Longest Path and Max Width in 0.007s – 0.010s.

# Implementation choice

We decided to implement an Adjacency List for the DAGMap, as opposed to an Adjacency Matrix. The Adjacency Lists also have lists for predecessor and successor nodes. To store the key value pairs, a set of keys is created.

A rootNode value is also created, which keeps track of the first parent node (a node with children).

There were points identified, to justify the decision, such as the pro’s and con’s of each data type. The time complexity for the standard operations these data types implemented, were also addressed and compared.

­Although the Adjacency List on the surface appears to be a less appealable data type, just via inspection of the general type of methods; it uses less space depending on the size of the graph. An adjacency matrix creates space for a graph with vertices that may not exist in the graph, thus occupying more space. On very sparse graphs, an adjacency matrix uses a lot of space that is not needed, and thus was our main argument for an adjacency list over an adjacency matrix.

From a non-scientific point of view, anecdotal evidence from Stack Overflow, would suggest that in most usual cases an adjacency list is preferable to an adjacency matrix implementation[[1]](#footnote-1)

The learning curve, for us, was trying to implement a recursive data type – to gain a better understanding of recursion and it’s role in data structures and algorithms.

|  |  |  |
| --- | --- | --- |
| **Operation** | **Adj Matrix Time Complexity** | **Adj List Time Complexity** |
| Adding/Deleting edge | O(1) | O(log (n)) |
| Does edge exist between vi & vj? | O(1) | O(log (n)) |
| Successors of vertex | O(n) | O(k) |
| Finding path between two vertices | O(n2) | O(n + m) , m <= n |

(2014, Algorithm of the Week). ‘*k’ refers to the length of lists containing the successors of a vertex, m refers to the number of edges.*

# The method complexities and implementation details

To note, methods that run in constant time, such as isEmpty, won’t have their time complexity detailed. The terms key, node and vertice are used interchangeably, and all mean the same thing. Value is also referred to as ‘label’.

**public void put (Key newKey, Value newValue)**

* The method calls the methods containsKey(from DAGMap), add(from java.util.Set). This method accepts two parameters, newKey and newValue. It checks to make sure newKey does not already exist in the set of keys in the DAGMap, and if not adds it to the set.

**public object get (Key k)**

* This method checks for the value of a given key. It makes sure that the key exists, else an illegal argument exception is thrown. This runs in O(n) time, as it peforms a single look-up operation, using the contains() method from the java.util.set.

**public void remove (Key k)**

* This method removes the given key from the keys in the DAGMap, as well as removing it from the list of predecessors and list of successors. Therefore the associated dependencies/requirements of this key are removed. It runs in o(n^2) time as it performs a lookup of all the keys in the predecessors that listed ‘k’ as a dependent, all the keys in the successors that listed ‘k’ as a requirement, and then removes ‘k’ correspondingly.

**public set<Key> getPredecessors(Key k)**

* This method returns the set of keys of predecessors (i.e. parents) of a given Key ‘k’. A single look up is performed with a boolean check, and the list of predecessors for k are returned. If the key is not defined an IllegalArgumentException is thrown. This method runs in O(n) time.

**public set<Key> getSuccessors(Key k)**

* This method returns the set of keys of successors (i.e child(ren)) of a given Key ‘k’. A single look up is performed with a boolean check, and the list of successors for k are returned. If the key is not defined an IllegalArgumentException is thrown. This method runs in O(n) time.

**public void addDependency (Key kReq, Key kDep)**

* This method adds an edge between a requirement key (kReq) and a dependent key (kDep), making kDep dependent on kReq. An ‘isDependent’ checkup is performed, making sure that kDep is not already dependent on kReq. It then adds kDep to the successor list for, and kReq to the predecessor list. It also removes the kDep key from the list of orphan keys (Keys with no predecessors). Runs in .. time

**public void removeDependency (Key kReq, Key kDep)**

* This method removes an existing edge between a requirement key (kReq) and a dependent key (kDep), making kDep no longer dependent on kReq. An isDependent check is performed and following, kReq is removed from the set of successors and kDep is removed from the set of predecessor. If no edge exists between kReq and kDep, an IllegalArgumentException is thrown.

**public Boolean containsKey (Key k)**

* This method checks the set of keys in the DAGMap for the given key k, if k exists, return true. It checks for the set of keys not equal to null, and then iterates through each key in the set of keys in the DAGMap. If the key is found, true is returned.

**public Boolean containsValue (Value value)**

* This method checks all of the keys in the DAGMap for a given value, for any key that the value is paired with. If the value doesn’t exist in the DAGMap, false is returned. As a standard lookup function, this runs in O(n) time, as it is simply performing an iteration.

**public Boolean isDependent (Key haystackKey, Key needleKey)**

* This refers to the finding a needle in a haystack metaphor. The method checks to make sure haystackKey and needleKey are not null, and if haystackKey equals needleKey it returns true, as this means that the haystackKey is dependent on itself. If not the case, haystackKey is checked against the list of successors, by performing a getSuccessors check, and then recursively checking for children keys that are dependent on the haystackKey through the needleKey. This way it is checking not just immediate dependencies, but rather all dependencies. If haystackKey or needleKey do not exist, an IllegalArgumentException is thrown.

**public Set<K> getKeySet()**

* This method performs a simple lookup and returns all of the keys in the DAGMap as a set. Because it is retrieving keys from a given set, it performs in O(n) time.

**public Object clone()**

* This method creates a shallow clone of the DAGMap. It copies all the dependencies and values. The time complexity of the clone method is O(n + m), where m is the number of edges in the DAG.

**public Object equals(Object o)**

* This method compares two DAGMaps for equality. Each node, value, predecessor and successor is checked. The time complexity comes to O(4\*n) which is just O(n) complexity

**public int getLongestPath()**

* This method traverses the entire DAGMap, looking for the longest possible path between any node. As an assumption, it can be assumed that a starting node will be an orphan, e.g. parentless node. Our implementation of getLongestPath involves a recursive implementation, with the use of a private method getPathToTop, running in O(n2) time.

**public int getMaxFlow( K source, K sink )**

* This method gets the maximum flow between two vertices, a source vertex and a sink vertex, where the max flow is defined as the maximum number of paths that do not share a vertex. This algorithm was implemented using a recursive DFS (Depth-First-Search) DFS-Visit based implementation. A DFS is performed on the DAGMap from a beginning source, and the DAG is traversed, marking nodes that are visited. This algorithm runs in O(2^n) time, as it involves a recursive algorithm lookup, at least once.

**public int getWidth()**

* The main focus of the getWidth algorithm was a DFS, which involved traversing the DAG from set of source nodes (orphans), and performing a depth-first-search recursive algorithm. The DFS would check for visited nodes, ignore them, until an unvisited child node was found, and the DFS would continue. Disconnected nodes gave the count of minimum paths an immediate addition of 1.

# Testing

Testing the DAGMap involved JUnit unit tests and a main method setup, where basic operations such as put, addDependency, clone, remove, equals.

The main setup involved creating various test data for put, addDependency, removeDependency, etc. Testing these functions was relatively easy.

The results were successful for all of those tests, meaning that the operations worked successfully.

Using JUnit, we were able to test getLongestPath and getMaxFlow, however due to a small test environment range (not using a large disparity of DAG size (eg. 1-10000)), the test data was not conclusive. The average time in seconds for a DAG of 6 vertices was recorded to be 0.006 – 0.007s. Although implemented, getWidth has not been tested extensively, although general tests suggest that the method returns the accurate width of the graph.

# Bibliography

Popov, Stoimen, Accessed on May 20th 2014, *Algorithm of the Week: Graphs and their Representation*, <http://java.dzone.com/articles/algorithm-week-graphs-and?utm_source=twitterfeed&utm_medium=twitter&utm_campaign=Feed%3A+javalobby%2Ffrontpage+%28Javalobby+%2F+Java+Zone%29>

Stack Overflow, Accessed on May 21st 2014, *What is better, adjacency lists or adjacency matrices for graph problems in c++?,*

http://stackoverflow.com/questions/2218322/what-is-better-adjacency-lists-or-adjacency-matrices-for-graph-problems-in-c

1. See Bibliography: What is better, adjacency lists or adjacency matrices.. [↑](#footnote-ref-1)