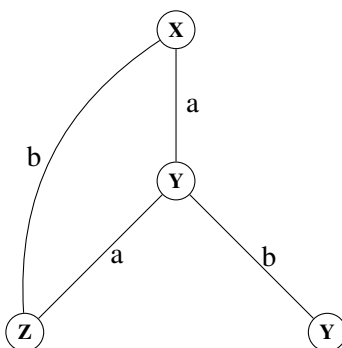


Knowledge Discovery in Databases II
 WS 2015/2016

Übungsblatt 13: Graph and Link Mining

Aufgabe 13-1 gSpan

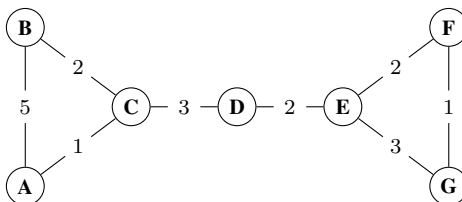
Given the following graph G :



- (a) List all DFS trees for G that start at node X .
- (b) Encode them as DFS codes and order them lexicographically.
- (c) Which code is used to represent the original graph? Why is the chosen DFS code unique? Which graph properties are necessary to apply this technique?
- (d) List all possible right-most extensions of the DFS corresponding to the selected graph.

Aufgabe 13-2 Betweenness Centrality

Calculate Betweenness Centrality of all nodes in the following weighted graph:



Aufgabe 13-3 Pagerank in Spark

In this exercise, the task is to implement pagerank for a graph given by a list of edges (e.g. (303,104) means that node 303 is connected to 104). After importing the graph from the given text file (fb0.edges) into a spark rdd and use the matrix multiplication from exercise 6-1 to compute the pagerank using the power iteration method.