

## 2003 Paper 12 Question 1

### Data Structures and Algorithms

- (a) A million singleton sets each containing a distinct integer are to be successively combined by calls of `union( $S_1, S_2$ )`. The result represents the union of the two disjoint sets represented by  $S_1$  and  $S_2$ . Interspersed among these calls are several calls of `inSameSet` where `inSameSet( $m, n$ )` yields `true` if and only if  $m$  and  $n$  are integers now in the same set. Describe in detail how you would implement `union` and `inSameSet` assuming they will be called about one million and five million times, respectively. Explain why your solution is efficient. [10 marks]
- (b) Describe in detail an implementation of Kruskal's algorithm for finding a minimum cost spanning tree of an undirected graph with positive integer costs on the edges that uses your version of `union` and `inSameSet`. [5 marks]
- (c) Explain why the spanning tree is unique if all the edge costs are distinct. [5 marks]