

## 2004 Paper 8 Question 9

### Database Theory

- (a) Define formally the semi-structured data (SSD) model. [5 marks]
- (b) Show how SSD expressions can be expressed in XML. [2 marks]
- (c) What are the main differences between the SSD and XML models? [2 marks]

When viewed graphically, simple SSD expressions denote trees. Consider a variant, d-SSD, where the edges emanating from any node in the tree must have a *unique* label, but where the labels may themselves be d-SSD expressions. (You may disregard object identities (oids); hence d-SSD expressions always denote trees.)

- (d) Define the syntax of d-SSD expressions. [3 marks]
- (e) Give a d-SSD expression to represent the following:
  - (i) the array ["Do", "Re", "Mi"];
  - (ii) the set {11, 52, 44};
  - (iii) the bag {{10, 10, 13, 42, 13, 10}}. [1 mark each]
- (f) Define the syntax for path expressions in the d-SSD model. [2 marks]
- (g) Hence describe precisely how the d-SSD model can be extended to represent graphs. (Answers that use oids will receive little credit.) [3 marks]