

1. (20 points) Of the following statements, identify all that hold about XSLT
- A. XSLT applies templates recursively by default, but only going downward in the XML document tree
 - B. In XSLT, templates can be applied to parent, children, or other descendants of a node, but not to its grandparent
 - C. XSLT allows only a proper subset of the XPath language to be used
 - D. In XSLT, the copy-of element expresses a command that provides a deep copy of an element on which it is applied
 - E. In XSLT, by default the attributes are ignored during recursive calls to apply templates
 - F. In XSLT, templates can be applied on attributes if you specify them appropriately
 - G. The empty XSLT transformation yields a (practically identical ignoring variations allowed by the XML Infoset) copy of a document on which it is applied
 - H. It is possible to construct a nonempty XML document that yields a copy (practically identical ignoring variations allowed by the XML Infoset) when transformed by the empty XSLT transformation
 - I. In XSLT, as in XPath, an attribute is not a child of its parent
 - J. In XSLT, as in XPath, a text node is not a child of its parent

Solution: A, D, E, F, I

2. (32 points) Of the following statements, identify all that hold about XML keys, integrity constraints, and other aspects of relating XML to databases:
- A. Referential integrity as captured by keyref means that an element that is referred to must exist
 - B. When XML structures are mapped to relational schemas, XML key constraints are lost
 - C. General XQuery expressions can be used to specify integrity constraints on XML documents
 - D. A key allows more than one selector subelement in order to construct a composite key
 - E. Contrary to its name, the data-centric view has little to do with relational database management
 - F. The document-centric view gains prominence in business settings where audit trails are required
 - G. Native XML databases support the storage of XML documents primarily by shredding
 - H. XML has yet to become popular for messaging because messaging middleware does not support complex structures
 - I. It is possible to write a key selector referring to a grandchild element of the given element
 - J. If the document schema states that a specific grandchild element of the given element may contain only text, it is possible to write a field expression referring to the text occurring within that grandchild element
 - K. It is possible to write a field expression referring to the first grandchild element of the given element
 - L. ../child is allowed in a selector provided child is a particular element that exists in the appropriate context
 - M. Any mapping of XML to relational schemas requires selecting one or more tuple-generating elements
 - N. In mapping XML to tables, sometimes you would create a table with exactly one column
 - O. In mapping XML to tables, sometimes you would be forced to introduce additional columns
 - P. The general representation of an XML document in a relational schema can lead to multiple self-joins even for simple queries on the XML document

Solution: A, C, F, I, J, N, O, P

B is false because key constraints can be expressed in relational DBMSs

E is false because the regular structures of the data-centric view are motivated by relational structures. Further they map more easily to relational schemas than documents that exemplify the document-centric view

G is false because shredding an XML document is what you need to do to map it to a relational DBMS

J is true because the text is extracted by simply specifying the element that contains the text

K is false because the subset of XPath allowed in this setting doesn't include the [] selector construct

L is false because the subset of XPath allowed in this setting doesn't include the .. construct; it doesn't matter if child is an element that exists in the given context

M is false because you can have general representations such as for graphs that don't use a TGE from the document

N is true because sometimes your XML document's TGE may only have one attribute or no attributes and some text

O is true because you may need additional columns, e.g., to make sure each relational tuple is unique. For example, consider an XML document whose TGE has three copies for some attribute values and seven copies for another attribute value. In the relational representation you can't have duplicates (even if you step outside the pure model and allow duplicates, you can't control how many duplicates would be stored)

P is true because multiple self-joins is how you would extract structure information from a general representation: for example, to solve a query for the descendants of an element