

## A.4 Addressing and Querying XML Documents

In relational databases, parts of a database can be selected and retrieved using query languages such as SQL. The same is true for XML documents, for which there exist a number of proposals for query languages, such as XQL, XML-QL, and XQuery.

The central concept of XML query languages is a *path expression* that specifies how a node, or a set of nodes, in the tree representation of the XML document can be reached. We introduce path expressions in the form of XPath because they can be used for purposes other than querying—namely, for transforming XML documents.

XPath is a language for addressing parts of an XML document. It operates on the tree data model of XML and has a non-XML syntax. The key concepts are path expressions. They can be

- absolute (starting at the root of the tree); syntactically they begin with the symbol `/`, which refers to the root of the document, situated one level above the root element of the document; or
- relative to a context node.

Consider the following XML document:

```
<?xml version="1.0" encoding="UTF-16"?>
<!DOCTYPE library PUBLIC "library.dtd">
<library location="Bremen">
  <author name="Henry Wise">
    <book title="Artificial Intelligence"/>
    <book title="Modern Web Services"/>
    <book title="Theory of Computation"/>
  </author>
  <author name="William Smart">
    <book title="Artificial Intelligence"/>
```