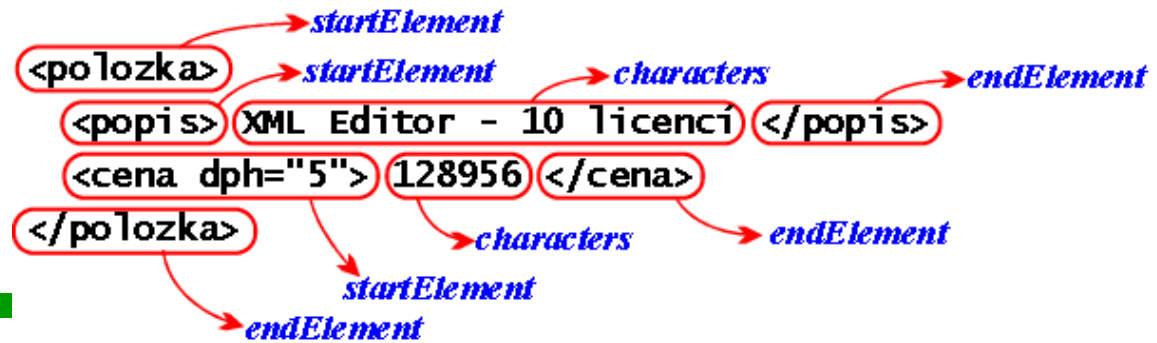

XML Technologies

DOM, SAX

SAX (1/2) – Overview



- Event based model
 - An event:
 - Begin/end of the document
 - Begin/end of the element
 - String data
 - ...
 - The parser raises a sequence of events based on the document order
 - It is not possible to jump to an arbitrary place or go back
-

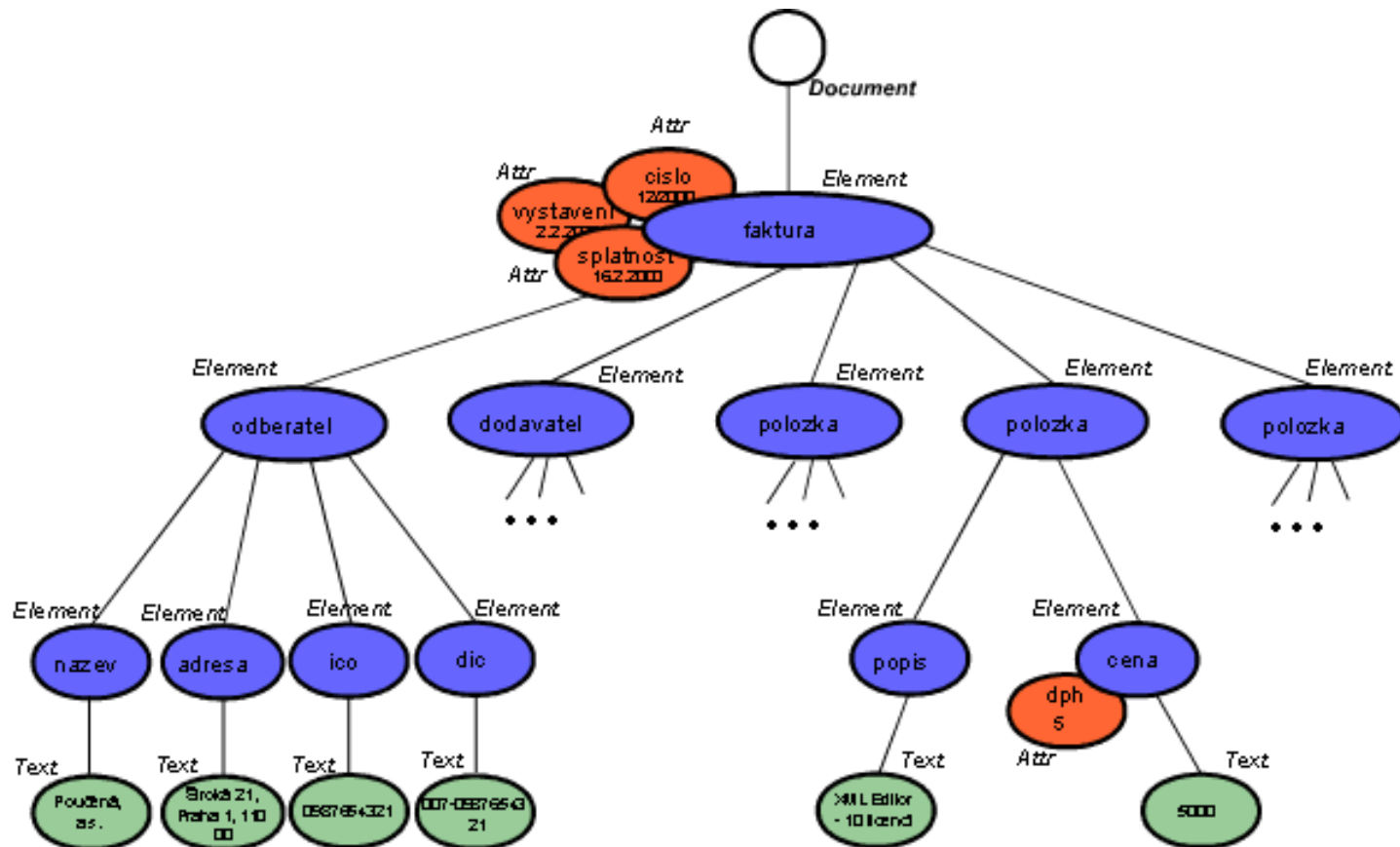
SAX (2/2) – Overview

- Begin/end of the document
 - **startDocument()**
 - **endDocument()**
 - Begin/end of an element
 - **startElement(String uri, String localName, String qName, Attributes atts)**
 - **endElement(String uri, String localName, String qName)**
 - String data
 - **characters(char[] ch, int start, int length)**
 - Array **Attributes atts**
 - **atts.getQName(i); atts.getValue(i);**
 - Returns attribute's name and value
-

DOM (1/4) – Overview

- Object model
 - Every logical part of the XML document is represented as an object
 - Element, attribute, text, ...
 - The objects form a tree
 - The whole XML document is (typically) loaded into memory
-

DOM Example



DOM (2/4) – Interface Node

- Represents a common node in XML document
 - It is further specialized into interfaces representing particular node types
 - Document, element, attribute, ...
 - **getAttributes()**
 - **getChildNodes()**
 - **getFirstChild()**
 - **getNodeName()**
 - **appendChild()**
 - **removeChild(Node oldChild)**
-

DOM (3/4) – Interface Document

- Represents the whole document
 - **getDocumentElement()**
 - **getElementById(String elementId)**
 - **getElementsByTagName(String tagName)**
 - **createAttribute(String name)**
 - **createElement(String tagName)**
-

DOM (4/4)

- Other interfaces *Element*, *Attr*, *Text*, ...
 - Google: "**java dom api**"
-

Other programming languages

- C#:
 - DOM: XmlDocument - <http://msdn.microsoft.com/en-us/library/azsy1tw2%28v=vs.100%29.aspx>
 - DOM: XDocument (Linq to XML) - <http://msdn.microsoft.com/en-us/library/system.xml.linq.xdocument.aspx>
 - SAX: XmlReader: <http://msdn.microsoft.com/en-us/library/cc189056%28v=vs.95%29.aspx>
 - Differences between SAX parser & XmlReader: <http://msdn.microsoft.com/en-us/library/sbw89de7.aspx>
 - C++:
 - DOM: <http://xerces.apache.org/xerces-c/>
 - DOM: <http://rapidxml.sourceforge.net/>
 - SAX: <http://xerces.apache.org/xerces-c/>
 - Python:
 - DOM: <https://docs.python.org/2/library/xml.dom.html>
 - SAX: <https://docs.python.org/2/library/xml.sax.html>
-