

# XML Programming in Java

Andy Clark



# Which API?

- Standard APIs for XML programming in Java
  - Document Object Model (DOM)
    - Defined by W3C
    - Generic tree model
  - Simple API for XML (SAX)
    - Defined by members of xml-dev mailing list
    - “Event” model
  - Others
    - DOM4J, JDOM, XNI, etc...

# DOM Design Premise

- Derived from browser document model
- Defined in IDL
  - Lowest common denominator programming lang.
    - Most methods packed into base Node object
    - Re-invents basic collection interfaces
- Factory model
  - Allows custom classes to be used for nodes
- Separation between required and optional modules

# DOM Modules

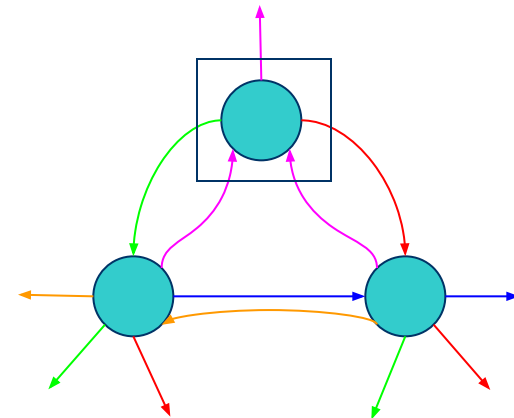
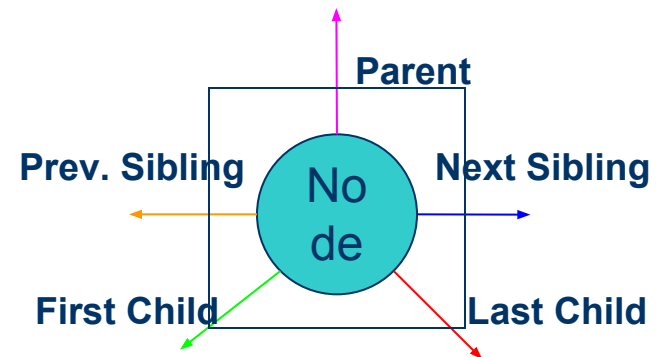
- DOM Level 1
  - Core (required)
  - HTML (optional)
- DOM Level 2
  - Core (required)
  - Views, Events, Style, Traversal, Range (optional)
- DOM Level 3 – *working draft*
  - Core (required)
  - Abstract Schema, Load/Save, Events, XPath (optional)

# Pros and Cons of DOM

- Pros
  - Entire document is in memory
  - Fully navigable and editable structure and content
  - Recognized standard for XML tree model
- Cons
  - Entire document is in memory
  - IDL mapping not intuitive to Java programmers

# Document Object Model

- Generic tree model
  - Node
    - Type, name, value
    - Attributes
    - Parent node
    - Previous, next sibling nodes
    - First, last child nodes
  - Collections
    - Lists
    - Maps



# SAX Design Premise

- Generic method of creating XML parser, parsing documents, and receiving document information
- “Streaming” information set
  - Application registers handlers
  - Parser “pushes” information to handlers
    - Serial (“as you see it”) notification
  - Application only uses information it needs

# Pros and Cons of SAX

- Pros
  - Simple API designed for Java
  - Uses very little memory
  - Recognized standard for XML event model
- Cons
  - Application forced to do *everything*



# Useful Links

- DOM Specification
  - Level 1: <http://www.w3.org/TR/REC-DOM-Level-1>
  - Level 2: <http://www.w3.org/TR/REC-DOM-Level-2>
- SAX
  - <http://sax.sourceforge.net/>

# **XML Programming in Java**

Andy Clark

---