BUSINESS INTELLIGENCE LABORATORY

Java and Eclipse - recap

Java

Java is a

- High-level
- Object oriented
- Functional (since Java 8)
- Multithreaded
- Architecture neutral
- Portable
- Robust
- Secure

computer programming language

www.oracle.com/technetwork/java/index.html

Java version history

From Wikipedia, the free encyclopedia

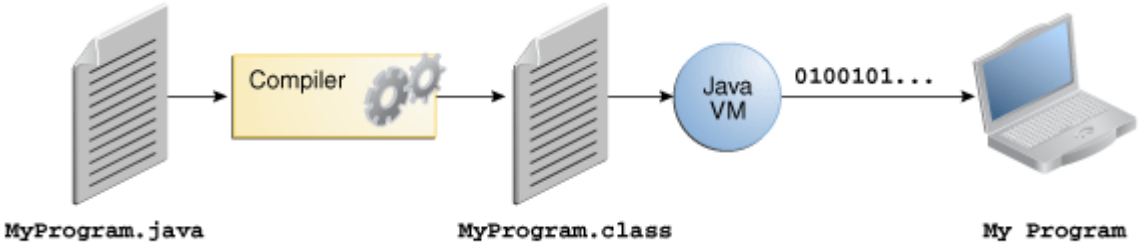
The Java language has undergone several ch
uses *Java Specification Requests* (JSRs) to p

In addition to the language changes, much m
have been introduced, and many of the origin
backporting tools).

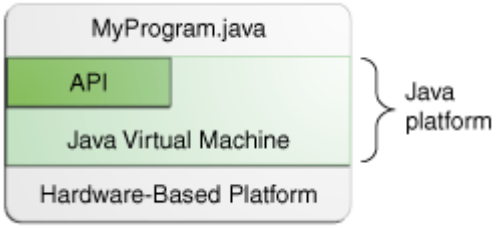
After the Java 7 release, Oracle promised to g

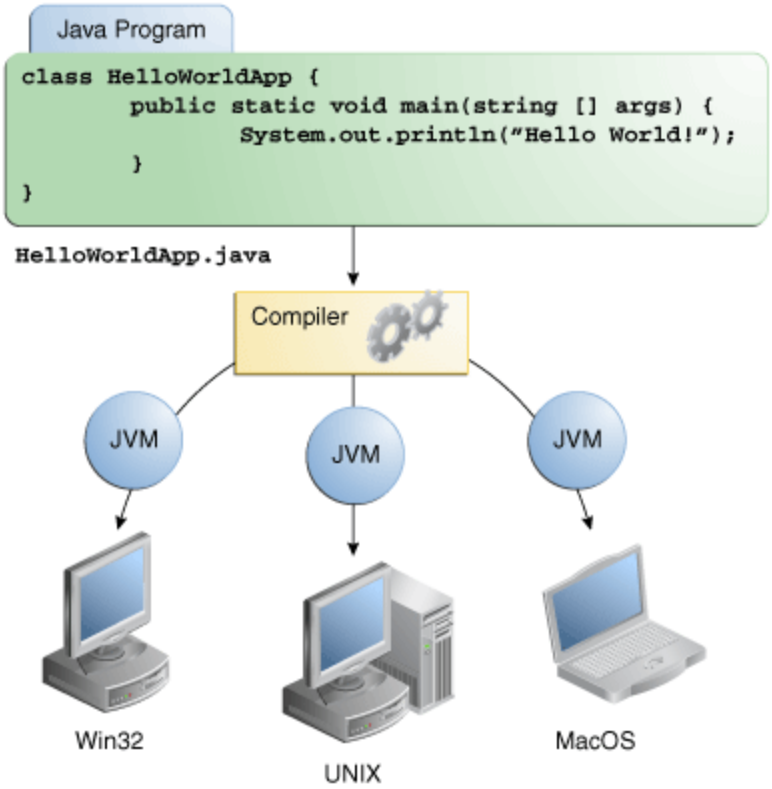
Contents [hide]

Java platform

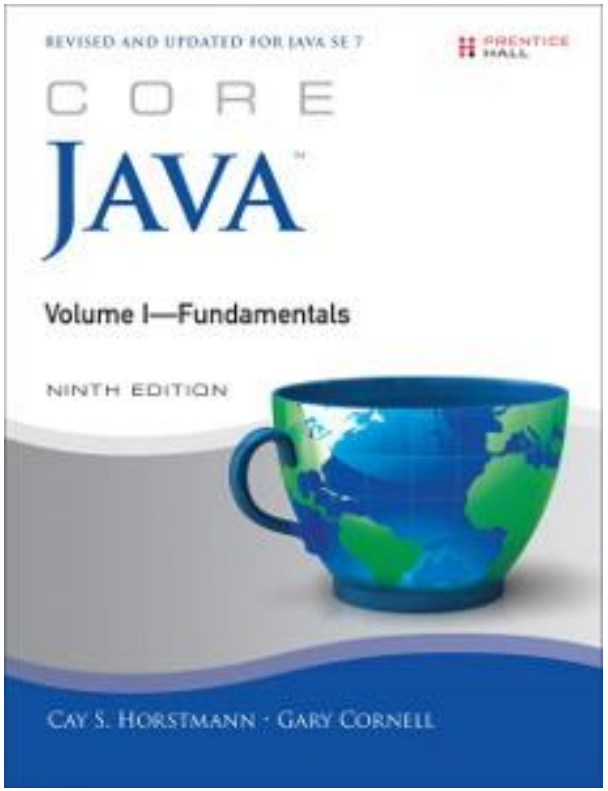


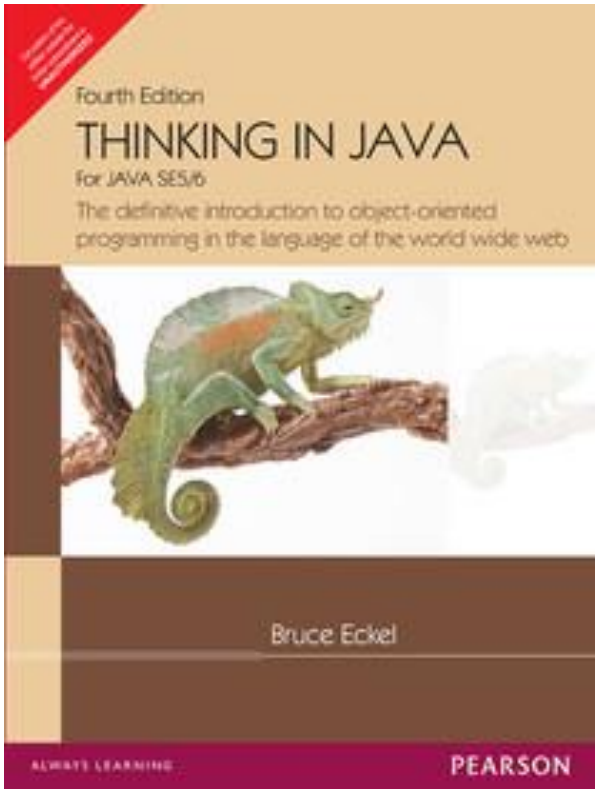
docs.oracle.com/javase/8/docs/api

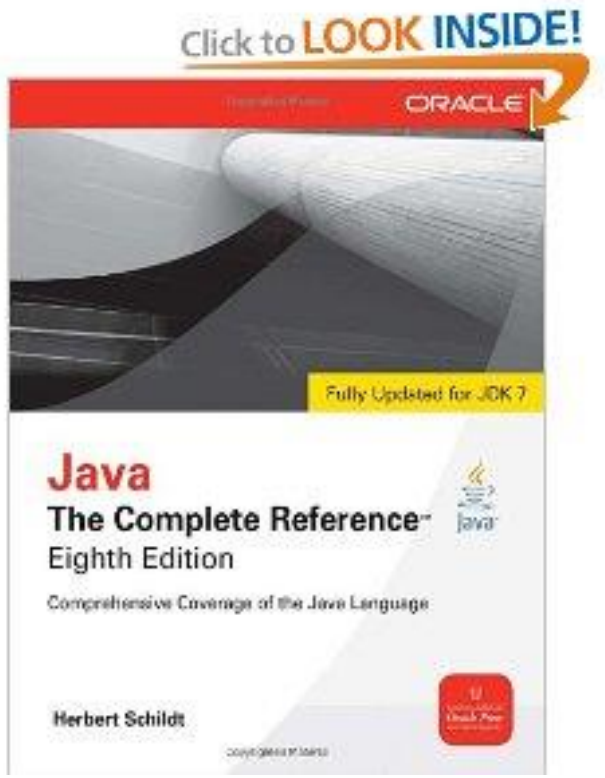




Java language: books and tutorials







docs.oracle.com/javase/tutorial

Eclipse

Eclipse is a multi-language software development environment

Version Name	Date	Platform Version
N/A	21 June 2004	3.0[18]
N/A	28 June 2005	3.1
Callisto	30 June 2006	3.2
Europa	29 June 2007	3.3
Ganymede	25 June 2008	3.4
Galileo	24 June 2009	3.5
Helios	23 June 2010	3.6
Indigo	22 June 2011	3.7
Juno	27 June 2012	3.8 and 4.2[25] [Notes 1]
Kepler	26 June 2013	4.3
Luna	25 June 2014	4.4
Mars	24 June 2015	4.5
Neon	22 June 2016	**4.6**
Oxygen	June 2017 (planned)	4.7

www.eclipse.org/downloads

Exercise: maximal subsequence

- Given an array of integers, e.g.
  - int []a = {−2, 1, −3, 4, −1, 2, 1, −5, 4 };
- and called
  - S(h, k) = $\sum_{i=h}^{k} a[i]$
- the sum of subsequence from k to k, find the maximal S(h, k)
  - max S(h, k)
- For the array above max S(h, k) = S(3, 6) = 4-1+2+1= 6

- Variants: array of integers
  - passed on the command line
  - read from a text file (one int per line)

OOP concepts

- An object-oriented program is made of objects.
  - Each object has a specific functionality that is exposed to its users, and a hidden implementation
    - Instance fields and methods
    - Encapsulation
- A *class* is the template or blueprint from which objects are made
  - Classes are cookie cutters. Objects are cookies

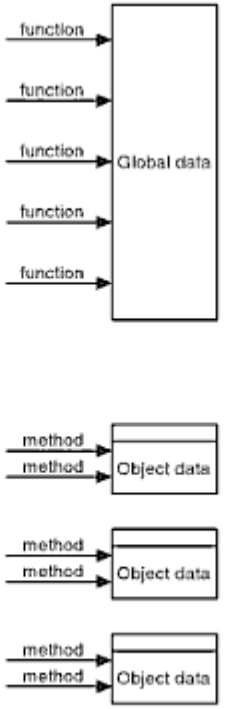


Figure 4–2: Procedural vs. OO programming

Objects in Java

Date birthday = **new Date();**

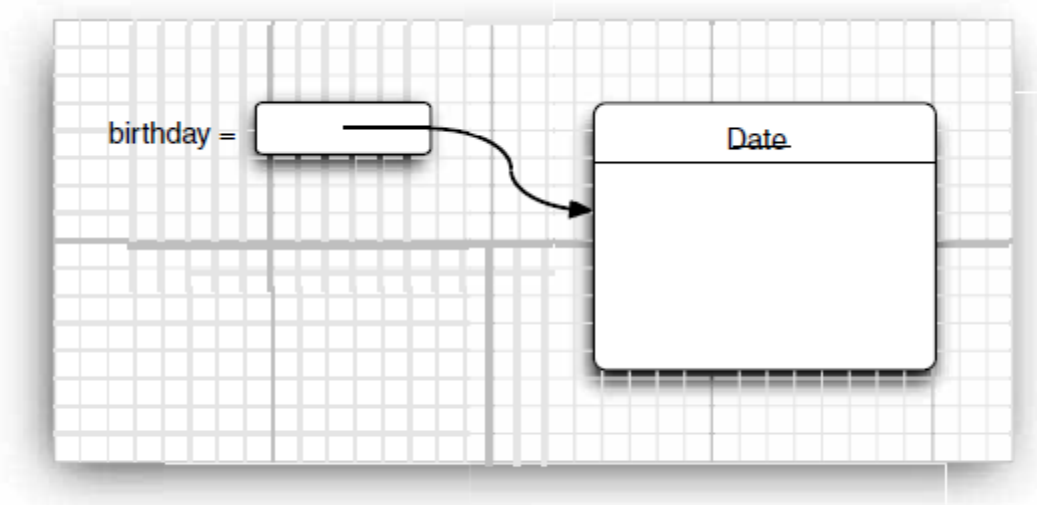


Figure 4–3 Creating a new object

Date today = birthday;

Date dealine; // initialized to null (no referenced object)

OOP concepts

- Classes can be built by *extending* other classes
  - Inheritance and polymorphism
    - The Object class in Java is at the top of the class hierarchy

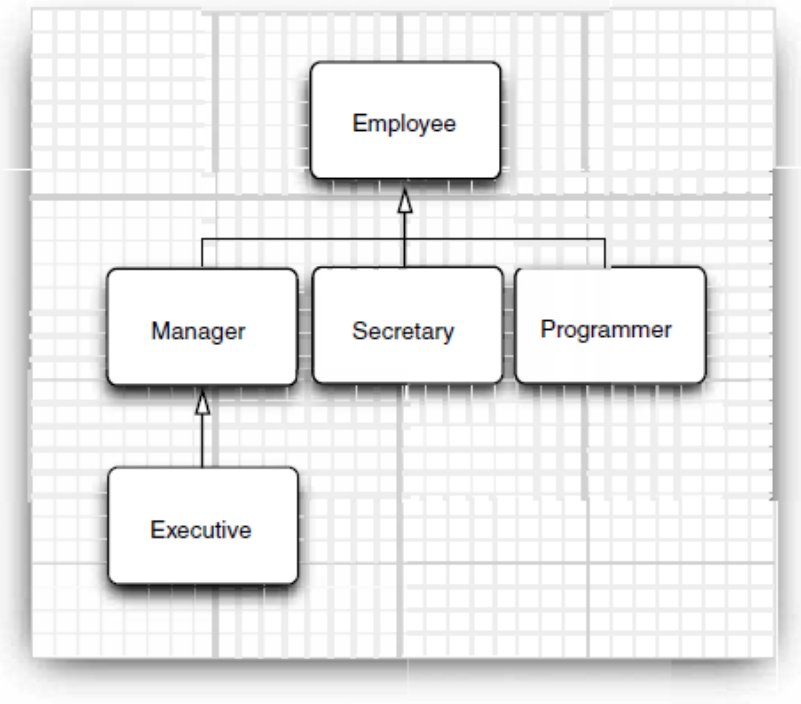


Figure 5–1 Employee inheritance hierarchy

Collections in Java

- https://docs.oracle.com/javase/tutorial/collections/

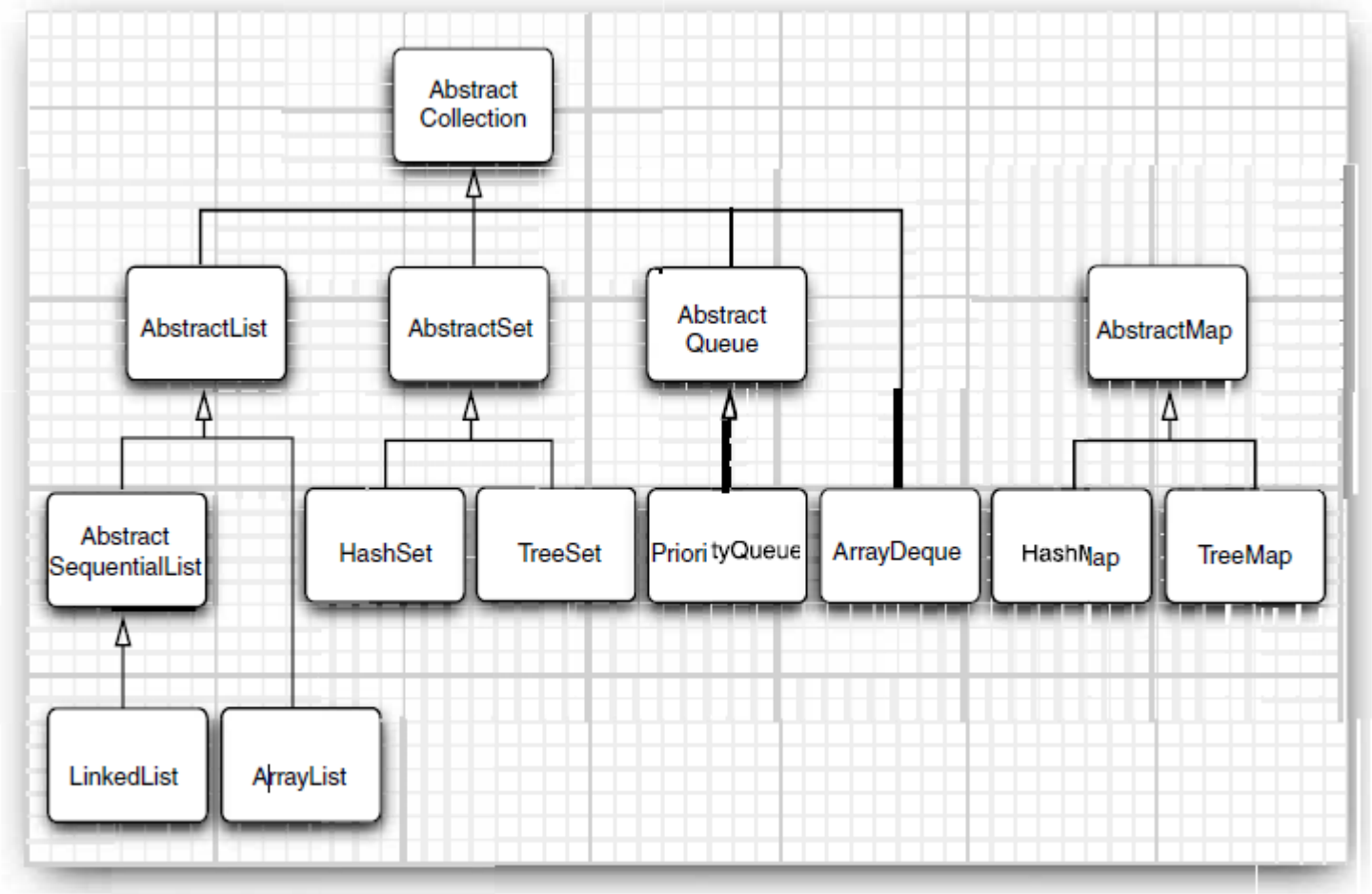


Figure 13–11 Classes in the collections framework

Exercise: Relational algebra

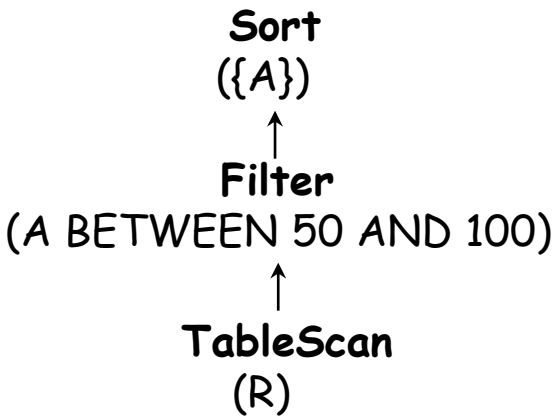
- Code a relational tuple as a Java class
  - Assume only integer and string data types
- Code a relational table as a Collection of tuples
- Code relational operators (as much as possible):
  - Selection: $\sigma_C(R)$
  - Projection: $\pi_A(R)$
  - Renaming: $\rho_{A \leftarrow B}(R)$
  - Distinct: $\delta(R)$
  - Ordering: $\tau_A(R)$
  - Grouping: ${}_A\gamma_{COUNT(*)\ AS\ B}$
  - Join: $R \bowtie_{A=B} S$
- Test them on two relations Sales(CustomerName, Product, Amount) and Customer(Name, Country, Age) for the query plans of the following queries:
  - SELECT Name FROM Customer WHERE Age > 20 AND Country='Italy' ORDER BY Age
  - SELECT Country, Count(*) AS TotSales FROM Sales JOIN Customer on CustomerName=Customer GROUP BY Country

Solution

- Class Tuple:
  - model pairs (attribute, value)
  - Using HashMap
- Class Table
  - model a collection of tuples
  - Using ArrayList
  - Operators return a new Table
    - Immutable class

Other approaches

- Using (parallel) streams to optimize performances
  - See e.g., http://www.oracle.com/technetwork/articles/java/ma14-java-se-8-streams-2177646.html
  - Map-Reduce programming pattern
- In .NET languages: LINQ
  - http://linqsamples.com/



Text File Data Access in Java

- Packages
  - java.io
- Classes
  - Scanner, BufferedReader, PrintWriter
- Utility package
  - lbi.utils
- Utility classes
  - IOUtil, Cronos, …

Exercise(s): format conversions (1)

- Write a Java program for converting a CSV file into the ARFF format
- Write a Java program for converting a CSV file into the JSON format
- Write a Java program for converting a XML-elements file into a CSV file using a DOM parser
  - Assume that CSV files have a first row of meta-data with the names of the columns

AGE
10
20
40
30
~~Bob~~ 33
43
45

METADATA

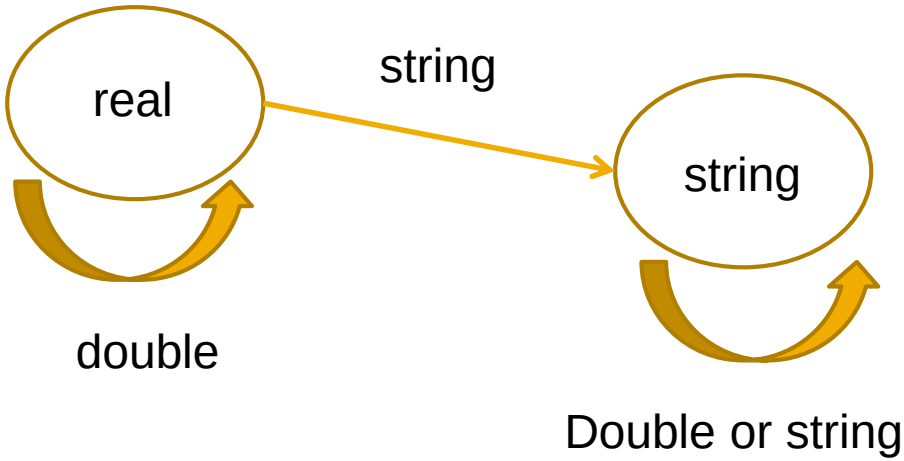
@relation ….
@attribute age ~~string~~ **numeric**

CSV2ARFF - solution

{ value1, value2, ..., valueK }

Set of values seen so far

enumerated $\leftarrow K \leq Max?$



real or string $\leftarrow K > Max?$

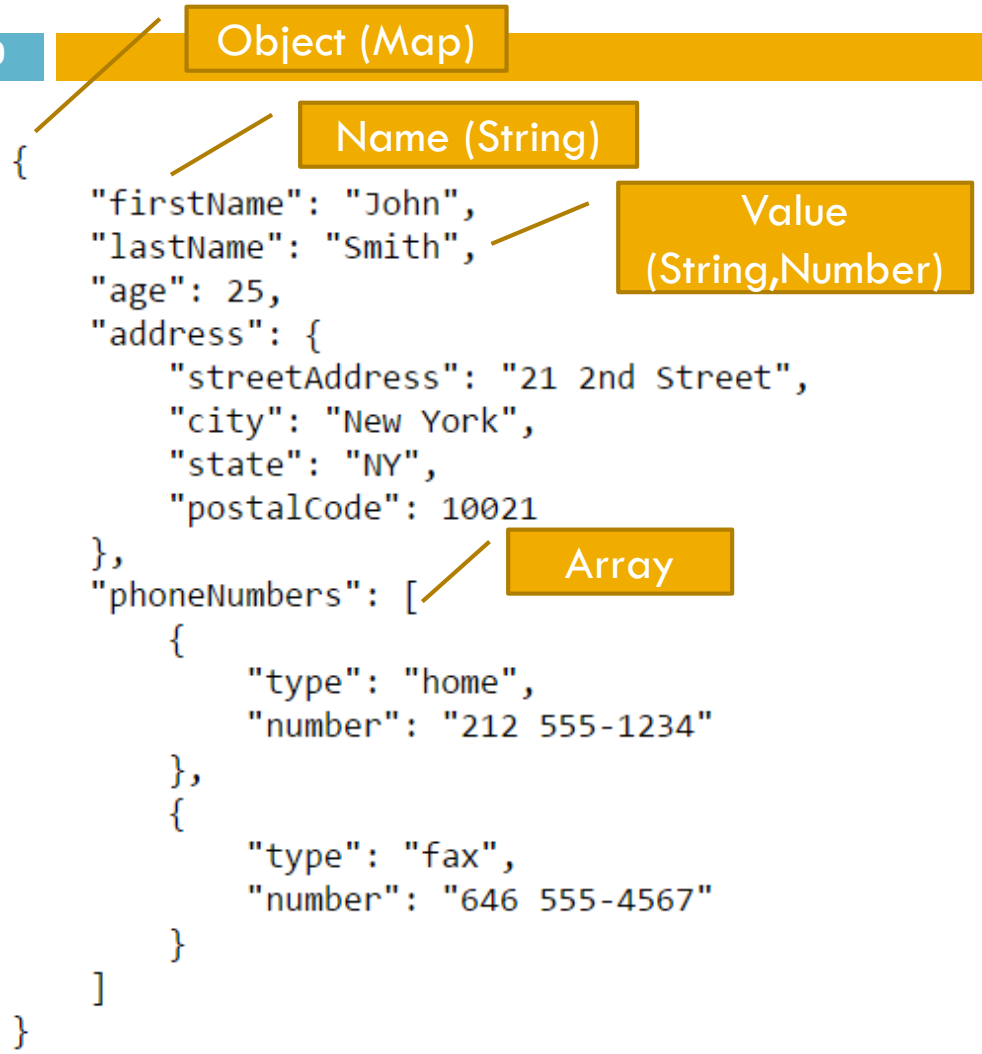
XML File Data Access in Java

- XML processors
  - DOM (Document Object Model) parser
    - it returns a tree data structure of the XML document
    - the tree is kept in main memory
      - **javax.xml.parsers.DocumentBuilder**
      - **org.w3c.dom.Document**
  - SAX (Simple API for XML) parser
  - StAX (Streaming API for XML) parser
    - SAX scans the document and make call-backs
    - StAX scans the document as a stream of events
    - No data structures in main memory
- Java API for XML Processing (JAXP)
  - Includes DOM, SAX, StAX and other XML processing features (XSLT, …)
  - http://docs.oracle.com/javase/tutorial/jaxp/

Exercise(s): format conversions (2)

- Write a Java program for converting a CSV file into the ARFF format
- Write a Java program for converting a XML-elements file into a CSV file using a DOM parser
- Write a Java program for converting a CSV file into the JSON format
  - Assume that CSV files have a first row of meta-data with the names of the columns

JSON File Data Access in Java



- JSONsimple library
  - https://code.google.com/archive/p/json-simple/
- Read/Write JSON
  - https://www.mkyong.com/java/json-simple-example-read-and-write-json/
- Java API for JSON (not yet part of JVM)
  - http://www.oracle.com/technetwork/articles/java/json-1973242.html

Exercise(s): format conversions (3)

- Write a Java program for converting a CSV file into the ARFF format
- Write a Java program for converting a XML-elements file into a CSV file using a DOM parser
- Write a Java program for converting a CSV file into the JSON format
  - Assume that CSV files have a first row of meta-data with the names of the columns