Class Announcements

- Database Assignment 1 Due 2/9 - TODAY!
- (regular) Assignment 3: Due 2/14
- Midterm: 2/16
Alibaba Discussion
Networked Computing in direct Procurement

- History predates Internet
- **Electronic Data Interchange** (EDI)
  - Exchange order information between firms involved in direct procurement
  - Usually large firms who could afford proprietary communication links
  - Initially order and invoice
  - Existed since 70’s
Business-to-Business Electronic Commerce: New Efficiencies and Relationships

• Electronic data interchange (EDI)
  
  • Major industries have EDI standards that define structure and information fields of electronic documents for that industry.
  
  • More companies increasingly moving away from private networks to Internet for linking to other firms.
In Class Exercise

- Use a laptop or phone and log in to Alibaba.com
- Pretend you are a small business in the US needing cheap products. Do a search, and say what your “story” is.
  - E.g. “We pretend to be a retailer aiming to sell American flags at marked up prices. We’ll search for American flags…”
  - E.g. “We pretend to be a contractor looking for cheap supplies. We’ll search for “toilet.”
  - E.g. “We pretend to be a beach shop selling tourist goods. We’ll search for “sunglasses.”

- How do the prices you get compare to retail prices? Are there minimum quantities to make an order?
- Are the first listings from “gold” suppliers?
- Do these suppliers seem to be large companies or small to medium sized?
Data and information

by

David G. Messerschmitt
©Copyright David G. Messerschmitt, 2000. This material may be used, copied, and distributed freely for educational purposes as long as this copyright notice remains attached. It cannot be used for any commercial purpose without the written permission of the author.
The key commodity manipulated by information technology is information.

To be manipulated in a computing/networking environment, information must be represented by data.
Information

From a user (human) perspective...
....recognizable patterns that influence you in some way
(perspective, understanding, behavior...)

In the computing infrastructure, information has a somewhat different connotation as structure and interpretation added to data
Data

A bit is “0” or “1” — the atom of the information economy

Data is a collection of bits, like

- “0101110111010110”
- “0000011”
- “11101110101101011011011011010”

Note: the terms data and information are not always used consistently!
Representation

- Take the place of the original
- Equivalent to, in the sense that the original can be reconstructed from its representation
- Often the original can only be approximately reconstructed, although it may be indistinguishable to the user
  - e.g. audio or video
### ASCII

<table>
<thead>
<tr>
<th>Alphabet</th>
<th>Hex</th>
<th>Binary</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>&lt;7&gt;</code></td>
<td>/x37</td>
<td>00110111</td>
</tr>
<tr>
<td><code>&lt;8&gt;</code></td>
<td>/x38</td>
<td>00111000</td>
</tr>
<tr>
<td><code>&lt;9&gt;</code></td>
<td>/x39</td>
<td>00111001</td>
</tr>
<tr>
<td><code>&lt;:</code></td>
<td>/x3A</td>
<td>00111010</td>
</tr>
<tr>
<td><code>&lt;;</code></td>
<td>/x3B</td>
<td>00111011</td>
</tr>
<tr>
<td><code>&lt;&lt;</code></td>
<td>/x3C</td>
<td>00111100</td>
</tr>
<tr>
<td><code>=&gt;</code></td>
<td>/x3D</td>
<td>00111101</td>
</tr>
<tr>
<td><code>&lt;/&gt;&gt;</code></td>
<td>/x3E</td>
<td>00111110</td>
</tr>
<tr>
<td><code>&lt;?&gt;</code></td>
<td>/x3F</td>
<td>00111111</td>
</tr>
<tr>
<td><code>&lt;At&gt;</code></td>
<td>/x40</td>
<td>01000000</td>
</tr>
<tr>
<td><code>&lt;A&gt;</code></td>
<td>/x41</td>
<td>01000001</td>
</tr>
<tr>
<td><code>&lt;B&gt;</code></td>
<td>/x42</td>
<td>01000010</td>
</tr>
<tr>
<td><code>&lt;C&gt;</code></td>
<td>/x43</td>
<td>01000011</td>
</tr>
<tr>
<td><code>&lt;D&gt;</code></td>
<td>/x44</td>
<td>01000100</td>
</tr>
</tbody>
</table>

Note that this representation is not unique…

….this one happens to be a **standard** (ANSI X3.110-1983)
A picture

This picture conveys information

This information is represented in this computer, but how?
Expanding a small portion of the picture, we see that it is represented by square pixels.

....300 tall by 200 wide.....

....with a range of 256 intensities per pixel

300 • 200 • 8 bits = 480,000 bits (but it can be compressed)
A color picture can be represented by three monochrome images…

At the expense of three times as many bits
Terminology

Information
---
Representation
---
Data processing
---
Communicate data to another user or organization
---

Data

Slide adapted from slides for *Understanding Networked Applications* by David G Messerschmitt. Copyright 2000. See copyright notice.
Representation needs to be standardized

If the representation is not standardized, the information is garbled!

Communicate data to another user or organization

Slide adapted from slides for *Understanding Networked Applications* by David G Messerschmitt. Copyright 2000. See copyright notice.
Regeneration

- Make a precise copy of the data (copy bit by bit)
- If you know the representation, this is equivalent to making a precise copy of the information
- Each such precise copy is called a generation
- Process is called regeneration
Replication of information

Anything that can be regenerated can be replicated any number of times

This is a blessing and a curse
Analog information cannot be regenerated

Analog information can be copied, but not regenerated

We will never know exactly what the original of this Rembrandt looked like
Discrete information can be regenerated

Regeneration can preserve data (but not its original physical form)

Regeneration is possible for information represented digitally (which is tolerant of physical deterioration)

\[0 + \text{noise} \rightarrow 0\]
\[1 + \text{noise} \rightarrow 1\]
Replication of information requires knowledge of representation.

Replication of information also presumes knowledge of its representation.

Replication preserves the integrity of the data, but that is not sufficient.

Every .xxx DOS file is a representation.
Implications

Digitally represented information can be preserved over time or distance in its precise original form by occasional regeneration

- digital library
- digital telephony

Replication of data is easy and cheap
Implications (con’t)

- Replication of information requires knowledge of the structure and interpretation
  - Standardization or some other means
- Extreme supply economies of scale
- You can give away or sell and still retain
- Unauthorized replication or piracy relatively easy
Architecture

by
David G. Messerschmitt
©Copyright David G. Messerschmitt, 2000. This material may be used, copied, and distributed freely for educational purposes as long as this copyright notice remains attached. It cannot be used for any commercial purpose without the written permission of the author.
What is Architecture?

How do you architect a solution?
A system is decomposed into interacting subsystems.

Each subsystem may have a similar internal decomposition.
Three elements of architecture

Decomposition
Organization
Functionality
Responsibility
Interaction
Cooperation

Slide adapted from slides for Understanding Networked Applications
By David G Messerschmitt. Copyright 2000. See copyright notice
System examples

Let’s quickly look at some system decomposition examples

- Quick tour of information technology systems
Time sharing

ASCII terminal
(no graphics)

Point-to-point wire
(no network)

Mainframe
(database and application server)
Two-tier client/server

Local-area network
Three-tier client/server
System integration

Architecture
    -> subsystem implementation
    -> system integration

Bring together subsystems and make them cooperate properly to achieve desired system functionality

- Always requires testing
- May require modifications to architecture and/or subsystem implementation
Emergence

Subsystems are more specialized and simpler functionality

Higher-level system functionality arises from the interaction of subsystems

**Emergence** includes capabilities that arise purely from that interaction (desired or not)

- e.g. airplane flies, but subsystems can’t
Why system decomposition?

- Divide and conquer approach to containing complexity
- Reuse
- Consonant with industry structure (unless system is to be supplied by one company)
- Others?
Networked computing infrastructure

by
David G. Messerschmitt
Layering

Existing layers

Elaboration or specialization

Services
Example of Layering: networking

- **Physical**
  - Bits
  - Frames
  - Packets
  - Messages

- **Link**
  - Bits

- **Network**
  - Frames
  - Packets
  - Messages

- **Transport**
  - Messages

- **Application**
Software Layering

- Application
- Middleware
- Operating System
Operating system functions

- Graphical user interface (client only)
- Hide details of equipment from the application
- Multitasking
- Resource management
  - Processing, memory, storage, etc
- etc
Middleware Functions

- Capabilities that can be shared by many applications, but that is not part of OS
  - Example: Database Management System (DBMS)
- Hide details of OS from application
  - Java Virtual Machine
- More purposes we’ll talk about later.
What’s a database?

Database

- File with specified structure
- Example: relational table
<table>
<thead>
<tr>
<th>Year</th>
<th>City</th>
<th>Accommodation</th>
<th>Tourists</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>Oakley</td>
<td>Bed&amp;Breakfast</td>
<td>14</td>
</tr>
<tr>
<td>2002</td>
<td>Oakley</td>
<td>Resort</td>
<td>190</td>
</tr>
<tr>
<td>2002</td>
<td>Oakland</td>
<td>Bed&amp;Breakfast</td>
<td>340</td>
</tr>
<tr>
<td>2002</td>
<td>Oakland</td>
<td>Resort</td>
<td>230</td>
</tr>
<tr>
<td>2002</td>
<td>Berkeley</td>
<td>Camping</td>
<td>120000</td>
</tr>
<tr>
<td>2002</td>
<td>Berkeley</td>
<td>Bed&amp;Breakfast</td>
<td>3450</td>
</tr>
<tr>
<td>2002</td>
<td>Berkeley</td>
<td>Resort</td>
<td>390800</td>
</tr>
<tr>
<td>2002</td>
<td>Albany</td>
<td>Camping</td>
<td>8790</td>
</tr>
<tr>
<td>2002</td>
<td>Albany</td>
<td>Bed&amp;Breakfast</td>
<td>3240</td>
</tr>
<tr>
<td>2003</td>
<td>Oakley</td>
<td>Bed&amp;Breakfast</td>
<td>55</td>
</tr>
<tr>
<td>2003</td>
<td>Oakley</td>
<td>Resort</td>
<td>320</td>
</tr>
<tr>
<td>2003</td>
<td>Oakland</td>
<td>Bed&amp;Breakfast</td>
<td>280</td>
</tr>
<tr>
<td>2003</td>
<td>Oakland</td>
<td>Resort</td>
<td>210</td>
</tr>
<tr>
<td>2003</td>
<td>Berkeley</td>
<td>Camping</td>
<td>115800</td>
</tr>
<tr>
<td>2003</td>
<td>Berkeley</td>
<td>Bed&amp;Breakfast</td>
<td>4560</td>
</tr>
<tr>
<td>2003</td>
<td>Berkeley</td>
<td>Resort</td>
<td>419000</td>
</tr>
<tr>
<td>2003</td>
<td>Albany</td>
<td>Camping</td>
<td>7650</td>
</tr>
<tr>
<td>2003</td>
<td>Albany</td>
<td>Bed&amp;Breakfast</td>
<td>6750</td>
</tr>
</tbody>
</table>
Storage Middleware example: DBMS

- Database Management System (DBMS)
  - Manage Multiple databases
  - Allow multiple applications to access common databases
  - Implement standard data “lookup” (query) functions.