Molecular Biomedical Informatics

Social Web Design & Research

Social Web Design & Research
Database
Database

- A database is an organized collection of data
- The data and their supporting data structures rather than the database management system (DBMS)
  - the database data collection with DBMS is called a database system
  - well known DBMSs include Oracle, IBM DB2, Microsoft SQL Server, Microsoft Access, PostgreSQL, MySQL, and SQLite
- Data models
  - relational model
  - entity-relationship model (for example, the food in a restaurant)
  - object model
  - object relational model
  - XML as a database data model
Table

- A row is an entry; a column is a property

(a) | Food | Place |
--- | --- | --- |
F1 | P1, P2 |
F2 | P1, P3 |

(b) | Place | Food |
--- | --- |
P1 | F1, F2 |
P2 | F1, F3 |

(c) | Food | Place | Name | Price |
--- | --- | --- | --- | --- |
F1 | P1 | F1 | 100 |
F1 | P2 | F2 | 200 |
F2 | P1 | F2 |  |  |
F2 | P3 | |  |  |

| Name | Phone |
--- | --- |
P1 | ... |
P2 | ... |
You prefer?
SQL (Structured Query Language)

- Different DBMSs can inter-operate by using standards (e.g. SQL) and ODBC (Open Database Connectivity) to support together a single application

- The most common operation in SQL is the declarative **SELECT** statement, which includes
  - a list of columns to be retrieved (an asterisk means all columns)
  - **FROM** clause to specify the table(s) from which data is to be retrieved
  - **JOIN** subclauses to connect tables
  - **WHERE** clause to specify the conditions
  - **ORDER BY** clause to specify the sorting scheme
SELECT

- SELECT *
  FROM food
  WHERE price > 100
  ORDER BY name

- SELECT place.phone
  FROM place
  JOIN f-p ON place.name = f-p.place
  JOIN food ON f-p.food = food.name
  WHERE food.price > 100

<table>
<thead>
<tr>
<th>Food</th>
<th>Place</th>
<th>Name</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>P1</td>
<td>F1</td>
<td>100</td>
</tr>
<tr>
<td>F1</td>
<td>P2</td>
<td>F2</td>
<td>200</td>
</tr>
<tr>
<td>F2</td>
<td>P1</td>
<td>F2</td>
<td></td>
</tr>
<tr>
<td>F2</td>
<td>P3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name</th>
<th>Phone</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>...</td>
</tr>
<tr>
<td>P2</td>
<td>...</td>
</tr>
</tbody>
</table>
INSERT, UPDATE and DELETE

- INSERT INTO food (name, price) VALUES (F1, 100)
- UPDATE food
  SET price = 100
  WHERE name = F1;
- DELETE FROM food
  WHERE name = F1;
Other SQL operations

- **Definition**
  - CREATE
  - ALTER
  - DROP

- **Control**
  - GRANT

- [SQL 教學 - SQL Tutorial](#)
Database technologies

- phpMyAdmin
- Other apps

- SQL

- Schema (normalization)

- Storage (B+ tree)
SQL in Perl

```perl
#!/usr/bin/perl -w
use DBI;
my $dbh = DBI->connect('dbi:mysql:db_name:host', 'user')
or die $DBI::errstr;

# insert
$dbh->do("INSERT INTO `food` ( `name`, `price` ) VALUES ( 'F1', 100 )")
or die $DBI::errstr;

# select
my $sth = $dbh->prepare("SELECT * FROM `food` WHERE `price` > 100");
$sth->execute() or die $DBI::errstr;
while ( my $row = $sth->fetchrow_hashref() ) {
    print $row->{name}, "\n";
}
```

- [Perl DBI - dbi.perl.org](http://dbi.perl.org)
- [DBI - search.cpan.org](http://search.cpan.org)

Social Web Design & Research
SQL in PHP

```php
<?php
$conn = mysql_connect( 'host', 'user', 'password' )
    or die( mysql_error() );
mysql_select_db('db_name') or die( mysql_error() );

# insert
mysql_query( "INSERT INTO `food` ( `name`, `price` ) VALUES ( 'F1', 100 )" )
    or die( mysql_error() );

# select
$result = mysql_query( "SELECT * FROM `food` WHERE `price` > 100" )
    or die( mysql_error() );
while ( $row = mysql_fetch_assoc( $result ) ) {
    echo $row['name'], "
";
}

mysql_free_result($result);
mysql_close($conn);
?>
```

PHP: Mysql - Manual
Any Questions?
HTTPS
Hypertext Transfer Protocol Secure
HTTPS

- A combination of the Hypertext Transfer Protocol (HTTP) with the SSL/TLS protocol

- HTTP operates at the highest layer of the OSI Model, the Application layer; but the security protocol operates at a lower sublayer, encrypting an HTTP message prior to transmission and decrypting a message upon arrival

- HTTP Secure - Wikipedia, the free encyclopedia
SSL/TLS

- Transport Layer Security (TLS) and its predecessor, Secure Sockets Layer (SSL), are cryptographic protocols that provide communication security over the Internet.
- TLS and SSL encrypt the segments of network connections at the Transport Layer, using asymmetric cryptography for key exchange.
- SSL was developed by Netscape for secure web communication with a great success; while IETF (www.ietf.org) made it a standard (RFC2246) and called it TLS.
- In technique, the SSL 3.0 and TLS 1.0 are very similar.
- Transport Layer Security - Wikipedia, the free encyclopedia.
In short, HTTPS is not a separate protocol, but refers to use of ordinary HTTP over an encrypted SSL/TLS connection.

加密 → 傳 → 解密
Asymmetric cryptography

- The basic of modern cryptography (密碼學)
- XOR could be a symmetric cryptography
  - there is a problem…
Suppose

you are a king and want to communicate with another king…
Public/private key

- Symmetric-key algorithms, used for thousands of years, use a single secret key for both encryption and decryption where the sender and receiver must securely share a key in advance.
- Two separate keys, one to lock or encrypt the plaintext, and one to unlock or decrypt the cyphertext.
- Neither key will do both functions, namely useless.
- One of these keys is published or public and the other is kept private.
  - if the lock/encryption key is the one published then the system enables private communication from the public to the unlocking key’s owner.
  - if the unlock/decryption key is the one published then the system serves as a signature verifier of documents locked by the owner of the private key.

Public-key cryptography - Wikipedia, the free encyclopedia
Creating asymmetric key pairs

- Heavily depends on large prime numbers \( \Rightarrow \) mathematics is important
- The ‘direct’ (encryption) operation is relatively easy, but the ‘inverse’ (decryption) operation must be significantly more difficult
- The multiplication of integers is easy, but finding the prime factors of the product is much harder
  - the famous RSA algorithm uses this concept
- The one-way function have many applications, such as the famous MD5 algorithm
- [MD5 - Wikipedia, the free encyclopedia](https://en.wikipedia.org/wiki/MD5)
Perfect

with asymmetric cryptography?
The man-in-the-middle attack
What exactly to do

- **Install**
  - `$ sudo apt-get install apache2`
  - `$ sudo apt-get install openssl`
  - `$ sudo apt-get install ssl-cert`
  - `$ sudo make-ssl-cert /usr/share/ssl-cert/ssleay.cnf /etc/ssl/private/host.pem`

- **In apache.conf**
  - `<VirtualHost *:443>`
    - `SSLEngine On`
    - `SSLCertificateFile /etc/ssl/private/host.pem`
  - `</VirtualHost>`

- **Start**
  - `$ a2enmod ssl` # enable ssl module of apache
  - `$ /etc/init.d/apache2 restart` # make the settings work

- 在 Ubuntu 上安装 Apache2 的 HTTPS Server
- So...
Nothing you can do,
unless you are the web administrator
The world is big
Actually there are things you can do (heard of malicious code?)
Any Questions?
Today’s assignment
今天的任務
Make your site more fluent

- Use database in your site

- Reference
  - Perl DBI - dbi.perl.org
  - DBI - search.cpan.org
  - PHP: Mysql - Manual

- Your web site (http://merry.ee.ncku.edu.tw/~xxx/cur/) will be checked not before 23:59 5/14 (Tue). You may send a report (such as some important modifications) to me in case I did not notice your features.
Appendix
Advance tips of MySQL

- BLOB is binary; TEXT is text
  - BLOB vs. TEXT
- |CHAR| is [0,255]; |VARCHAR| is [0,65535]
  - CHAR vs. VARCHAR
- DATE vs. DATETIME vs. TIMESTAMP
  - DATE from '1000-01-01' to '9999-12-31'
  - DATETIME from '1000-01-01 00:00:00' to '9999-12-31 23:59:59'
  - TIMESTAMP from '1970-01-01 00:00:01' UTC to '2038-01-19 03:14:07' UTC.
- email: VARCHAR(320)
  - Maximum length of a valid email id
- utf8
  - $dbh→do('SET NAMES utf8');
  - UTF8, MySQL, Perl and PHP
通訊協定堆疊總結

- 應用層支援網路的各種應用軟體
  - FTP, SMTP, HTTP
- 傳輸層負責兩端點之間資料傳輸的虛擬連線
  - UDP, TCP
- 網路層負責傳送端和接收端之間的路徑
  - routing protocol, IP
- 連結層負責相臨端點之間的傳輸
  - Ethernet, PPP
- 實體層為實體線材上的訊號
實際上的通訊

Application
Transport
Network
Link
Physical

Application
Transport
Network
Link
Physical