

CS122B Projects in Databases and Web Applications
Department of Computer Science, UC Irvine, Prof. Chen Li
Winter 2017, Quiz 2

Initial Score (out of 100)

Name

Student ID

In accordance with both the letter and spirit of the Honor Code, I have neither given nor received assistance on this examination.

Signature _____

Make sure to write succinct answers to each question.

Question 1: (40 points)

For each of the following statements, mark **True ("T")** or **False ("F")**:

1. _____ The reCAPTCHA service by Google provides subscribing websites with images that are hard for a computer program to analyze.
2. _____ In MySQL, if we set call `Connection.setAutoCommit(false)` for a JDBC connection, then every update to the database using this connection will be made permanent immediately.
3. _____ Most DBs support native full-text search.
4. _____ XML is a semi-structured data model because it has some structure, but less flexible than the structured model. ☐
5. _____ Using a search architecture with a middleware layer such as Lucene/Solr has a disadvantage of keeping two copies of the same data.
6. _____ In project 3, one way to optimize the process of parsing the XML files is to store the parsed records as in-memory hash tables.
7. _____ One of the full-text search challenges is stemming, which tokenizes a string into multiple words.
8. _____ Each thread has one or more processes.

Question 2: (20 points)

There are two commonly used methodologies in parsing XML, namely SAX and DOM. Match each parsing method with its properties.

SAX _____	1. More control to the implementer;
	2. Stores the entire XML document in memory before further processing;
DOM _____	3. Represent an XML document as a hierarchy;
	4. Event-based API.

Question 3: (40 points)

For each question, select one and only one correct answer.

1- Which statement describes the Java Servlet architecture compared to the CGI architecture?

- A. More expensive to create a new process for each request;
- B. Stateless, no history remembered;
- C. No interaction with the Web server once the process has started;
- D. **Portable across different operating systems.**

2- Which statement describes how an AJAX-based Web application works?

- A. On detecting a user action, the browser on the client downloads Java code from the server and runs it on the client machine's JVM. It uses the results to refresh the HTML in the page. If no JVM is found, the client shows an error;
- B. **On detecting a user action, the browser makes a request to the backend server. The returned XML/HTML content is passed to a Javascript method. This method uses the data to refresh certain part of the current page;** ☒
- C. On detecting a user action, the user is provided a URL that contains a page with new content. The user is expected to paste this URL into a browser window and view the new page;
- D. On detecting a user action, the browser is requested to open a different page (with a different URL). The new page is downloaded and shown in the browser. The state of the original page is lost. ☐

3- Which option is **false** regarding XML?

- A. Uses a tree-structured representation for storing data;
- B. May introduce high redundancy ☐;
- C. Is a self-describing data format; ☐
- D. **Is not usable unless ported along with a DTD file.**

4- Consider the following DTD definition of an element named **product** in XML:

```
<!ELEMENT Product (Id, note*, spec?, price+)
```

Which statement about a **Product** element is true?

- A. A product must have an **Id** element, need not have any **note** element, must have exactly one **spec** element, and must have at least one **price** element;
- B. A product must have an **Id** element, must have more than one **note** element, need not have any **spec** element, and can have several **price** elements;
- C. A product must have an **Id** element, may have a **note** element, can have at most one **spec** element, and must have at least one **price** element;
- D. The only child a product must have is the **Id**. All other children are optional.

5- Which option is **not** a good choice for a secure Web application?

- A. Using an HTTP GET method;
- B. Using an HTTP POST method;
- C. Using a PASSWORD text input;
- D. Using HTTPS.

6- Suppose MySQL has an inverted index on an attribute called “review”. For a full-text search query that looks for records whose review includes both “apple” and “orange”, the database engine does the following:

- A. Do a union of the “apple” inverted list and the “orange” inverted list;
- B. Do an intersection of the “apple” inverted list and the “orange” inverted list;
- C. Do a difference operation of the “apple” inverted list and the “orange” inverted list;
- D. Scan all the records.

7- Which SQL statement is correct to search for the “title” attribute with both keywords “good” and “day”?


- A. `SELECT id FROM movies WHERE MATCH (title) AGAINST ('+good +day' IN BOOLEAN MODE);`
- B. `SELECT id FROM movies WHERE MATCH (title) AGAINST ('good day' IN BOOLEAN MODE);`
- C. `SELECT id FROM movies WHERE MATCH (title) AGAINST ('+good -day' IN BOOLEAN MODE);`
- D. `SELECT id FROM movies WHERE MATCH (title) AGAINST ('-good -day' IN BOOLEAN MODE);`

8- Which SQL statement is correct for prefix search for the “title” attribute with “pro” as a keyword prefix:

- A. `SELECT id FROM movies WHERE MATCH (title) AGAINST ('pro+' IN BOOLEAN MODE);`
- B. `SELECT id FROM movies WHERE MATCH (title) AGAINST ('pro*' IN BOOLEAN MODE);`
- C. `SELECT id FROM movies WHERE MATCH (title) AGAINST ('pro');`
- D. `SELECT id FROM movies WHERE MATCH (title) AGAINST ('pro-');`

Question 1: (40 points)

For each of the following statements, mark **True ("T")** or **False ("F")**:

1. **T** The reCAPTCHA service provided by Google provides subscribing websites with images that are hard for a computer program to analyze.
2. **F** In MySQL, if we set call Connection.setAutoCommit(false) for a JDBC connection, then every update to the database will be made permanent immediately.
3. **T** Most DBs support native search.
4. **F** XML is a semi-structured data model because it has some structure, but less flexible than the structured model. 
5. **T** Using a search architecture with a middleware such as Lucene/Solr has a disadvantage of keeping two copies of the same data.
6. **T** In project 3, one way to optimize the process of parsing the XML files is to store the parsed records as an in-memory hash table.
7. **F** One of the full-text search challenges is stemming, which tokenizes a string into multiple words.
8. **F** Each thread has one or more processes.

Question 2: (20 points)

There are two commonly used methodologies in parsing XML, namely SAX and DOM. Match each parsing method with its properties.

SAX <u> 1,4 </u>	5. More control to the implementer
	6. Stores the entire XML document into memory before further processing
DOM <u> 2,3 </u>	7. Represent an XML document as a hierarchy
	8. Event-based API

Question 3: (40 points)

1	2	3	4	5	6	7	8
D	B	D	C	A	B	A	B

1- Which statement describes the Java Servlets architecture compared to the CGI architecture?

- A. Expensive to create a new process for each request
- B. Stateless, no history remembered
- C. No interaction with the web server once the process started
- D. portable, across operating systems**

2- Which statement describes how an AJAX-based Web application works:

- A. On detecting a user action, the browser on the client downloads Java code from the server and runs it on the client machine's JVM. Use its results to refresh the HTML in the page. If no JVM is found, the client shows an error. ☒
- B. On detecting a user action, the browser makes a request to the backend server. The returned XML/HTML content is passed to a Javascript method. This method uses the data to refresh certain part of the current page. ☑
- C. On detecting a user action, the user is provided a URL that contains a page with new content. The user is expected to paste this URL into a browser window and view the new page. ☒
- D. On detecting a user action, the browser is requested to open a different page (with a different URL). The new page is downloaded and shown on the browser. The state of the original page is lost. ☒

3- Which option is false regarding XML?

- A. Uses a tree-structured representation for storing data.
- B. May introduce high redundancy☒
- C. Is a self-describing data format. ☒
- D. Is not usable unless ported along with a DTD file

4- Consider the following DTD definition of an element named **product** in XML:

`<!ELEMENT Product (Id, note*, spec?, price+)`

Which statement about a **Product** element is true?

- A. A product must have an **Id** element, need not have any **note** element, must have exactly one **spec** element, and must have at least one **price** element.
- B. A product must have an **Id** element, must have more than one **note** element, need not have any **spec** element, and can have several **price** elements.
- C. A product must have an **Id** element, may have a **note** element, can have at most one **spec** element, and must have at least one **price** element.
- D. The only child a product must have is the **Id**. All other children are optional

5- Which option is not a good choice for a secure web application?

- A. Using HTTP GET method
- B. Using HTTP POST method
- C. Using PASSWORD text input
- D. Using HTTPS

6- Suppose MySQL has an inverted index on an attribute called "review". For full text search query that looks for records whose review includes both "apple" and "orange", the database engine will do the following:

- A. Do a union of "apple" inverted list and "orange" inverted list.
- B. Do an intersection of "apple" inverted list and "orange" inverted list.

- C. Do a difference operation of “apple” inverted list and “orange” inverted list.
- D. Scan all the records

7- Which SQL statement is correct to search for the “title” attribute with both keywords “good” and “day”?

- A. `SELECT id FROM movies WHERE MATCH (title) AGAINST ('+good +day' IN BOOLEAN MODE);`
- B. `SELECT id FROM movies WHERE MATCH (title) AGAINST ('good day' IN BOOLEAN MODE);`
- C. `SELECT id FROM movies WHERE MATCH (title) AGAINST ('+good -day' IN BOOLEAN MODE);`
- D. `SELECT id FROM movies WHERE MATCH (title) AGAINST ('-good -day' IN BOOLEAN MODE);`

8- Which SQL statement is correct for prefix search for the “title” attribute with “pro” as a keyword prefix:

- A. `SELECT id FROM movies WHERE MATCH (title) AGAINST ('pro+' IN BOOLEAN MODE);`
- B. `SELECT id FROM movies WHERE MATCH (title) AGAINST ('pro*' IN BOOLEAN MODE);`
- C. `SELECT id FROM movies WHERE MATCH (title) AGAINST ('pro');`
- D. `SELECT id FROM movies WHERE MATCH (title) AGAINST ('pro-');`