**Final Exam Page 1** 

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | |  | | --- | | **1.**(TCO A) Which of the following is not a numeric data type? (Points : 5) |    binary         integer         float **Boolean**         decimal |  |  |  | | --- | --- | | |  | | --- | | **2.**(TCO B) Mary is told she is working at a dumb terminal (or network station). (Points : 5) |    She should be insulted that the firm thinks so little of her work         She has a menu system and a GUI         She is unable to backup her system from her terminal         **She is probably in a host/terminal environment**         Her station can store data, but she has to do her processing remotely |  |  |  | | --- | --- | | |  | | --- | | **3.**(TCO C) The purpose of normalization is to: (Points : 5) |    Reduce data redundancy         eliminate modification anomalies         Determine candidate keys         Identify dependencies **all of the above** |  |  |  | | --- | --- | | |  | | --- | | **4.**(TCO D) XML is: (Points : 5) |    **SGML compatible**         HTML compatible         SQL compatible         based on objects and cases         JavaScript compliant |  |  |  | | --- | --- | | |  | | --- | | **5.**(TCO E) To revoke update privileges on the DEPT table from Re you use theSQL command (Points : 5) |   **REVOKE ALL ON DEPT FROM Ren**         REVOKE SELECT, INSERT, DELETE ON DEPT FROM Ren         REVOKE INSERT, UPDATE, DELETE ON DEPT TO Ren         REVOKE INSERT, UPDATE, DELETE ON DEPT FROM Ren         ALTER USER Ren REVOKE INSERT, UPDATE, DELETE ON DEPT |  |  |  | | --- | --- | | |  | | --- | | **6.**(TCO E) What is used to undo unwanted database changes? (Points : 5) |    **backward recovery**         forward recovery         UNCOMMIT         reprocessing transaction         REVOKE |  |  |  | | --- | --- | | |  | | --- | | **7.**(TCO F) OLAP typically works with what type of data? (Points : 5) |    object oriented         direct         continuous         drill-down **multidimensional** |  |  |  | | --- | --- | | |  | | --- | | **8.**(TCO F) a web application is an example of (Points : 5) |    a universal standards         **a client/server system**         a content management system         a matrix organizational model         a Java applet |  |  |  | | --- | --- | | |  | | --- | | **9.**(TCO G) Users can update the files on your file server. This architecture should only be used when (Points : 5) |    your files have a lot of graphics or video in them         many people utilize the same files         you have a highly secured LAN         **files are small and not used often**         files are frequently exchanged |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | |  | | --- | | **10.**(TCO A) A data dictionary is used : (Points : 5) |    to access external databases         as a journaling facility **to track the items in an corporations databases and their components**         to resolve deadlocks         to check the format and spelling of data as they are entered   |  | | --- | | **1.**(TCO A) A primary key is composed of three attributes (i.e. three separate fields). Can one of these 3 fields be NULL as long as the other two are not NULL? State your reason why or why not. (Points : 20)        NO. All columns that are part of a table’s a primary key must contain aggregate unique values other than NULL.  The chances that the other two keys have simliar data down the line can be higher.  NULL indicates non-existence of value. When Primary Key contains NULL values it affects the performance as well as keys are not indexed due to NULL values of the column.  All the issues of the NULL operations are now attached to the primary key which is unique indication of the row.  Any operations on NULL results in NULL (besides ISNULL, COALESCE), JOIN clause on NULL values filters out all the NULL values. Index seek suffers due to NULL values in Primary Key. |  |  | | --- | | **2.**(TCO B) You are the company's head network system engineer. The CIO is investigating the feasibility of moving the current centralized database to distributed database architecture. Analyze the impact on the network infrastructure and write an impact statement for his review. Be sure to cover advantages, disadvantages, issues, and cost impacts. (Points : 20)        First a large geographically distributed database would complicate the data modeling analysis because the database customers would be from different locations making it harder to gather everyone’s requirements. Additionally, the customers would probably be from many different departments with many diverse and possibly competing requirements on how the database should be used.  Secondly we have to be concerned with  Data Integrity: Is the data be replicated successfully among the databases? Do we have any anomalies ? When we have a distributed system we are dealing with multiple copies of the data and maintenance is hard.  Security: Data security and integrity will be a cause and all systems need to be monitored. The chances of data corruption and intruding are higher.  Administration:: With the data residing in one location, adminstration over the data is simplistic. Now when we distribute the data, we have a copy of the data residing at multiple locations and we are having to add the additional cost to administrate the data. |  |  | | --- | | **3.**(TCO C) What is a domain? Why is it useful? (Points : 20)        A domain is a set of values all of the same data type. For example, the domain of nation name is the set of all possible nation names. The domain of all stock prices is the set of all currency values in, say, the range $0 to $10,000,000. You can think of a domain as all the legal values of an attribute.  The practical value of a domain is to define what comparisons are permissible. Only attributes drawn from the same domain should be compared; otherwise it is a bit like comparing bananas and strawberries. For example, it makes no sense to compare a stock’s PE ratio to its price. They do not measure the same thing; they belong to different domains. Although the domain concept is useful, it is rarely supported by relational model implementations. |  |  | | --- | | **4.**(TCO C) what are the different types of integrity constraints that can be enforced in the database and what is the purpose of each constraint?(Points : 20)  ntity Integrity: entity integrity rule, ensures that every row has a unique, non - null primary key.  referential integrity : you cannot define a foreign key without first defining its matching primary key  Domain Integrity: every element from a relation should respect the type and restrictions of its corresponding attribute  User Defined Integrity: intended to assert business structure or to control or influence the behavior of the business. E.g.: Age>=18 && Age<=60 |  |  | | --- | | **5.**(TCO C) What is normalization? List and explain at least three goals of normalization. (Points : 20) | | |

Normalization is a method for increasing the quality of database design. It is also a theoretical base for defining the properties of relations. The theory gradually developed to create an understanding of the desirable properties of a relation.

goal 1: eliminating redundant data (for example, storing the same data in more than one table)

goal 2: ensuring data dependencies make sense (only storing related data in a table).

goal 3. Perserve data integrity. Data definitions should remain the same through out the database accross tables.

|  |
| --- |
| **1.**(TCO A) Why does a DA need to have very good communication skills? (Points : 20)        Because unlike the DBA, a DA would be communicating with clients and business clients. The DA's job is to translate technical terms to business terms and gather business terms and translate them into technical terms. The DA therefore, serves as the bridge between the technical and business staff. |

|  |
| --- |
| **2.**(TCO D) Explain why XML would reduce network traffic given that XML documents can be up to 6 times larger than other formats. (Points : 20)  The introduction of XML will see a shift of processing from the server to the browser. At present, most processing has to be done by the server because that is where knowledge about the data is stored. The browser knows nothing about the data and therefore can only present but not process. However, when XML is implemented, the browser can take on processing that previously had to be handled by the server.  Also XML can be application specific and is much more compact than html. |

|  |
| --- |
| **3.**(TCO E) Analyze the different methods databases use to implement access controls in order to increase security in the database. Include in your discussion:  - accounts  - passwords  - privileges  - views  - roles (Points : 20)  let start with accounts. First of all not everyone needs an account to access the database. also not all accounts have the same priveleges. However, anyone who needs to have access to the db needs to have a seperate account so that we can easily track their behavior in the system.  with passwords: passwords are never shared or distributed they are kept confidential and each account has a different password.  with each account and password, the user only has certain privileges to access the dat a. some users may have access to all tables, but can only select, some may have read and write access and some may not have any access to any tables but can only have access to views.  Views are used to hide the data. the DBA creates views to only display what he or she wants the user to see. It also simplifies the selecting of data in the application side.  With roles, each user as a role. some have root admin role where they can do whatever they want to the db and some have a user roll where some operations are restricted. |

|  |
| --- |
| **4.**(TCO F) What are the significant differences between an OLTP database and a data warehouse? What are their purposes that make them so dissimilar. (Points : 20)  The OLTP database records transactions in real time and aims to automate clerical data entry processes of a business entity. Addition, modification and deletion of data in the OLTP database is essential and the semantics of the application used in the front end impact on the organization of the data in the database.  The data warehouse on the other hand does not cater to real time operational requirements of the enterprise. It is more a storehouse of current and historical data and may also contain data extracted from external data sources.  Data warehouse database: Designed for analysis of business measures by categories and attributes Optimized for bulk loads and large, complex, unpredictable queries that access many rows per table.  Loaded with consistent, valid data; requires no real time validation  Supports few concurrent users relative to OLTP Supports thousands of concurrent users. |

|  |
| --- |
| **5.**(TCO G) What is the most expensive part of accessing data, and what are some of the methods of reducing its performance impact? (Points : 20)  A DDBMS introduces a need for a data store containing details of the entire system. Information must be recorded about the structure and location of every database, table, row, and column and their possible replicas. The traditional system catalog is extended to include such information. The system catalog must also be distributed; otherwise, every request for information would have to be routed through some central point. For instance, if the systems catalog were stored in Tokyo, a query on Sydney data would first have to access the Tokyo-based catalog. This would create an expensive bottleneck.  The solution would be to have replication of tokyo's data in sydney.  However, replication is expensive because of the need to synchronize inserts, updates, and deletes across the replicated tables. When one table is altered, all the replicas must also be modified. A compromise is to use partial replication by duplicating the indexes only. This will increase the speed of query processing. The index can be processed locally, and then the required rows retrieved from the remote database. |