SAMPLE QUESTIONS FOR PhD QUALIFYING EXAMINATION

SE/CS 6329 Object-Oriented Software Engineering

System under Study:
The software to be designed will control a simulated automated teller machine (ATM) having a magnetic stripe reader for reading an ATM card, a customer console (keyboard and display) for interaction with the customer, a slot for depositing envelopes, a dispenser for cash (in multiples of $20), and a printer for printing customer receipts. The ATM will communicate with the bank's computer over an appropriate communication link. (The software on the latter is not part of the requirements for this problem.)
The ATM will service one customer at a time. A customer will be required to insert an ATM card and enter a personal identification number (PIN) - both of which will be sent to the bank for validation as part of each transaction. The customer will then be able to perform one or more transactions. The card will be retained in the machine until the customer indicates that he/she desires no further transactions, at which point it will be returned - except as noted below.
The ATM must be able to provide the following services to the customer:

- A customer must be able to make a cash withdrawal from any suitable account linked to the card, in multiples of $20.00. Approval must be obtained from the bank before cash is dispensed.
- A customer must be able to make a deposit to any account linked to the card, consisting of cash and/or checks in an envelope. The customer will enter the amount of the deposit into the ATM, subject to manual verification when the envelope is removed from the machine by an operator. Approval must be obtained from the bank before physically accepting the envelope.
- A customer must be able to make a transfer of money between any two accounts linked to the card.
- A customer must be able to make a balance inquiry of any account linked to the card.
- A customer must be able to abort a transaction in progress by pressing the Cancel key instead of responding to a request from the machine.

The ATM will communicate each transaction to the bank and obtain verification that it was allowed by the bank. Ordinarily, a transaction will be considered complete by the bank once it has been approved. In the case of a deposit, a second message will be sent to the bank indicating that the customer has deposited the envelope. (If the customer fails to deposit the envelope within the timeout period, or presses cancel instead, no second message will be sent to the bank and the deposit will not be credited to the customer.)
If the bank determines that the customer's PIN is invalid, the customer will be required to re-enter the PIN before a transaction can proceed. If the customer is unable to successfully enter the PIN after three tries, the card will be permanently retained by the machine, and the customer will have to contact the bank to get it back.
If a transaction fails for any reason other than an invalid PIN, the ATM will display an explanation of the problem, and will then ask the customer whether he/she wants to do another transaction.
The ATM will provide the customer with a printed receipt for each successful transaction, showing the date, time, machine location, type of transaction, account(s), amount, and ending

and available balance(s) of the affected account ("to" account for transfers).

Sample Questions

- Define a detailed (fully dressed) use case for one of the problem scenarios
- Define a high-level architecture diagram for the system
- Define a sequence diagram for the use case, using proper design patterns
- Define a design class diagram for the objects identified in the sequence diagram

CS 6352-Performance Computer Networks

1. Into a single queue with unlimited buffer size, customers **always** arrive in **pairs** but line up one after another in the waiting line for service. One of an arriving pair joins the line ahead of the other, by a random choice. The single server takes up one customer at a time from the head of the queue and serves. Each customer leaves immediately upon completion of his/her service (that is, without waiting for the completion of the partner's service). Arrivals of **pairs** are Poisson with a rate of 1 **pair** per hour. Service times of **individual customers** are independent and identically distributed Exponential random variables with a mean of 0.25 hour for each customer (and not for a pair). Determine the expected response time of an **individual customer.**

2. Consider a two-processor server where each processor in enabled with technology that allows the processors to operate at a higher speed under heavy loads, and to operate at a reduced speed under low loads to conserve power. Users submit jobs according to a Poisson process with rate jobs per second. If there is one job in the system, only a single processor is busy, and the processor operates at a speed of 500 MHz (500 million cycles per second). If there are two jobs in the system, both processors are busy, and each operates at a speed of 1000 MHz. If there are three or more jobs in the system, both processors are busy, and each processor operates at a speed of 1500 MHz. The length of a job is modeled as exponentially distributed and requires an average of 500 million CPU cycles. Buffer size is unlimited.

(a) Draw the state diagram for the system, clearly labeling transition rates.
(b) Find the steady-state probabilities for the number of jobs in the system.
(c) Find the condition for in order for the system to be stable.

CS6360 - Database Systems

Problem 1

Consider the following relations:

Student (<u>*snum: integer*</u>, sname:string, major:string, level:string, age:integer) Class (<u>*name:string,*</u> meets_at:time, room:string, fid:integer)
Enrolled (<u>*snum:integer, cname: string*</u>)
Faculty (<u>*fid:integer*</u>, fname:string, deptid:integer)

The meaning of these relations is straightforward; for example, Enrolled has one record per student-class pair such that the student is enrolled in the class.

Write the following queries in ***SQL.***

a. Find the age of the oldest student who is either a History major or enrolled in a course taught by I. Teach.

b. Find the names of all classes that either meet in room R128 or have five or more students enrolled.

c. Find the names of all students who are enrolled in two classes that meet at the same time.

d. Find the names of faculty members who teach in every room in which some class is taught.

e. Find the names of students who are enrolled in the maximum number of classes.

f. For each age value that appears in Students, find the level value that appears most often. For example, if there are more FR level students aged 18 than SR, JR, or SO students aged 18, you should print the pair (18, FR).

CS 6362-Software Architecture and Design

1. Consider the following declarations:

```
module M1
        provides: w, n
        requires: m
        consist-of: module M12, function F11

module M12
        provides: n;
        requires: m;
        has-access-to: module M2
        int n, Boolean m
end M12

function F11 provides:
        w; requires:
        n;
        boolean w, int n
end F11

module M2
        provides: m,q;
        int m, char q
end M2
```

1.1 List the set of varables that M12 can access

1.2 The specification above has one inconsistency. What is the inconsistency?

2. Suppose you are to build a system to help buyers and sellers, in a real estate market, who are interested only in the price and location of the real estate.

2.1 Depict a diagram of a software architecture in an implicit invocation style with control for the system.

2.2 Depict graphically how the architecture in 4.1 can be implemented using the Java Event Model.

CS 6363-Algorithms

1. Describe a polynomial time reduction from CLIQUE to VERTEX COVER. Prove that the reduction is correct*, i.e.* prove that x CLIQUE if and only if f(x) VERTEX COVER, where f is the function you define for the reduction.

2. Give a simple example of a directed graph with negative weight edges for which Dijkstra's algorithm produces incorrect answers. Explain.

3. Does the standard Ford-Fulkerson network algorithm for computing maximum flow from a given source to a given sink, where the edges have capacity constraints, allow an arbitrary selection for an augmenting path from the source to the sink in each successive residual network and still manage to guarantee eventual termination with maximum flow? Explain why or why not.

CS 6364-Artificial Intelligence

1. The knowledge base KB has one Boolean formula:

(x1 x2) (x2 x3)

Find an equivalent knowledge base in Conjunctive Normal Form (CNF).

2. Consider a student that passes the PhD qualifiers test on AI, and the PhD qualifiers test on ALGORITHMS. Use Naïve Bayesian reasoning to decide if the student would produce a good dissertation based on the following information:

1. 10% of all students can produce a good dissertation.
2. Among students that produce good dissertations 90% can pass the AI qualifiers, and 80% can pass the ALGORITHMS qualifiers.
3. 50% of the students that cannot produce a good dissertation are able to pass the AI qualifiers test, and 30% are able to pass the ALGORITHMS qualifiers test.

Explain your conclusion.

CS 6367-Software Testing

1. Given a complicated software system to be tested, assume you have all the source code as well as the final executable system. Assume also you have some documentation even though it might be incomplete and/or out-of-date.

Assume that testing techniques such as equivalence partitioning, boundary-value analysis, control flow-based testing (e.g., branch testing), dataflow-based testing (e.g. "all-uses"), fault-based testing are available. Make a suggestion on how to conduct unit testing, integration testing, and system testing so that the selected testing techniques complement each other to provide a **cost-effective** approach for detecting as many faults as possible.

Justify your answer and explain all the assumptions you make, it any.

2. (a) Given the pseudo-code below:

```
integer x,y;
read x;
read y;
x = abs(x) +
2; z = 3;
if (even(x/y)){
   for( int i = 0 ; i < (x+y) ;
      i++){ if ((i mod x) == 0)
         z = z * x ;
      else
```

```
              z = z * y;
        }
        print z
       }
     print x,y
```

1) draw a control flow graph showing all the def, c-use, p-use actions.
2) construct the set of items to be covered for "all-uses" (i.e., the

   (def,c-use), (def, p-use) pairs).
3) Ensuring that each test increases the coverage, create test cases to satisfy the "all-uses" strategy (explicitly state each def-p-use, def-c-use pair covered by each test case).

(b) An exception is a signal that indicates that some sort of exceptional condition (e.g., a run-time error) has occurred. When an exception occurs, an "exception-handler" is invoked to handle the exception. Explain if, how the possibility of exceptions would affect the way you carry-out "all-uses".

CS 6375-Machine Learning

Naïve Bayes Learning

Table below indicates training example for the target concept *Play Tennis.*

Day	*Outlook*	*Temperature*	*Humidity*	*Wind*	*PlayTennis*
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

Use Naïve Bayes Learning Algorithm to decide how the new instance should be classified. Show all the computation steps.

	Outlook	Temp	Humidity	Wind	
<	Sunny	Cool	Normal	Strong	>

CS 6378-Advanced Operating Systems

Question 1:

(a) The Schiper-Eggli-Sandoz protocol for causally ordered message delivery states that the communication channels can be non-FIFO. However, the description does not state anything about the reliability of channels. What will be the impact of message losses on the operation of the Schiper-Eggli-Sandoz protocol? Provide explanations to support your answer.

(b) Assume a system of N nodes. If all communication is broadcast (all nodes are destinations of every message) is it still necessary for every message to carry $O(N^2)$ dependency information? Would it be possible to ensure causally ordered delivery while having each message carry less than $O(N^2)$ dependency information? If yes, how much dependency information should each message carry and why? Explain your answer.

Question 2:

(a) The three-phase commit protocol is said to be non-blocking and resilient to single site failure. This means that inspite of the failure of a single site, all the operational sites agree on the outcome of the transaction by examining their local states. Is the three-phase commit protocol resilient to the failure of more than one site? If so, prove your answer. If not, provide a counter-example to support your answer.

(b) The state transition and the resultant sending of message(s) in the three-phase commit protocol is atomic (either the message is sent to all the intended recipients or to none). Would the correctness of the protocol be affected by relaxing the atomicity requirement? Explain.

CS 6390-Advanced Computer Networks

1. **Network Calculus**

1. Let A be the cumulative arrival function into a system, B be the cumulative departure function of the system. Give the definition of a service curve, i.e., if is the service curve, express (*t*) as a function of A(*t*) and B(*t*).

2. Assume you have two systems S1 and S2 concatenated to each other, where 1 is the service curve of S1 and 2 is the service curve of S2. What is the service curve of the entire system S1; S2? (i.e., the service curve of S1 concentrated with S2).

2. **FCFS**

What is the importance of the delay theorem for FCFS by Chlamtac, Farago and Fumagalli (the parper entitled "A Deterministic Approach to the End-to-End Analysis of Packet Flows in Connection-Oriented Networks") as opposed to the original results of FCFS multiplexors by R. L. Cruz (paper title "A Calculus for Network Delay Part I and II")? i.e., how does the former improve upon the results of the latter?

3. What is quality of service (QoS)?
Give an example of QoS defintion in ATM network;

4. Describe and compare integrated service architecture (ISA) and differentiated service (Diffserv) model.

5a. Assume that a router in the path between two hosts communicating using TCP can sniff all the IP packets and send them to a distant host C realiably. If host A is downloading a file f from B, how can C get the entire contents of the file?

5b. If the sequence number field in TCP is to be increased to 64 bits and the data rate of the communication path between the two end hosts is 100 Mbps, how long does it take for the sequence number to wrap around?

6. Route optimization is mobile IP is used when we want to avoid the inefficient routing of packets from the source to the home agent and then tunneling the packets from the home agent to the destination host at the foreign network. Explain how this is done and what are the disadvantages of this scheme?