

**UNIVERSITY OF DUBLIN
TRINITY COLLEGE**

**Faculty of Engineering, Mathematics & Science
School of Computer Science and Statistics**

B.A. (Computer Science)

Trinity Term 2011

SENIOR FRESHMAN EXAMINATION

CS2031 & CS2041

Telecommunications and Information Management

Saturday 7 May

Examination Hall

14.00 – 17.00

Declan O'Sullivan, Stefan Weber

Answer **two** questions from each section.

All questions carry equal marks (25 marks).

Use separate answer books for each QUESTION.

Materials permitted for this examination:

Calculator (non-programmable)

SECTION A**Question 1**

The data link layer is responsible for the communication between two stations that are connected by a physical medium. Data compression, error detection and error correction are part of the data link layer.

(1a) Hamming Code, Parity and Cyclic Redundancy Checksum (CRC) and Checksums are four common methods of error detection. Analyse the following 3 cases of communication over a given link:

I) 7000 bytes are to be transferred over a serial line between two computers. The program that transmits the data, issues them in units of 7 bit. The errors in this communication link are limited to individual bits.

II) 1000 bytes are to be transferred from a computer to a switch over an Ethernet. The errors on the link are very rare in comparison to other types of links; but if errors occur, they result in the corruption of a number of subsequent bits, rather than the corruption of a single bit.

III) 1000 bytes are to be transferred between a base station on earth and a satellite. The errors in this case are generally limited to single bits. The retransmission of data is relatively expensive compared to communication over other types of links because the large transmission delays between the base station and the satellite.

For each of the 3 cases above: Contrast the suitability of the methods for error correction listed and suggest the most suitable method. Consider in your answer characteristics such as processing efforts required by the methods and ability of the individual methods to detect different types of errors.

(12 marks)

Question 1 is continued on the next page

Question 1 continued from the previous page

- (1b)** Demonstrate the encoding and decoding of the following 7 bits using Hamming code: 1011011 and explain how the result of the decoding may be used by a receiver.

(6 marks)

- (1c)** Huffman encoding uses statistical analysis of texts to determine the optimal codes for individual symbols. Table 1 lists the distribution of symbols from a text. Give the binary tree for Huffman encoding as a diagram and use this binary tree to determine the code for the following sequence of symbols: "communication".

Symbol	Count
C	2
O	6
T	3
I	8
M	7
N	4
A	3
U	2

Table 1: Distribution of symbols

(7 marks)

(Total 25 marks)

Question 2

The communication between two nodes on possibly different networks is governed by network layer protocols such as the Internet Protocol (IP) and by routing protocols such as Open Shortest Path First (OSPF).

(2a) The Internet Protocol is used to provide connectivity between nodes that are connected by a variety of physical networks. Assume that node A wants to transfer a packet of 5000 bytes to node B. The network that node A is connected to has a maximum transfer unit (MTU) of 4000 bytes. Router C connects this network to another network that has a MTU of 1500 bytes. Router D connects this network to yet another network that has a MTU of 3000. Router E in turn connects this network to the network on which node B is located. The network that router B is connected has a MTU of 1500 bytes.

Analyse the behaviour of IP in terms of the transport of packets over networks with varying MTUs in the given scenario, highlight the advantages and disadvantages of the approach taken by IP and propose an alternative approach. Your analysis should be accompanied by a diagram that visualises the infrastructure and the transmission of the data between node A and node B.

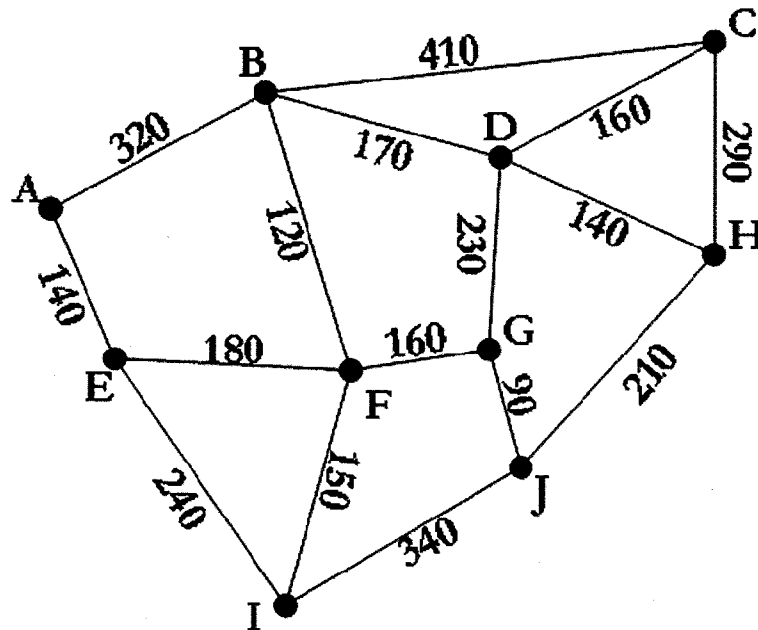
(13 marks)

Question 2 is continued on the next page

Question 2 continued from the previous page

(2b) Link State Routing (LSR) represents one of the major routing concepts. Describe the LSR concept in your own words. The description should be accompanied by diagrams where appropriate.

Dijkstra's Shortest Path algorithm is used in LSR to determine the routing table of individual nodes. Describe this algorithm and explain it with the help of the following diagram. Contrast the approach of LSR to Distance Vector Routing.



(12 marks)

(Total 25 marks)

Question 3

The Transport Control Protocol (TCP) is widely used to transfer data between computers. In order to fulfil this task efficiently a number of mechanisms have been developed that adapt TCP to limitations of networks and computers.

(3a) 4000 bytes are transferred between two hosts, A and B. In one case the Transport Control Protocol (TCP) is used for this transfer and in a second case the User Datagram Protocol (UDP) is used. Describe the packets that are involved in the transfer of the data in each case and compare the overhead that is involved in the transfer. The description of the communication should include a description of the reaction to transmission errors. Discuss the advantage and disadvantage of each protocol using the transfer of the 4000 bytes as an example. Use diagrams to visualise the communication for each case.

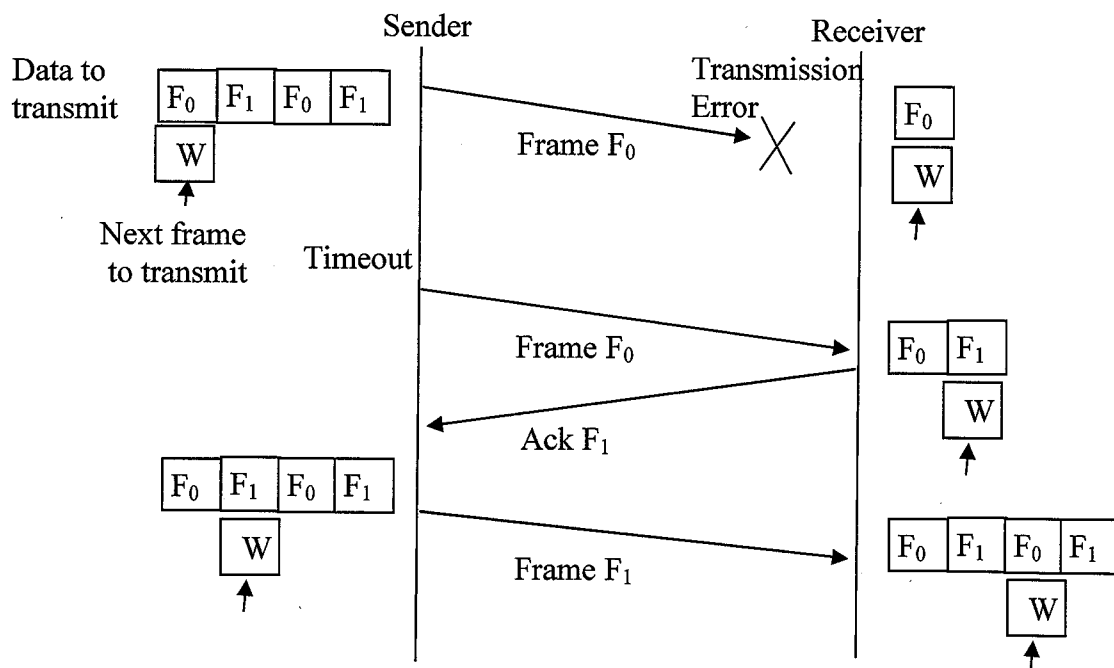
(13 marks)

Question 3 is continued on the next page

Question 3 continued from the previous page

(3b) The terms “Stop-and-Wait ARQ”, “Go-Back-N ARQ” and “Selective Repeat ARQ” specify well-known Flow and Error Control mechanisms. Following the example of the description of “Stop-and-Wait ARQ” below, explain the “Go-Back-N ARQ” and “Selective Repeat ARQ” mechanisms in turn, including the state that needs to be kept at the sender and receiver side of a transfer, the reaction of each mechanism to transmission errors in the transfer of the various packets and the advantages and disadvantages of each mechanism. The explanation of each mechanism should be accompanied by diagrams that visualise the transfer of the packets.

Stop-and-Wait ARQ:



The sender sends a frame at a time, waiting for an acknowledgement before progressing to send the next frame. After sending a frame, the sender starts a timer. An acknowledgement may be delayed or lost or the original frame may be corrupted due to a transmission error. The sender resends the frame, if the timer expires before an acknowledgement is received. Each, the sender and the receiver, have a transmission window of one frame.

(12 marks)

(Total 25 marks)

Question 4

Medium access control is used to coordinate the communication over a shared medium. A number of protocols have been developed that provide various degrees of flexibility in the access to the shared medium.

- (4a)** An access point coordinates the communication of 4 laptops using the Point Coordination Function (PCF) of 802.11. Assume that 2 laptops have data to transfer and 2 laptops receive data from the access point.

Describe the frames that are exchanged by the access point and the laptops and the inter frame spaces that are involved in this exchange. Use diagrams to visualise the chronological exchange of the frames and the inter frame spaces.

(13 marks)

- (4b)** Assume that four stations use a wired medium to communicate. All stations may intend to transmit data at the same time as all other stations. The access to the medium is controlled by a Carrier Sense Multiple Access (CSMA) scheme with Collision Detection (CD) or a Time Division Multiple Access (TDMA) scheme with a reservation protocol. For the CSMA/CD approach, assume times as you see fit. For the TDMA approach, assume that every reservation slot is 10ms long and every contention free slot is 100ms. Contrast the two access control schemes with each other considering 2 scenarios:

- I. High network load i.e. every station would like to send data at any given time.
- II. Low network load i.e. stations only have data to be transmitted every few seconds.

Your analysis should be accompanied by diagrams that visualise the behaviour of the two schemes in each of the scenarios.

(12 marks)

(Total 25 marks)

SECTION B

SOME QUESTIONS REFER TO THE FOLLOWING
FIGURE A:

```
<?xml version="1.0" encoding="ISO-8859-1"?>

<purchaseOrder>
  <shipTo exportCode="1" type="UKAddress">
    <name>Helen Zoe</name>
    <street>47 Eden Street</street>
    <city>Cambridge</city>
    <postcode>CB1 1JR</postcode>
  </shipTo>

  <billTo type="Address">
    <name>Robert Smith</name>
    <street>8 Oak Avenue</street>
    <city>Old Town</city>
    <county>cork</county>
    <zip>95819</zip>
  </billTo>

  <items>

    <item partNum="833-AA" >
      <productName>Lapis necklace</productName>
      <quantity>1</quantity>
      <EuroPrice>99.95</EuroPrice>
      <comment>Want this for the holidays!</comment>
      <comment>Must send by express</comment>
      <shipDate>2006-01-10</shipDate>
    </item>

    <item partNum="123-BA" manufacturer="Duracell">
      <productName>Batteries</productName>
      <quantity>10</quantity>
      <shipDate>2006-01-10</shipDate>
      <description>
        <colour>black</colour>
        <material>metal</material>
      </description>
    </item>

  </items>
</purchaseOrder>
```

Question 5

- a) Describe in detail what is meant by a “Solid State Drive” and name two advantages they have over “Hard Disk Drives”.

(4 marks)

- b) Describe the parts (hardware and software) of a computer involved in moving a piece of data from an application and encoding it onto a hard disk. In other words describe “the journey of a byte” during the **writing** process.

(8 marks)

- c) Consider a hard disk with the following characteristics: block size $B = 512$ bytes, interblock gap size $G = 128$ bytes, number of blocks per track=25, number of tracks per surface=400. A disk pack consists of 15 double sided disks.

- i. What is the total capacity of a track and what is its useful capacity (excluding interblock gaps)?
- ii. What is the total capacity and useful capacity of a cylinder?
- iii. What is the total capacity and the useful capacity of a disk pack?
- iv. Suppose an average seek time is 20 msec, average rotational delay of 10msec, and block transfer time of 1 msec. How much time does it take (on average) in msec to locate and transfer a single block given its block address?
- v. Calculate the average time it would take to transfer 20 random blocks and compare it with the time it would take to transfer 20 consecutive blocks using double buffering to save seek time and rotational delay.
- vi. Assuming 500 contiguous blocks make up a sequential file..Calculate the amount of time needed to do a binary search on the file.

(13 marks)

(Total 25 marks)

Question 6

- a) What is XML Schema and what is it used for? Describe at least two of the statements in XML Schema that have no equivalence in DTDs.

(4 marks)

- b) Give two reasons for using the “entity” declaration in a DTD. From the evidence given in the XML document shown in Figure A, propose a DTD for the document. Explain your design decisions.

(8 marks)

- c) Describe what is meant when describing XQuery as consisting of “FLWOR” expressions. In addition, define and explain XQuery statements for each of the following queries posed over the document in Figure A. Also show expected results and explain your design decisions.

- i) Return the <quantity> element for each <item> element.
- ii) Return in one element called <commentary>, all the <comment> elements.
- iii) For each <item>, create an "item summary" element. The content of the item summary is the values of all the sub elements of the <item> element and the attribute values, all separated by “;”.
- iv) Return all the <purchaseOrder> elements where one of the subelements contain an attribute called “type” and the value of that attribute contains the word “Address”.

(13 marks)

(Total 25 marks)

Question 7

a) Describe why hashing techniques have limitations with respect to indexing.

(4 marks)

b) Describe what is a B-Tree and explain the rules involved in **inserting** a value into a node.

(8 marks)

c) Given the B-Tree in Figure B (with $m=5$), show the growth of the B-Tree as you add the keys in the following order

16, 18, 21, 3, 22, 25, 14, 30, 35, 42, 53, 23, 55, 67, 70

- Show the tree at each stage and provide an explanation for any splits
- What is the average number of node splits per new key inserted?
- What is the average search time?

(13 marks)

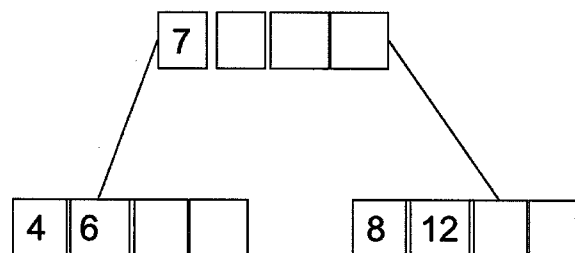


FIGURE B

(Total 25 marks)

Question 8

The introduction of the "semantic web" will radically change the effectiveness of searching on the web. Discuss. Include at least in your answer: how Information Retrieval on the web is currently organised; a diagram and description of the preprocessing that is involved in creating an inverted file index; how effectiveness of information retrieval is measured; the idea behind the "semantic web" itself; the role of the XML RDF vocabulary and illustrate its usage; outline at least three advantages of the semantic web and three obstacles to the introduction of the semantic web.

(Total 25 marks)