

UNIVERSITY OF DUBLIN

TRINITY COLLEGE

Faculty of Engineering, Mathematics and Science

School of Computer Science & Statistics

B.A. (Computer Science)
SENIOR FRESHMAN EXAMINATION

2BA5

Telecommunications and Information Management

Thursday May 22 2008

Exam Hall

09:30 – 12:30

John Keeney and Stefan Weber

Answer **five** questions, at least **two** from each section.

Use separate answer books for each section.

Section A

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.

- (a) 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) Several data units – each consisting of 1000 bytes – are to be transferred from a computer to a switch over an Ethernet connection. 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 single bits.

ii) 1400 bytes are to be transferred between a base station on earth and a satellite in orbit around earth. 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 due to the large transmission delay between the base station and the satellite.

For each of the 2 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.

(10 marks)

- (b) Flow control is used to prevent a sender from overflowing a receiver while aiming to maximize the use of the available bandwidth. Contrast a protocol using Stop-And-Wait ARQ against a protocol using Selective Repeat ARQ. Your analysis should examine factors such as requirements on the terminal, considerations in the design of the protocol header and the throughput of the protocol.

(6 marks)

- (c) Demonstrate the process of LZ77 encoding on the following text:
"To be, or not to be (from Hamlet)"

(4 marks)

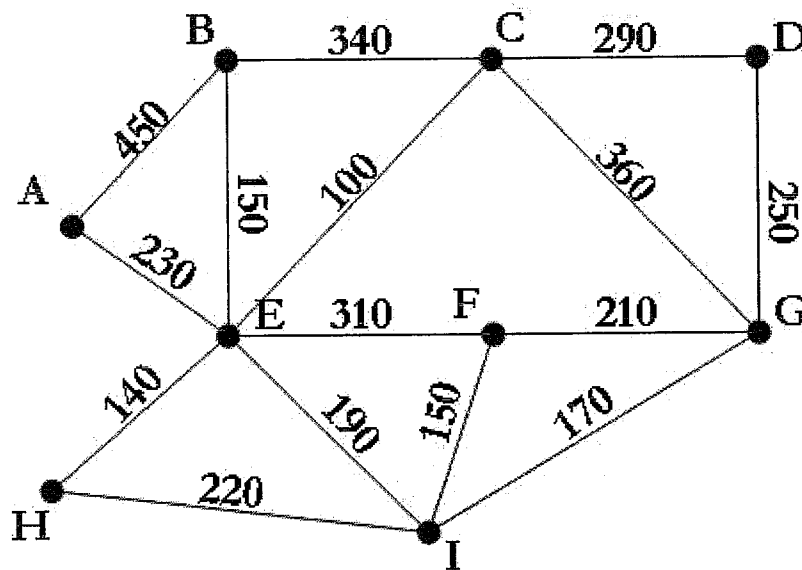
(Total 20 marks)

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).

(a) The initial use of IP addresses suggested the use of classful addresses. During the mid 90s, it was feared that by using classful addresses, the number of IP addresses would not be sufficient for the number of computers on the Internet. Explain the cause for the shortage of IP addresses and describe 2 techniques that were conceived to resolve this problem.

(8 marks)

(b) Describe and contrast the two concepts of distance vector routing and link state routing using the network graph shown below. The description should include diagrams that visualise the process that each concept follows to establish routing tables.



(8 marks)

(c) Describe what is understood under the term “count-to-infinity”. The description of the term should be accompanied by a diagram.

(4 marks)

(Total 20 marks)

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 adopt TCP to limitations of networks and computers.

- (a) A web browser on computer A requests a web page from a web server on computer B. Explain the life cycle of a TCP connection including connection establishment, data exchange and connection termination. Consider in your explanation the loss of packets and the reaction of the TCP protocol on these losses.

The explanation should include a diagram with a list of segments that are exchanged between the client and the server and explanation of the fields of the TCP header that are used during this exchange. A figure of the TCP header is given below in figure 1.

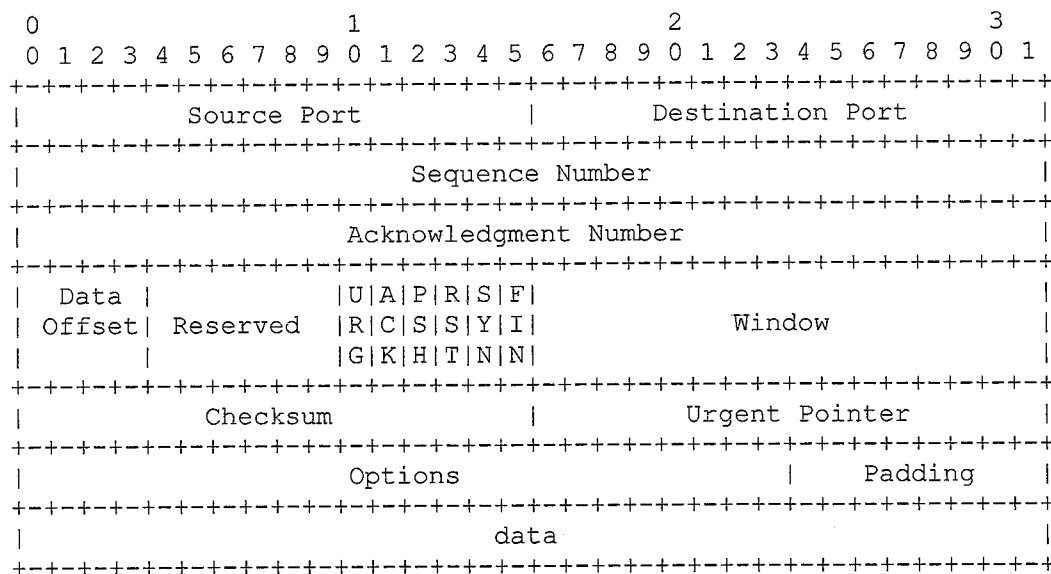


Figure 1: TCP Header Format

(12 marks)

- (b) A TCP connection uses Additive Increase/Multiplicative Decrease (AIMD) for congestion control. The link over which the connection is established has a roundtrip time of 20msec and no congestion at the beginning. Note that in the following: Each graph should be accompanied by a description that gives an analysis of the respective congestion control algorithm.

i) Draw a graph that shows the progression of the AIMD algorithm from 0 to 500msec with timeouts for packets occurring at 200msec, 300msec, and 440msec.

ii) Consider the TCP slow start algorithm with fast recovery for the connection described above and draw a graph for the progress of this algorithm from 0 to 500msec with timeouts for packets occurring at 100msec, 200msec, 300msec, 380msec and 460msec.

(8 marks)

(Total 20 marks)

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.

- (a) Assume that six terminals intend to communicate over a medium. Each station will want to communicate with every other station at some time. Suggest 2 schemes to allocate the medium to the individual stations and discuss the advantages and disadvantages of each scheme with respect to the scenario described above. At least one scheme should allocate the access to the medium dynamically and at least one scheme should use a fixed allocation of the medium.

Your analysis should be accompanied by diagrams that visualize the behaviour of each of the three schemes.

(9 marks)

- (b) Explain the terms "hidden terminal problem" and "exposed terminal problem". Each explanation should be accompanied by a diagram that visualises the term. How does the design of the 802.11 MAC layer protocol address these problems?

(7 marks)

- (c) Explain the term "Carrier Sense Multiple Access (CSMA)" and its two variants Collision Avoidance (CA) and Collision Detection (CD).

(4 marks)

(Total 20 marks)

Section B

5.

- (a) Compare B-Trees to B+Trees. In particular mention the different node structures of the two tree types.

(4 Marks)

- (b) Describe the concept of fan-out (fo) in relation to B-Trees. In particular make reference to the relationship between fan-out and blocking factor (bfr) and how fan-out is related to the search time of a B-Tree.

(6 Marks)

- (c) A PARTS file with Part# as key field includes records with the following Part# values:

23, 65, 37, 60, 46, 92, 48, 71, 56, 59, 18, 21, 10, 74, 78, 15, 16, 20, 24, 28

Suppose the search field values are inserted in the given order into a **B+Tree** of order $m=4$. Show how the tree will expand and what the final tree will look like.

(10 Marks)

(Total 20 marks)

6. A simple hashing algorithm for alpha-numeric keys may be described in the following three stages –

1. Represent the key in numerical form
2. Fold and add
3. Divide by the size of the address space and use the remainder as an address

- (a) Using the simple hashing function described above show how a key of value TRINITY may be hashed into an index of size 13. The ASCII key length is 12 characters. What position in the index does the key hash to?

NOTE The ASCII value of 'A' is 65; The ASCII value of a blank space is 32.

(5 Marks)

- (b) Describe the advantages and disadvantages of hashing techniques by comparison to indexing.

(5 Marks)

- (c) Hashing techniques invariable suffer from collisions, i.e. where two or more keys hash to the same location in the index. Briefly describe how each of the following techniques may be used to resolve collisions.

- (i) Buckets
- (ii) Linear Probing
- (iii) Double Hashing

(10 Marks)

(Total 20 marks)

7.

(a) You have 4 documents ($D_1 - D_4$) that contain the following index terms:

D_1 : alpha bravo charlie delta echo foxtrot golf

D_2 : golf golf golf delta alpha

D_3 : bravo charlie bravo echo foxtrot bravo

D_4 : foxtrot alpha alpha golf golf delta

- (i) Use an incidence matrix of terms to calculate a similarity matrix for these four documents, with no term weighting.
- (ii) Use a frequency matrix of terms to calculate a similarity matrix *with weights* for these documents. Weights are proportional to the term frequency and inversely proportional to the document frequency.

(8 Marks)

(b) Define and describe the terms recall and precision.

(6 Marks)

(c) Q is a query. D is a LARGE collection of 1,000,000 documents.
When the query Q is run, a set of 200 documents is returned.

- (i) How in a practical experiment would you calculate the precision?
- (ii) How in a practical experiment would you calculate the recall?
- (iii) Suppose that, by some means, it is known that 100 of the documents in D are relevant to Q . Of the 200 documents returned by the search, 50 are relevant.
What is the precision?
What is the recall?

(6 Marks)

(Total 20 marks)

8.

- (a) What is an XML vocabulary? (2 marks)
- (b) Mention the similarities and differences of XML DTDs and XML Schemas. From the evidence given in the XML document "*PurchaseOrder.xml*" shown in Figure X, propose a DTD for the document. Explain your design decisions. (6 marks)
- (c) An XPath expression is made up of a series of "location steps", where each individual location step is separated by either a "/" or "//". In general each individual location step has three parts. Describe the three parts referred to. (4 marks)
- (d) Provide an XPath expression that represents the following queries upon the XML document "*PurchaseOrder.xml*" shown below in figure 2. Explain the rationale for each XPath expression used.
- (i) Retrieve all streets from ShipTo elements;
 - (ii) Retrieve all items where quantity ordered is greater than 1;
 - (iii) Retrieve the second comment from the first item;

(8 marks)

(Total 20 marks)


```

<?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>

```

Figure 2: PurchaseOrder.xml