[C054/SQP176]

Advanced Higher Information Systems

Specimen Question Paper

Time: 2 hours

NATIONAL QUALIFICATIONS

Attempt both questions in Section I.

Choose **one** of the parts (Part A **or** Part B **or** Part C) in Section II and attempt all of the questions in your chosen part.

Read all questions carefully.

Write your answers in the answer book provided. Do not write on the question paper.

Use a separate page for each answer.

Marks

SECTION I

Attempt both questions in this section.

1. A conference is being organised. It is divided into sessions. At each session several papers are presented. The papers are written by authors. Several authors may share in the writing of one paper but only one paper can be submitted from any given author. Each paper has a presenter who is one of the authors of the paper.

The organiser of the conference has devised two forms for recording data about the authors and the papers. One form relates to an author and one form relates to an author's paper. Each author is allocated a unique number and in addition this form shows the author's personal details (full name, sex and job title) and details of the author's employer (name, address and phone number). Each paper is allocated a unique number and in addition this form shows the details of the paper being presented (paper title, subject category, length in pages and date paper received), session details (session date, session time and session venue) and the author number of the presenting author.

You may assume that there are the following dependencies between the items:

the personal details are functionally dependent on the unique author number;

the employer address and phone number are functionally dependent on employer name;

the paper details are functionally dependent on paper number;

the session details are functionally dependent on the combination of session date and session time.

(*a*) Explain why the two primary entities, Author and Paper, are already in second normal form (2NF). **1**

(*b*) Complete the normalisation of these entities by producing entities in third normal form (3NF). **4**

(*c*) Construct an entity relationship diagram for these entities. Indicate the nature of each relationship (such as many-to-one) between the 3NF entities. **2**

(*d*) A rule change permits a single author to submit more than one paper. Modify your entity relationship diagram to reflect this change. State any changes to the attributes of existing entities and include any new entities that this change might cause. Indicate the nature of each relationship. **3**

(10)

Marks

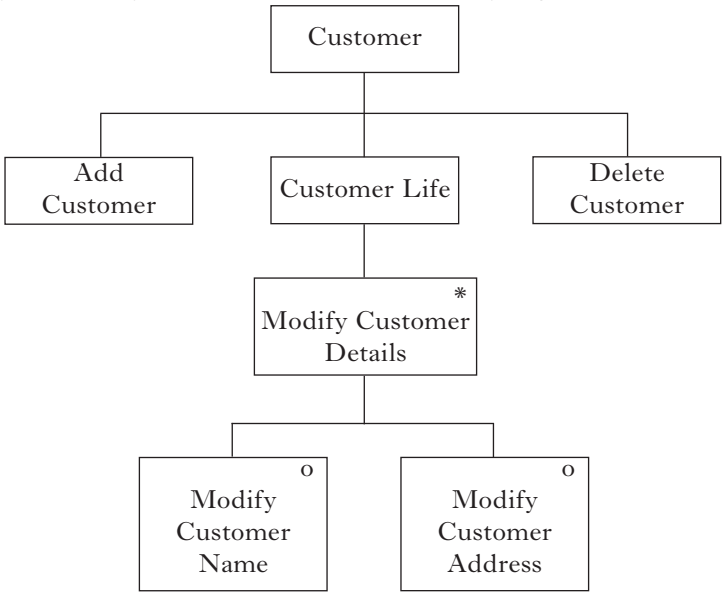
2. (*a*) The following table is an excerpt from a data dictionary that contains errors. Identify **four** errors in this data dictionary and explain each error.

Item	**Entity**	**Description**	**Type/ Size**	**Range/ Validation**	**Required**	**Key**
Athlete_Nbr	Athlete	Unique number for athlete	A(6)	999999	Y	Y
Name	Athlete	Full name of athlete	A(40)		Y	N
Address	Athlete	Address of athlete	A(100)		N	N
Athlete_Club_No	Athlete	Number of club in which athlete trains	A(4)	Existing Club_Nbr	Y	N
Club_Nbr	Club	Unique number of club	Number	In 0..999	N	Y
Name	Club	Name of club	A(32)		Y	N
Club_Captain_No	Club	Athlete_Nbr of captain of club	A(6)	Existing Athlete_Nbr	Y	N

4

2. (continued)

(*b*) The entity life history (ELH) of the Customer entity is given below.



(i) Classify the following sequences as **legal** or **illegal** according to the diagram and for any illegal sequence explain why the sequence is illegal.

(1)	(2)	(3)	(4)
Add Customer Delete Customer	Add Customer Customer Life Delete Customer	Modify Customer Address Modify Customer Address Delete Customer	Add Customer Modify Customer Name Modify Customer Address Delete Customer

(ii) Show how to change the ELH if there are two different events that can create a new customer: Add By Postal Enquiry, Add by Electronic Enquiry. **4**

(*c*) An existing information system is currently based on paper records. A decision is taken to computerise the system using a database. At what stage should the decision be taken as to which database product to use? Justify your answer. **2**

(10)

Total for section: 20

[*END OF SECTION I*]

Marks

SECTION II

Attempt Part A or Part B or Part C

Part A—Multimedia

Answer ALL of the questions in this part.

3. You have been asked to create a multimedia product for use in a primary school by pupils aged 7–10 years old. The product should reinforce the pupils' knowledge of fractions through the use of graphics, text, animation and sound. The product will provide graded exercises that test skills such as fraction recognition and fraction equivalences.

(*a*) Describe the proposed target audience. Your description should identify the requirements of the audience and their likely computing experience. **2**

(*b*) Suggest an appropriate medium for this product. Justify your choice. **2**

(*c*) The specification of this multimedia product must clearly define the *structure* of the product and identify any *interactivity* within the product.

How would the product specification define the structure of the product and identify any interactivity within the product? **3**

(*d*) Explain what use is made of *storyboards* in the production of a multimedia product. **3**

(10)

4. When a multimedia product is being developed, many important decisions are made at the planning stage.

(*a*) (i) Explain how decisions made at the planning stage can affect the machine specification for the planned product.

(ii) Describe the information that would be specified in the plan for the development of the product. **7**

(*b*) Once the multimedia product has been assembled, the development team complete a product evaluation.

What factors would be considered by the development team when evaluating their multimedia product? **3**

(10)

5. You have produced a multimedia CD-ROM for local businesses to train employees on health and safety issues.

(*a*) After producing the package you discover that some companies have a lower specification of equipment than originally envisaged. How could you adapt your package to deal with the following situations?

(i) The machines have a slower processor, less memory and a lower capacity hard disk than originally anticipated.

(ii) The machines have smaller monitors than originally anticipated with lower graphics resolution. **4**

(*b*) You consider transferring the product from CD-ROM to the World Wide Web. Explain how this change in medium might affect the product's content and structure. **2**

(*c*) Once the product has been developed it has to be tested before being released to customers. Outline an appropriate test strategy for this multimedia product. **4**

(10)

[*END OF PART A*] **Total for section: 30**

Marks

Part B—Natural Language Processing

Answer ALL of the questions in this part.

6. Natural language processing techniques are being used increasingly in the development of:

machine translation systems

text analysis systems

question answering systems

command and control systems.

(*a*) Outline **two** main goals of any natural language processing system. **2**

(*b*) Provide a description of any **two** of the applications of natural language processing techniques mentioned above. **2**

(*c*) *Parsing* is a technique commonly used in natural language processing.

What is the purpose of parsing? Outline how parsing is accomplished. **2**

(*d*) (i) Many modern database systems make use of natural language techniques other than parsing.

Provide a description of an alternative technique to parsing. Your description should indicate how the technique is applied in database systems. **2**

(ii) Outline the implications of natural language processing techniques in database systems in terms of user productivity and development costs. **2**

(10)

7. (*a*) Consider the following newspaper headline:

Giant waves in tunnel.

Now consider the sentence below:

"I cooked the cabbage in the pot and then I ate it."

Discuss any problems of *ambiguity* illustrated in each of these examples. **2**

(*b*) Specify a grammar that represents a subset of English that has the following constituents and lexicon.

Constituents
sentence
nounphrase
verbphrase

Lexicon	
noun	cat, mouse, cheese, pear
propernoun	David, Kirsty
verb	ate, saw, caught, sketched
determiner	the, a, an
adverb	slowly, quickly

Examples of sentences that your grammar should be able to generate are the following.

The mouse slowly ate the cheese.

Kirsty saw the cat.

The cat caught a mouse.

David quickly sketched a cat. **3**

Marks

7. (continued)

(*c*) Specify **two** test sentences, in addition to those given above, for your grammar. You should ensure that one sentence would succeed when parsed by your grammar and one sentence should fail. Indicate the expected success or failure in each case. **2**

(*d*) Parse the sentence below, showing the full grammatical structure of the sentence. Use bracketing to indicate the structure.

"The mouse slowly ate the cheese." **2**

(*e*) Extend your grammar to include adjectives. **1**

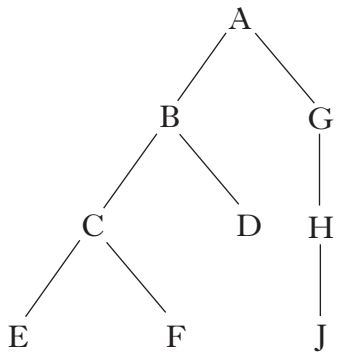
(10)

8. (*a*) State the difference between top-down and bottom-up parsing strategies.

Which one is the better strategy if the text contains many words like "can" and "bank" which are both nouns and verbs? Justify your answer. **3**

(*b*) Explain why it is sometimes necessary to search a parse tree. **2**

(*c*) Describe the methods for performing depth-first search and breadth-first search. State the order in which the nodes are searched using each method on the following tree.



3

(*d*) Give a brief description of a natural language technique that does not depend on parsing using a grammar. **2**

(10)

Total for section: 30

[*END OF PART B*]

Marks

Part C—Systems Analysis & Design

Answer ALL of the questions in this part.

9. You have been asked to analyse the paper-based system your school currently uses to keep records of its science laboratory equipment and put forward proposals for a computerised system to assist in this task.

(*a*) Describe the **five** stages involved in the analysis and design process. Relate these stages to the proposed laboratory system. **4**

(*b*) Explain why the stages in the analysis and design process do not always follow a simple sequence. **1**

(*c*) Explain the difference between the *logical* design and the *physical* design of an information system. Relate your answer to the laboratory system. **2**

(*d*) What is a *system specification*? Give **three** examples of what might be included in the system specification for the laboratory system. **3**

(10)

10. Your local video store has asked you to analyse its existing manual system with a view to designing a new computerised system to record issues and returns.

(*a*) Describe **three** methods you might use to obtain information about the existing system. Give **one** disadvantage of each of your methods of obtaining information. **6**

(*b*) The existing system can be described in many ways. Explain why alternatives to standard English (narrative) are sometimes necessary when describing an information system.

Suggest an alternative to standard English for this system and justify your choice. **4**

(10)

11. *Film Support Services* is a small film production company which is used by film-makers to provide a range of production services such as props for film-sets and costumes for the cast. The company is responsible for ensuring that all the resources required for a particular scene are available when the scene is scheduled to be filmed. In order to do this a Production Assistant completes a form detailing the resources required for the scene. A sample form is shown below.

Film Support Services Resource Request Form					
Film Title	Star Trip 6		Date required	31/07/2001	
Scene no.	ST6/048		Time required	08:00	
Director ID	RT02	Name	Scott Kubrick	Location	Hangar 1, Glasgow Airport
Producer ID	JB01	Name	Meg Chambers	Prod. Asst. ID	AD

Resource code	Description	Cost/hour	Hours required	Cost
CM06	Camera	100.00	5	500.00
LT01	Lighting	30.00	6	180.00
PR62	Lightsword	10.00	3	30.00
PR66	Blaster	15.00	3	45.00
			Total Cost	755.00

11. (continued)

Marks

A copy of each Resource Request Form is filed after it is completed. The original form is sent to the company's suppliers. The supplier will provide the resources and deliver them to the appropriate location at the specified date and time. The supplier will then send *Film Support Services* an invoice. On receipt of the invoice the Production Assistant will issue payment to the supplier. Invoices are filed for future reference.

(*a*) Draw a level 0 and level 1 data flow diagram to describe the above procedure. **3**

(*b*) The company wishes to issue each Production Assistant with a notebook computer so that the form can be completed electronically and sent to the supplier by e-mail. The completed request forms will then be stored in an electronic database. To aid computerisation you have been asked to normalise the system to third normal form. Perform this normalisation. **5**

(*c*) Using your answers to (*a*) and (*b*), create a data dictionary for this application. Suggest appropriate field names, sizes and data types. State any assumptions made. **2**

(10)

Total for section: 30

[*END OF PART C*]

[*END OF SECTION II*]

[*END OF QUESTION PAPER*]

[C054/SQP176]

Advanced Higher Information Systems

NATIONAL QUALIFICATIONS

Specimen Marking Instructions

SCOTTISH QUALIFICATIONS AUTHORITY

Commentary on paper

Structure of paper

The layout of the paper is modelled on layout of the Intermediate and Higher level papers. It has two sections as follows.

Section I Questions covering the core units (Database Systems and Information Systems Project)
Section II Questions covering the optional units

Section II is sub-divided into three parts (A, B and C). Each part relates to an optional unit as follows.

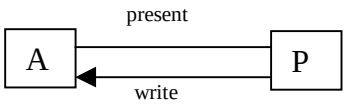
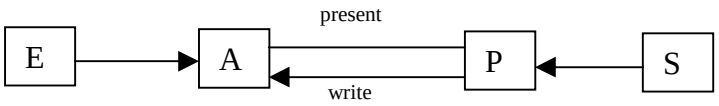
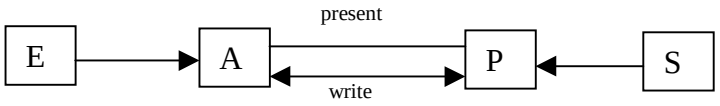
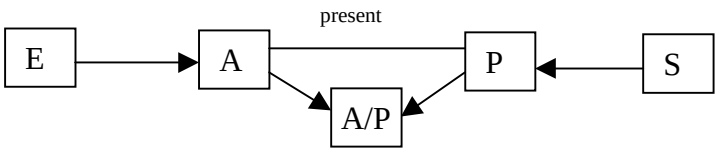
Part A: Multimedia
Part B: Natural Language Processing
Part C: Systems Analysis & Design

There are two questions in Section I and three questions in each part of Section II. Each question is worth 10 marks. Candidates are required to answer both questions in Section I and all the questions from one of the optional parts of Section II. This represents five questions in total, worth 50 marks. The remaining marks (50) are awarded for the coursework.

Distribution of marks

The paper is marked out of 50. Twenty marks are available in Section I (core units) and 30 marks in the optional section. This distribution of marks between core and option (20:30) is biased towards the optional units. This is to counter-balance the marks awarded to the coursework (50) which are drawn exclusively from the core units. An important objective within the paper is to integrate knowledge and understanding and the optional sections draw on knowledge and understanding gained as part of the core units.

A table is supplied at the end of the paper to relate each question to its source unit. This explains where each question comes from in terms of units, outcomes, performance criteria and range statements.

No.	Commentary
1	This question covers Outcome 1 with emphasis on the parts that are not in Higher Databases, namely 2NF and 3NF. The structure of the question is shown below where A is Author, P is Paper, E is Employer, S is Session, A/P is Authorship and arrow heads are used for the "many" end of a relationship. Original manual model  Model after (b)  Model after (d), With many-to-many  Model after (d), With many-to-one  **These data model diagrams are not part of the required answer.**

Type	Question	Marking scheme
PS	A conference is being organised. It is divided into sessions. At each session several papers are presented. The papers are written by authors. Several authors may share in the writing of one paper but only one paper can be submitted from any given author. Each paper has a presenter who is one of the authors of the paper. The organiser of the conference has devised two forms for recording data about the authors and the papers. One form relates to an author and one form relates to an author's paper. Each author is allocated a unique number and in addition this form shows the author's personal details (full name, sex and job title) and details of the author's employer (name, address and phone number). Each paper is allocated a unique number and in addition this form shows the details of the paper being presented (paper title, subject category, length in pages and date paper received), session details (session date, session time and session venue) and the author number of the presenting author. You may assume that there are the following dependencies between the items: • the personal details are functionally dependent on the unique author number; • the employer address and phone number are functionally dependent on employer name; • the paper details are functionally dependent on paper number; • the session details are functionally dependent on the combination of session date and session time. (a) Explain why the two primary entities, Author and Paper, are already in second normal form (2NF).	The key of the entity Author is "Author number". Since this is a single item there can be no partial dependencies. Since it is in 1NF already (no repeating items or groups), it is also in 2NF. The key of the entity Paper is "Paper number". Since this is a single item there can be no partial dependencies. Since it is in 1NF already (no repeating items or groups), it is also in 2NF. **1 mark for similar answer up to max. of 1 mark. Both parts of this answer must be provided for the award of the mark**
PS	(b) Complete the normalisation of these entities by producing entities in third normal form (3NF).	Remove from Author to a separate entity the indirectly dependent items in employer details. Key known from functional dependencies. Author (Author number, Full name, Sex, Employed by, Paper written) Employer (Employer name, Employer address, Employer phone number) Remove from Paper to a separate entity the indirectly dependent items in session details. Key known from functional dependencies. Paper (Paper number, Paper title, Subject category, Length,

		Date received, Presenting author, Session date, Session time) Session (Session date, Session time, Session venue) As an alternative to the composite key in Session an item "Session number" may be added as an artificial key, the pair of foreign keys in Paper is then reduced to a single item, say, "Presenting session". **1 mark for each of four correct entities (to include keys and foreign keys), maximum 4 marks**
PS	(c) Construct an entity relationship diagram for these entities. Indicate the nature of each relationship (such as many-to-one) between the 3NF entities.	The relationships are • One-to-many from Session to Paper • One-to-one from Author to Paper for presenting author • Many-to-one from Author to Paper for authoring of paper • Many-to-one from Author to Employer They are represented by the foreign keys. Foreign keys in Author • Link to Paper with "Paper written" • Link to Employer with "Employed by" Foreign keys in Paper • Link to Author with "Presenting author" • Link to Session with "Session date", "Session time" (or the artificial key "Presenting session") **2 marks for giving all four relationships correctly and identifying that foreign keys are used. The foreign keys need not be listed explicitly. They may have been identified in the answer to (b).** **1 mark for 2 or 3 correct relationships (with foreign keys) otherwise 0**
PS	(d) A rule change permits a single author to submit more than one paper. Modify your entity relationship diagram to reflect this change. State any changes to the attributes of existing entities and include any new entities that this change might cause. Indicate the nature of each relationship.	The relationship between Author and Paper is changed to many-to-many. This requires a new entity, say, Authorship, that is in a many-to-one relationship with each of Author and Paper. It contains the two foreign keys, "Author of" and "Paper by". Author loses its foreign key "Paper written". No change to any other entity. Authorship has composite key consisting of the two foreign keys, so is "all key" and must be in 3NF. Author (Author number, Full name, Sex, Employed by) Authorship (Author of, Paper by) **1 mark for new entity and its key, 1 mark for change to Author, 1 mark for showing 3NF, maximum of 3 marks**

No.	Commentary	
2	This question focuses on outcome 2 with some material integrating with the information systems project.	
Type	Question	Marking Scheme
PS	(a) The following table is an excerpt from a data dictionary that contains errors. Identify **four** errors in this data dictionary and explain each error.	1. The data item "Name" is duplicated, it appears in each of the entities Athlete and Club. 2. The abbreviation for "number" is not consistent (both Nbr and No are used). 3. The type of the foreign key Club_Captain_No is different from the type of the key to which it refers (A(4) and Number, respectively). 4. The key of Club is shown as not required, but all key items have required values. **1 mark for each error correctly identified and for which a reason is given.** **0 marks for error with no explanation**

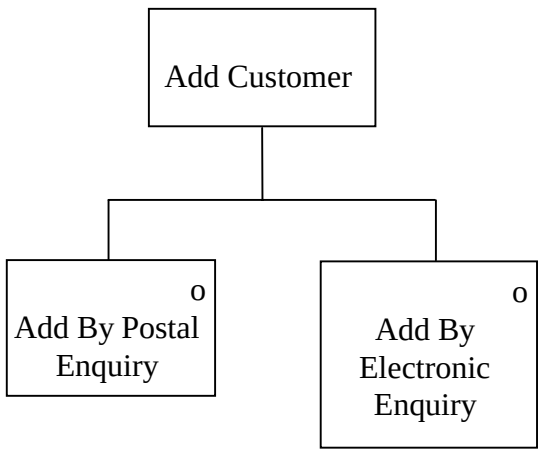
Item	**Entity**	**Description**	**Type/ Size**	**Range/ Validation**	**Required**	**Key**
Athlete_Nbr	Athlete	Unique number for athlete	A(6)	999999	Y	Y
Name	Athlete	Full name of athlete	A(40)		Y	N
Address	Athlete	Address of athlete	A(100)		N	N
Athlete_Club_No	Athlete	Number of club in which athlete trains	A(4)	Existing Club_Nbr	Y	N
Club_Nbr	Club	Unique number of club	Number	In 0..999	N	Y
Name	Club	Name of club	A(32)		Y	N
Club_Captain_No	Club	Athlete_Nbr of captain of club	A(6)	Existing Athlete_Nbr	Y	N

PS	(b) The entity life history (ELH) of the Customer entity is given below.	
	(i) Classify the following sequences as **legal** or **illegal** according to the diagram and for any illegal sequence explain why the sequence is illegal.	(1) is legal (2) is illegal since "Customer Life" is not an event, being an internal node (3) is illegal since there must be an occurrence of "Add Customer" at the start of each sequence (4) is legal **2 marks for all 4 correctly classified and reasons given.** **1 mark for 2 or 3 errors. Otherwise 0 marks**

(1)	(2)	(3)	(4)
Add Customer Delete Customer	Add Customer Customer Life Delete Customer	Modify Customer Address Modify Customer Address Delete Customer	Add Customer Modify Customer Name Modify Customer Address Delete Customer

(ii) Show how to change the ELH if there are two different events that can create a new Customer: Add By Postal Enquiry, Add By Electronic Enquiry.	Change the "Add Customer" node to be an internal node with two successors as shown below. **1 mark for keeping Add Customer, 1 mark for showing the alternatives correctly, maximum of 2 marks**

KU	(c) An existing information system is currently based on paper records. A decision is taken to computerise the system using a database. At what stage should the decision be taken as to which database product to use? Justify your answer.	The decision should be taken as late as possible. Reasons for an early decision would be lack of any available alternative to one database system, need for early training of implementors if unfamiliar with use of databases. Advantages of a late decision are being able to choose the most appropriate alternative (cost, ease of use, complexity, sufficient power for future changes and expansions), keeping the analysis and design free of any bias of a particular product. **2 marks for a well-argued case, 1 mark only for an incomplete argument**



No.		Commentary	
3	This question combines knowledge and understanding gained from the study of multimedia systems.		
Type		**Question**	**Marking scheme**
PS	a)	Describe the proposed target audience. Your description should identify the requirements of the audience and their computing experience.	• Requirements of the target audience are identified correctly *e.g. product must be informative, easy to use, interactive, fun* • IT experience of the target audience is clearly described *e.g. audience will have limited IT experience* **1 mark each for correctly described requirements and IT experience up to a maximum of 2 marks**
PS	b)	Suggest an appropriate medium for this product. Justify your choice.	• Proposed medium meets the needs of the target audience *e.g. product should be viewed from CD ROM* • Proposed medium is suitable for the product content *e.g. product may be used in various classrooms and will need to be portable; product will have to run off a stand-alone PC* **1 mark each for correctly identifying suitable medium and 1 mark for providing appropriate justification**
KU	c)	The specification of this multimedia product must clearly define the *structure* of the product and identify any *interactivity* within the product. How would the product specification define the structure of the product and identify any interactivity within the product?	Specification should include: • Detail of all navigation and interactivity features of the product *e.g. main pathways through the product should be detailed along with user choices* • Detail of the range of constructs to be included *e.g. any repetition, selection or sequencing should be shown* • Detail of any interactivity associated with media elements *e.g. any videos that play automatically or audio sequences that play once selected should be shown* **1 mark each for each suitably described purpose to a maximum of 3 marks**
KU	d)	Explain what use is made of *storyboards* in the production of a multimedia product.	Storyboards are used to: • Show suitable user interface *e.g. for pupils in P4-6, interface should be colourful with large clear buttons* • Show the sequence of media elements *e.g. the order of introduction of media elements should be shown* • Show any layering and animation paths *e.g. storyboard should include any layering sequences and animations* **1 mark each for each use correctly identified to a maximum of 3 marks**

No.		Commentary	
4	This question combines knowledge and understanding gained from the study of multimedia systems.		
Type		**Question**	**Marking scheme**
KU	a i)	Explain how decisions made at the planning stage can affect the machine specification for the planned product.	Machine specification should include: • Hardware for creating the product *e.g. hardware to be used to create the product is specified in terms of input, processing, output and storage devices* • Authoring and production software *e.g. any software to be used should be identified; this is likely to include authoring package as well as graphics and audio software* • Hardware for viewing the product *e.g. minimum level of hardware and software required to view the product should be stated* **1 mark for each component of machine specification above correctly identified and described to maximum of 3 marks**
PS	a ii)	Describe the information that would be specified in the plan for the development of the product.	Product development plan would include information on the following: • All required tasks are identified *e.g. all tasks required to complete the product are stated* • Timescales for completion of each task are identified *e.g. the estimated time required for each task should be identified in the plan* • Dependencies between tasks are identified and described *e.g. any tasks which depend on the prior completion of other tasks should be identified since this will affect the order of completion* • Critical path is identified *e.g. the critical path should be identified since this will determine which tasks are vital to the completion of the product* **1 mark for each correct description of each item included in the product development plan to a maximum of 4 marks**

Type		Question	Marking scheme
KU	b)	Once the multimedia product has been assembled, the development team complete a product evaluation. What factors would be considered by the development team when evaluating their multimedia product?	• Progress is compared with initial plan and timescales *e.g. the time taken to complete the various tasks is compared with the original estimated timescales and reasons provided for why tasks ran over or under the estimated time* • Comparison of product with initial proposal *e.g. the finished product should be compared with what was initially intended and any discrepancies should be identified* • Further development work is identified *e.g. any parts of the product that require further work should be identified together with details of any appropriate future enhancements* • Sources of media elements are acknowledged *e.g. sources of any media elements should be acknowledged appropriately* **1 mark for an appropriate description of any 3 factors considered when writing a product evaluation to a maximum of 3 marks**

No.		Commentary	
5	This question relates to the candidates' knowledge and understanding of multimedia in a problem solving context.		
Type		**Question**	**Marking scheme**
	(a)	After producing the package you discover that some companies have a lower specification of equipment than originally envisaged. How could you adapt your package to deal with the following situations?	
PS		(i) The machines have a slower processor, less memory and a lower capacity hard disk than originally anticipated. (ii) The machines have smaller monitors than originally anticipated with lower graphics resolution.	• Eliminate unnecessary graphics. • Reduce quality (resolution) of graphics. • Reduce sound and video components. • Reduce quality of sound and video components. • Reduce quality (resolution) of graphics elements. • Increase font size of text. • Change typeface of text to aid clarity. • Other appropriate answer. **1 mark for each bullet point.** **2 marks for each part to max. of 4.**
PS	(b)	You consider transferring the product from CD-ROM to the World Wide Web. Explain how this change in medium might affect the product's content and structure.	• Structure should conform to WWW standards (hyperlinks, frames etc.) • Content will have to be reduced to conserve bandwidth. **1 mark for each bullet point. Max. of 2 marks.**
	(c)	Once the product has been developed it has to be tested before being released to customers. Outline an appropriate text strategy for the original (CD-ROM) multimedia product.	Testing strategy should relate to: • Hardware – ability to run on a range of hardware platforms • Software – compatibility to operating systems • Performance – speed of operation on typical hardware/software platforms • Reliability – system testing to ensure robustness of final product. **1 mark for each bullet point to max. 4.**

No.		Commentary	
6	This integrative question combines knowledge and understanding gained from the study of multimedia systems and database systems.		
Type		**Question**	**Marking scheme**
KU	a)	Outline **two** main goals of any natural language processing system.	• To build computer systems that humans can communicate with using natural language • To use computers as tools to help us understand better how humans use language through modelling it **Award 1 mark for a correct description of each goal to max. 2 marks**
KU	b)	Provide a description of any **two** of the applications of natural language processing techniques mentioned above.	• Text analysis – the use of natural language techniques to evaluate texts or documents by looking at their structure and style • Machine translation – automatic conversion of texts from one language into another language • Command and control – natural language interaction with devices using speech rather than typing • Question answering – the use of natural language techniques to enable users of interactive systems to request information from the system and supply the system with information on request **Award 1 mark each for a correct description of 2 of the applications of NLP listed**
KU	c)	*Parsing* is a technique commonly used in natural language processing. What is the purpose of parsing? Outline how parsing is accomplished.	• Purpose is to analyse a sentence in terms of grammar and provide a grammatical description of the words in the sentence • A natural language sentence is analysed in terms of the grammar of the language and broken down into its component parts **Award 1 mark for description of purpose and 1 mark for description of what is involved**

Type		Question	Marking scheme
KU	d) i)	Many modern database systems make use of natural language techniques other than parsing. Provide a description of an alternative technique to parsing. Your description should indicate how the technique is applied in database systems.	• Description of an alternative technique such as transition networks, key phrase matching, use of logic *e.g. logic involves representing the knowledge in the system in terms of facts and rules; this knowledge can be searched in order to find solutions to user queries* • Description refers to use of the technique within database systems *e.g. logic can be used in question answering in a database system* **Award 1 mark for a suitable description of an alternative technique and 1 mark for indicating how this technique could be used in a database system**
PS	d) ii)	Outline the implications of natural language processing techniques in database systems in terms of user productivity and development costs.	• Implications in terms of user productivity are described *e.g. user productivity is likely to increase since the system is more intuitive and easier to use – the user of such a system will therefore be able to complete more work in less time.* • Implications in terms of costs are described *e.g. database systems with a natural language 'front end' are likely to take longer to produce and test; this will therefore increase development costs* **Award 1 mark for an appropriate description of implications for user productivity and 1 mark for description of implications for development costs**

<table>
<tr><th colspan="2">No.</th><th colspan="2">Commentary</th></tr>
<tr><td>7</td><td colspan="3">This question combines knowledge and understanding and problem solving gained from the study of natural language processing.</td></tr>
<tr><th>Type</th><th></th><th>Question</th><th>Marking scheme</th></tr>
<tr><td>PS</td><td>a)</td><td>Consider the following newspaper headline:

Giant waves in tunnel.

Now consider the sentence below:

“I cooked the cabbage in the pot and then I ate it.”

Discuss any problems of ambiguity illustrated in each of these examples.</td><td>• Giant waves in tunnel – problem with interpretation. It isn’t clear from the headline whether there is a giant waving inside the tunnel or whether there are huge waves in the tunnel. An example of word class ambiguity.

• I cooked the cabbage in the pot and then I ate it – problem with referential ambiguity; we don’t know what the pronoun ‘it’ refers to – the pot or the cabbage.

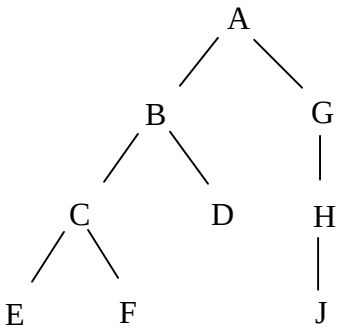
1 mark each for a clear description of the ambiguity or problem with interpretation in each case. Max. 2 marks.</td></tr>
<tr><td>PS</td><td>b)</td><td>Specify a grammar that represents a subset of English that has the following constituents and lexicon.

<table><tr><th>Constituents</th></tr><tr><td>sentence</td></tr><tr><td>nounphrase</td></tr><tr><td>verbphrase</td></tr></table>
<table><tr><th colspan="2">Lexicon</th></tr><tr><td>noun</td><td>cat, mouse, cheese, pear</td></tr><tr><td>propernoun</td><td>David, Kirsty</td></tr><tr><td>verb</td><td>ate, saw, caught, sketched</td></tr><tr><td>determiner</td><td>the, a, an</td></tr><tr><td>adverb</td><td>slowly, quickly</td></tr></table>
Examples of sentences that your grammar should be able to generate are the following.
The mouse slowly ate the cheese.
Kirsty saw the cat.
The cat caught a mouse.
David quickly sketched a cat.</td><td>• Sentence
nounphrase, verbphrase
• Nounphrase
determiner, noun
• Nounphrase
propernoun
• Verbphrase
verb, nounphrase
• Verbphrase
adverb, verb, nounphrase

3 marks for complete grammar described above; deduct 1 mark for each omission</td></tr>
<tr><td>PS</td><td>c)</td><td>Specify two test sentences, in addition to those given above, for your grammar. You should ensure that one sentence would succeed when parsed by your grammar and one sentence should fail. Indicate the expected success or failure in each case.</td><td>• Any sentence that would fail when parsed e.g. The black cat caught the mouse.
• Any sentence that would succeed when parsed e.g. Kirsty slowly ate an apple.

1 mark each for an appropriate sentence correctly stated as a success or failure. Max. 2 marks.</td></tr>
</table>

Type		Question	Marking scheme
PS	d)	Parse the sentence below, showing the full grammatical structure of the sentence. Use bracketing to indicate the structure. "The mouse slowly ate the cheese."	Sentence (nounphrase (det (the), noun (mouse)), verbphrase (adverb (slowly), verb (ate), nounphrase (det (the), noun (cheese)). **2 marks for a complete parse of the sentence as indicated above; award 1 mark for a maximum of 2 omissions; otherwise award 0 marks.**
PS	e)	Extend your grammar to include adjectives.	Extend the grammar in part (b) above by adding: • Nounphrase determiner, adjective, noun **1 mark for additional nounphrase required**

No.		Commentary	
8		This question combines knowledge and understanding and problem solving gained from the study of natural language processing.	
Type		**Question**	**Marking scheme**
KU/PS	a)	State the difference between top-down and bottom-up parsing strategies. Which one is the better strategy if the text contains many words like "can" and "bank" which are both nouns and verbs? Justify your answer.	• Top-down parsing begins with a rule for the sentence and proceeds down through phrases and categories to words. • Bottom-up parsing begins with words and their categories and combines into phrases, then into higher-level phrases and so into sentences. • Top-down would suit this better since bottom-up leads to trying all possibilities including those that would never lead to a well-formed sentence. **1 mark each for a clear description of the two methods. 1 mark for choice with a reason. Maximum 3 marks**
KU	b)	Explain why it is sometimes necessary to search a parse tree.	• The parser has to try which of several possible rules can be used to expand/rewrite a given category (such as noun-phrase). • To search through the possibilities it may do a depth-first search, backtracking when a match with a category fails. **2 marks for a complete reason, 1 mark if only partial (e.g. no mention of matching).**
KU	c)	Describe the methods for performing depth-first search and breadth-first search. State the order in which the nodes are searched using each method on the following tree.  	• Depth-first search: Visit top node, then for each successor node search that node's successors in order. • Breadth-first search: Visit top node, visit its successors in order, then visit the successors of each of these nodes at the next level. • Depth-first: ABCEFDGHJ • Breadth-first: ABGCDHEFJ **1 mark each for the two methods, 1 mark for stating orders. Maximum of 3 marks.**
KU	d)	Give a brief description of a natural language technique that does not depend on parsing using a grammar.	Techniques include: • Key phrase matching • Transition networks • Use of templates • Use of logic for semantics **2 marks for a good description, 1 mark if correct but incomplete, 0 marks otherwise.**

No.		Commentary	
9	This question assesses the candidate's knowledge of the systems analysis and design process.		
Type		**Question**	**Marking scheme**
KU	(a)	Describe the **five** stages involved in the analysis and design process. Relate these stages to the proposed laboratory system.	• Analysis: investigation of current system - observing / interviewing users, examining documentation. • Design: specification of new system, files, processes, inputs, outputs. • Implementation: software development, changeover, user training. • Testing: program testing, integration testing, acceptance testing. • Maintenance: ad hoc, planned **Deduct one mark for each omitted stage.** **Max. 4 marks.**
KU	(b)	Explain why the stages in the analysis and design process do not always follow a simple sequence.	Information uncovered at a later stage may lead to revision of an earlier stage, e.g. a problem uncovered during testing which may lead to changes in design. **Max. 1 mark.**
KU	(c)	Explain the difference between the *logical* design and the *physical* design of an information system. Relate your answer to the laboratory system.	Logical: design of the system concentrating on its functionality with no reference to implementation details. Physical: includes implementation issues such as file size, platforms. **1 mark for each answer to max. 2.**
KU	(d)	What is a *system specification*? Give **three** examples of what might be included in the system specification for the laboratory system.	Detailed description of the new system and the products expected from the designer. **1 mark.** Examples: • Data structures • Processes/events • Screen layouts. • Other appropriate example. **2 marks for three appropriate examples.** **1 mark for two appropriate examples.** **0 marks otherwise.**

No.	Commentary
10	This question assesses the candidate's knowledge of analysis of an existing (manual) information system.

Type		Question	Marking scheme
KU	(a)	Describe **three** methods you might use to obtain information about the existing system. Give **one** disadvantage of each of your methods of obtaining information.	**2 marks for each description. 1 mark for description, 1 mark for disadvantage. Max. 6 marks.**
KU	(b)	The existing system can be described in many ways. Explain why alternatives to standard English (narrative) are sometimes necessary when describing an information system. Suggest an alternative to standard English for this system and justify your choice.	Standard English is imprecise and misleading – open to interpretation. **2 marks for explaining why alternatives to standard English are required.** • Decision tables • Algorithms • Structured English • Pseudocode **1 mark for naming alternative and 1 mark for justifying choice. Max. 2 marks.**

No.	Commentary
11	This question assesses the candidate's ability to analyse an existing information system.

Type		Question	Marking scheme
PS	(a)	Draw a level 0 and level 1 data flow diagram to describe the above procedure.	**1 mark for correct notation.** **2 marks for correct diagrams (1 mark for level 0 and 1 mark for level 1).** **Max. 3 marks.**
PS	(b)	The company wishes to issue each Production Assistant with a notebook computer so that the form can be completed electronically and sent to the supplier by e-mail. The completed request forms will then be stored in an electronic database. To aid computerisation you have been asked to normalise the system to third normal form. Perform this normalisation.	**Correct entities 1 mark** **Correct UNF 1 mark** **Correct 1NF 1 mark** **Correct 2NF 1 mark** **Correct 3NF 1 mark** **Max. 5 marks.**
PS	(c)	Using your answers to (a) and (b), create a data dictionary for this application. Suggest appropriate field names, sizes and data types. State any assumptions made.	**Consistent attributes 1 mark.** **Correct sizes/types 1 mark.** **Max. 2 marks.**

[END OF MARKING SCHEME]

SOURCE OF QUESTIONS

Question	Part	Section	Unit(s)	Outcome(s)	PC	Range
1	a		Database Systems	1	c	
	b		Database Systems	1	d, e	
	c		Database Systems	1	d	
	d		Database Systems	1	b, c, d, e	
2	a		Database Systems	2	b	
	b	i-ii	Database Systems	2	e	
	c		Database Systems Information Systems Project	1, 2 1	*passim* a, d	
3	a		Multimedia	1	b	
	b		Multimedia	1	e	
	c		Multimedia	1	d	
	d		Multimedia	1	d	
4	a	i	Multimedia	2	a	Hardware – all Software – all
		ii	Multimedia	2	b	
	b		Multimedia	4	b, c, d	
5	a	i	Multimedia	3	b	
		ii	Multimedia	3	b	
	b		Multimedia	1	c, d, e	
	c		Multimedia Information Systems Project	3d 2c		

Question	Part	Section	Unit(s)	Outcome(s)	PC	Range
6	a		Natural Language Processing	1	b	
	b		Natural Language Processing	1	c	All
	c		Natural Language Processing	2 or 3?	a b?	
	d	i	Natural Language Processing Database Systems	4	a, b	?
		ii	Natural Language Processing Database Systems	4	b	
7	a		Natural Language Processing	2	c	Word sense ambiguity, referential ambiguity
	b		Natural Language Processing	3	a	
	c		Natural Language Processing	3	d	
	d		Natural Language Processing	3	c	
	e		Natural Language Processing	3	e	
8	a		Natural Language Processing	3 2	b c	Top-down, bottom-up
	b		Natural Language Processing	3	c	
	c		Natural Language Processing	3	b	Depth-first, breadth-first
	d		Natural Language Processing	4	a	Choice from whole range
9	a		Systems Analysis and Design	1	a	Stages
	b		Systems Analysis and Design	1	c	
	c		Systems Analysis and Design	3	c	
	d		Systems Analysis and Design	1 3	a a	System specification
10	a		Systems Analysis and Design	2	a	
	b		Systems Analysis and Design	2	d	
11	a		Systems Analysis and Design	3	c	
	b		Systems Analysis and Design Database Systems	3 1	c b, c, d, e	
	c		Systems Analysis and Design	3	c	System specification