

- 1 A company runs a Virtual Learning Environment (VLE). Schools can register students to use the VLE. The students get their own account and the school can view and monitor their students who are registered. There are currently over 10 000 schools registered, each with up to 1000 students.

The students can watch videos, take quizzes and communicate using forums and online chat tools.

The company gathers a large amount of data and wants to use data mining to help them decide how to improve the VLE in the future.

Discuss how the company can use data mining to decide how to improve the VLE.

You should include the following in your answer:

- the characteristics of data mining
- the benefits of data mining in this scenario
- the drawbacks of data mining in this scenario.

This image shows a blank sheet of white paper with ten horizontal dashed lines, evenly spaced from top to bottom. The lines are thin and black, typical of primary or elementary school writing paper. There are no margins, text, or other markings on the page.

[12]

- 2 A computer game has a building containing 7 rooms. There are secret passages between each room. Fig. 3 shows the rooms and the passages between the rooms represented as a graph data structure.

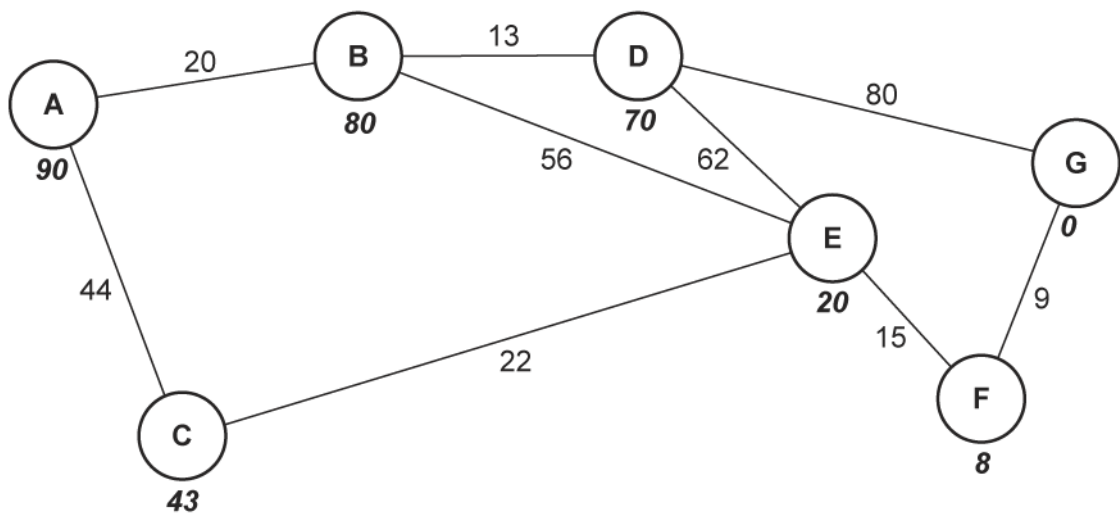


Fig. 3

The final game will involve multiple computer-controlled characters and interactive elements that make use of artificial intelligence to determine the moves they will make.

The artificial intelligence will use heuristics to determine where the computer-controlled characters will move in the game.

Discuss how heuristics can be used in algorithms.

You should include the following in your answer:

- the purpose of heuristics
- the benefits and drawbacks of heuristics
- the suitability of heuristics in algorithms within a computer game.

[9]

- a description of both problem recognition and decomposition

- how each method can be used when designing the solution
- the benefits of using each method when designing the solution.

Blank lined paper for writing.

4 Amy’s processor makes use of pipelining during the fetch-decode-execute cycle.

The processor’s pipeline consists of the following stages:

- Fetching the instruction from memory
- Decoding the instruction
- Executing the instruction.

Instructions A, B, C and D need to be processed.

Identify the stage(s) and instruction(s) run during each pipeline below.

Pipeline 1 -----

Pipeline 2 -----

Pipeline 3 -----

Pipeline 4 -----

[4]

5 * A game will have thousands of users. Taylor will store data about the users and their actions while playing the game in a large database.

Evaluate how Taylor can use data mining to inform future changes to improve his circus game.

[9]

6 Mabel is a software engineer. She is writing a computer game for a client. In the game the main character has to avoid their enemies. This becomes more difficult as the levels of the game increase.

Mabel uses decomposition to design the program.

Explain how decomposition can aid the design of this program.

[2]

8(a) A company merger is joining five e-commerce retailers under one company, OCRRetail. Each retailer has a different sales system and OCRRetail wants to develop one computer system that can be used by all the retailers.

Mary's software development company has been employed to analyse and design a solution for the company.

Mary plans to use data mining to generate information about OCRRetail's customers. Mary will use this information to benefit the company.

i. Define the term 'data mining'.

[1]

ii. Identify **two** pieces of information that data mining could provide OCRRetail about sales, and state how OCRRetail could make use of this information.

1

2

[4]

(b) Mary has developed the program and is considering using performance modelling before installing the system.

i. Define the term ‘performance modelling’.

----- [1]

ii. Identify **one** way performance modelling could be used to test the new system.

----- [1]

(c) Mary created the program as a series of sub-programs that can be reused.

Describe **one** benefit of Mary creating reusable program components.

----- [2]

END OF QUESTION PAPER

Mark Scheme

Question	Answer/Indicative content	Marks	Guidance
1	Mark Band 3 - High level (9-12 marks) The candidate demonstrates a thorough knowledge and understanding of data mining; the material is generally accurate and detailed. The candidate is able to apply their knowledge and understanding directly and consistently to the context provided. Evidence/examples will be explicitly relevant to the explanation. <i>There is a well-developed line of reasoning which is clear and logically structured. The information presented is relevant and substantiated.</i> Mark Band 2 - Mid level (5-8 marks) The candidate demonstrates reasonable knowledge and understanding of data mining; the material is generally accurate but at times underdeveloped. The candidate is able to apply their knowledge and understanding directly to the context provided although one or two opportunities are missed. Evidence/examples are for the most part implicitly relevant to the explanation. The candidate provides a reasonable discussion, the majority of which is focused. Evaluative comments are, for the most part appropriate, although one or two opportunities for development are missed. <i>There is a line of reasoning presented with some structure. The information presented is in the most part relevant and supported by some evidence.</i> Mark Band 1 - Low Level (1-4 marks) The candidate demonstrates a basic knowledge of data mining with limited understanding shown; the material is basic and contains some inaccuracies. The candidate makes a limited attempt to apply acquired knowledge and understanding to the context provided. The candidate provides a limited discussion which is narrow in focus. Judgements if made are weak and unsubstantiated. <i>The information is basic and communicated in an unstructured way. The information is supported by limited evidence and the relationship to the evidence may not be clear.</i>	12	AO1: Knowledge and Understanding Indicative content • Analysis of patterns and anomalies in large data sets • Turns large quantities of data into useful information. These may not be immediately obvious to a casual reader. • Resulting information is used to make predictions, to increase revenue, to target advertising and improve services. AO2: Application • Identify the amount of time students spend on the system • Identify the days / times when is it used most? Least? • Identify the features/tools students use most and least • Identify which questions find the most difficult. • Identify the time / day of the week when students learn the most • Identify which schools are performing better or which areas of the country are performing better. • Identify which courses students are enjoying more AO3: Evaluation • Data mining can spot patterns/trends, however it cannot explain them. Company may still need to do extra research. • Students may have privacy concerns about their activities being logged. • Requires powerful computers with a lot of processing power to process huge amounts of data. • Need to make sure data collection is legal and in terms and conditions • Analysis can be costly e.g. may need an external company • Security - data being collected requires protection • Include more features students use so they will use it more, and gain more revenue • Identifying features not used means resources and development can be moved from these onto other areas

Mark Scheme

Question			Answer/Indicative content	Marks	Guidance
			0 mark No attempt to answer the question or response is not worthy of credit.		<u>Examiner's Comments</u> Many Level 2 responses demonstrated an ability to define data mining and give relevant examples of its use. Few candidates were able to give an evaluative critique as to the effectiveness of data mining for this scenario. Many candidates struggled to structure their responses. Where candidates did produce a clear structure, they often used paragraphs effectively, starting with a definition, considering benefits and drawbacks, then going on to evaluate within the parameters of the scenario.
			Total	12	

Mark Scheme

Question	Answer/Indicative content	Marks	Guidance
2	Mark Band 3 – High level (7-9 marks) The candidate demonstrates a thorough knowledge and understanding of heuristics; the material is generally accurate and detailed. The candidate is able to apply their knowledge and understanding directly and consistently to the context provided. Evidence/examples will be explicitly relevant to the explanation. There is a well-developed line of reasoning which is clear and logically structured. The information presented is relevant and substantiated. Mark Band 2 – Mid level (4-6 marks) The candidate demonstrates reasonable knowledge and understanding of heuristics; the material is generally accurate but at times underdeveloped. The candidate is able to apply their knowledge and understanding directly to the context provided although one or two opportunities are missed. Evidence/examples are for the most part implicitly relevant to the explanation. The candidate provides a reasonable discussion, the majority of which is focused. Evaluative comments are, for the most part appropriate, although one or two opportunities for development are missed. There is a line of reasoning presented with some structure. The information presented is in the most part relevant and supported by some evidence. Mark Band 1 – Low Level (1-3 marks) The candidate demonstrates a basic knowledge of heuristics with limited understanding shown; the material is basic and contains some inaccuracies. The candidates makes a limited attempt to apply acquired knowledge and understanding to the context provided. The candidate provides a limited discussion which is narrow in focus. Judgements if made are weak and unsubstantiated. The information is basic and communicated	9	AO1: Knowledge and Understanding Indicative content • Heuristics are used to reduce time taken to solve a problem • It is a general 'rule of thumb' or an educated guess. • It finds a solution which is 'good enough' / close to the best solution • Heuristic is a weight added to a node/decision • E.g. Description of use such as in A* algorithm as estimate of distance to destination AO2: Application • Heuristics reduce the time complexity as every possibilities within the game does not need to be examined. • Heuristics require skill to implement effectively • Used in AI when the exact steps cannot be pre-programmed and decisions need making • Due to time-saving, they are not always accurate, the solution e.g. shortest path might not be the most efficient. AO3: Evaluation • Heuristics are more appropriate with complex time-critical tasks - some aspects of game may require faster searching/decisions - current graph is not complex or time-critical so not required • Heuristics are more appropriate with large-scale tasks - game could be large scale and AI algorithms may need to be shortened • Games are not life-critical, so a good answer is likely enough, a perfect answer is not necessarily required. • Avoid programs running indefinitely - in a computer game there could be too many possibilities so will terminate with a solution faster <u>Examiner's Comments</u> Candidates demonstrated reasonable knowledge of heuristics in terms of

Mark Scheme

Question			Answer/Indicative content	Marks	Guidance
			in an unstructured way. The information is supported by limited evidence and the relationship to the evidence may not be clear. 0 mark No attempt to answer the question or response is not worthy of credit.		definition and application to the game. Many candidates identified some benefits and drawbacks and the need to have an accurate heuristic value. Few candidates were able to give a detailed evaluation, or expand to scalability for example, to achieve a Level 3 response.
			Total	9	

Mark Scheme

Question	Answer/Indicative content	Marks	Guidance
3	Mark Band 3 – High level (7–9 marks) The candidate demonstrates a thorough knowledge and understanding of both computational thinking methods; the material is generally accurate and detailed. The candidate is able to apply their knowledge and understanding directly and consistently to the context provided. Evidence/examples will be explicitly relevant to the explanation. <i>There is a well-developed line of reasoning which is clear and logically structured. The information presented is relevant and substantiated.</i> Mark Band 2 – Mid level (4–6 marks) The candidate demonstrates reasonable knowledge and understanding of both computational thinking methods; the material is generally accurate but at times underdeveloped. The candidate is able to apply their knowledge and understanding directly to the context provided although one or two opportunities are missed. Evidence/examples are for the most part implicitly relevant to the explanation. The candidate provides a reasonable discussion, the majority of which is focused. Evaluative comments are, for the most part appropriate, although one or two opportunities for development are missed. <i>There is a line of reasoning presented with some structure. The information presented is in the most part relevant and supported by some evidence.</i> Mark Band 1 – Low Level (1–3 marks) The candidate demonstrates a basic knowledge of both computational thinking methods with limited understanding shown; the material is basic and contains some inaccuracies. The candidates makes a limited attempt to apply acquired knowledge and understanding to the context provided. The candidate provides a limited discussion which is narrow in focus. Judgements if made are weak and unsubstantiated. <i>The information is basic and communicated</i>	9	AO1: Knowledge and Understanding Indicative content • Problem recognition is identifying that there is a problem to be solved, to determine exactly what the problem is from a description/scenario and to determine if the problem can be solved with computational methods • Decomposition is splitting the problem down into subproblems that can be solved independently AO2: Application • Problem recognition: identifying the need for the scheduling system, what it will take as its inputs, what will need to be output etc. • Decomposition: subproblems could include: ◦ inputting the requirements ◦ generating possible routes ◦ evaluating the routes ◦ outputting the schedule AO3: Evaluation Computational methods allow the e.g.: • programmer to determine what the problem, what the challenges may be and what additional information is required before starting to code the solution • identification of the key features for programmers to focus on • splitting of the task into smaller, more manageable/solvable problems which allows for a solution to be developed quicker • design of an effective/efficient solution that makes best use of a processor • splitting of a task to allow programmers to focus on areas they specialise in. <u>Examiner's Comments</u> Many Level 1 responses gave simplistic definitions of problem recognition and/or decomposition, often with more successful descriptions for decomposition. Relatively few candidates could clearly state that problem recognition starts with identifying inputs and outputs. Some degree of

Mark Scheme

Question			Answer/Indicative content	Marks	Guidance
			<i>in an unstructured way. The information is supported by limited evidence and the relationship to the evidence may not be clear.</i> 0 mark No attempt to answer the question or response is not worthy of credit.		application to the question scenario context was required for Level 2. Most candidates' marks clustered around Level 2 with 4 or 5 marks. Many generic answers focused on problem recognition and/or decomposition as learnt terms, with associated benefits given, but little more than that, which limited marks to Level 2. For Level 3 an evaluation of ideas relevant to the context of the scenario given was expected. Some good responses indicated recognition of a pathfinding type problem with a need for shortest path/A* algorithm, approximation to travelling salesman type solutions or the need to employ heuristics. Candidates giving this level of insight often accessed Level 3.
			Total	9	

Mark Scheme

Question			Answer/Indicative content	Marks	Guidance
4			1 mark for each interval Interval 1 • A is fetched Interval 2 • A is decoded • B is fetched Interval 3 • A is executed • B is decoded • C is fetched Interval 4 • B is executed • C is decoded • D is fetched	4 AO1.2 (4)	
			Total	4	

Mark Scheme

Question	Answer/Indicative content	Marks	Guidance
5	Mark Band 3 – High level (7-9 marks) The candidate demonstrates a thorough knowledge and understanding of data mining; the material is generally accurate and detailed. The candidate is able to apply their knowledge and understanding directly and consistently to the context provided. Evidence/examples will be explicitly relevant to the explanation. The candidate is able to weigh up the use of data mining which results in a supported and realistic judgment as to whether it is possible to use them in this context. <i>There is a well-developed line of reasoning which is clear and logically structured. The information presented is relevant and substantiated.</i> Mark Band 2 – Mid level (4-6 marks) The candidate demonstrates reasonable knowledge and understanding of data mining; the material is generally accurate but at times underdeveloped. The candidate is able to apply their knowledge and understanding directly to the context provided although one or two opportunities are missed. Evidence/examples are for the most part implicitly relevant to the explanation. The candidate makes a reasonable attempt to come to a conclusion showing some recognition of influencing factors that would determine whether it is possible to use data mining in this context. <i>There is a line of reasoning presented with some structure. The information presented is in the most part relevant and supported by some evidence</i> Mark Band 1 – Low Level (1-3 marks) The candidate demonstrates a basic knowledge of data mining with limited understanding shown; the material is basic and contains some inaccuracies. The candidates makes a limited attempt to apply acquired knowledge and understanding to the context provided. The candidate provides nothing more than an unsupported assertion. <i>The information is basic and communicated in an unstructured way. The information is supported by limited evidence and the relationship to the evidence may not be clear.</i>	9 AO1.1 (2) AO1.2 (2) AO2.1 (2) AO3.3 (3)	AO1: Knowledge and Understanding Indicative content • Data mining searches large amounts of data • Searches for relationships between facts/components/events that may not be obvious • May include pattern matching algorithms • May involve anomaly detection algorithms • Used for business modelling • Used to plan for future eventualities AO2: Application • Can look for how people use the website e.g. visiting times, what they click on, how long they spend on certain features, what they do first, which elements are used least • Give recommendations for future changes to the game e.g. features to add, or remove AO3: Evaluation Candidates will need to evaluate the benefits and drawbacks of using data mining. e.g. • Can use to introduce new features • Increase use from users • Increase revenue by selling features used more often • Make the game more appealing • Remove features people don't use • Use to target advertising • Privacy concerns from users • Misuse of information

Mark Scheme

Question			Answer/Indicative content	Marks	Guidance
			0 marks No attempt to answer the question or response is not worthy of credit.		
			Total	9	
6			1 mark per bullet • She can split the problem down into sub problems • It will creates a more manageable problem / simpler to understand / maintain • can tackle each sub problem independently	2 AO1.2 (1) AO2.2 (1)	<u>Examiner's Comments</u> Decomposition was clearly understood and most candidates achieved at least one of the two marks. Those that did not gain full marks did not go beyond decomposition breaking the task into sub-tasks.
			Total	2	

Mark Scheme

Question	Answer/Indicative content	Marks	Guidance
7	Mark Band 3 – High level (7-9 marks) The candidate demonstrates a thorough knowledge and understanding of data mining; the material is generally accurate and detailed. The candidate is able to apply their knowledge and understanding directly and consistently to the context provided. Evidence/examples will be explicitly relevant to the explanation. The candidate is able to weigh up the context which results in a supported and realistic judgment as to whether it is possible to use data mining in this context. <i>There is a well-developed line of reasoning which is clear and logically structured. The information presented is relevant and substantiated.</i> Mark Band 2 – Mid level (4-6 marks) The candidate demonstrates reasonable knowledge and understanding of data mining; the material is generally accurate but at times underdeveloped. The candidate is able to apply their knowledge and understanding directly to the context provided although one or two opportunities are missed. Evidence/examples are for the most part implicitly relevant to the explanation. The candidate makes a reasonable attempt to come to a conclusion showing some recognition of influencing factors that would determine whether it is possible to use data mining. <i>There is a line of reasoning presented with some structure. The information presented is in the most part relevant and supported by some evidence.</i> Mark Band 1 – Low Level (1-3 marks) The candidate demonstrates a basic knowledge of data mining with limited understanding shown; the material is basic and contains some inaccuracies. The candidates makes a limited attempt to apply acquired knowledge and understanding to the context provided. The candidate provides nothing more than an unsupported assertion. <i>The information is basic and communicated in an unstructured way. The information is</i>	9 AO1.1 (2) AO1.2 (2) AO2.1 (2) AO3.3 (3)	AO1: Knowledge and Understanding Indicative content • Extracting data from databases • Using large data sets • Looking for patterns/specific occurrences of data • Gathering data that can be analysed and used to inform decisions AO2: Application e.g. • Use to find out what his users do • Find features that are used most often • Find features that are not used • Find out what people in his target age group do on other sites • Find out characteristics of people who use the site AO3: Evaluation e.g. • Can identify areas to focus attention • Save time and money by identifying areas that are not popular/used • New features targeted at specific groups could bring in new business e.g. advertising • But care would need to be applied to privacy issues / GDPR and potential impact on the users <u>Examiner's Comments</u>

Mark Scheme

Question			Answer/Indicative content	Marks	Guidance
			<i>supported by limited evidence and the relationship to the evidence may not be clear.</i> 0 marks No attempt to answer the question or response is not worthy of credit.		Many candidates were familiar with the context of social networking and could provide examples of how data mining could present useful data to assist in facilitating customer connections or targeted advertising. However, on occasion, a number of candidates did forget to actually define what they actually meant by data mining or illustrate the type of data that could be collected for processing. A pleasing number of candidates did go on to discuss and evaluate the computational resources that would be required and/or the impact on users' personal liberties and freedoms that needed to be weighed against the potential benefits to the company.
			Total	9	

Mark Scheme

Question			Answer/Indicative content	Marks	Guidance
8	a	i	Turning large quantities of data into useful information / Finding patterns within large quantities of information	1 AO1.1 (1)	Must refer to large quantities of data <u>Examiner's Comments</u> Whilst many candidates had some familiarity with the term data mining there were a number of very vague answers. Candidates need to know that very large volumes of data are collected with the intention of finding patterns and trends.
		ii	1 mark per identifying data, 1 for use e.g. - Identify customer trends - To identify items to sell/offers to send customers - Identify which stores are making the most profit - To identify what the other stores are doing well - Which items are not selling well - To replace them with other items	4 AO2.2 (4)	Accept any valid responses <u>Examiner's Comments</u> Many candidates could give a good response that explained how data trends identified could be applied in a sales scenario, but some candidates only gave answers related to stock control which was not relevant.
	b	i	Simulate/test the behaviour of the system before it is used	1 AO1.1 (1)	
		ii	e.g. - Testing it with a large number of simultaneous orders (stress testing) - Testing it with a large number of customers/items/orders	1 AO2.2 (1)	<u>Examiner's Comments</u> Many candidates had clearly not come across the term 'performance modelling' which is in the specification, and they struggled to relate it to the context given. However, some very good responses were given, that included the use of Big O notation analysis and mathematical modelling to determine algorithmic performance.
	c		1 mark per bullet to max 2 e.g. - the components can be used in a future program... - they do not need to be rewritten / saves time - they have already been tested... ...it will save time	2 AO1.1 (1) AO2.1 (1)	<u>Examiner's Comments</u> Whilst many candidates could identify that reusable components would save time in future projects, fewer could identify specific reasons why this would be the case.

Mark Scheme

Question			Answer/Indicative content	Marks	Guidance
			Total	9	