Data Representation

1.3 Data Storage and file compression

Marking Scheme

Q1)

(c) Any **three** from:
- company calculation is based on 1 GByte = 1000 MByte
- so (500 × 1000)/8 = 62 500 files
- customer calculation based on 1 GByte = 1024 MByte
- so (500 × 1024)/8 = 64 000 files
- giving the difference of 1500 files

[3]

Q2)

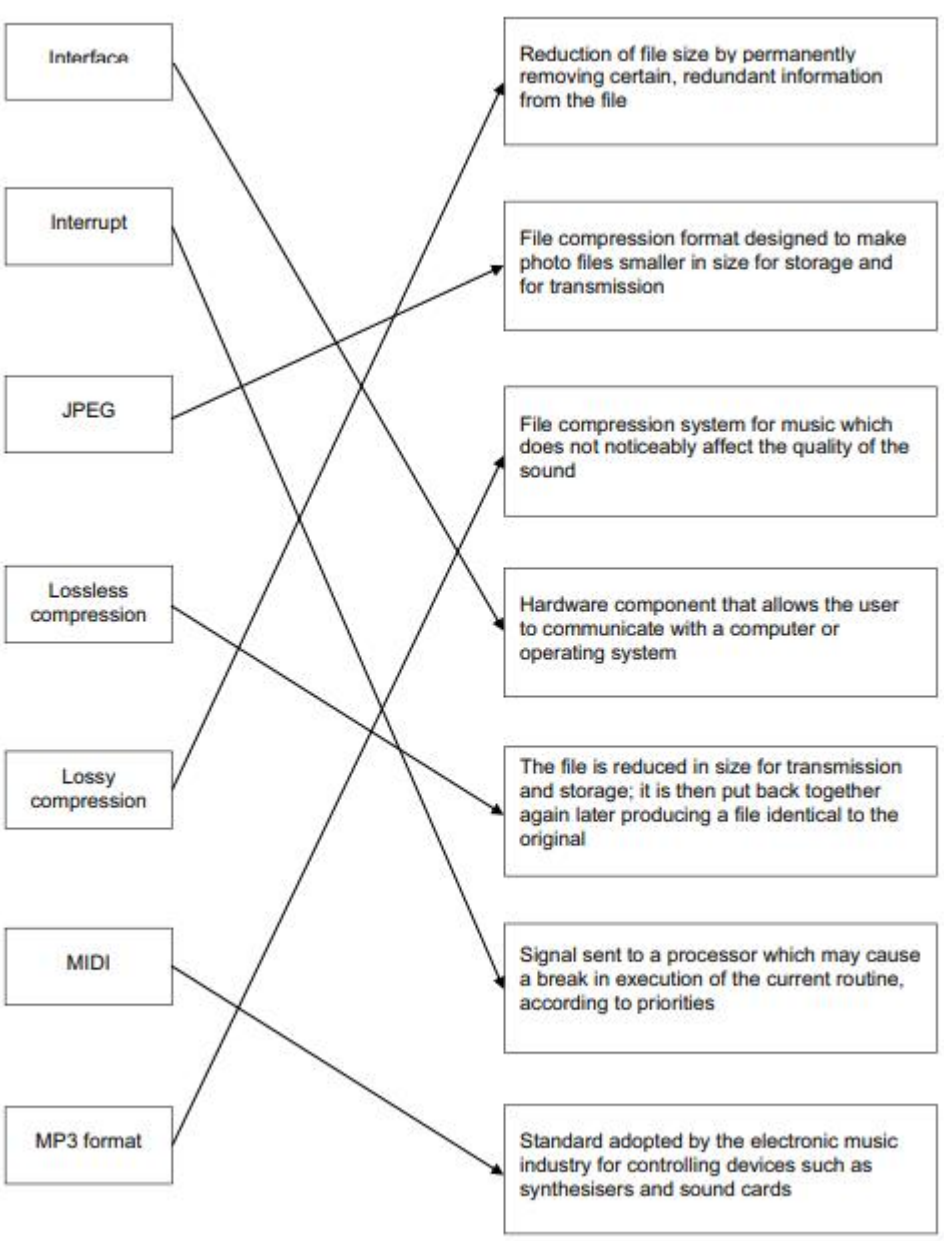
(a) Any **one** from:
- verification is being described
- validation is when data follows a set of rules, e.g. length/range/type check

[1]

(b) Any **one** from:
- send as JPEG files
- carry out a file compression first

[1]

Q3)


Interface
Interrupt
JPEG
Lossless compression
Lossy compression
MIDI
MP3 format
Reduction of file size by permanently removing certain, redundant information from the file
File compression format designed to make photo files smaller in size for storage and for transmission
File compression system for music which does not noticeably affect the quality of the sound
Hardware component that allows the user to communicate with a computer or operating system
The file is reduced in size for transmission and storage; it is then put back together again later producing a file identical to the original
Signal sent to a processor which may cause a break in execution of the current routine, according to priorities
Standard adopted by the electronic music industry for controlling devices such as synthesisers and sound cards

[6]

Q4)

(a) – Memory card / SSD / HDD / magnetic tape
– Suitable description of device given [2]

(b) 2 hours = 120 minutes
120 × 180 = 21 600
21 600 / 1024 (or 21 600 / 1000)
= **21.1 GB (or 21.6 GB)**

(1 mark for correct answer and 1 mark for correct calculation) [2]

Q5)

(a) 8 MB
100 [2]

(b) **(i)** Any **two** from:
– removes sounds human ear can't hear very well
– if two sounds played at same time, softer sound removed
– uses perceptual music shaping [2]

(ii) Lossy [1]

(iii) **One** from, for example:
– jpeg
– MP4
– zip
– gif [1]

Q6)

Question	Answer	Marks
	1 mark per correct tick	4

Statement	true (✓)	false (✓)
47KB is larger than 10MB.		✓
250bytes is smaller than 0.5MB.	✓	
50GB is larger than 100MB.	✓	
1TB is smaller than 4GB.		✓

Q7)

Question	Answer	Marks
(a)	**Two** from: ∞ Smaller file to transmit ∞ The file is transmitted quicker ∞ Uses / requires less bandwidth	**2**
(b)(i)	∞ Lossless (compression) … ∞ … It is important the code must be (exactly) the same as the original file ∞ … If it does not match the original file it will not work	**3**
(b)(ii)	∞ Lossy (compression) … ∞ … It would make the file smaller than lossless compression / the file would stream faster than lossless compression ∞ … The quality of the video can be reduced but it can still be viewed	**3**

Q8)

Question	Answer	Marks
	1 mark for each correct file format e.g.	**3**

File type	File format
Pictures	.JPEG
Text	.doc, .txt, .rtf, .docx, .odt .pdf
Sound	.mp3, .wav, .aif, .flac, .mid
Video	.mp4, .flv, .wmv

Q9)

Question	Answer	Marks
(a)	1 mark for correct calculation method, 1 mark for correct answer: ∞ 2048/1024 (or 1024 × 2) ∞ 2 GB	**2**

Q10)

Question	Answer	Marks
(a)	1 mark for each correct answer: Lossy (compression) Lossless (compression)	2
(b)	1 mark for correct compression, 3 marks for description: – Lossless (compression) Any **three** from: – The file can be restored/decompressed to the exact same state it was before compression/ to original – (It is a computer program so) no data can be lost // Lossy would remove data – Will not run correctly (with any other compression) – (Lossless) will give repeating words/sections of word a value// RLE is used // Other valid examples of methods of lossless compression – Value is recorded in an index	4

Q11)

Question	Answer	Marks
	1 mark for each unit, in the given order: – nibble – byte – megabyte (MB) – gigabyte (GB)	4

Q12)

Question	Answer	Marks
(a)	Any **four** from: – Image is converted from analogue to digital (using ADC) – Image is turned into pixels – Each pixel is given a binary value – Pixels form a grid (to create the image) – Each pixel has a colour – Pixels are stored in sequence (in a file) – Meta data is stored (to describe the dimensions/resolution of the image) // It stores the dimensions/colour depth .etc. – An example of a suitable photo file format e.g. JPEG	4
(b)	1 mark for correct compression, 3 marks for explanation: – Lossy Any **three** from: – Lossy would reduce the file size **more** (than lossless) – The **redundant** data can be removed from the files // by example (must be about redundant data) – Images can still be a similar quality – There is no requirement for the files to be exactly the same as original file – Photos can be sent quick**er** // fast**er** to upload // fast**er** to download	4

Q13)

Question	Answer	Marks
(a)	1 mark for each correct line (to a maximum of 3) **File format** → **File type** .jpeg → Image file .mp3 → Audio file .mp4 → Video file .txt → Text file	3
(b)	2 marks for working, 1 mark for correct answer ∞ 150*100 = 15 000 ∞ 15 000/1024 ∞ 14.65kB	3
(c)	**Three** from: ∞ a compression algorithm is used ∞ no data is lost in the process ∞ repeated words/patterns can be indexed // repeated sections of words/patterns can be indexed // given by example ∞ The indexed words/patterns can be replaced with numerical values // given by example	3

Question	Answer	Marks
(d)	1 mark for each correct tick (✓)	4

File format	Lossy (✓)	Lossless (✓)
.jpeg	✓	
.mp3	✓	
.mp4	✓	
.zip		✓

Q14)

Question	Answer	Marks

	Five from:
	• A (compression) algorithm is used
	• No data is removed in the process // original file can be restored
	• **Repeated** words (are identified) // **Patterns** in the data (are identified)
	• … and are indexed/put into a table // by example
	• … and are replaced with their index // by example
	• … and their positions are stored (in the table) // by example
	• … and the number of times the word/pattern appears is stored (in the table) // by example
	
	NOTE: Other valid methods of lossless compression can be awarded marks
		5

Q15)

Question	Answer	Marks
	One mark for each correct tick **Statement** / **True (✓)** / **False (✓)** 25 kB is larger than 100 MB /	✓ 999 MB is larger than 50 GB /	✓ 3500 kB is smaller than 2 GB / ✓ / 2350 bytes is smaller than 2 kB /	✓	**4**

Q16)

Question	Answer	Marks
(a)	**Four** from: ∞ A compression algorithm is used ∞ Discards any unnecessary sounds … ∞ … using perceptual musical shaping ∞ … such as removing background noise / sounds humans can't hear // or other suitable example ∞ Reduces sample size / resolution // by example ∞ Reduces sample rate // by example ∞ Sound is clipped ∞ The data is permanently removed	**4**
(b)(i)	**One** from: ∞ The file size will be smaller than lossless ∞ Requires less storage space ∞ Requires less time to transmit	**1**
(b)(ii)	**One** from: ∞ The quality of the sound will be reduced ∞ The original file cannot be restored	**1**

Question	Answer	Marks
(c)(i)	**Four** from: ∞ Musical Instrument Digital Interface file ∞ Stores a set of commands / instructions for how the sound should be played ∞ Does not store the actual sounds ∞ Data in the file has been recorded using digital instruments ∞ Specifies pitch of the note // specifies the note to be played ∞ Specifies when each note plays and stops playing // Specifies key on/off ∞ Specifies duration of the note ∞ Specifies volume of the note ∞ Specifies the tempo ∞ Specifies the type of instrument	**4**
(c)(ii)	**Four** from: ∞ It uses a single wire … ∞ … therefore, it is cheaper to manufacture / buy / install ∞ … therefore, less likely to have interference // no crosstalk ∞ … therefore, can be used over longer distances ∞ Data is sent a bit at a time … ∞ … therefore, less chance of data being skewed // data is received in order ∞ Transmission can be synchronised … ∞ … can reduce rate of errors	**4**

Q17)

Question	Answer	Marks
(a)	**Four** from: ∞ A compression algorithm is used ∞ Discards any unnecessary sounds … ∞ … using perceptual musical shaping ∞ … such as removing background noise / sounds humans can't hear // or other suitable example ∞ Reduces sample size / resolution // by example ∞ Reduces sample rate // by example ∞ Sound is clipped ∞ The data is permanently removed	**4**

Q18)

Question	Answer	Marks
	Four from (Max **two** per format): MIDI – Musical Instrument Digital Interface (file) – Stores a set of commands/instructions (for how the sound should be played) – Does not store the actual sounds – Data in the file has been recorded using digital instruments // produced by synthesizer – Specifies pitch of the note // specifies the note to be played – Specifies when each note plays and stops playing // Specifies key on/off – Specifies duration of the note – Specifies volume of the note – Specifies the tempo – Specifies the type of instrument – Individual notes can be edited MP3 – MP3 is a format for digital audio – MP3 is an actual recording of the sound – MP3 is a (lossy) compression format – Recorded using a microphone	**4**

Q19)

Question	Answer	Marks
(a)(i)	**Four** from: – (Compression) **algorithm** is used – No data will be removed // original file can be restored – Example of type of algorithm that would be used e.g. RLE – Repeated patterns in the music are identified – … and indexed NOTE: If another lossless method is described, marks can be awarded.	**4**
(a)(ii)	Any **one** from: – To provide the highest quality of music file (that compression will allow) – The user is able to listen to the original sound file – No loss of quality for the sound file provided	**1**
(a)(iii)	Any **one** from: – Allow for quicker streaming speed – Would not require as much bandwidth (to stream) – Does not need as much RAM – Smoother listening experience // less lag – Will not use as much of data allowance	**1**

Question	Answer	Marks
(a)(iv)	**Two** from: – Streaming speed may be slower – … and may affect listening experience // buffering may occur – User may need more bandwidth to stream – … that could be more expensive – It would be a larger file size – … so may take longer to upload – … so will take up more storage space … – … on webserver	**2**

Q20)

Question	Answer	Marks
.(a)	Any **three** from: – A compression algorithm is used – Redundant data is removed – Reduce colour depth – Reduce image resolution – Reduce sample rate – Reduce sample resolution – Reduce frame rate – Use perceptual music shaping – Data is **permanently** removed	3
.(b)	Any **two** from: – Lossy decreases the **file size** more – Take up less storage space on webserver/users' computer – Quicker to upload/download – May not need to be high quality – Website will load faster for users – Less lag/buffering when watching – Takes up less bandwidth to download/upload – Uses less data allowance	**2**

Q21)

Question	Answer	Marks
(a)	**Two** marks for any **two** correct workings and **one** mark for the correct answer. Working: – 100 × 50 = 5000 bits – 5000 × 8 = 40,000 bits – 40,000 / 8 = 5,000 bytes – 5,000 × 10 = 50,000 bytes – 50,000 / 1024 Answer: 48.83 kB // 49 kB **NOTE:** Alternative correct methods of working can be credited. Answer can be given to any number of dp.	**3**
b)	**One** mark per correct method, **two** marks per justification. – Lossless – Lossy would remove data **permanently** // lossless would not remove any data **permanently** // File could be restored to original … – … that could affect the quality (lossy) // … to maintain the quality (lossless)	**3**
(c)	– Light – Lens – Charge-coupled – Analogue-to-digital – Pixel	**5**

Q22)

Question	Answer	Marks
(a)	Any **three** from: – Password – Add a biometric device to the laptop // set biometric password – Use two-step verification // Use two factor authentication – Physically lock the laptop away in a secure cupboard // Taking laptop with him at all times	**3**
(b)(i)	Any **three** from: – A compression algorithm is used – The resolution could be reduced – **Colour** depth could be reduced // bits per pixel reduced – **Sounds** not heard by human ear could be removed // Perceptual music shaping can be used – Repeating frames could be removed	**3**
(b)(ii)	Any **one** from: – Quality may be reduced – Data is lost // **original** file cannot be reconstructed	**1**
(c)(i)	Any **one** from: – Maintains quality // quality better than lossy – Original file is retained // Data is not **permanently** lost – A significant reduction in file size is not required	**1**

Question	Answer	Marks
(c)(ii)	Any **two** from: – Takes more time to transmit file // Takes more time to upload **to web server** // Takes more time to download **to customer** // Web page will load slower – Takes up more **storage** space – Data usage would be increased – Uses more bandwidth	**2**

Q23)

<table>
<tr><th>Question</th><th>Answer</th><th>Marks</th></tr>
<tr><td>(a)</td><td>One mark for the correct tick
<table><tr><th>File Size</th><th>Tick (✓)</th></tr><tr><td>999 kB</td><td></td></tr><tr><td>1 MB</td><td>✓</td></tr><tr><td>850 000 bytes</td><td></td></tr></table></td><td>1</td></tr>
<tr><td>(b)</td><td>One mark for the correct tick
<table><tr><th>File Size</th><th>Tick (✓)</th></tr><tr><td>4000 MB</td><td></td></tr><tr><td>2 GB</td><td>✓</td></tr><tr><td>2 500 000 kB</td><td></td></tr></table></td><td>1</td></tr>
</table>

Q24)

Question	Answer	Marks
(a)	Any **four** from: – Creates an executable file – … so, would not release source code – … so, the source code cannot be stolen/edited. – … so, would not need to be translated every time // so, translator is not required – … making it machine independent	4
(b)(i)	Any **three** from: – Compression algorithm used – …, e.g. RLE – Repeating frames/pixels are identified – … and are collated/indexed – No data is **permanently** removed – It just records the changes between frames/pixels	3
(b)(ii)	Any **one** from: – Maintains quality // quality better than lossy – Original file is retained // Data is not **permanently** lost – A significant reduction in file size is not required	1

Q25)

Question	Answer	Marks
(a)(i)	• Sound	1
(a)(ii)	• Lossy compressed file	1

Q26)

Question	Answer	Marks
	Two marks for two correct stages of working, **one** mark for correct final answer • 100 × 150 • 15 000 × 16 // 15 000 × 2 • 240 000 / 8 • 30 000 bytes	3

Q27)

Question	Answer	Marks
(a)	• Image	1
(b)	• Lossy compressed file	1
(c)	Any **four** from: • A light is shone onto the surface of the document • The light is moved **across/down/under** the document • The reflected light is **captured** (using mirrors and lenses) • The reflections are converted to binary	4
(d)	• Lossless compression	1

Q28)

Question	Answer	Marks
	One mark for the correct answer • 262 // 250 **Three** marks for three stages of working • 100 × 100 • 10 000 * 16 then / 8 // 10 000 *2 • 20 000 / 1024 or 1000 = 19.5 kB // 20 kB • 5 × 1024 = 5120 // 5 × 1000 = 5000 • 5120 / 19.5 // 5000 / 20	4

Q29)

(b)	• MP4	1
(c)	Any **two** from: • Reduces the size of the **file** • Takes up less storage space • Quicker to transmit to device • Use less bandwidth • Less buffering	2

Q30)

(b)(i)	• It reduces the **file** size	1
(b)(ii)	Any **four** from: • A compression algorithm is used • … such as RLE/run length encoding • **Repeating** words/characters/phrases are identified // Patterns are identified • … and indexed • … with number of occurrences • … with their position	4
(b)(iii)	Any **two** from: e.g. • To save **storage** space • To make it quicker to transmit • To make it small enough to attach to an email • To reduce the bandwidth needed to transmit	2

Q31)

Question	Answer	Marks
(a)	(1) byte	1
(b)	8192	1
(c)	(1) Tebibyte // TiB	1
(d)	**One** mark for each correct stage of working (**max** 2): • 512 × 512 • 262 144 * 2 // multiplied by 16 and divided by 8 • 524 288/1024 **One** mark for the correct answer: 512 (KiB)	3

Q32)

Question	Answer	Marks
(a)	8	1
(b)	2048	1
(c)	nibble	1

Q33)

Question	Answer	Marks
(a)	B	1
(b)(i)	• Lossy • Lossless	2
(b)(ii)	Any **three** from: • The file requires less storage space • Takes less time to transmit • A lower bandwidth can be used to transmit the file • Less data usage for data allowance • More likely to meet file size limits set by email clients/apps	3

Q34)

Question	Answer	Marks
(a)	bit	1
(b)	4	1
(c)	Any **two** from: • 22 016 × 8 then divided 8 • 22 016 × 10 • 220 160 / 1024 **One** mark for: 215 KiB	3
(d)(i)	Reducing the **size** of a file	1
(d)(ii)	It will take up/use less **storage** space	1

Q35)

Question	Answer	Marks
(a)	The file **size** will be reduced	1
(b)	Any **two** from: • It will be under the file size limit for the email • It will be uploaded/transmitted/downloaded faster • It will take less storage space (on computer) • It will use less data allowance (if mobile data used) • Requires less bandwidth	2
(c)	Any **three** from: • Data will be **permanently** removed • … and that could be **important/necessary** data • (The report will have text in it and) lossy is not suitable for **text files** • … as it will **damage/corrupt** the file • The report may have images in it and the quality of these will be reduced	3