Accredited

AS and A LEVEL

Delivery Guide

H046/H446

COMPUTER SCIENCE

Theme: Compression, Encryption and Hashing

September 2015

OCR
Oxford Cambridge and RSA

We will inform centres about any changes to the specification. We will also publish changes on our website. The latest version of our specification will always be the one on our website (www.ocr.org.uk) and this may differ from printed versions.

Copyright © 2015 OCR. All rights reserved.

Copyright
OCR retains the copyright on all its publications, including the specifications. However, registered centres for OCR are permitted to copy material from this specification booklet for their own internal use.

Oxford Cambridge and RSA Examinations is a Company Limited by Guarantee.
Registered in England. Registered company number 3484466.

Registered office: 1 Hills Road
Cambridge
CB1 2EU

OCR is an exempt charity.

CONTENTS

Introduction

Delivery guides are designed to represent a body of knowledge about teaching a particular topic and contain:

- Content: A clear outline of the content covered by the delivery guide;
- Thinking Conceptually: Expert guidance on the key concepts involved, common difficulties students may have, approaches to teaching that can help students understand these concepts and how this topic links conceptually to other areas of the subject;
- Thinking Contextually: A range of suggested teaching activities using a variety of themes so that different activities can be selected which best suit particular classes, learning styles or teaching approaches.

If you have any feedback on this Delivery Guide or suggestions for other resources you would like OCR to develop, please email resourcesfeedback@ocr.org.uk.

KEY

Click to view associated resources within this document.

Click here

Click to view external resources

Curriculum Content

Compression, Encryption and Hashing

a) Lossy vs Lossless compression.

b) Run Length Encoding and dictionary coding for lossless compression.

c) Symmetric and asymmetric encryption.

d) Different uses of hashing.

Thinking Conceptually

Approaches to teaching the content
The overarching theme of this section is the application of algorithms to everyday data processing via the use of patterns. With Lossy vs Lossless compression, learners need to understand that the media that we consume is full of repetitive patterns, eg when 24 frames of video flash by our eyes every second, not everything is changing between the consecutive frames: the background is likely to remain static, so it makes no sense to retransmit it. We can send only the things that have changed and nobody will notice. This is a very common example of lossy compression. Learners often confuse the two, due to them sounding similar to each other. It is important to highlight that the psychological foundations of lossy compression are just as interesting as the algorithms. (A good link is here: File Formats and Compression by JISC Digital Media http://www.jiscdigitalmedia.ac.uk/infokits/).

When discussing Run Length Encoding and dictionary coding, it is important to consider the limitations of both. Run Length is easier to implement in pseudocode and is part of the commonly used JPG image format. Learners should be able to code both algorithms in their choice of a programming language and to process a text of limited length.

Encryption is best demonstrated in practical exercises. Starting with Ceasar cipher might be better for weaker learners, such as those without GCSE Computing experience (where it is discussed in practical assessments). A number of links to free online encryptors and decryptors will help with understanding of what is done to the source text to encrypt it.

A simple example shown in the following link: http://www.online-toolz.com/tools/text-encryption-decryption.php;

A more sophisticated example is shown in the following link: https://www.infoencrypt.com/

Hashing is one of the most common mechanisms for organising data addresses. From storing files on disk, to looking up domain names, to holding passwords; hashing is ubiquitous. It lends itself to a number of practical examples and needs to be contrasted with other ways of organising data.

Common misconceptions or difficulties students may have
The biggest barrier is understanding public and private key encryption (asymmetric). It might take a few lessons to internalise that concept. It is important to consider current affairs in light of Snowden's revelations and recommendations on the use of encryption for personal communication.

Conceptual links to other areas of the specification – useful ways to approach this topic to set students up for topics later in the course.
There is also a cross-curricular link to Maths through the use of prime numbers for encryption. A good illustration is here: Prime Numbers and Public Key Cryptography by Simon Pampena (http://www.youtube.com/watch?v=56fa8Jz-FQQ).

Thinking Conceptually

Here is another: Encryption and HUGE numbers narrated by Dr James Grime, published by Numberphile (http://www.youtube.com/watch?v=M7kEpw1tn50)

There is also a link to Algorithms – the Numberphile video (link above) also contains an algorithm for RSA encryption algorithms.

Hashing has strong links both to the Algorithms (such as File System), as well as the Data Protection Act. In random binary files records addressed are calculated using a hashing algorithm often based on transformation of the record's primary key.

The Data Protection Act holds all collectors of private information to keep the collected data secure, which is accomplished with encryption, amongst other things, like secure physical access or not sending the data outside the EU, etc.

The use of hashing in security is illustrated here: Encoding vs. Encryption vs. Hashing by Daniel Miessler:

(http://danielmiessler.com/study/encoding_encryption_hashing/).

It is particularly important for examining whether the data received has been tampered with, which is further explained here: An Illustrated Guide to Cryptographic Hashes (Steve Friedl's Unixwiz.net http://www.unixwiz.net/techtips/iguide-crypto-hashes.html) and in the following article: Why salted hash is as good for passwords as for breakfast by William Jackson (http://gcn.com/articles/2013/12/02/hashing-vs-encryption.aspx).

ACTIVITIES

Looking at the examples at http://www.vorbis.com/ vs https://xiph.org/flac/ will demonstrate the difference between lossy and lossless audio compression. Audacity can manipulate and convert between these formats.

Snippets of Python code from the Interactive Python book will give a practical dimension to the topics in this section. Encryption can be either done through pupilcoded programs or some of the freely available programs, the links to which are provided here.

It might be interesting to get learners to investigate the difference between the HD and SD quality on YouTube. Filming something and putting it up on YouTube will allow learners to compare the quality and comment on the nature of the compression YouTube uses.

An additional topic for advanced pupils might be the use of vector graphics, such as Flash or SVG to keep low file sizes combined with very high image/video quality. This can be contrasted with raster-based formats.

Thinking Contextually

Activities	Resources
Hashing The corresponding chapter in the Interactive Python book is very helpful in getting some practical coding done for the topic. http://interactivepython.org/runestone/static/pythonds/SortSearch/Hashing.html	Click here
Online converter Use Online converter with various settings: zoom in a viewer, describe, pixilation, colours, contrast, shades, file size. Same for sound. Wired magazine has a good opinion piece on the use of encryption: http://www.wired.com/2014/10/fbi-is-wrong-apple-encryption-is-good/. The teacher could facilitate a debate with the learners around the issues raised in the article.	Click here
Encryption in Python A practical exercise on encryption in Python can follow this very detailed (and industry-strength) video Python Encryption Tutorial with PyCrypto by Sentdex: http://youtu.be/8PzDfykGg_g?list=PLQVvvaa0QuDfhTF3Zfyzc_yD-Mq9iTp4G. If this is too involved, the teacher could give learners a list of 10 numbers ranging between 0 and 100. Using a hash function index=item mod 10, learners could generate a list of addresses that these 10 numbers can be uniquely stored at. **Lossy and lossless compression activity** Provide learners with Learner Resource 1 activity sheet so they can investigate the topic of compression.	Click here to see the clip Learner Resource 1
Using online encryption activity Provide learners with Learner Resource 2 activity sheet to enable them to explore the different types of encryption.	Learner Resource 2

Learner resource 1 Compressing the compression

See page 9

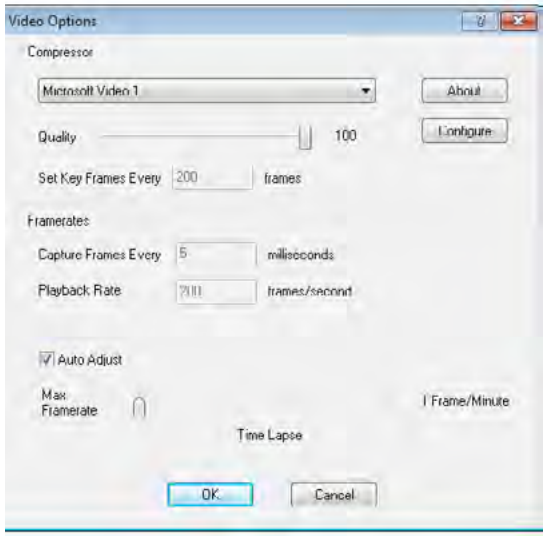
1. **Using screen capture software (like free CamStudio), record a part of your screen. You can move a mouse about, run a slide show, etc.**
   View the captured clip. How does it quality compare with the original? Could you confuse one for the other?

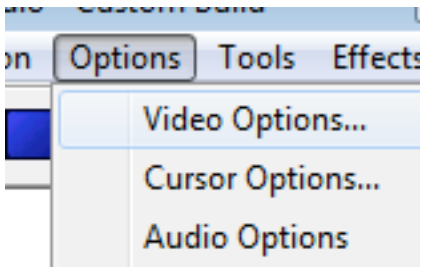
2. **Play the resulting video clip on your chosen media player, while capturing the screen again.**
   Note down the quality in this "second generation". Why do you think it happened?

Learner resource 1 Compressing the compression

See page 9

3. In CamStudio, we can trade-off quality for disk space by adjusting Video Options:





Learner resource 1 Compressing the compression

See page 9

There is also a selection of "codecs" – coders/encoders available in most Windows installations:

- Microsoft Video 1
- Intel IYUV codec
- Cinepak Codec by Radius

Research the names of the codecs above.

What are their strengths and weaknesses?

Are they lossy or lossless?

Why do we need so many of them?

Learner resource 1 Compressing the compression

See page 9

4. **Experiment with the different compression schemes and comment on the trade-off between quality and disk size.**

5. **Research the three sound formats MP3, FLAC and Ogg-Vorbis. Obtain the same recording in both formats and comment on the quality and disk size.**

What are the similarities and differences between these compression formats and the ones for video we have looked at in previous tasks?

Learner resource 1 Compressing the compression

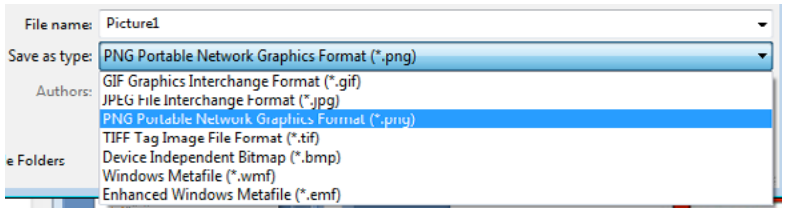
See page 9

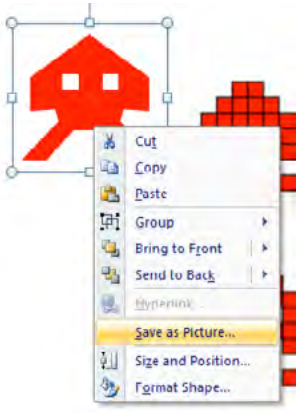
6. Microsoft Powerpoint has an ability to convert between a few image formats.
Find an image, insert it into a slide and then save in a different format.

Comment on the quality and file size.

Try all the formats that it works with, for example:

- GIF Graphics Interchange Format
- JPEG File Interchange Format
- PNG Portable Network Graphics Format
- TIFF Tag Image File Format
- Device Independent Bitmap
- Windows Metafile
- Enhanced Windows Metafile





Learner resource 2 Using online encryption

See page 9

1. **Compare the following websites by putting the same message through their encryption mechanisms. Comment on the differences.**
   a) [Simple] http://encryption.online-toolz.com/tools/text-encryption-decryption.php

   b) [More sophisticated:] https://www.infoencrypt.com/ , https://www.infoencrypt.com/details

Example output:
encrypting "t" with pass "123":
First run:
To decrypt following message use https://www.infoencrypt.com
<Encrypted>
lNL6kd5k+71jC4+EKsnAs3iWhQ9q8Y2effBA3WXBOogPw5ZkKsrMIOJQbpeldsabVjlebLReKwwyHstg
GHVoSbtu6yguyz9vnCAn5SzoRAI=
</Encrypted>

Second run:
To decrypt following message use https://www.infoencrypt.com/
<Encrypted>
0p053oJ3ayCXqjyFZX/kQDg+pl+d3c37K2sJzEyPON4YYuRgFLi4zlHm6LCaLezUu3/+jjnytufVZd/C
z/zEdpxwjIqOFptqga4avkQFsso=
</Encrypted>

Compare with the same on other encryption sites.

Learner resource 2 Using online encryption

See page 9

2. **In one of the above, the encoded message is always the same but in another it changes every time. Look at the help section of the site to find out how they achieve this effect.**

3. **With the website that encrypts differently every time, explain how it always decrypts to the original message, despite the varying encrypted message.**


We'd like to know your view on the resources we produce. By clicking on the 'Like' or 'Dislike' button you can help us to ensure that our resources work for you. When the email template pops up please add additional comments if you wish and then just click 'Send'. Thank you.

If you do not currently offer this OCR qualification but would like to do so, please complete the Expression of Interest Form which can be found here: www.ocr.org.uk/expression-of-interest

OCR Resources: *the small print*

OCR's resources are provided to support the teaching of OCR specifications, but in no way constitute an endorsed teaching method that is required by the Board and the decision to use them lies with the individual teacher. Whilst every effort is made to ensure the accuracy of the content, OCR cannot be held responsible for any errors or omissions within these resources. We update our resources on a regular basis, so please check the OCR website to ensure you have the most up to date version.

© OCR 2015 - This resource may be freely copied and distributed, as long as the OCR logo and this message remain intact and OCR is acknowledged as the originator of this work.

OCR acknowledges the use of the following content:
Thumbs up and thumbs down Alex_White/Shutterstock.com

Please get in touch if you want to discuss the accessibility of resources we offer to support delivery of our qualifications: resources.feedback@ocr.org.uk

OCR customer contact centre
General qualifications
Telephone 01223 553998
Facsimile 01223 552627
Email general.qualifications@ocr.org.uk

For staff training purposes and as part of our quality assurance programme your call may be recorded or monitored.

©OCR 2015 Oxford Cambridge and RSA Examinations is a Company Limited by Guarantee. Registered in England.
Registered office 1 Hills Road, Cambridge CB1 2EU. Registered company number 3484466. OCR is an exempt charity.