UNIVERSITY OF LONDON

GOLDSMITHS COLLEGE

Department of Computing

B. Sc. Examination 2018

IS52020B
Perception & Multimedia Computing

Duration: 3 hours

Date and time:

Audio track *students should answer ALL questions from part A, ONE question from part B, ALL questions from part C, and ONE question from part D.*

Graphics track *students should answer ALL questions from part A, ONE question from part B, ALL questions from part E, and ONE question from part F.*

Information for all students: *Parts A, C, and E each carry 20 marks, and parts B, C, and F each carry 30 marks. The marks for each part of a question are indicated at the end of the part in* [.] *brackets.*

THIS PAPER MUST NOT BE REMOVED
FROM THE EXAMINATION ROOM

Part A

Term 1: Multiple Choice
Answer all questions
(All students)

Answer all questions; only one answer to each question is correct.

Question 1 Which one of the following statements is true about the cones and rods in the human eye?

[2]

(a) The cones have three types of pigments, and is more sensitive to lights than the rods.

(b) The cones have three types of pigments, and is less sensitive to lights than the rods.

(c) The rods have three types of pigments, and is more sensitive to lights than the cones.

(d) The rods have three types of pigments, and is less sensitive to lights than the cones.

Question 2 If the RGB white is defined as (255,255,255), what is the Hue value of the RGB colour (100,0,0), when converted to HSB?

[2]

(a) 0 degree

(b) 90 degrees

(c) 100 degrees

(d) 180 degrees

Question 3 The most common colour blindness is deuteranomaly. People suffer from this condition have:

[2]

(a) Mutated L pigment

(b) Mutated M pigment

(c) Mutated S pigment

(d) No cone cells

Question 4 Which one of the following statements about printing and computer display is true?

[2]

(a) Printing uses the additive colour model, computer display uses the subtractive colour model.

(b) Both printing and computer displays use additive colour model.

(c) Both computer display and printing uses the RGB colour model.

(d) Printing uses the CMYK colour model, computer display uses the RGB colour model.

Question 5 Given two vectors A and B, which of the following operations generates a scalar?

[2]

(a) $A + B$

(b) $A \bullet B$

(c) $A \times B$

(d) $A - B$

Question 6 Which one of the following statements is true about dot and cross products?

[2]

(a) Dot product is usually used to calculate the angle between two vectors, cross product is usually used to calculate surface normal.

(b) Dot product is usulally used to calculate surface normal, cross product is usually used to calculate the angle between two vectors.

(c) Cross product cannot be used to calculate the angle between two vectors

(d) Both dot product and cross product can be used to calculate surface normal

Question 7 If sound A has twice the amplitude of sound B, what is the difference in power (measured in decibels) between sounds A and B?

[2]

(a) 2

(b) 4

(c) $20 \log_{10} 2$

(d) $10 \log_{10} 2$

Question 8 In sound localisation, which statement is true about ITD and ILD?

[2]

(a) ITD works well with low frequency sounds, ILD works well with high frequency sounds.

(b) ILD works well with low frequency sounds, ITD works well with high frequency sounds.

(c) Both ITD and ILD work well with high frequency sounds.

(d) Both ITD and ILD work well with low frequency sounds.

Question 9 When using the sampling frequency of 44k hz, what is the Nyquist frequency?

[2]

(a) 11k hz

(b) 22k hz

(c) 88k hz

(d) 176k hz

Question 10 Which one of the following is lossless compression?

[2]

(a) JPEG

(b) PNG

(c) AAC

(d) MP3

Part B

Term 1: Answer one of the two questions (All students)

Question 11 Animation

(a) This is the pseudo code of a lerp function:

```
vector lerp(vector p0, vector p1, float t)
{
      return p1*t + p0*(1-t);
}
```

What does the lerp function do? What is the range of t?

[4]

(b) Suppose we set two keyframes of an object: it is at height 20 at frame 3, and at height 40 at frame 8. How would you use the lerp function to calculate the height of the object at frame 4 (i.e., write down p0, p1, and t). What will be the return value?

[4]

(c) What is the problem of keyframe animation generated with the lerp function? Discuss in terms of the trajectory and velocity.

[4]

(d) What is a particle? Why is it useful in physical simulation?

[3]

(e) We have introduced two ways to control *motion* in an animation, namely *kinematics* and *dynamics*. Explain the two terms and point out which one is used in physical simulation.

[6]

(f) Briefly explain, in computer animation, what physical simulation is, and give two examples.

[9]

Question 12 Human Perception

(a) In which colour space, distances correlate with perceived colour differences? [2]

i. sRGB
ii. CIE XYZ
iii. CIE LAB
iv. CIE xyY

(b) Pre-attentive Processing is part of: [2]

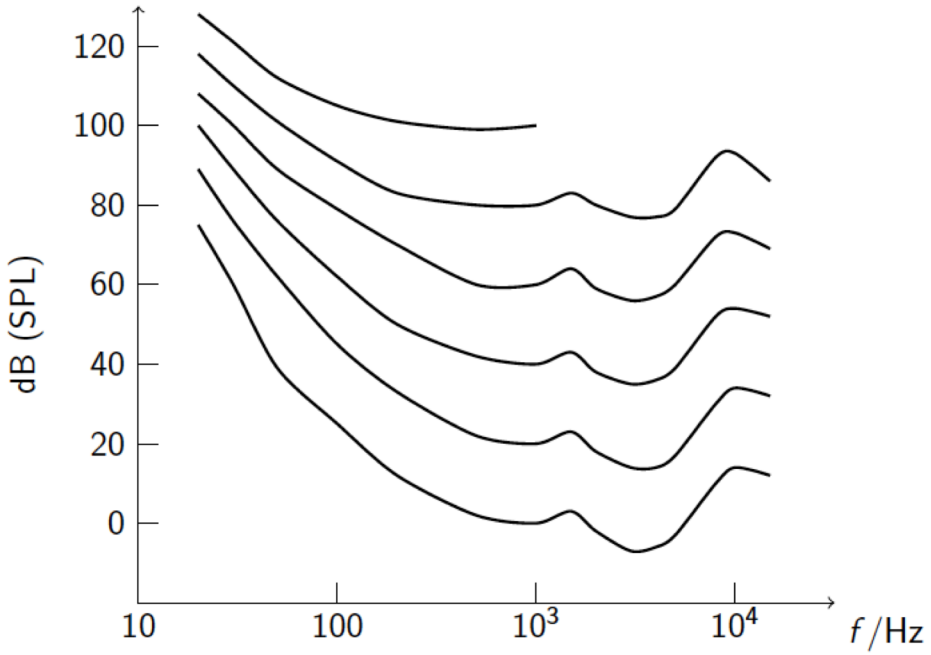
i. Low level visual processing
ii. Pattern perception
iii. Working memory

(c) Gestalt laws can be useful in web design to create more friendly user interfaces. Give an example of a user interface design where this is the case. List and explain how two Gestalt laws could be used in the example. [8]

(d) The frequency of Middle C is 261.6 HZ. What is the frequency of the C one octave above Middle C? And what is the frequency one octave below middle C? [4]

(e) What is the human hearing frequency range? [4]

(f) The figure shows the Fletch-Munson curves. What do the curve show and how this could be useful in application design? Explain it with examples. [10]



Part C

Term 2 Audio: Multiple Choice
Answer all questions
(Audio track students only)

Question 13

Answer all questions; more than one answer to each question might be correct, and all must be given for full credit. [20]

(a) A note is played on a pitched instrument, and its fundamental frequency is 200Hz. Which of the following frequencies are also likely to be present in the sound? Select all that apply.

i. 150 Hz
ii. 250 Hz
iii. 600 Hz
iv. 1000 Hz
v. None of the above

(b) Which of the following would probably be a good distance metric to use to compare two sounds in a sound effect database? Select all that apply

i. Hamming
ii. Levenshtein
iii. Kilometers
iv. Euclidian
v. None of the above

(c) You compute an FFT on 4096 samples of audio. Which statement is true? Select the best answer.

i. The FFT computes 2048 bin values, each of which is a real number
ii. The FFT computes 4096 bin values, each of which is a real number
iii. The FFT computes 4096 bin values, each of which is a complex number
iv. The FFT computes 8192 bin values, each of which is a complex number

(d) Which of the following statements are true about the Fourier Transform? Select all that apply.

i. The time complexity of the Fast Fourier Transform is $O(n \log n)$
ii. The Fast Fourier Transform is fastest when applied to a number of samples which is a power of 2.
iii. When plotting magnitude spectrum output by applying the FFT to an audio signal, frequency will appear on the x-axis of the plot.
iv. When plotting magnitude spectrum output by applying the FFT to an audio signal, time will appear on the x-axis of the plot.
v. The output of the FFT gives us information about both the magnitude and phase of each constituent frequency
vi. None of the above

(e) You are designing a new digital reverb plug-in using the impulse response you have measured for a reverberant cathedral. With this impulse response, you can apply reverb to any new input by doing what? Select only one answer.

i. Delaying the input by one sample

ii. Adding the input to the impulse response of the system

iii. Multiplying the input with the impulse response of the system

iv. Convolving the input with the impulse response of the system

v. None of the above

(f) Which of the following is **not** true about mel-frequency cepstral coefficients? Select just one:

i. They are widely used in speech analysis systems

ii. They provide a lower-dimensional representation than the FFT

iii. They can be used as an analysis feature in music recommendation systems

iv. They are widely used in pitch tracking systems

v. They use the mel scale to model human volume perception

(g) Which of the following statements are true about audio perception? Select all that apply, assume that the waves are audible.

i. Changing the phase of a sine wave will change its pitch.

ii. Increasing the amplitude of a sine wave will increase its volume.

iii. Increasing the frequency of a sine wave will raise its pitch

iv. Changing the shape of a waveform will change its timbre / tone colour.

v. None of the above.

(h) The diagram below shows the magnitude response and the phase response for a digital filter. What type of filter is this?

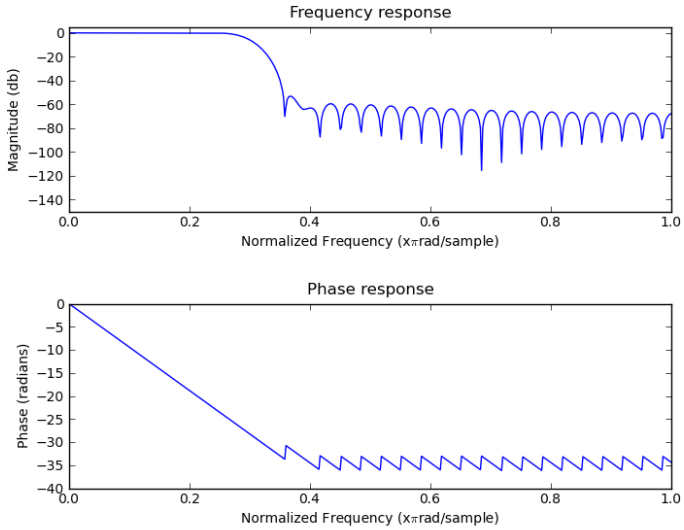
i. A band-pass filter

ii. A comb filter

iii. A low-pass filter

iv. A high-pass filter

v. An all-pass filter



(i) A digital filter has an impulse response $h = [-1, 1]$. What type of filter is it?

i. A low-pass filter

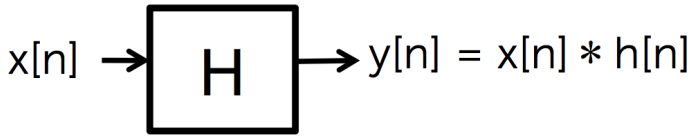
ii. A high-pass filter

iii. A band-pass filter

iv. A band-stop filter

v. An all-pass filter

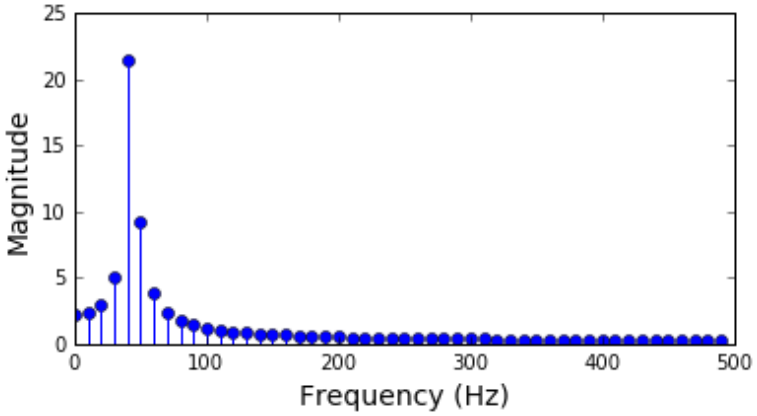
(j) What is H in the diagram below? Select one answer.



i. A signal
ii. An impulse
iii. A system
iv. An FFT
v. A duck

(k) You apply an FFT to a signal that you know is a single sine wave, which does not change in frequency over time. You get the magnitude spectrum below. Which **one** of the actions below is most likely to help you improve your ability to detect the exact frequency of this sine wave?

i. Use a longer FFT (i.e., more bins)
ii. Use a shorter FFT (i.e., fewer bins)
iii. Use the STFT instead
iv. Look at the complex values output by the FFT instead of just the magnitudes



(l) What is the amplitude of the sine wave with the equation below?
$y = 0.6\sin(2\pi 300t + \pi/2)$

i. 0.6
ii. 2
iii. 300
iv. $\pi/2$
v. None of the above

(m) Which of the following are potential advantages of IIR filters over FIR filters? Select all that apply.

i. IIR filters can be applied to both audio and sensor data, but FIR filters can only be applied to audio
ii. IIR filters might never stop making sound
iii. FIR filters can "explode"
iv. IIR filters are usually more computationally efficient

(n) The FFT computes the imaginary value "$5 + 12j$" for a particular bin. What is the magnitude of this bin?

i. 11
ii. 13
iii. 17
iv. j
v. None of the above

(o) Which of the following audio features would be most useful if you wanted to build a search engine to find sounds with a similar timbre/tone colour? Select the one best answer.

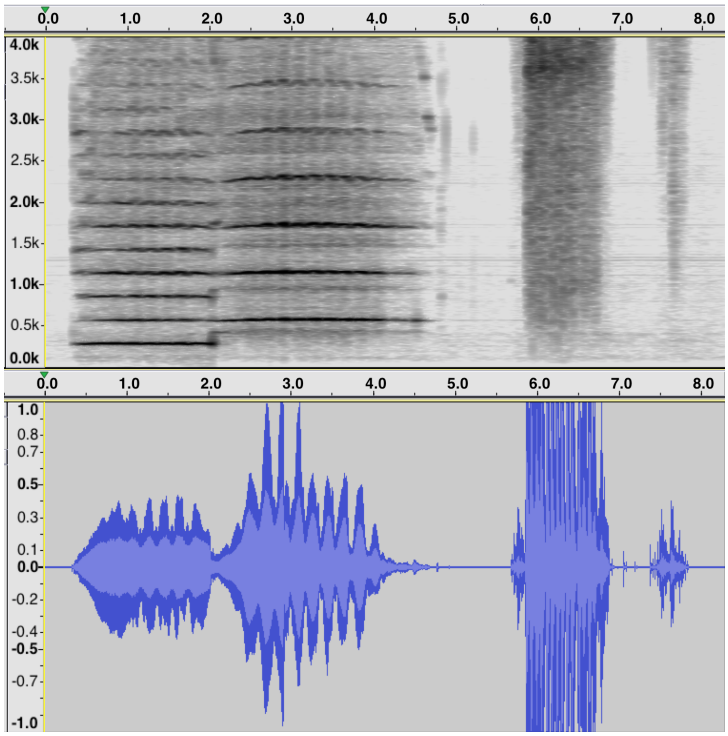
i. RMS
ii. Zero-crossing rate
iii. Fundamental frequency
iv. Pitch histogram
v. Spectral centroid

Part D

Term 2 Audio: Answer one of the two questions
(Audio track students only)

Question 14 Listening and Hearing

(a) The diagram below shows the spectrogram (top) and waveform (bottom) for the same audio signal. Time is on the x-axis, labeled in seconds on the top. Describe as precisely as possible what you will hear if you listen to this sound. For instance, you might comment on possible source(s) of the sound, pitch (if any), volume, tone quality/timbre, the number of sound sources you hear, and how these all change over time. [7]



(b) Musical expectation

i. What factors typically cause listeners to have certain musical expectations when listening to a piece of music? [2]

ii. Give a specific example of how a piece of music might satisfy musical expectations at one moment in time, then break expectations at another moment in time. [2]

iii. Describe why a composer may want to satisfy musical expectations. That is, how does it affect the listener when their expectations are satisfied? [2]

iv. Describe why a composer may want to break musical expectations. That is, how does it affect the listener when their expectations are not satisfied? [2]

(c) What is melody? [2]

(d) What is the harmonic series? Also, why is knowing about the harmonic series relevant to human perception of pitch and timbre, and to understanding consonance and dissonance of musical tones? [6]

(e) You have been asked to create a new music visualiser program, in which the colour of the visualisation changes according to the current musical chord, and the size of the visualisation changes according to the volume. The input to this visualiser will be audio only: that is, you won't have access to MIDI or other data about the music; you must compute chord and loudness information directly from the audio samples.

i. Describe how you would compute information about the current musical chord. For instance, what musical feature(s) would you use, and why? [3]

ii. How well would you expect your computational approach described above to work? For instance, are there circumstances in which it is likely to fail? [2]

iii. What feature(s) would you compute from the audio in order to capture information about the current volume? Why? [2]

Question 15 Audio Signal Processing

(a) Impulse Response

i. What is an impulse response? [2]

ii. Why is it useful to know the impulse response of a system? [3]

iii. What does it mean to say a system is time invariant? [2]

iv. The the system T in the following equation demonstrates superposition:

$$a\mathrm{T}\{\mathrm{x}[n]\} + b\mathrm{T}\{\mathrm{y}[n]\} = \mathrm{T}\{a\mathrm{x}[n] + b\mathrm{y}[n]\}$$

Why is this a useful property of linear systems? [3]

(b) You are given the task of designing a filter to remove a low-frequency hum from a recording of a music concert.

i. What sort of filter might you use for this task? [1]

ii. What steps will you take to figure out how to build a good filter for this task? For example, what will you do to determine good cutoff frequency/frequencies, and to determine a good filter order? Be as specific as you can. [4]

iii. Sketch a realistic frequency response for a filter that you might build for this task. [3]

(c) A waveform has the following equation:

$y = 0.1\sin(2\pi 150t) + 0.3\sin(2\pi 450t) + 0.2\sin(2\pi 700t)$

It is sampled at 4000Hz. Sketch its spectrum, with frequency on the x-axis and magnitude on the y-axis. (Do label the x-axis with specific frequencies, but don't worry about the exact values on the y-axes; just try to get the relative magnitudes approximately correct.) [5]

(d) Consider the Python signal processing example below, which begins by loading an audio waveform from a file into `signal1`.

```
signal1 = wavReadMono("w1.wav")
signal2 = hamming(2048)
tmp = signal1[0:2048]*signal2
fft_result = fft.fft(tmp)
```

i. What, exactly, is `signal2`? [1]

ii. Why, precisely, is it a good idea to multiply `signal1` by `signal2` before computing the FFT? [3]

(e) Let x be a digital signal: $x[n] = 0.5, -1, 1]$. What is the result of convolving x with the impulse response $h = [1, -1]$? Show your work. [3]

Part E

Term 2 Graphics: Multiple Choice
Answer all questions
(Graphics track students only)

Question 16

Answer all questions; more than one answer to each question might be correct, and all must be given for full credit. [20]

(a) Which of the following can be considered the same as a rectangular coordinate system?

i. Cartesian coordinate system
ii. Polar coordinate system
iii. Homogeneous coordinate system
iv. Circular coordinate system
v. Spherical coordinate system

(b) Which of the below generates the X coordinate in a cartesian to polar conversion?

i. $r * sin(\theta)$
ii. x=atan2(y/angle)
iii. $r * cos(\theta)$
iv. angle * cos(radius)
v. Something else

(c) In which of the below graphics operations might you find the following code?

```
x=(zx*zx)-(zy*zy);
y=2*(zx*zy);
```

i. Drawing a square
ii. Drawing a mandelbrot fractal
iii. Plotting the square root of a complex number z where $z=(x,y)$
iv. Getting the distance between to pixels
v. Computing a fractal tree

(d) Why won't the below code calculate the distance between two 3D points properly?
$d(a,b) = (ax - bx)^2 + (ay - by)^2 + (az - bz)^2$

i. Because d is not equal to the distance
ii. Because a and b are not normalised
iii. Because it does not calculate the square root of the squared distance
iv. Because the vertices are being calculated to a power of 2
v. Something else

(e) Given a Square, a Sphere, a Rhombus, a Triangle and a Cube, how many are platonic solids?

i. Five
ii. Four
iii. Three
iv. Two
v. One

(f) What kind of output would the below code generate?

```
m = 100;
n1 = 20;
n2 = 1;
for (var i = 0; i < 100; i++) {

    r1 = mag * Math.pow(Math.pow(Math.abs(Math.
       cos((m * spacing * i) / 4) / a), n1) +
       Math.pow(Math.abs(Math.sin((m * spacing *
        i) / 4) / b), n2), -(1 / n1));

    x = r1 * Math.cos(spacing * i);
    y = r1 * Math.sin(spacing * i);
    context.beginPath();
    context.strokeStyle = "rgb(0,0,0)";
    context.moveTo(lastX, lastY);
    context.lineTo(x, y);
    context.stroke();
    context.closePath();
    lastX = x;
    lastY = y;
}
```

i. It would draw a face
ii. It would draw a cone

iii. It uses a 2D superformula to create geometry
iv. It uses a series of rings of radius r1 to generate a sphere
v. something else

(g) Which of the following statements is true?

i. A vector indicating the direction that a surface should face is called a unit vector.
ii. A point on a surface that should be used as the normal light distribution for other objects in a computer graphics scene is called an inverted normal.
iii. A vector perpendicular to the surface of a given object that can be used to calculate how the object should be lit is defined as a normal.
iv. A function that is used to check whether or not a surface is normal is called a scalar.
v. A vector that represents a normal surface is called a unit normal.

(h) What is a Skybox?

i. A single piece of geometry that can be used to set the ground plane of a 3D scene
ii. A texture for making a 3D world appear more realistic
iii. A method for texturing spheres or cubes
iv. A piece of textured geometry used for representing the environment surrounding a 3D scene
v. something else

(i) In 3D graphics, what do the coordinates S and T commonly represent?

i. The contents of an image being applied to 3D geometry
ii. The coordinates of a texture from a 2D graph
iii. The texture coordinates for a specific piece of geometry
iv. The surface normals of a vector
v. The Unit Vector

(j) If we get the difference V between points B and A, then divide V by its magnitude, what do we get?

i. The distance
ii. The velocity
iii. The unit vector
iv. The texure coordinates for geometry described by A and B
v. Something else

(k) What is the correct term for a global variable that is used in a GLSL fragment shader?

i. A global
ii. A Uniform
iii. A Varying
iv. vec2
v. bvec4

(l) In GLSL, what would the below code do?

```
vec2 position = gl_FragCoord.xy;
```

i. Get the coordinates of the window
ii. Set the coordinates of the current fragment
iii. Get the color of the current fragment
iv. Make the screen go green
v. Get the coordinates of the current fragment

(m) In a fragment shader, which of the below gets a normalised value for the current fragment based on the screen resolution?

i. $vec2p = gl_FragCoord.xy * 0.001$
ii. $resolution.xy = 1$
iii. $vec2p = gl_FragCoord.xy/resolution.xy$
iv. $vec2p = gl_FragCoord.xy/resolution.xy * mouse.xy$
v. $vec3p = vec3(0, 1, 1)$

(n) In a vertex shader, how do we refer to a variable that contains information about a specific vertex?

i. A global variable
ii. An Attribute
iii. A Uniform
iv. A Normal
v. An initialisation

(o) What does the below matrix do if multiplied with a vector containing x,y,z,w values?

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

i. Rotates a 3D vector
ii. Scales a 3D vector
iii. Translates a 3D vector
iv. Nothing
v. Something else.

(p) What does the below matrix do if multiplied with a vector containing x,y,z,w values?

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

i. Scales a 3D vector
ii. Rotates a 3D vector
iii. Translates a 3D vector
iv. Rotates and translates a 3D vector
v. Nothing

(q) Which of the below is the correct order for computing a specific matrix multiplication when manipulating 3D coordinate systems in openGL?

i. rotation after translation
ii. translation before scaling
iii. rotation before translation
iv. scale * rotation * translation
v. something else

(r) Which of the below statements is true?

i. By deforming geometry in the vertex shader, we can create a bump map
ii. By modifying the surface normals of 3D objects, we can create a bump map
iii. By adding noise functions to a texture in a fragment shader, we can create a bump map
iv. None of the above statements are true.

(s) In GLSL, what does the smoothstep function return?

i. Either a vector or a float
ii. An interpolated output between two values
iii. 1 or 0
iv. a logarithmic gradient
v. a time parameter

(t) Which of the below best describes what a Geometry shader does?

i. Creates a textured surface
ii. Generates additional geometry using existing geometry as an example
iii. Subdivides patches of vertex data into smaller primitives
iv. Sets the resolution of the screen
v. Divides the screen up into a grid of triangles.

Part F

Term 2 Graphics: Answer one of the two questions (Graphics track students only)

Question 17 OpenGL : Fragment Shader Programming

Answer the following questions in as much detail as you can.

i. Describe in as much detail as possible a technique for generating a procedural texture in a fragment shader using trigonometry. [5]

ii. You are asked to program a method for generating a surface texture of a planet. What kind of approach will you take? Describe as fully as you can what algorithmic approach you will use, and give an example of how you would implement this in a fragment shader. [10]

iii. What is the main reason for using a Uniform in a fragment shader? [2]

iv. If you have a Uniform called *time* declared in your management code, how would you initialise this Uniform in your fragment shader? [3]

v. Describe as precisely as you can how you might use a time uniform to make a procedural texture more interesting. [5]

vi. Say how you might create a function that allows you to draw a straight line between two points in a fragment shader. [5]

Question 18 OpenGL Shader Programming Exercise: Distances

Distance calculations are essential to the large majority of shader operations. Answer the following questions both by describing relevant processes, providing written GLSL code examples where appropriate.

i. In as much detail as possible, describe a method for drawing a circle using a fragment shader. [10]

ii. Write a paragraph on the simplest method you can think of for creating a 2D lighting effect that can be used to shade a shape from top to bottom. [5]

iii. Describe a method for generating a 3D lighting effect without using a vertex shader. [5]

iv. You are asked to create a sequence of circles in a way that requires the least code possible, and without using a bespoke function. Describe how you will achieve this. [5]

v. You are asked to fill the screen with a series of circles of different sizes and colours by using a function. Describe how you will write a function in order to do this. [5]