COSC 4P98 Topics in Digital Audio

Instructor: B. Ross
Date: March 13, 2014

COSC 4P98 TERM TEST

KEY

NAME (print): ________________________________

STUDENT NUMBER: __________________________

A 3" by 5" card of hand-written (not photocopied or digitally reproduced) notes may be used. A calculator may be used. No other aids are permitted. Please answer all questions on the exam paper. Use the backs of pages if necessary. Keep written answers brief and to the point. Write neatly and legibly. Carefully read all questions!

<table>
<thead>
<tr>
<th>Question</th>
<th>Total</th>
<th>Mark</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>Total:</strong></td>
<td><strong>108</strong></td>
<td></td>
</tr>
</tbody>
</table>
Question 1 [24] Multiple choice: select the best answer to each question

[+2 correct, 0 nothing, -0.5 wrong guess; one free wrong guess permitted without penalty]

1. You define CSound instruments in a:
   a. SCO file
   b. ORC file
   c. MIDI effect
   d. VST file
   [B]

2. This kind of noise can be used for generating musically interesting random note sequences:
   a. pink noise
   b. white noise
   c. grey noise
   d. burnt umber noise
   [A]

3. MIDI files contain:
   a. highly compressed audio recordings
   b. lossless audio samples
   c. harmonics in the audible spectrum
   d. note on/off event data
   [D]

4. Removes high frequencies from an audio signal:
   a. high-pass filter
   b. low-pass filter
   c. comb filter
   d. delay filter
   [B]

5. The decibel scale is:
   a. logarithmic because the human ear is sensitive to ratios of sound power differences
   b. logarithmic because the human ear is sensitive to absolute differences in sound power
   c. parabolic because the human ear is sensitive to fractal patterns
   d. finite because the human hearing range is between 20 Hz to 20 kHz
   [A]

6. An uncompressed 5-second 24-bit stereo recording using 44 kHz sampling rate has a size of approximately:
   a. 1,320,000 bytes
   b. 660,000 bytes
   c. 5,280,000 bytes
   d. 10,560 bytes
   [A]
7. For sample playback, you have a sample table with 1,000 samples of a periodic wave. You are using a sampling frequency of 44,000 samples/sec. You wish to generate a sound having a frequency of 440 Hz. The sample index increment you should use is:

a. 1
b. 10
c. 100
d. 1000

\[ \text{inter } = \frac{C \times \text{freq}}{\text{sam. freq.}} \]

8. Using a similar sample playback system as above, the computed index increment turns out to be "50.4". The most practical way to deal with such fractional indices in order to create an accurate pitch is to:

a. use more bits per sample
b. truncate the index
c. round the index
d. interpolate the index

9. After analyzing examples of music, Clark and Voss found that the frequency differences between successive notes conformed to this kind of distribution:

a. \(1/f\) distribution
b. uniform (random)
c. Bell curve
d. C-major chord

10. Granular synthesis needs to create an envelope around each grain:

a. in order to isolate specific sounds
b. to remove sudden amplitude changes at start/end of grains, which causes distortion
c. to simulate sine wave harmonics
d. because grain clouds have soft edges

11. Grammars can be used in music to:

a. formally model musical structure
b. generate music via generative grammar interpretation
c. parse musical passages in order to ascertain their "correctness" to rules
d. all of the above

12. One could use a genetic algorithm to compose sequences of notes in a melody. Fitness evaluation might conceivably be done via:

a. goodness of fit to a \(1/f\) noise distribution model
b. grammar-based model analysis
c. neural network analysis (trained earlier)
d. interactive human listener who rates passages
e. all of the above

\[ \text{any answer is OK} \]
Question 2 [15] Define each term. Include a brief discussion of each as well.

a) periodic wave

repeating wave, repeat over a time period $T$

$\therefore \quad f(t) = f(t + T)$

b) Quantization noise

- round-off or aliasing error due to different amplitude and bit-resolution of samples
  - "his" / "distortion"

c) Short-time Fourier Transformation (STFT)

- apply Fourier analysis to a short interval $T$
  - number of points
- apply window (cosine, etc.)
- overlap to connect for envelope
- used to show changes in music over time

d) Nyquist frequency

- $\frac{1}{2}$ the sampling rate
- highest frequency that can be sampled at sampling rate

- should not be exceeded

- e) LFO

- low-frequency oscillators
  - sine, triangle, square
  - sub-sustainer frequency (≤ 30 Hz?)
  - used to control eq, frequency, modulation, mix, etc.
a) [12] Draw a diagram showing the complete processing for sampling and playback of digital audio. Include a descriptive discussion of all the major steps in the diagram.

Mike: - currentleichung:
  - umwandlung von audio signals (voltage)
  - pro-amp: linearizes the undriven voltage
  - DAC: digital-analog converter
    - converts source voltage to number at sampling freq

- Impair: roughness of voice
  - Smooth: smooth, roughness of voice
  - Loud speaker to speaker level
  - Speaker: display wave air to generate sound

\[ \frac{1}{2} = \text{label/figure} \]
\[ \frac{1}{2} = \text{short description} \]
b) [4] Human hearing has a dynamic range of approximately 120 dB. Illustrate how 16-bit audio is inadequate for capturing audio for this dynamic range.

\[ 16 \text{ bit gain} \quad 2^{16} = 65536 \text{ levels of amplitude measured} \]

\[ 20 \log_2(2^{16}) = 96.3 \text{ dB} \]

which is less than 120 dB of human ear.

c) [6] Audio equipment with high sampling rates of 192 kHz and higher are controversial. Briefly discuss 2 points against the usefulness of such high sampling rates.

- Higher frequency air below 22 kHz so that kHz is far above minimum audible rate (44 kHz)

- 192 kHz means more data
  - Storage
  - Bandwidth
  - System H/W
Question 4 [30] Consider the discrete Fourier transformation formula:

\[ f(t) = \sum_{n=1}^{T/2} \left( a_n \cos\left( \frac{2\pi nt}{T} \right) + b_n \sin\left( \frac{2\pi nt}{T} \right) \right) + a_0 \]

\[ a_0 = \sum_{k=0}^{T-1} \frac{x[k]}{T} \]

\[ a_n = \left( \sum_{k=0}^{T-1} x[k] \cos\left( \frac{2\pi nk}{T} \right) \right) \cdot \frac{2}{T} \]

\[ b_n = \left( \sum_{k=0}^{T-1} x[k] \sin\left( \frac{2\pi nk}{T} \right) \right) \cdot \frac{2}{T} \]

a) [6] Define / describe the following terms:

- **t**: time (continuous, approx)
- **T**: total # samples (Time)
- **x[k]**: sample at order k (time “k”)
- **n**: harmonic #
- **a_0**: average angle value (or coeff of cos term at n=0)
- **f(t)**: amplitude of wave at time t

b) [4] Given the above equations, what is the frequency of the kth harmonic (ie. n=k)? (You don’t need to give the exact frequency in kHz, but rather, a description of what it is, given the information available)

\[ k \times \text{fundamental freq (ie of } f(t) \text{), or } n=1 \]
c) [6] A Fourier Series is said to be the sum of sine waves, each with an amplitude and phase shift. Define “phase shift”. Explain how it is represented in the above equations. How can the phase shift angle be calculated from the DFT equations? Also mention how you can tell whether there is no phase shift arising for a DFT-transformed audio signal.

\[ \sin(2\pi n t + \phi_n) \]

- phase shift: a time delay, namely

\[ \theta_n \] - represented by use of cos and sin terms

- cos & sin are "phase shift" terms; if \( \phi_n \neq 0 \), a phase shift exists

- angle calculator: \( \phi_n = \tan^{-1} \left( \frac{a_n}{b_n} \right) \) for each \( n \).

- no phase shift: all \( a_n = 0 \)

d) [6] What is the difference between the Fourier Series (FS) and Discrete Fourier Transformation (DFT)? Mention an advantage of each.

- FS: continuous time
- DFT: discrete time

- adv FS: - exact solution
- adv DFT: - easy to compute
  - works well with samples (digital audio)
f) [2] Why do we not need to compute more than T/2 harmonics?

Because we need at least one full cycle of samples to define a signal of frequency f.

- not enough samples to define any f > \( \frac{1}{2} \) f

(Same principle as Nyquist frequency, but applied to sample table + FFT.)

appendix!! \( e^{ix} \) (my bad!)

---

g) [2] The DFT is derived to work with periodic waves, rather than aperiodic waves. What assumption lets us use it with periodic waves?

- Simply assume aperiodic wave repeated!

(And it as periodic)

---

h) [4] What does the acronym FFT stand for? What is the relationship of the DFT and FFT?

1. Fast Fourier Transform
2. FFT is a more efficient implementation of DFT
   \( O(n \log n) \) vs \( O(n^2) \)

FFT
DFT
Question 5 [16]


- Synthesis technique that tests speed as short sections of grains
- Grains can be created and controlled in variety of ways
- With sample table, grains chunks of sample table, envelope gain (to prevent distortion), and play
- Can move forward/backward in table
- Change grain pitch

b) [8] Identify and discuss 4 parameters that can be defined to control aspects of granular synthesis.

- Duration of grain
- Pitch of grain
- Location of grain in table
- Duration of grain repeated
- Frequency of grain variation (density)
- Stereo grains

...lots!!
Question 5 (cont)

c) [4] Describe a way in which you might use a graphical visualization to define parameterization changes for a granular performance.

- lot of possible

\[ \text{eq 1: } 30 \text{ surface: map } x,y,z \text{ to 3 parameters} \]

- can control cloud

\[ \text{eq 2: } 20 \text{ control envelope} \]

\[
\begin{array}{c}
\text{part} 2 \\
\text{time}
\end{array}
\]

\[ \Rightarrow \text{scm param representation w/ time}. \]

(I accepted GUI design answers, although I was thinking of param. changes over time).

Question 6 [1] Briefly describe what you are thinking of doing for the course project.

\[ ? \Rightarrow (\text{any answer}) \]

*** The End. ***