15-323/623 Spring 2019
Homework 4
Due April 5

1. Audio Concepts

1. What is a unit generator?
2. Why does audio software process blocks rather than samples?
3. What is an audio plug-in?
4. The WinMM API on Windows has a high latency, typically 100ms or more. Typically, audio is delivered to WinMM in buffers with 1 to 20 milliseconds of audio, so Windows must be queuing up *many* buffers of audio to result in such a long latency. Why would Windows do this? What would happen if the buffer queue is shorter?
5. Suppose you implement an audio system with a circular buffer and discover that while the *computation time is negligible*, the process that computes audio is not run by the operating system for up to 50ms after it becomes ready to run (your process becomes ready to run when space is available in the audio output buffer for more samples). If the sample rate is 44,000 samples per second, how many samples should be in the output buffer that holds samples until the DAC can convert them? (Another way to look at this is as follows: We want enough samples in the output buffer to allow audio playback to continue for 50ms. How big is the buffer?)
6. Assume that you have a double buffer where the block size is 64, the number of sample periods to process each block is 37, and the maximum pause is 70 sample periods. What is the total audio latency (in sample periods)?
7. Assume that you have a queue of buffers (of size 64 frames), but the other parameters are the same as above. What is the total audio latency (in sample periods)?

2. Dynamic Programming

1. Complete this table as in the slides, where the number in each cell is the #matches - #insertions - #deletions - #substitutions.

melody:		key:	G	F	A	D
		0	-1	-2	-3	-4
C	0					
A	0					
D	0					
G	0					
F	0					
A	0					
D	0					
G	0					

Circle the cell of the last column with the largest value.

2. Complete this table, but compute the edit distance: #insertions + #deletions + #substitutions. Note that the "row 0" values (1 2 3 4) are increasing because each skipped note of the key (or query) is considered a deletion. Also note that this is now a measure of *distance*, so smaller is better.

melody:		key:	G	F	A	D
		0	1	2	3	4
C	0					
A	0					
D	0					
G	0					
F	0					
A	0					
D	0					
G	0					

Circle the smallest (best) value in the right-hand column.

3. Readings

1. In the paper "Towards the Digital Music Library: Tune Retrieval from Acoustic Input," what number of notes of melody are required to uniquely identify a tune in their experiment with 10,000 melodies in the database? (State a range, since it depends on the representation.)
2. What representation requires the least number of notes for a unique match?

3. What representation requires the most notes for a unique match?
4. Give and example of what the authors call the "contour" representation.
5. In the paper "Making Sounds with Numbers…," what are three steps performed by Analysis/Resynthesis packages?
6. We know that computing in blocks is more efficient than sample-at-a-time computation. According to the paper, what are two disadvantages of block-based computation?