University of Pennsylvania
Department of Electrical and System Engineering
Digital Audio Basics

ESE1500, Spring 2023 | Final | Friday, May 5

- Exam ends at 5:00PM; begin as instructed (target 3:00PM)
- Do not open exam until instructed to begin exam.
- Problems weighted as shown.
- Calculators allowed.
- Closed book = No text or notes allowed.
- Provided reference materials on next to last page.
- Show work for partial credit consideration.
- Unless otherwise noted, answers to two significant figures are sufficient.
- Sign Code of Academic Integrity statement (see last page for code).

____

I certify that I have complied with the University of Pennsylvania's Code of Academic Integrity in completing this exam.

Name: Solution

1			2	3	4		5			6		7			8
a	b	c			a	b	a	b	c	a	b	a	b	c	
2	5	5	12	12	6	6	3	5	5	8	4	5	3	4	15

1. Robots communicate in sequences of multi-frequency tones. Use a combination of 16 tones at 100 Hz intervals between 1000Hz and 2500Hz. The robot can send any subset of the 16 tones simultaneously. A 500 Hz tone is also sent simultaneously to indicate a communication is in progress. Tones are sent in pulses with 5ms pulse of tones and 5ms of silence before the next pulse of tones.

   (a) How many bytes per second can be sent with this scheme?

   16b/10ms = 1600b/s = 200 B/s

   (b) What is the minimum sampling rate to recover the tones?

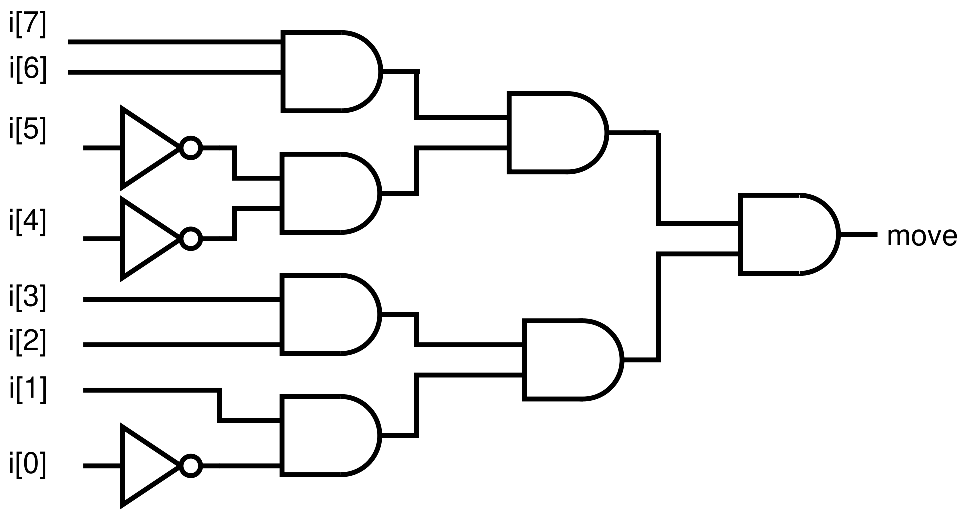
   5000 Hz – Nyquist sampling at twice the maximum frequency of 2500 Hz

   (c) Using the techniques from class, at the identified sampling rate, how many multiply accumulates/second are required to decode presence/absence of all 16 tones and the communication-in-progress tone?

$5000 \times 2 \times 17 = 170{,}000$ multiply-accumulates/second

5000 samples, sine and cosine term per frequency, 17 frequencies to extract

2. The first byte of a robot message indicate its type, similar to an instruction opcode. The following logic is used to decode if a message is a movement message. What 8b binary value is it selecting?



1101110

3. Consider a UI to task a robot to go from its current location to a destination. There may be obstacles or limiting terrain between the current source and the destination that the robot must avoid. Rank the following in order of usability (1 – easiest for humans, 4 – hardest) and identify the strenghths and weakness of each for human usability.

- [3] User provides the final GPS coordinates for the destination by filling in two text boxes with lattitude and longitude and pressing a [submit] button.

  Lattitude: [ ] Longitude: [ ] [Submit]

  weak: tedious, error prone, maybe obscure to have GPS coodinates and type in

  Strength: only type 2 numbers; not burden user with intermediate route, only destination

- [2] User is shown a map and uses a mouse or touch pad to select the final destination on the map.

  Strength: Only point and click.

  Weak: maybe hard to pinpoint on map? so get low precision.

  Accepted this as most usable, too.

- 4 User provides a sequence of moves by selecting a distance in meters and a heading in radians. Input is a series of text box pairs the user completes with a [submit] butteo at the bottom. Before the submit button is a button to add [More moves] in sequence.

1	Distance:		Heading:	
2	Distance:		Heading:	
3	Distance:		Heading:	
4	Distance:		Heading:	
5	Distance:		Heading:	
		More moves		Submit

  Weak: Tedious, lots of typing, high cognitive load as user must compute all steps in route.

- 1 User points at the destination with a dedicated laser pointer and presses a button on it to select the destination.

  Strength: easy point and select based on physical world.

  Weak: only works with line-of-sight

  Allowed to swap this and map. Could be interpretation as to which is more accurate.

4. Assuming the robot can "see" the destination (can point its forward camera at the destination and the obstacles do not preclude it from showing up in the camera screen). Asume further the robot can identify where in the camera image the destination is (an X, Y pixel). Assume the robot can rotate left or right and move forward or backward.

   (a) How can the robot rotate so that the destination is in the center of the image (has a center point with an X pixel location around half the width of the camera image)?

   Rotate robot left or right to move destination to center of camera.

```
while (X pixel for destination $\neq$ center X)
{
     if (X pixel for destination > center X) rotate robot right
     if (X pixel for destination < center X) rotate robot left
  }
```

   (b) Assuming no obstacles, how can the robot use the ability from part (a) to travel to the destination?

   Orient robot to destination (part a) and move forward.

```
while (robot not at destination)
   {
     Rotate robot to put destination in center of camera (part a)
     Move forward
   }
```

5. In order to operate, the robots needs to perform the following simultaneous tasks:

   (a) listen for input tone sequences – 1 task
   (b) "speak" output tone sequences – 1 task
   (c) read GPS data points – 1 task
   (d) process input from 2 cameras (front, back) – 2 tasks
   (e) actuate two drive motors (left and right wheels) – 1 tasks
   (f) actuate two arms – 2 tasks
   (g) read local proximity sensors – 1 task
   (h) plan movements – 1 task

   Initially, the robot is designed to run on 10 processors with each task given a dedicated processor that can complete 10^8 instructions per second and has 16 MB of memory.

   (a) Assuming that locating the destination in the front camera (Problem 4) takes 10^7 instructions on a single processor dedicated to the camera, at what rate (number of times per second) can this processor perform destination locations?

   10 locations per second

   (b) Describe how the single processor can support these 10 tasks?

   Store instructions for all 10 tasks in memory. Interleave on one faster processor, saving the state (PC, memory) of each task periodically and switching to run other tasks.

```
while (true)
  {
  for each task x
    {
      restore state of X (memory, PC)
      run for a number of cycles T
      store state of X back to memory
    }
  }
```

   (c) How fast would the single processor need to run to guarantee to provide the same level of computation as the original 10 processor design?

   10^9 instructions/second

6. A controller can send a set of intermediate destinations to the robot. The more it sends, the less computation the robot needs to perform, but the more time it takes to send instructions.

- Assume on-robot planning can perform 10^8 instructions per second.
- Continue to assume communication to the robot as in Problem 1.
- Assume an intermediate destination requires 64 bits.
- Assume the processor can compute a sequence of n missing intermediate destination between two present destinations in $n^2 \times 4 \times 10^5$ instructions.

Sequence:	sent intermediate	compute intermed. 1	compute intermed. 2	...	compute intermed. n	sent intermed.	compute intermed. 1	...

(a) How many computed intermediates should be used between sent intermediates to minimize total compute and communication time?

Want to minimize total communication and send time. We must send a first intermedaite. Then we have a sequence of n computed before each next communicated intermediate. So, we want to minimize the time for computing n and sending 1, realizing we send (Total$/n+1$) of these sequences.

n =3

Minimize $\frac{40\text{ms}+n^2\times 4\text{ms}}{n+1}$

Minimize $\frac{40\text{ms}}{n+1} + \frac{n^2}{n+1} \times 4\text{ms}$

Minimize $\frac{40\text{ms}}{n+1} + \frac{n^2-1+1}{n+1} \times 4\text{ms}$

Minimize $\frac{40\text{ms}}{n+1} + \left(n-1+\frac{1}{n+1}\right) \times 4\text{ms}$

Minimize $\frac{41\text{ms}}{n+1} + (n-1) \times 4\text{ms}$

Derivative $\frac{-41\text{ms}}{(n+1)^2} + 4\text{ms} = 0$

$(n+1)^2 = \frac{41}{4}$

$n \approx$2 or 3

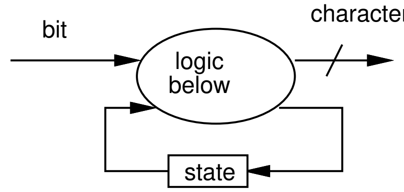
(b) Based on the answer above, how long will it take the robot to receive the intermediate destinations (communicate) and compute the missing intermediate destinations for a sequence with 20 intermediate destinations?

$n=3$: 1st destination, then 19/4 means 5 more sent destinations + 5 groups of 3 ...can be 4 groups of 3 and one of two

$6 \times 40 + 4 \times 4 \times 9 + 1 \times 4 \times 4$=400ms

$n = 2$: 1st destination, then 19/3 means 7 more sent destinations + 6 groups of 2 and one group of one.
$8 \times 40 + 6 \times 4 \times 4 + 1 \times 4 \times 1$=420ms

7. Consider the following FSM for decoding serial, Huffman encoded data into decimal digits.




```
int state=START;
while(true) {
  int bit=nextInputBit();
  switch(state) {
  case START:  if (bit==0) state=S0; else state=S1; break;
  case S0:  if (bit==0) {state=START; putc('0');} else {state=START; putc('1');} break;
  case S1:  if (bit==0) state=S10; else state=S11; break;
  case S10: if (bit==0) {state=START; putc('2');} else {state=START;
putc('3');} break;
  case S11:  if (bit==0) state=S110; else state=S111; break;
  case S110: if (bit==0) state=S1100; else state=S1101; break;
  case S111: if (bit==0) {state=START; putc('8');} else {state=START;putc('9');} break;
  case S1100: if (bit==0) {state=START; putc('4');} else {state=START;
putc('5');} break;
  case S1101: if (bit==0) {state=START; putc('6');} else {state=START;
putc('7');} break;
               }
  }
```

(a) Give the bit level encoding for 1500:
01110010000

(b) What is the minimum number of state bits needed to encode state for a hardware implementation?
4

(c) What digits can be encoded with each of the following number of bits:

bits to encode	encoding
1	none
2	0,1
3	2,3
4	8,9
5	4,5,6,7
6	none

8. Consider sending audio data over a UDP (Unreliable Datagram Protocol) link. Assume the receiver is aware of which packets are lost and will make the best of the information it does receive. As in lab, the frequency packets are based on Fourier Transforms of time windows. Rank in order of subjective quality for humans based on your knowledge of audio processing and human psychoacoustics from this course (1-best, 5-worst). Justify your answers by explaining the impact of losing a packet in each case.

   (a) [5] Packets contain PCM (raw, quantized time-sample data)

   Every lost data packet will result in silence for the length of time covered by the packet.

   (b) [4] Packets come in pairs for a time window, where the highest amplitude frequencies from each critical band are in one packet and the lowest in the other.

   It will depend on which packet is lost. If the packet with the low amplitude is lost, the result will be unnoticeable, as the high amplitude frequencies will mask the low amplitude frequencies. If the packet with the high amplitude frequencies is lost, the sound will not be reproduced properly. The important components are lost. The remaining components are likely not noise, but they are sounds that would typically not have been heard, so the result is only slightly better than silence.

   (c) [1] Packets come in pairs for a time window where both the first and second packet in a pair contains the top one-third frequencies (by amplitude) in each critical band. The first packet also contains the middle one-third of the highest frequency critical bands, while the second packet contains the middle one-third of the remaining critical bands.

   Loss is likely unnoticeable. The highest amplitude frequencies in a critical band will mask the lower frequencies. Since the highest (one-third) frequencies exist in both packets in a pair, reconstruction will be able to produce these frequencies even if one of the packets in the pair is lost.

(d) [3] Packets come in pairs for a time window with the highest frequency critical bands in one packet, and the rest in the second.

In the atypical case in which there are only high or low frequencies, this will be perfect half the time (when the empty frequency set is lost). Typically, there will be both. If the high frequencies are lost, the sound is likely understandable but lower quality; telephone audio only represents data up 4K Hz. If only the high frequency packet is received, this time period will sound odd and any spoken words may be hard or impossible to understand.

(e) [2] Packets come in pairs for a time window with the even timeslot PCM samples in one packet of the pair, and odd timeslot PCM samples in the other.

Losing one packet of a pair will be equivalent to sampling at half the frequency for that time interval. As long as the sound has no frequency content higher than the one-quarter the original sampling rate, the sound will be perfectly reproduced (half for the half lost, half since the Nyquist sampling rate is twice the highest frequency component). If there are higher frequencies present, we will get aliasing effects for the frequencies between one-quarter and half the original sampling rate (assuming the original data was filtered to remove frequencies above half the sampling rate).

This page intentionally left mostly blank for pagination.

Human auditory critical bands:

Band Number	Low	High
1	20	100
2	100	200
3	200	300
4	300	400
5	400	510
6	510	630
7	630	720
8	720	920
9	920	1080
10	1080	1370
11	1270	1480
12	1480	1720
13	1720	2000
14	2000	2320
15	2320	2700
16	2700	3150
17	3150	3700
18	3700	4400
19	4400	5300
20	5300	6400
21	6400	7700
22	7700	9500
23	9500	12000
24	12000	15500

Code of Academic Integrity

Since the University is an academic community, its fundamental purpose is the pursuit of knowledge. Essential to the success of this educational mission is a commitment to the principles of academic integrity. Every member of the University community is responsible for upholding the highest standards of honesty at all times. Students, as members of the community, are also responsible for adhering to the principles and spirit of the following Code of Academic Integrity.*

Academic Dishonesty Definitions

Activities that have the effect or intention of interfering with education, pursuit of knowledge, or fair evaluation of a student's performance are prohibited. Examples of such activities include but are not limited to the following definitions:

A. Cheating Using or attempting to use unauthorized assistance, material, or study aids in examinations or other academic work or preventing, or attempting to prevent, another from using authorized assistance, material, or study aids. Example: using a cheat sheet in a quiz or exam, altering a graded exam and resubmitting it for a better grade, etc.

B. Plagiarism Using the ideas, data, or language of another without specific or proper acknowledgment. Example: copying another person's paper, article, or computer work and submitting it for an assignment, cloning someone else's ideas without attribution, failing to use quotation marks where appropriate, etc.

C. Fabrication Submitting contrived or altered information in any academic exercise. Example: making up data for an experiment, fudging data, citing nonexistent articles, contriving sources, etc.

D. Multiple Submissions Multiple submissions: submitting, without prior permission, any work submitted to fulfill another academic requirement.

E. Misrepresentation of academic records Misrepresentation of academic records: misrepresenting or tampering with or attempting to tamper with any portion of a student's transcripts or academic record, either before or after coming to the University of Pennsylvania. Example: forging a change of grade slip, tampering with computer records, falsifying academic information on one's resume, etc.

F. Facilitating Academic Dishonesty Knowingly helping or attempting to help another violate any provision of the Code. Example: working together on a take-home exam, etc.

G. Unfair Advantage Attempting to gain unauthorized advantage over fellow students in an academic exercise. Example: gaining or providing unauthorized access to examination materials, obstructing or interfering with another student's efforts in an academic exercise, lying about a need for an extension for an exam or paper, continuing to write even when time is up during an exam, destroying or keeping library materials for one's own use., etc.

* If a student is unsure whether his action(s) constitute a violation of the Code of Academic Integrity, then it is that student's responsibility to consult with the instructor to clarify any ambiguities.