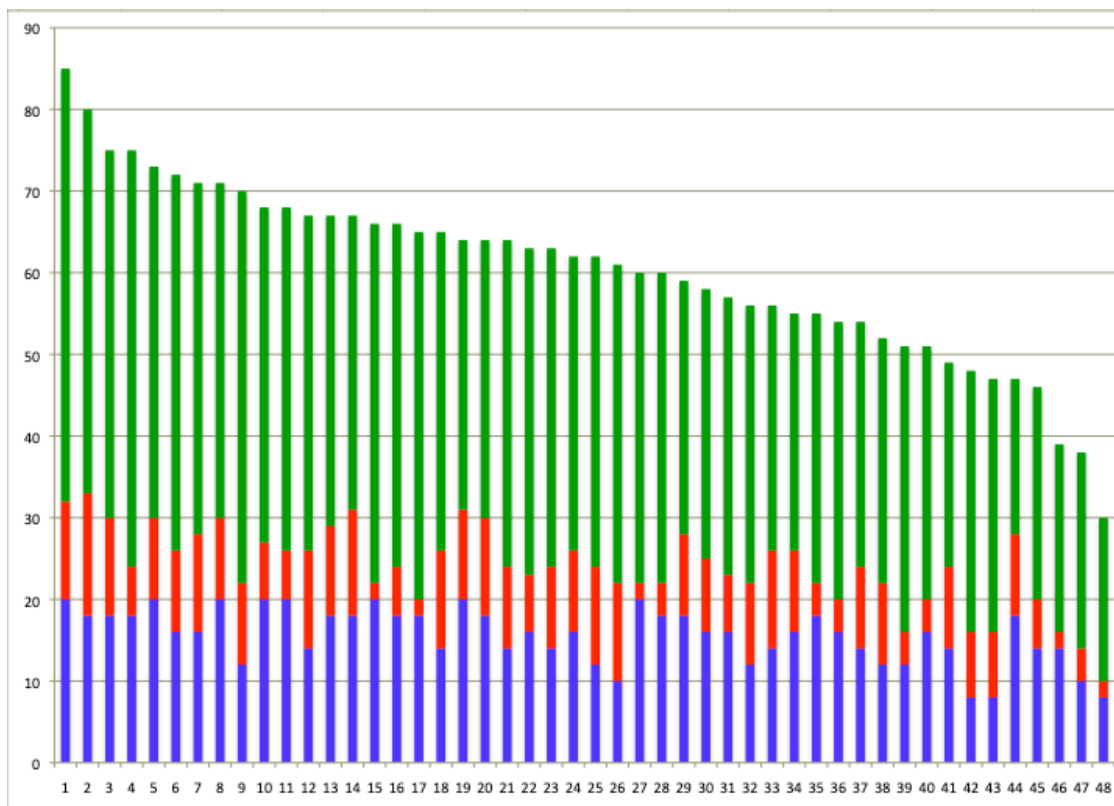


## COS 109 Midterm Exam, Fall 2023

I graded this myself. The median was 62, and the upper and lower quartiles were 67 and 54 this year, compared to 57, 67 and 38 last year. This year’s class definitely did better, which I hope means that everyone learned the material better, though it could also be ascribed to an easier exam, more generous grading, the phase of the moon, or all of the above.

Lest you get too cocky, however, (quoting from last year) the troublesome topics are the same: how to represent information in bits and bytes, how many of those would be required, notations like hexadecimal, and generally anything that involves arithmetic in bases other than 10 (and some problems with base 10 too). Many, indeed most, are unsure about how the Toy machine works. A significant number are shaky on how area and volume relate to linear dimensions.

The three values in each column below are for parts 1, 2 and 3, reading from the bottom.



**1. (20 points, 2 each) Short Answers. Circle the right answer or write it in the space provided.**

- (a) Unix systems distinguish upper case letters from lower case letters in filenames: lab3.html, Lab3.html and LAB3.HTML are different names (as you might have learned while trying to upload your labs). How many different ways are there to write the filename **lab3.html** in mixtures of upper and lower case letters?

**128, or 2<sup>7</sup>.** The 3 remains unchanged, and contrary to what a surprising number of people said, html does change; observe that both html and HTML appear in the question itself.

- (b) The first character of the Unicode code chart for the Gothic alphabet has code **10330** and the last character has code **1034F**. How many characters are there in the Gothic code chart?

**32.** During the Q/A, there was an extended discussion of an almost identical variant from an earlier year, where we talked about being sure to include the last value in the count. I hope those who attended got the right answer.

- (c) In the following list of 24-bit RGB colors, expressed in hexadecimal, which one has the smallest amount of blue? (The character **0** is a zero.)

**ACCEDE    BAOBAB    BOBBED    COFFEE    DECODE    DOODAD    EFFACE    FAÇADE**

**BAOBAB.** These are base 16 numbers -- no arithmetic is needed. AB is less than AD, and all other blues are bigger. Most people got this.

- (d) The speeds of supercomputers are measured in floating-point operations per second, or “flops”. Which one of these would be the most representative speed for the fastest of today’s supercomputers?

**1 Mflop    1 Gflop    1 Tflop    1 Pflop    1 Eflop    1 Zflop    1 Yflop**

**1 Eflop.** We talked in class about the exaflop speed attained by the latest supercomputer. Most people got it.

- (e) Modern computers can efficiently process integers of several sizes, usually 1, 2, 4, and 8 bytes long. Which of these is the least number of bytes that could be used for storing a binary number that represents the population of California?

**1 byte    2 bytes    4 bytes    8 bytes    something bigger    no way to tell**

**4 bytes.** 2 bytes is only 65K; 8 bytes is  $2^{64}$ , so vastly more than needed. Mostly well done.

- (f) Archie reports a headline from *The Guardian*: “WhatsApp increases group chat size limit to 256 people. It’s not clear why WhatsApp settled on such an oddly specific number.” In no more than half a dozen words, explain the likely reason for this specific number.

**It’s the largest size that can be encoded in one byte if the range is from 1 to 256, so the choice is probably for speed and space efficiency. Not as clear a question as I had hoped, so I was generous.**

- (g) “Daily active users for [Facebook] Threads on Android dropped from 49 million on July 7 to 23.6 million on July 14, and then to 12.1 million on July 21”, says a web analytics company. If this decline had continued at the same exponential rate, how many weeks would it have been before the number of users was below 1 million?

**4 weeks further, or 7 weeks from the beginning, or 6 more weeks from the beginning. Ambiguously worded, so graded generously.**

- (h) What is the decimal value of the binary number **111001.11**?

**57.75.** A common wrong answer was 57.3. Place value works on both sides of the decimal point; **.11** is (in decimal)  $0.5 + 0.25$ .

- (i) Suppose that a secretive spy agency stores the name, address, phone number and social security number for every person in the USA. *Very roughly*, how many gigabytes would be required to hold all this information, without compression? State your assumptions clearly.

**20 GB?** 20 bytes for name, 30 for address, 10 for phone, 10 for SSN is 70 bytes, and a population of 330 million, so roughly 20-25 GB.

- (j) Computer pioneer Grace Hopper (1906-1992) said, “The instruction code should use symbols which are easily learned and identified with the operations by already existing mental associations: ‘a’ for add, etc. Replacing a sequence of binary numbers with a single letter to represent an operation [simplifies] the coding process and makes it much more intuitive for users.” What kind or level of programming language is Hopper describing? One or two words is enough.

**Assembly language.**

## 2. (15 points) Machines

Here is a program in the Toy assembly language, with reminders about what the instructions do.

```

Foo   GET           get a number from keyboard into accumulator
        IFZERO Bar    if accumulator is zero, go to Bar
        IFPOS Prn     if accumulator is positive, go to Prn
        STORE N       load accumulator value into location N
        LOAD 0        load 0 into accumulator
        SUB N         subtract value in location N from accumulator
Prn   PRINT        print value in accumulator
        GOTO Foo      go to instruction labeled Foo
Bar   STOP
N     0    reserve a memory location called N, set its initial value to 0
    
```

(a) If this program is given the sequence of inputs **3 -1 4 1 -5 9 2 -6 -5 0** *exactly* what does it print?  
**3 1 4 1 5 9 2 6 5.** It's printing the absolute values of the inputs.

(b) It is possible to simulate the Toy instruction **GOTO Foo** with a sequence of two other Toy instructions, in several ways. Show one such sequence.

```

LOAD 0
IFZERO Foo
    
```

One of many possibilities. Others include **IFPOS Foo**; **IFNEG Foo**. But not **IFPOS Foo**; **IFZERO Foo**, since that sequence doesn't work if the accumulator holds a negative value.

(c) Imagine that Alan Turing and John von Neumann are having an argument. Turing says "I can simulate *any* of your computers on my Turing machine." Von Neumann replies "So what? I can simulate your silly Turing machine on *all* of my computers." Who is right?

**neither one          only Turing          only von Neumann          both**

**Both.**

(d) In his 1946 paper, John von Neumann said "We are therefore forced to recognize the possibility of constructing a hierarchy of \_\_\_\_\_, each of which has greater capacity than the preceding but which is less quickly accessible." Which of the following is the proper word to fill in the blank

**accumulators    controls    instructions    memories    orders    organs    processors**

**memories.**

(e) Von Neumann also said "It is convenient to group the binary digits into tetrads, groups of 4 binary digits." What synonym or alternative terminology might be used today instead of tetrads?

**Hex(adecimal) digits.** A few people said **nibble**, a fine alternative, though the word never caught on.

**3. (55 points, 5 each) Miscellaneous**

(a) A recent biography of John von Neumann says that a particular computation produces a number "less than a trillionth of a trillionth of a billionth of 1 joule-second, far too small to be noticed in everyday life but large enough to be significant at the atomic scale."

(i) What power of ten is "a trillionth of a trillionth of a billionth"?

**10^-33**

(ii) What power of two is nearest to this power of ten?

**2^-110**

Small penalty for getting the sign(s) wrong.

- (b) Last month I did some experiments with a new algorithm. The measured running times were 15.0 milliseconds to process 4 items, 30.1 msec for 5 items, 59.9 msec for 6 items, and 120.2 msec for 7 items.

(i) Given these times, what is the likely running time for 10 items?

**960.** This really is approximate; watch out for excessive precision.

(ii) Which of these is the most likely description of how the algorithm's running time grows in proportion to the number of items?

**logarithmic   linear   n log n   quadratic   cubic   exponential   none of these**

**exponential.** It's doubling each time, the very definition of exponential growth. Missed by many.

- (c) The hexadecimal value **FF00FF** can be interpreted as an RGB color, but it is really just a 24-bit integer. Suppose that we add 1 to this integer value, that is, compute **FF00FF + 1**.

(i) What is the resulting value in hexadecimal?

**FF0100.**

(ii) If the resulting value is interpreted as an standard 24-bit RGB color, which of these colors is it closest to?

**red   green   blue   yellow   cyan   magenta   black   white**

**red.** Maximum red, almost no green, no blue.

- (d) Suppose that Princeton wants to encode certain information about current Princeton undergrads *in as few bits as possible*. The information for each person is: birthday (like October 13), age (assume everyone is between 17 and 24 inclusive), class year (2024 through 2027), and donor potential on a scale of 1 through 4. What is the *minimum* number of bits needed per person, and why?

**16 = 9 + 3 + 2 + 2.** Missed by many. Just figure out how many bits are needed for each piece, add them up.

- (e) Suppose that we fill Friend 008 with old vacuum tubes like the ones that were passed around in class. Ignoring chairs, desks, people, and everything else, *very roughly* how many vacuum tubes would the room hold? You must base your answer on sound estimates and quantitative reasoning. Be brief but clear about your assumptions and computations.

**10 million?** The (guessed) dimensions of the room were on a previous midterm, maybe  $40 \times 30 \times 20 = 24000$  cubic feet. A vacuum tube like the one(s) passed around in class, and like the one in the book (same one, in fact) is maybe 4 cubic inches. There are 1728 cubic inches in a cubic foot ( $12^3$ ). So  $25000 \times 1728 / 4$ , which is about 10 million. Graded generously, though a lot of people had implausible dimensions for both room and tubes. Significant penalty for using area instead of volume, or (even worse) linear, like " $10000$  cubic feet =  $120000$  cubic inches."

- (f) The *NY Times* said (4/7/23) that "Eleanor Catton's Booker Prize-winning novel *The Luminaries* imposes a precise numerical rule on its chapters, each of which is half the length of the [previous one]." The book has 12 chapters. I do not believe this story, but assuming that it is true and that the final chapter is only one quarter of a page long ...

(i) Approximately how many pages are in the first chapter?

**512.** Maybe easiest seen by just counting from the back; I got it wrong first time by thinking it was  $2^{12}$ .

(ii) Approximately how many pages are in the whole book?

**1024,** approximately. The sum of  $1 + 2 + 4 \dots + 2^n$  is  $2^{(n+1)} - 1$

- (g) Suppose that Thomas Sweet has a special on ice cream cones: they will double the diameter of the scoop for only 4 times the price. Is this a good value for an ice-cream lover, a bad value, or not special at all? Explain your answer by quantitative reasoning.

**Good deal.** Volume goes up as the cube of the linear dimension like diameter, so it's 8 times the ice cream for 4 times the price. Not well handled; many people either thought this was about area, which appeared to lead to "nothing special," or somehow got things upside down, hence "bad value." Questions like this are about linear

dimensions, areas (the square), and volume (the cube).

(h) In 2008 there were about 25 COS majors per year. Today, 15 years later, there are about 200. Assume that this represents a smooth exponential growth that will continue into the future.

(i) Very roughly, what is the percentage rate of increase of the number of COS majors each year?

**14%.** It doubled 3 times in 15 years, so it's doubling once in 5 years; the rule of 72 gives about 14%.

(ii) If this trend continues (let's hope not!), and if Princeton remains at its current size, in about how many years will all students be COS majors?

**15 years or so.** Five more years would be 400, then 800, then 1600, which is the size of the graduating class. The question says "majors per year", not total, but this seems to have been missed, leading to answers like 25 years.

(i) Many years ago, Pat Programmer wrote a C program for a computer that now no longer exists. She still has both the C source code and the compiled code for the ancient computer. She wants to work on the program again, but will have to run it on her brand new laptop. For each of the following, circle the most appropriate answer.

- She could run the original compiled code, unchanged, on her laptop **likely** unlikely
- She could compile the C program and run that compiled code on the laptop likely **unlikely**
- She could write a simulator in C for the old computer, and run it on the laptop likely **unlikely**
- She could run the original compiled code on this simulator on the laptop likely **unlikely**
- The simulated computer could run faster than the old physical computer it simulates likely **unlikely**

*We discussed a variant of this question in the Q/A; hopefully that was helpful*

(j) Random quickies:

- An assembler has to be written in assembly language **true** false
- The Turing Award goes to mathematicians for contributions to theoretical computer science **true** false
- Leibnitz advised Babbage to use binary arithmetic for his mechanical computing devices **true** false
- A prox card is powered by a tiny embedded battery **true** false
- The binary representation of ☹️, whose hex representation is 1F641, fits in 2 bytes **true** false

(k) The picture on the left is a close-up of a seriously geeky t-shirt from Thinkgeek.com. Exactly what does it say? Write your answer clearly and unambiguously.

```
0100100001100101
0110110001101100
0110111100100000
0101011101101111
0111001001101100
0110010000100001
```



	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	NUL	SOH	STX	ETX	EOT	ENQ	ACK	BEL	BS	HT	LF	VT	FF	CR	SO	SI
1	DLE	DC1	DC2	DC3	DC4	NAK	SYN	ETB	CAN	EM	SUB	ESC	FS	GS	RS	US
2	SPC	!	"	#	\$	%	&	'	(	)	*	+	,	-	.	/
3	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5	P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	^	_
6	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7	p	q	r	s	t	u	v	w	x	y	z	{		}	~	DEL

**Hello World!** Don't forget case.