

COS 109 Final Exam, Fall 2022

December, 2022

3 hours 180 points total

You may do this exam in any 3-hour period. When you are finished, return the exam to room 311 in the CS building, or scan it (or just the answers) and email it to bwk@cs.princeton.edu. You must return it by 5:00pm EST on Thursday December 22. I would appreciate it if you could tell me how you plan to return it so I know where and when to look, but it's not a requirement.

Please PRINT your name here _____

Honor Pledge: "I pledge my honor that I have not violated the Honor Code during this examination."

Please write the pledge in full and sign it:

This examination is open-book and open-note:

- you may use the textbook, course notes, your own notes, corrected problem sets and solutions, old exams and answer sheets from the course web page, lab instructions, etc.
- you may use a calculator.
- you may not use anything else; specifically, you may not use a computer, phone or tablet (except that you can use a calculator program on one of these, and you can use your computer to view course notes, answer sheets, etc.). No other Internet use is allowed.

There are 180 points for the questions; use the point values for each question to allocate your time (one point per minute). If you're writing or calculating a lot on a question, you may be off on the wrong track.

Write your answers directly on these pages; use the back if necessary. In general, be brief, but if you need more space, attach extra pages and make sure your name is on every extra page. Please write legibly -- I can't grade it if I can't read it. [You can submit just the answers if that makes scanning easier, though showing your work can help with part credit.]

Good luck.

1. (50 pts)

2. (30 pts)

3. (100 pts)

Total

	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

1. **(50 points, 2 each) Short Answers.** Circle the right answer or write it in the space provided.

(a) If **m** and **n** are positive integers, how many 1-bits (that is, bits whose value is 1) are there in the binary representation of $2^{m+n} - 1$?

GIF JPG Lempel-Ziv MP3 MPEG PNG Zipf

(b) The source code for the Linux kernel is about 1.3 GB of text files in C and the like, uncompressed. Which of these techniques would be most suitable for compressing it for faster downloading?

(c) If Alice and Bob are discussing the significance of Etain Shrdlu, which of these is the most likely topic of their conversation?

compression error detection machine translation passwords programming languages

(d) Which one of these would be the most appropriate name for someone working at a Certificate Authority?

Alice Bob Carol David Eve Mallory Trent

(e) The NSA's TAO (Tailored Access Operations) program surreptitiously "installs beacon devices into our targets' electronic devices" before they are shipped to overseas purchasers, to enable network monitoring and other effects. What kind of attack is this an example of?

DoS IoT MITM MS-DOS spear-phishing Trojan horse

(f) Braille represents each letter, number or other symbol as 3 rows and 2 columns of dots that are individually raised or not raised, like this:



How many different characters can there be in Braille?

(g) On Dec 11, 2022, ChatGPT reported that “We're experiencing exceptionally high demand. Please hang tight as we work on scaling our system,” and was otherwise non-responsive. What kind of attack is this an example of?

DoS IoT MITM MS-DOS spear-phishing Trojan horse

(h) In *The Innovators*, Walter Isaacson quotes _____ as saying “It is possible to invent a single machine which can be used to compute any computable sequence.” Whose name belongs in the blank?

(i) “These digital tokens use _____ technology, in which computers contribute to a shared ledger that can be used to track digital assets.” What word or phrase belongs in the blank?

(j) Name one example of the kind of a digital asset mentioned in the previous question (part (i)).

(k) Ransomware encrypts all the files on a victim’s computer, and the victim has to pay the bad guys for a password that will decrypt them. If you were (just hypothetically) a bad guy, which of these mechanisms would be most suitable for implementing your ransomware attack?

AES DES MD5 Rot13 RSA SHA-512 Tor

(l) Circle all of the following that must be kept secret from an adversary to protect a message encrypted with AES:

AES algorithm AES key AES source code compiler compiled AES code

(m) In November 2022, the Supreme Court agreed to hear the case *Jack Daniel’s Properties v. VIP Products*. VIP sells dog chew toys shaped like a Jack Daniel’s whiskey bottle with the label “Bad Spaniel’s.” The toy includes dog-related changes to the original label, like replacing alcohol content with “43% Poo by Vol.” and replacing “Old No. 7” with “Old No. 2.” What specific kind of intellectual property does this case deal with?

(n) What distinguished Princeton alum was recently named the university's new provost?

Jeff Bezos '86

Christopher Eisgruber '83

Robert Kahn *64

Jennifer Rexford '91

Eric Schmidt '76

Peter Wendell '72

(o) The US national debt is about 31 trillion dollars in December of 2022. How many *bytes* are needed to store this value in binary?

(p) Suppose that over the 12 days of Christmas I plan to give the following presents. Circle all those that *do not* use radio.

Apple airpods

barcode reader

GPS receiver

prox card

smart watch

US passport

(q) Put these names into chronological order of when they made the contribution(s) that caused them to be mentioned in COS 109, by writing the numbers 1 through 5 on them.

Tony Hoare

John von Neumann

Guido van Rossum

Bjarne Stroustrup

Ken Thompson

(r) RSA-250, the largest currently solved factoring challenge number, has 250 decimal digits and is 829 bits long. About how many bits long would you expect RSA-500 to be?

(s) Knowing that binary search doesn't work on unsorted data, a zealous programmer modifies a binary search function by adding statements at the beginning to verify that the data values being searched are in order. What is the likely running time of the revised function, expressed in terms of **n**, the number of data items?

log n

n

n log n

n²

2ⁿ

(t) “Though _____ software is designed to be shared freely among coders and companies, this sharing is governed by licenses designed to ensure that it is used in ways to benefit the wider community of programmers.” What word or phrase belongs in the blank?

(u) We counted the number of people in the classroom one day by an algorithm that paired standing students; one of each pair then sat down while the other retained the combined counts. Suppose that we want to do the same thing in a crowd of a million people, but instead of pairs, at each stage people will combine in groups of 4 at a time. If done perfectly, about how many stages will this procedure require to produce a final count?

(v) Suppose that one afternoon during an interminable COS 109 lecture you use Safari to visit Amazon, Facebook and Google from your laptop, and Amazon sends you a cookie. Which of the following sites will receive that cookie later that day when you visit Amazon from your phone?

only Amazon Amazon+Facebook Amazon+Google all three of them none of them

(w) In the *Jeopardy* category “Let’s Get I.T. On,” the clue was “This programming language isn't a little worse than B minus; it's a 1980s improvement of a language called C.” What is the language?

(x) “Companies make cordless mice that use radio signals instead of Bluetooth to communicate with the computer.” (*NY Times*, 7/23/05) At best this is sloppy; at worst it’s actively misleading. In no more than 5 or 6 words, identify and correct the imprecise statement or technical misinformation conveyed by this quote.

(y) If I were a high-school guidance counselor who wanted to use a computer to generate plausible college recommendation letters for my students, which one of these AI techniques or systems would be the most directly useful?

ChatGPT DALL-E2 deep learning neural network recommendation engine
reinforcement learning supervised learning unsupervised learning

2. (30 points) Understanding Programs

- (a) [6] The Python function `pow(m, n)` is supposed to raise a positive integer `m` to a positive integer power `n` by repeated multiplication; that is, it computes m^n . For example, `pow(2, 3)` should return 8 and `pow(3, 2)` should return 9. Unfortunately, this version has three small errors. Fix the errors: either rewrite `pow` or state clearly what the errors are and how you are fixing them. (This is a question about correct logic; don't worry about syntactic trivia, but make your corrected code clear.)

```
def pow(m, n):
    p = 1
    i = 1
    while i < n:
        p = p + m
        i = i + 1
    return m
```

- (b) [2] Once it has been properly fixed, how does the running time of this algorithm depend on `m`?

`log m` `m` `m log m` `m2` `m3` `2m` independent of `m`

- (c) [2] Once it has been properly fixed, how does the running time of this algorithm depend on `n`?

`log n` `n` `n log n` `n2` `n3` `2n` independent of `n`

- (d) [1] Once fixed, does this function work properly when `m` is not an integer? **Yes** or **No**

- (e) [1] Once fixed, does this function work properly when `n` is not an integer? **Yes** or **No**

- (f) [4] Here's a Python function, with various parts identified by line numbers; the line numbers are not part of the function.

```
1: def AbsoluteValue(v):
2:     if v >= 0:
3:         return v
4:     else:
5:         return -v
```

- (i) Which part is the API for this function? Identify the line or lines by number.

- (ii) Which part is the implementation? Identify the line or lines by number.

(g) [6] The Python function `random.randint(1,100)` produces an endless sequence of random integers between 1 and 100 inclusive; any number is as likely as any other, so over a long enough period, any number will occur about as often as any other number. If the following Python loop is executed, approximately how many lines of each type of output would you expect to see?

```
for i in range(0,10000):
    num = random.randint(1,100)
    if num > 60:
        print("big")
    elif num > 30:
        print("middling")
    elif num > 5:
        print("small")
```

big:

middling:

small:

(h) [6] Suppose that the Toy machine version 2.0 has a new instruction **rshift N**, which shifts the contents of the accumulator **N** bit positions to the right (and discarding the bits that “fall off the end”). What does this program print when given the number **37** as input?

	GET		<i>get a number from user, place it in accumulator</i>
TOP	PRINT		<i>print content of accumulator</i>
	IFZERO	DONE	<i>if accumulator content is zero, go to location DONE</i>
	RSHIFT	1	<i>shift accumulator content one bit to the right</i>
	GOTO	TOP	<i>go to instruction labeled TOP</i>
DONE	STOP		<i>stop execution</i>

(i) [2] How does the running time of this program depend on the size of the input number **N** that it is given?

logarithmic linear N log N quadratic cubic exponential independent of N

3. (100 points, 5 each) Miscellaneous

(a) In November 2022, the SI units were updated by adding names for the new biggest and smallest units: zetta and yotta are now followed by **ronna** and **quetta** as the largest numbers, and there are now matching **ronto** and **quecto** as the smallest.

(i) How many quectograms are there in a quettagram, expressed as a power of ten?

(ii) What power of two is closest to this number, the number of quectograms in a quettagram?

(b) *Base64 encoding* is a technique that represents arbitrary binary data in a printable form. It's similar to hexadecimal, but uses 6-bit chunks instead of 4: each possible 6-bit input combination is encoded with a unique 8-bit ASCII letter (a-z, A-Z) or digit (0-9) or other character (+, /) so the result of the encoding is a sequence of ASCII characters that is longer than the input sequence of bits.

(i) If an IPv4 packet is 3000 bytes long, how long is the Base64 encoding of the packet?

(ii) If instead of Base64, we write the packet contents in hexadecimal, how many hex digits would it take to write out the packet contents?

(c) When I create an online grocery order at Shoprite, every time I add an item to my cart, nearly two dozen trackers try to monitor me. Fortunately my defenses block them all (or so I believe).

(i) What programming language are the trackers most likely to be written in?

(ii) Name two tools that you or I could use to block explicit trackers.

(d) “_____ stinks!”, says a top secret PowerPoint slide produced by _____ and revealed to the world in 2013 by _____. Fill in the blanks with the appropriate names.

(e) “Morse decided to puzzle his brain no more on how 23 in base 10 could be expressed in base 5.” (From the 1976 Inspector Morse novel *Last Seen Wearing*, by Colin Dexter.)

(i) What is 23 base 10 expressed in base 5?

(ii) What is 23 base 10 expressed in base 2?

(iii) What is 23 base 10 expressed in base 16?

(f) Princeton’s new Stadium Drive parking garage has spaces for 1,560 cars. Suppose that a license-plate reader records the plate number as text, the arrival time and the departure time for each car.

(i) Estimate approximately how many bytes would be needed to store the plate number, arrival time, and departure time for any given car, reasonably compactly.

(ii) Estimate approximately how many megabytes of disk space would be needed to store all this data for one year, making sensible assumptions about the amount of traffic in and out of the garage.

(g) Joe College has 1,000 files on his computer, of which 50 are correctly labeled Word .docx files and 25 are correctly labeled Excel .xlsx files.

(i) How many times does Joe have to run Word to compute the total number of bytes in all of those .docx and .xlsx files?

(ii) How many times does he have to run Word and Excel to determine whether the largest Word file is larger than the largest Excel file?

Word _____

Excel _____

(h) *Supreme Conflict*, a 2008 book on the Supreme Court, describes how before each session each of the nine justices shakes hands with each of the others.

(i) What is the total number of handshakes?

(ii) If there are N people doing handshakes, how does the total number of handshakes grow in proportion to N ?

(i) Netflix still distributes physical DVDs, but the number of users has been shrinking steadily, from about 16 million in 2010 to about 1 million today in 2022.

(i) If this decline has been a smooth exponential process, what is the percentage rate of decrease *per month*?

(ii) If the decline continues at the same rate in the future, in what year will there be only 1,000 users left? (Clearly this is a very over-simplified model, so excess precision is not appropriate.)

(j) In December 2022, TSMC announced plans to build two new integrated circuit fabrication plants in Arizona, at a cost of \$40B! The first plant was going to use a 6 nm technology but will now use 4 nm; the second plant will open later and use 3 nm technology. (I have simplified the numbers a bit.)

(i) If a 12-inch wafer using 6 nm technology has 200 chips, how many would it have with the 4 nm technology, if nothing else changes?

(ii) If a 12-inch wafer using the 6 nm technology has 200 chips, how many would it have with the 3 nm technology, if nothing else changes?

(k) A network address translator (NAT) maps internal IP addresses from a specified range into a single external IP address and vice versa.

(i) A home wi-fi router acts as a NAT. It usually uses the IPv4 address range 192.168.0.0 to 192.168.255.255 internally, assigning a unique IP address in this range to each device in the home. In principle, how many devices could this range support? Express it as a power of 2.

(ii) A large corporation might use a similar mechanism but with the internal IPv4 address range 10.0.0.0 to 10.255.255.255. How many devices would this address range support? Express it as a power of 2.

(l) The late John Lions, author of *A Commentary on the Unix Operating System*, once said of the source code listing of 6th Edition Unix, “The whole documentation is not unreasonably transportable in a student’s briefcase.” For each of the following, would it be reasonably transportable in your backpack?

A hundred terabytes of laptop SSD disks	yes	no
A terabyte of magnetic core memory from the 1960s	yes	no
A listing on paper of a C implementation of the AES algorithm	yes	no
A listing on paper of a C implementation of the TCP/IP protocol	yes	no
10 kilometers of bare fiber optic cable	yes	no

(m) An article about the NYC marathon says “Every competitor will wear a shoe with a chip that will record their progress, and can send e-mail updates every five kilometers to spectators who subscribe to the service.” For each of the following inferences that a non-technical reader might make from this quotation, assess whether they are likely to be correct or likely to be incorrect. (A marathon is about 26 miles or 45 km long.)

the chip has enough memory to store at least a dozen time measurements	correct	incorrect
the chip uniquely identifies the runner who wears it	correct	incorrect
the chip uses GPS to determine how far the runner has run so far	correct	incorrect
the chip sends e-mail messages to a server	correct	incorrect
the chip’s memory determines the maximum number of e-mail subscribers	correct	incorrect

(n) The hex value **00FFFF** can be interpreted as an RGB color. Suppose that instead this value is interpreted simply as a 24-bit integer, stored in a variable **v**, and incremented by 1 with the Python statement **v = v+1**.

(i) What is the resulting value of **v** in hexadecimal?

(ii) What color is the resulting value closest to?

red green blue yellow cyan magenta black white

(o) If I use my phone to send mail to a friend in England, as the mail goes from me to his laptop, which of these mechanisms (A) is almost sure to be used? (B) might be used but need not be? (C) is very unlikely to be used? Circle the best answer.

TCP/IP	almost sure	might be	very unlikely
Ping	almost sure	might be	very unlikely
Fiber-optic cable	almost sure	might be	very unlikely
NAT	almost sure	might be	very unlikely
Wi-fi base station	almost sure	might be	very unlikely

(p) A *NY Times* article about E-ZPass, the electronic highway toll system, says, “A list of valid and invalid tag numbers is sent every day to computer drives in every toll booth. As a vehicle drives through an E-ZPass lane, a high-speed optical reader almost instantly identifies the tag mounted to a dashboard or windshield and matches it against the list to see if the holder has enough money set aside to pay the toll.” Identify three technical “facts” in this quotation that are almost surely wrong, or at least badly misleading. *Be brief*— a few well-chosen words should be adequate for each.

(i)

(ii)

(iii)

(q) Refer to the ASCII chart on the cover page of the exam.

(i) If **ch** is a variable that contains the ASCII character [(left bracket), how many bits must be changed to convert **ch** into the ASCII character { (left brace)?


(ii) If **ch** is a variable that contains an arbitrary ASCII character, explain in at most half a dozen words what this test is trying to determine. *DO NOT* just repeat the code in words.

```
if ch >= 65 and ch <= 90 ...
```

(r) To coordinate their romantic activities, Alice and Bob naturally use public-key cryptography to exchange encrypted email. Suppose that Eve learns Alice’s private key. What can Eve now do?

Eve can convince Bob that she (Eve) is really Alice	true	false
Eve can convince Mallory that she (Eve) is really Alice	true	false
Eve can convince Alice that she (Eve) is really Bob	true	false
Eve can read an encrypted message from Bob to Alice	true	false
Eve can read an encrypted message from Alice to Bob	true	false

(s) [10 pts] Random quickies: Circle the best answers.

The GDPR applies primarily to residents of Germany	true	false
Trans-oceanic Internet traffic is transmitted with communications satellites	true	false
The Turing machine predates von Neumann's <i>Johnniac</i> computer at IAS	true	false
"An IP address is like a zip code: it tells where your computer is located"	true	false
End-to-end encryptions prevents the NSA from knowing what your browser connects to	true	false
<code>/* You are not expected to understand this */</code> comes from Unix kernel source code	true	false
I would have to purchase the domain kernighan2024.com directly from ICANN	true	false
"The S&P500 returned 9% annually so your investment doubled in value in 8 years"	true	false
Alan Turing was the first recipient of the ACM Turing Award	true	false
The Cuneiform character  whose hex representation is 1236E, fits in 2 bytes	true	false