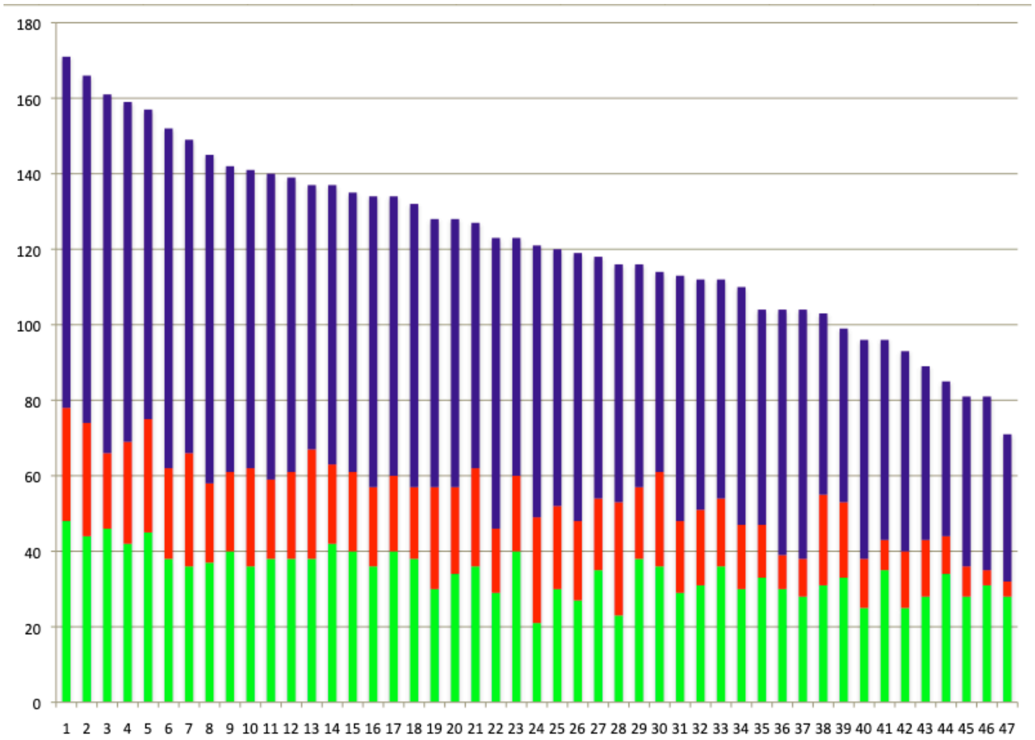
COS 109 Final Exam, Fall 2023

I graded this myself. This year's median was 121 and the quartiles were 139 and 104. For comparison, last year the median was 125 and the quartiles were 138 and 103, which is pretty close to identical. Make what you will of that. The colored bars are for parts 1, 2 and 3, reading up from the bottom.



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

(a) I'm thinking of a floating-point number between 1 and 1,000. About how many guesses would it take you to figure it out to within 3 decimal places?

20. There are a million possible numbers, so log2(10^6). A surprising number of people thought it was 10.

(b) Shrdlu! If I apply a static Huffman coding algorithm to this string of words, it will not compact it as tightly as it might on a not so statistically unusual bunch of words. What's going on? What's odd about all of this? What do you think?

Exactly one person noticed that this sequence of contrived sentences did not contain any occurrences of the letter "e", the most common letter of English, something noted in an extended classroom discussion of Etaion Shrdlu, and books like *A Void* that didn't contain any "e". A compression algorithm based on the statistics of normal text would not do as well on text like this. So much for close reading? Or was it just too cute?

(c) What is a Frankenpine? (Or, if you prefer, a Frankenpalm?)

A cell tower disguised as a tree. A gift to all.

(d) If it takes one eon to break a message encrypted with AES-128 using brute force, how many eons will it take to break a message encrypted with AES-256?

1 2 128 256 128^2 256^2 2^{128} 2^{256}

2^{128}. Each extra bit is another power of two.

(e) Which of these people were knighted for their contributions to computing? Circle all who were.

Charles Babbage **Anthony Babington** **Tim Berners-Lee** **Tony Hoare** **Alan Turing**

Berners-Lee, Hoare.

(f) Each of the 50,000 runners in the New York City marathon has an RFID tag to identify them specifically. How many bits does the tag need to use so that each runner has a unique ID?

16, since 50,000 is more than 2^15 and less than 2^16. Most people got this.

(g) The country code for Tuvalu is **`.tv`**. How many web servers does Yahoo need to install in Tuvalu itself so that it can provide some web service for the domain **`yahoo.tv`**?

Zero. Most people got it.

(h) A 2001 article about Moore's Law says that Intel "took three decades to produce a chip that ran at one gigahertz, but only 18 months to double that speed." Was Intel's pace of development unusually slow for the three decades, unusually fast for the last year and a half, or just about what would be expected throughout the entire period?

unusually slow **unusually fast** **just about expected**

Just about expected. Practically the definition of Moore's Law, especially at the time.

(i) The Great Cannon is a Chinese cyber-warfare tool that "injects malicious JavaScript into pages served from behind the Great Firewall. These scripts, potentially served to millions of users across the Internet, hijack the users' connections to make multiple requests against the targeted site." What kind of attack is this on targeted sites?

DDoS **IoT** **MITM** **MS-DOS** **Trojan horse**

DDos, a distributed denial of service attack on the targets.

(j) Which of these historical personages were victims of a cryptographic failure. Circle all that were.

Auguste Kerckhoffs **Isoroku Yamamoto** **Julius Caesar** **Mary Queen of Scots** **Queen Elizabeth I**

Yamamoto, Mary. Caesar was not a victim of cryptographic failure!

(k) In an IPv4 address, the first part is the network id and the rest is the host id on that net. If there are **N** bits in the network id part, what is the maximum number of host ids that could be on that network?

2 ^ (32-N)

(l) The current US Department of Justice vs Google case deals with which one of these legal issues?

antitrust **API copyrights** **patents** **spoliation** **trademarks**

antitrust

(m) If **n** is a positive integer, how many 1-bits (that is, bits whose value is 1) are there in the binary representation of

2^2^2 … n times, that is, $2^{2^{2^{2}}}$ … to a height of **n** ?

1, as with any power of 2.

(n) The Luhn algorithm for error-checking a credit card number or a phone IMEI starts at the rightmost digit, and multiplies successive digits alternately by 1 or 2. If the result is > 9 then subtract 9. Add the resulting digits. For a valid number, the sum must be a multiple of 10. How does the running time of this computation grow in proportion to the number of digits in the original number?

log n **n** **n log n** **n^2** **n^3** **2^n** **no way to predict**

n. Just loops over each digit.

(o) In November 2023, the ______________________ issued proposed regulations that would require telecom companies to protect their customers better against SIM-swapping attacks. Which one of these entities did this?

FCC **FDA** **FTC** **ITU** **NHTSA** **SEC** **TIAA**

FCC. Most people got this.

(p) From a cyber-security bill in the US House of Representatives: "_______________________ devices have at least one transducer (sensor or actuator) for interacting directly with the physical world, have at least one network interface, and are not conventional Information Technology devices, such as smartphones and laptops." Which of these belongs in the blank?

5G **Bluetooth** **Cyber-warfare** **Drone weapons** **Integrated circuit**

Internet of Things **Radio frequency identification** **Two-factor authentication**

Internet of Things.

(q) Alice says "A crypto algorithm is much more likely to be secure if everyone knows exactly how it works." Bob says "Nonsense! The only way to make a cryptographic algorithm really secure is to keep how it works a secret." Eve says "The real security is in the key you use." Who is right?

none of them **only Alice** **only Bob** **only Eve** **both Alice & Eve** **both Bob & Eve**

Alice and Eve

(r) If I use my cellphone camera to make a movie of daytime traffic on Nassau Street in Princeton, it has enough memory for about 20 minutes. If instead I make a movie of the night sky while looking for meteor showers, what will I likely discover about the length of movie I can make?

longer movie at night **shorter movie at night** **about the same** **no way to predict**

longer at night. It's a lot easier to compress black with no motion.

(s) In whose collected works would you most likely find the sequence ***1, 2, 3, …, ut, re, mi, fa, sol, la***?

Babbage **Bach** **Goethe** **Hertz** **Leibniz** **Mahler** **Newton**

Leibniz. We spent quite a while talking about his development of binary numbers and his hexadecimal encoding that uses musical notes instead of letters for 10..15. "Bach" was by far the most common (but wrong) answer.

(t) Modern processors like those in current PCs and Macs have multiple "cores," that is, two or more individual CPUs on a single chip. Assuming that all the potential processing power can be perfectly utilized, how does that processing power increase in proportion to **n**, the number of CPUs on a chip?

logarithmic **linear** **n log n** **quadratic** **cubic** **exponential**

linear.

(u) Whose picture appears on the United Kingdom's newish 50 pound note, on the opposite side from Queen Elizabeth II?

Turing. Those who attended the lecture where Archie and I passed around real ones probably got it right.

(v) Which pair of these acronyms are most closely related? Circle the two closest.

CSS GCC GPT GPU LLM MD5 NAT NDA

GPT, LLM. Probably easy for those who attended the lecture where we talked about these, though it's also explicit in the AI/ML lab.

(w) A technical white paper says "5G phone systems should be able to support 1 million connected devices per square kilometer." If supported devices were spread around uniformly, how many of them could you fit into a square meter?

1. A square kilometer is a million square meters.

(x) What decimal *integer* is the infinitely long binary number `1101100.1111111111`… closest to?

109. A simple conversion from binary to decimal, with a familiar number to confirm the answer.

(y) If I want to create a new top-level Internet domain called `.bwk`, analogous to `.biz`, `.info`, etc., which one of these would have to authorize its creation?

DNS ICANN ITU registrar root server TLA W3C WIPO

ICANN.

2. (30 points) Understanding Programs

(a) [10 pts] The following Python code is supposed to print a 3-column table that shows each integer from 1 to 100 inclusive, together with its square and its cube; there should also be a line at the beginning of the table that labels the three columns. Sadly, the program doesn't work. Fix the errors by rewriting the code or clearly showing the changes you would make. (This is a question about correct logic, not syntax, but be clear about indentation so I can tell what you mean.)

```
n = 0
while n < 100:
    print("n     n squared      n cubed")
    print(n, n*n, n*3)
```

```
n = 1
print("n    n squared    n cubed")
while n <= 100:
   print(n, n*n, n*n*n)
   n = n + 1
```

or any number of variants. The most common error was failing to move the heading print outside the loop.

(b) [6 pts] Suppose that the Toy machine is augmented with a new instruction `ABS` that replaces the value in the accumulator by its absolute value. That is, if the accumulator value is negative it becomes positive, and if the accumulator is positive it is unchanged. This program uses the `ABS` instruction, with reminders about what the instructions do.

```
MORE   GET                get a number from user, place it in accumulator
       IFZERO END         if accumulator value is 0, go to END
       IFPOS  MORE        if accumulator value is >= 0, go to MORE
       ABS                replace value in accumulator by its absolute value
       ADD    FOO         add value in location FOO to value in accumulator
       STORE  FOO         store value in accumulator in location FOO
       GOTO   MORE        take next instruction from location MORE
END    LOAD   FOO         load value in location FOO into accumulator
       PRINT              print value in accumulator
```

```
        STOP
FOO     0
```

If this program is given the sequence of input numbers **2 1 –7 3 –8 –4 –6 5 9 0**, what does it print?

25. It's adding up the absolute values of the negative input numbers.

(c) [3 pts] The fourth line (**ABS**) in this program could be moved to one other place and the program would produce the same answers. Where is that place?

After END LOAD FOO. That converts the sum of the negatives numbers to its absolute value rather than converting each individual number on input.

(d) [4 pts] The Python function **weird** takes two arguments, a list **A** and a value **x**, and returns an integer. In no more than about a dozen words, state clearly what it computes.

```
def weird(A, x):
   i = 0
   while i < len(A):
      if A[i] == x:
        return i
      i = i + 1
   return -1
```

Returns position where x occurs in A, or -1 if it doesn't.

(e) [2 pts] What is the value returned by **weird([-2,-1,0,1,2], 0)**? The expression in **[** brackets **]** is a list.

2. A 0 appears in the 3rd position of the list, but counting starts at zero.

(f) [2 pts] What is the value returned by **weird([0,1,2,4,8,16], 3)**?

-1. 3 doesn't appear in the list.

(g) [3 pts] Modify the implementation of **weird** in any non-trivial way that preserves its API and correct operation. (fiddling with spacing or changing the names of variables is too trivial.) You only need to indicate clearly what you would change.

i += 1 is probably easiest, but a range(…) would work if you get the limits right, i.e., (0, len(A)).

3. (100 points, 5 each) Miscellaneous

(a) Molly White, author of the blog web3isgoinggreat, uses the Twitter handle **@molly0xFFF**. Suppose that Joe Green and Susan Black decide to copy Molly's idea.

(i) What would Joe Green use in place of **0xFFF** ?

0x0F0

(ii) What would Susan Black use?

0x000

(iii) How many potential choices would Earl Gray have that do not collide with Molly, Joe or Susan?

14. There are 16 shades of gray in this format (not 4096!), but **000** and **FFF** are taken.

I thought this would be a fun and easy variant on the usual color questions, but it was not as well handled as I had hoped.

(b) Alice and Bob, bored out of their minds in a COS 109 lecture in Friend 008, are exchanging messages with each other using Gmail from their laptops and texts from their phones. Alice has an Android and uses AT&T; Bob has an iPhone with Verizon. Alice uses Windows while Bob uses macOS. For each of the following statements, assess its likely accuracy.

Their mail messages will use TCP/IP and HTTPS	[✓ **likely**]	**unlikely**
Their text messages will use the router in the ceiling	**likely**	[✓ **unlikely**]
Their text messages will go through different base stations on campus	[✓ **likely**]	**unlikely**
Their mail addresses will be logged by servers at AT&T and Verizon	**likely**	[✓ **unlikely**]
Their text phone numbers will be logged by servers at Google	**likely**	[✓ **unlikely**]

(c) The US International Trade Commission ruled in October 2023 that some versions of the Apple Watch violated pulse-oximetry patents owned by Masimo Corp. For each of these statements, is it likely to be true or false?

Masimo can prevent Apple from importing these watches into the US	[✓ **true**]	**false**
Masimo can prevent Apple from selling these watches in the US	[✓ **true**]	**false**
Masimo can refuse to license its patents to Apple	[✓ **true**]	**false**
Masimo can prevent other watch makers from using its patented technology	[✓ **true**]	**false**
Apple could invent or buy some technology that does not infringe Masimo's patents	[✓ **true**]	**false**

This story was in the news a lot during exam week. I hope everyone recognized it.

(d) Suppose that the signal received by a cell phone at a distance of one mile from a base station is 100 milliwatts.

(i) How many milliwatts will it receive at a distance of two miles?

25. Radio strength follows an inverse square law, so twice as far is 1/4 of the power.

(ii) How many milliwatts will it receive at a distance of five miles?

4. 1 / (5*5)

(iii) Why are they called "cell phones"? Mark the right answer.

A biological metaphor, of cells communicating with each other and proliferating as in biological systems

[✓ **A geometrical metaphor**], of honeycomb-like cells that fill a given geographical area

A security metaphor, of locked-down systems that must be jail-broken to access more services

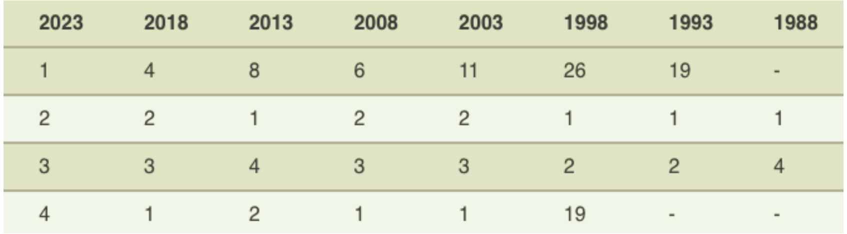
None of these

Probably easier for those who were in class when cell phones were discussed, though it's also in the book.

(e) *Ars Technica* says "OpenAI estimates that it took more than 300 billion trillion floating point calculations to train GPT-3. That's months of work for dozens of high-end computer chips." ***Very roughly***, how many months would that same computation have taken on your laptop? (A month is about 3 million seconds.)

(f) **10^8?** 300 billion trillion is 3 x 10^23 (not 21, as many people thought). A typical laptop like the ones that appeared in problem sets and lectures are roughly 10^9 ops/sec. Some people have faster laptops, including a few that claim a trillion ops/sec, probably because they have GPUs. That's ok as long as the arithmetic was done correctly.

(g) The Tiobe Index, a somewhat flaky measure of language popularity, reports this long-term history:



2023	2018	2013	2008	2003	1998	1993	1988
1	4	8	6	11	26	19	-
2	2	1	2	2	1	1	1
3	3	4	3	3	2	2	4
4	1	2	1	1	19	-	-

What are the top four languages of 2023, in order? Hint: none are primarily aimed at the web.

Python, C, C++, Java. The hint is pretty explicit about excluding JavaScript.

(h) At the end of the New York City marathon, the organizers report each runner's name, place, and finish time, in order of finish time.

(i) If you know a friend's finish time and you want to find out from this list, as efficiently as possible, what her place was, how long would it take in proportion to **n**, the number of runners?

log n **n** **n log n** $\mathbf{n^2}$ $\mathbf{2^n}$

log n. use binary search.

(ii) If you don't know anything about your friend's time, how long will it take?

log n **n** **n log n** $\mathbf{n^2}$ $\mathbf{2^n}$

n. linear search

(i) A recent story at Statista.com predicts that the number of IoT devices will grow from 15 billion in 2020 to 30 billion in 2030.

(i) Assuming that this is a smooth exponential growth, what is the ***approximate*** growth rate per year of the number of Internet-connected devices?

7%. Trivial application of the Rule of 72. Most people just did that.

(ii) If growth continues at the same rate, in what year will there be 15 trillion connected devices?

2120. There are ten doublings from 15B to 15T, at 10 years/doubling.

(j) A story in *Advertising Age* says that T-Mobile is suing Lemonade, an insurance company, over the latter's use of the color magenta, which T-Mobile claims it owns.

(i) What kind of intellectual property is at issue here?

trademark

(ii) What specific kinds of intellectual property might the eponymously-named lawyers Phosita and Eula be associated with?

Phosita **patents** Eula **licenses**

No credit for just expanding the acronyms.

(k) A pixel is a picture element and a voxel is a volume element. Suppose you wanted to attach tiny probes all over your body to serve as "touchels", that is, units of touch. (Whether these might be used for sensing or stimulation we will leave to your imagination.) If each touchel is 0.1 inch by 0.1 inch, estimate ***very roughly*** how many touchels there would be on your body. You can use metric units if you prefer; if so, assume that touchels are 1 mm by 1 mm. *You must reason quantitatively*. Be sure to state your assumptions clearly.

A million, give or take a factor of something? Watch out for excess precision. And it is about surface area, not volume! Most well done, except for occasional failures to convert square dimensions.

(l) Somewhat surprisingly, with 23 people in a room the odds are about 50% that two people will have the same birthday; with 50 people, it's 99.99%. Suppose that each person in the class writes their name and birthday on an index card.

(i) Describe an ***efficient*** algorithm to determine whether any two people in the class have the same birthday. (Don't worry about multiple duplicates or triples.) ***Be clear but brief***; two or three short sentences is enough.

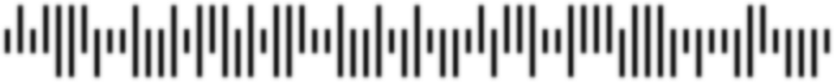
Use quicksort to get the dates in order, then look for adjacent duplicates. Alternatively, sort one list with quicksort, then use binary search for the dates in the other list.

Hardly anyone said anything like this straightforward, so it was really hard to see what people had in mind.

(ii) If there are **N** people, how does the time that your algorithm takes vary in proportion to **N**?

N log N.

(m) The US postal service encodes address information in "Intelligent Mail" barcodes like the one in the picture below. There are 65 vertical bars.



(i) How many different possible addresses could this encoding represent? Just give an expression.

2^130 or 4^65. A remarkable number of people said 2^65, apparently missing the fact that there are four possible values at each position (as mentioned in class and clearly visible in the image).

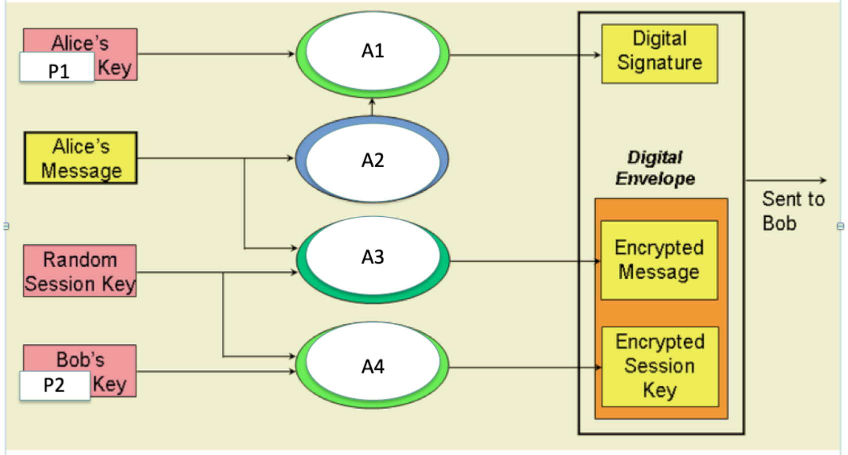
(ii) What is the closest power of 10 to this number?

10^39

(n) In the diagram below, various cryptographic algorithms and terms are used to describe the process of encrypting and digitally signing a message from Alice to Bob. What are valid words or terms to insert in the spaces marked P1, P2, A1, A2, A3, and A4?

P1 __________**private**__________________ P2 ___________**public**________________

A1 **public key / RSA** A2 **crypto hash / SHA** A3 **secret key / AES** A4 **public key / RSA**



Not very well done, unfortunately, though I could probably have made the question clearer.

(o) The *NY Times* described a system that allows parents to monitor teenagers who are driving the family car, by visiting a web site that displays the current car location on a map. Assuming that the system is implemented in the most technically feasible and sensible way, assess the likelihood of the following statements:

The position of the car could be monitored by satellite imaging **likely** [✓ **unlikely**]

A GPS receiver in the car could broadcast the car's location to a GPS satellite **likely** [✓ **unlikely**]

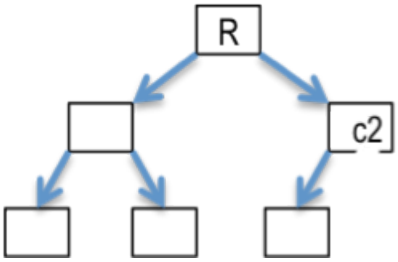
A GPS-enabled cell phone in the car could report its location to a cell phone base station [✓ **likely**] **unlikely**

A cell phone could only report its location when a conversation is in progress **likely** [✓ **unlikely**]

RFID would be a viable alternative to a cell phone-based location system **likely** [✓ **unlikely**]

GPS is a broadcast-only system, as noted in class and the book.

(p) A particular computer network is organized as a *balanced binary tree*: the root computer **R** is directly connected to two child computers, each of which is in turn connected directly to two other child computers, and so on, with no duplicates. "Balanced" means that the connections are made so that as much as possible each computer has exactly two children. For instance, if another computer were added to the network below, it would be added as the right child of **c2**.



(i) Which of these terms best describes how the number of connecting wires will grow in proportion to **n**, the number of computers on the network?

logarithmic [✓ **linear**] **n log n** **quadratic** **cubic** **exponential**

(ii) Which of these terms best describes how the maximum distance from any computer to any other computer in the network will grow in proportion to **n**?

[✓ **logarithmic**] **linear** **n log n** **quadratic** **cubic** **exponential**

(q) UTF-8 is a variable-length encoding of Unicode that represents ASCII characters as one byte and other characters as two or more bytes. It uses one or more of the leftmost bits of each byte to indicate the length of the character: 1-byte characters are represented as `0xxxxxxx`, and 2-byte characters are represented as `110xxxxx 10xxxxxx`, where each `x` is either `0` or `1`. (This encoding saves space if the text is mostly ASCII.)

(i) How many one-byte characters can this format encode? (An expression is ok.)

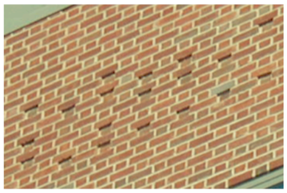
2^7 = 128

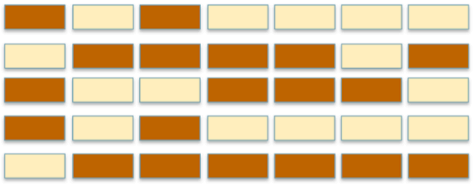
(ii) How many two-byte characters can this format encode? (An expression is ok.)

2^11 = 2048

(r) The picture on the left shows a pattern of bricks protruding from the wall of a campus building. I've drawn it more clearly on the right.

(i) ***Exactly*** what does the pattern say?





P=NP? A gratifying number of people got this, apparently remembering similar examples from previous problem sets and exams.

(ii) Bonus: What is the answer? **Yes** **No** [✓ **We don't know**]

Still an open problem. There's a million-dollar prize if you resolve it.

(s) An IPv4 address is a 32-bit integer.

(i) If the IPv4 address `63.254.255.255` is stored in a 32-bit integer variable `v` and incremented by the statement `v = v+1`, what is the resulting value, also expressed in dotted decimal notation?

63.255.0.0. We did almost this exact problem in the Q/A; I hope the attendees benefited.

(ii) What is that resulting value expressed in hexadecimal?

`3F FF 00 00`. I accepted **`0 0`** instead of **`00 00`** though technically both zeros are necessary.

(t) [10 pts] Quickies. Circle the best answers:

Fei-Fei Li '99 won a Turing award for her development of ImageNet	**true**	[✓ **false**]
"Satellites track your cellphones and can tell 911 operators where you are"	**true**	[✓ **false**]
Intel dominates Nvidia in the GPU marketplace	**true**	[✓ **false**]
A two-factor device is hardware used to do efficient prime testing for the RSA algorithm	**true**	[✓ **false**]
In supervised learning, a human evaluates each decision made by an ML system	**true**	[✓ **false**]
An ML word vector encodes directional information like "north" and "south"	**true**	[✓ **false**]
Bitcoin transactions are anonymous to the US Internal Revenue Service	[✓ **true**]	**false**
"AI Winter" refers to an artificial intelligence funding cycle that ends every December	**true**	[✓ **false**]
It is possible for there to be more static web pages than IPv4 addresses	[✓ **true**]	**false**
Javascript code can easily monitor where you move your mouse on your laptop screen	[✓ **true**]	**false**