

Final Pre-Exam

Read this page and fill in your name, pledge, and email ID.

Name: _____

UVA Email ID: _____

For this pre-exam, you must **work alone**. You are not permitted to obtain help from people other than asking clarifying questions of the course staff. You are not permitted to provide help to others taking the exam.

You are permitted to use any resources you want to complete this pre-exam including course materials, anything you find on the Internet or in a library, and any generative AI tools. You should expect at your scheduled exam, though, to be asked to explain your answers, and answer follow-up questions that probe your understanding of the topics.

Sign below to indicate that you understand these expectations and can be trusted to behave honorably:

Signed: _____

Representing Information

1. Solve these arithmetic problems, you should express your answer using the same number notation as used in the equation.

(a) Binary: $11111 + 1 =$ _____

(b) Binary: $1010101 \times 10 =$ _____

(c) Hexadecimal: $FE + 1 =$ _____

(d) Hexadecimal: $FFF + 2 =$ _____

2. In a protein sequence, there are 22 different amino acids. How many bits are needed to represent one amino acid?

3. Why is nearly all computing done on digital data instead of analog data? (You don't need to write anything for this, but should be prepared to answer this question, which was on the first Exam.)

The Internet and the Web

4. Explain why Google searches are so fast even though the amount of data on the Internet is so huge. (You don't need to write anything for this, but should be prepared to answer this question, which was on the first Exam.)

5. These questions are about understanding latency and bandwidth, and how the Internet works. Your answers to not need to be practical.

(a) What is something that could be done to improve (increase) *bandwidth* on the Internet, but that is unlikely to improve (reduce) *latency*?

(b) What is something that could be done to improve (decrease) *latency* on the Internet, but that is unlikely to improve (increase) *bandwidth*?

Language

Consider the BNF Grammar below (with starting non-terminal S):

$$S ::= R T R$$

$$R ::= R R$$

$$R ::= 0$$

$$R ::= \epsilon$$

$$T ::= 1 T$$

$$T ::= 1$$

6. Show how the grammar can derive 10 starting from S .

7. Write a simpler BNF grammar that describes the same language as the one above. Your grammar should use as few rules as possible, and produce exactly the same set of strings as is produced by the language above.

Programming

Recall this (believed to be correct) JavaScript code for computing edit distance from Class 20 and the previous Exam:

```
1 function edit_distance(s, t) {
2     if (!s && !t) { return 0 }
3     if (!s) { return t.length }
4     if (!t) { return s.length }
5
6     if (s[0] == t[0]) {
7         score = 0
8     } else {
9         score = 1
10    }
11
12    return Math.min(
13        score + edit_distance(s.substr(1), t.substr(1)),
14        1 + edit_distance(s.substr(1), t),
15        1 + edit_distance(s, t.substr(1))
16    )
17 }
```

8. For a spell-checking application, we will be passing in the word the user typed as the first input (**S**), and the dictionary candidate word as the second input (**t**). A very common spelling mistake is to leave out one of a repeated letter. For example, if the user typed `mispell` it is much more likely that they meant to type `misspell` than they meant to type `mipell` or `mistpell` (assuming both of those are dictionary words). Explain how you would modify the `edit_distance` code to reflect this. (Hint: it is more complicated than just doing the change that was done on the exam, since that would make `mistpell` and `misspell` equally likely. Note that we are not expecting you to write the correct code for this, but to demonstrate enough that you understand what changes would be needed and have some idea how to implement them.)

Artificial Intelligence and Machine Learning

9. How did Turing define a machine, and why such a simple formal notion of a machine is a reasonable model for the complexity of modern computers. (You don't need to write anything for this, but should be prepared to discuss this question.)

10. Explain why Large Language Models like ChatGPT are prone to producing plausible text that is not grounded in facts (sometimes called hallucinations). (You don't need to write anything for this, but should be prepared to discuss this question.)

End of Pre-Exam

Bring your completed Pre-Exam to your scheduled exam and be ready to discuss your answers.