

Project

CSE20S24

Due: 5/8/24 at 5pm (late submission until 8am next morning) - Extended by 24 hours

In the project component of this class, you will extend your work on assignments and explore applications of your choosing. *Why?* To go deeper and explore the material from discrete math and how it relates to Computer Science. You will watch some videos and read some articles, and then connect them to our work in CSE 20. There are two tasks in the project, and for each one you will submit a short video and a PDF document, each addressing specific questions.

As you work on the project, keep in mind our three high-level goals for CSE 20:

- Model systems with tools from discrete mathematics and reason about implications of modelling choices. Explore applications in CS through multiple perspectives, including software, hardware, and theory.
- Know, select and apply appropriate computing knowledge and problem-solving techniques. Reason about computation and systems. Use mathematical techniques to solve problems. Determine appropriate conceptual tools to apply to new situations. Know when tools do not apply and try different approaches. Critically analyze and evaluate candidate solutions.
- Clearly and, unambiguously communicate computational ideas using appropriate formalism. Translate across levels of abstraction.

What resources can you use?

This project must be completed individually, without any help from other people, including the course staff (other than logistics support if you get stuck with screencast). You can use any of this quarter's CSE 20 offering (notes, readings, class videos, homework feedback) and videos and articles explicitly referenced in the project description. These resources should be more than enough. If you are struggling to get started and want to look elsewhere online, you must acknowledge this by listing and citing any resources you consult (even if you do not explicitly quote them), including any large-language model style resources (ChatGPT, CoPilot, etc.). Link directly to them and include the name of the author / video creator, any search strings or prompts you used, and the reason you consulted this reference.

If you get stuck on any part of the project, we encourage you to focus on communicating what you think the question might mean, including referring to an example from class or homework you think might be relevant, and include in your submission a discussion of any aspect where you're unsure. Clear communication about these theoretical ideas and their applications is one of the main goals of the project.

Submitting the project

You will submit a PDF plus a video file for each of the two tasks. All file submissions will be in Gradescope. One way to record the video is to record your screen (this is sometimes called screencast). You can use any software you choose. One option is to record yourself with Zoom; a tutorial on how to use Zoom to record a screencast (courtesy of Prof. Joe Politz) is here: https://drive.google.com/open?id=1KR0MAQuTck40zwrEFot1YSJJQdcG_GUU. The video that was produced from that recording session in Zoom is here: <https://drive.google.com/open?id=1MxJN6CQcXqIb0ekDYMxjh7mTt1TyRVM1> Please send an email to the instructor (minnes@ucsd.edu) if you have concerns about the video / screencast components of this project or cannot complete projects in this style for some reason.

Task 1: Exploring an application

In CSE 20 this quarter, we will be exploring the applications of discrete mathematics for core Computer Science and Engineering topics. **Pick one** of the following videos about work done here at UC San Diego (or by people associated with UC San Diego) and complete both steps of the task described below.

[Video] Bioinformatics and virology *Niema Moshiri* <https://www.youtube.com/watch?v=PrAoks70hE8>

[Video] Human robotics interaction *Healthcare Robotics Lab at UCSD* <https://www.youtube.com/watch?v=bS0-asHDXPc>

[Video] Natural Language Processing *Taylor Berg-Kirkpatrick* <https://www.youtube.com/watch?v=8zMfAdPZKnk>

[Video] Cryptography and Complexity *Russell Impagliazzo* <https://youtu.be/RjzSFa03i2U>

[Video] Machine learning (and surfing) for climate science *Jasmine Simmons and Engineers for Exploration* <https://www.uctv.tv/computer-science/search-details.aspx?showID=34350>

1. Watch the video from above that you selected. **Record a new video** where you present your answers to the following three questions:
 - Which video did you watch?
 - Why did you choose the video you watched?
 - What are three kinds of data or information that are related to the project described in the video?

Your video for this task should be 1-3 minutes. Start with your face and your student ID visible for a few seconds at the beginning, and introduce yourself audibly while on screen. You don't have to be on camera for the rest of the video, though it's fine if you are. We are looking for a brief confirmation that it's you creating the video and doing the work submitted for the project. When you are explaining the three kinds of data or information (third part of question 1), we recommend you show them in the video in some form.

2. In the document part of this task, you will explore how to use CSE 20 techniques to model each of the three kinds of data or information that you identified in your video. In particular, answer the following questions in a document that you will submit to Gradescope:

- Describe how each of the three kinds of data or information can be modelled using the data types that we discussed in class (sets, ordered n -tuples, strings, or functions). Explain why you are choosing to model this data or information with this data type: what are the benefits and what are the limitations of this model?
- Write a set using roster method or set builder notation or recursive description that has at least three elements and that demonstrates your model. Include a description of the data or information in English and also include how it is represented in your model.

Type out your answers to these questions for all of the three kinds of data you identified, how you could model each one, and example sets, and upload your PDF to Gradescope.

Task 2: Errors and multiple representations

Sometimes, the way we represent data leads to imprecision or outright mistakes. Watch the video and read the articles below and then complete the task described.

[Video] Minecraft mysteries: https://youtu.be/ei58gGM9Z8k?si=oWZQtM_9-7WTGuR0

[Article] Excel bug causes a wide-spread problem in published genomics papers. <https://www.nature.com/articles/d41586-021-02211-4> (You may need to be on the UCSD network to access this article.)

[Article] IEEE profile of Katherine Johnson, a NASA “computer” who calculated trajectories for early space exploration and who passed away in 2020

<https://spectrum.ieee.org/the-institute/ieee-history/katherine-johnson-the-hidden-figures-mathem>

[Article] NASA report about the unsuccessful 1999 Mars Climate Orbiter mission <https://solarsystem.nasa.gov/missions/mars-climate-orbiter/in-depth/>

[Article] Article about NASA Voyager 1 data corruption

<https://www.sciencealert.com/nasa-has-finally-identified-the-reason-behind-voyager-1s-gibberish>

1. Record a video where you discuss your answers to these questions:

- What are examples in the video or articles above where computers or Computer Science were used to help **avoid** an error?
- Give an example where **you** used computers or Computer Science to help you **avoid** an error?
- What are examples in the video or articles above where the use of computers or Computer Science **caused** an error.
- Give an example where **your** use of computers or Computer Science **caused** an error.
- What do you do to increase your confidence in the results of your own human and digital (i.e. machine) computation? Why do you think these are sufficient?

Focus on your communication clarity in the video for this task. Imagine that your audience is a high school student who is exploring the benefits and drawbacks of using computers to solve problems. Your video for this task should be 1-3 minutes. Start with your face and

your student ID visible for a few seconds at the beginning, and introduce yourself audibly while on screen. You don't have to be on camera for the rest of the video, though it's fine if you are. When you are giving examples, you should speak about them as well as having them displayed on the screen (written or typed using clear and correct notation if relevant, or screen shots from the video if relevant).

2. In the document part of this task, you will explore what mistakes our choice of representations can cause by doing the following:
 - Pick one of the definitions we've used in CSE 20 for representing **numbers**.
 - Copy the definition for this representation into your writeup, and cite which page of which week's notes you're using.
 - Describe, using roster method or set builder notation, the set of numbers that can be represented with this number representation definition.
 - Give an example of a limitation of this number representation by showing what error could be introduced when using this number representation for an application of your choosing.

Type out your work above and upload your PDF to Gradescope.

Grading

Your work on the project will be assigned a letter grade.

- To earn at least a C on the project, most parts of the project should be attempted and your submission should correctly demonstrate some of the tools, techniques, and formalisms we used in class.
- To earn at least a B on the project, almost all of the parts of the project should be substantially complete, and a significant amount of detail and correct notation should be used throughout your examples and explanations.
- To earn at least an A on the project, all of the parts of the project should be substantially complete, with correct and appropriate use of CSE 20 concepts and notation and clear and detailed explanations.

Since the project is also used to add +/- modifiers at the end of the quarter, you can consider going beyond the requirements as well.

- For example: in your PDF writeup for task 1, you could propose several alternate modelling choices and discuss the tradeoffs (advantages and disadvantages) between them.
- Another example: in your PDF writeup for task 2, you could discuss what could be done to detect or correct possible errors resulting from choices of data representations.
- You are welcome to explore other extensions too. Keep in mind, however, that resource limitations means we will be limited to grading no more than 4 minutes of video and 2 pages for each task.