CS166 Final Project

Task: Build a simulation of your choice to evaluate the effects of different strategies, different experiments, or different actions taken within your simulation.

Topics: Two topics for scenarios to model and simulate are provided below. You can choose either of these topics or come up with your own topic. If you choose to do your own topic, follow the instructions in the Final Project Proposal assignment to submit your proposal.

Main goal: Use a model/simulation to evaluate different strategies, actions, or choices in the scenario being modeled. Your model and simulation have to incorporate some randomness. This is usually not difficult since most processes contain some uncertainty or inherent randomness. As part of your assignment, you will create a Monte Carlo simulation of the random process and interpret the distribution over results generated by your simulation.

Table of contents

Feedback and grading

In a short paragraph near the beginning of your project report, describe which part or parts of your work you would like feedback on most. Your request can be specific or vague but you should expect the feedback you get to match your request. Specific requests will get feedback that is detailed and narrow in scope. Vague requests will get feedback that is general and broad in scope. Both types of feedback can be useful. You should decide and explain what you want. Your instructor has limited time to provide feedback (and you have limited time to read and process it) so use this as an opportunity to get what you need out of your instructor.

You will get 1 grade on each of the 6 course learning outcomes plus some HC grades. Be sure to address all 6 LOs as well as the foregrounded HCs. All course LOs are graded using the grading policy you were given at the start of the course.

Project outline

- Describe the scenario you are modeling.
    a. What do you want your simulation to be able to do and what information do you want to be able to measure from it?
    b. Describe the rules of your simulation and how they capture your scenario.
    c. Identify any modeling assumptions and explain under what circumstances these assumptions may and may not be valid.
    d. Describe your model parameters and how they affect the behavior of the simulation.
    e. Describe the output/measurements of your simulation and how they relate to the scenario you are modeling. What are the quantities of interest?
- Implement your model in well-structured Python code using classes as appropriate. Demonstrate that your code is correct and produces the expected behavior of your model. Make sure your code is easy to read.
- Describe and interpret your simulation results, and provide advice on the best course of action to take or strategy to employ in your scenario.
    a. What is the distribution over the outputs of your simulation? Visualize, describe and analyze your results.
    b. Remember that you need to compare at least 2 different strategies, actions, experiments, or other choices in the scenario being modeled.
    c. You should use histograms to show the distribution of your simulation outcomes. Since this is a Monte Carlo simulation, different simulation runs will give different, random results.
    d. Write a paragraph or two to advise someone on the best action to take based on your results. Think of this as writing an executive summary to the client or employer for whom you built the simulation. Write this paragraph in a style that is appropriate for your specific audience.
- Analyze the uncertainty in your results. In a Monte Carlo simulation, the more simulation runs there are, the more certain you are about the average results. Do all of the following.
    a. Provide 95% confidence intervals of the expected values of your metrics.
    b. Comment on whether these confidence intervals are wide (high uncertainty) or narrow (low uncertainty).
    c. Comment on how many more simulations or simulation steps would you need to reduce the widths of your confidence intervals.

Topics

Rainfall and flooding

Use a 2-dimensional cellular automaton for a rainfall and flooding model. Each cell has a particular elevation, depending on the local topology of the landscape. Water can infiltrate the soil unless the soil is waterlogged. Water can run off to lower cells, where "lower" means smaller (cell elevation + cell water depth).

Come up with (or research online) realistic rainfall patterns for the area or type of area you choose to model. When it rains, which areas flood? And, how might you mitigate or eliminate floods in important areas – improving or building canals or dams, for example?

You may choose to model flooding at the scale and type of landscape you choose. Do you want to model rainfall and floods in a town, over a region containing multiple towns and farms, for a whole country? Use this Wikipedia article on floods for inspiration about the different types of floods that can occur.

Use your model to provide advice about how likely floods are to occur and what can be done to mitigate the effects of floods.

Resources

You do not have to read or use any of these resources – they are provided for support and inspiration. Feel free to look around for other resources. If you do decide to use a resource, be sure to cite it in your project report.

- Cirbus, J., Podhoranyi, M. (2013). Cellular Automata for the Flow Simulations on the Earth Surface, Optimization Computation Process.
- Coppola, E., et al. (2007). Cellular automata algorithms for drainage network extraction and rainfall data assimilation.

Waste removal

A waste removal company has a truck (or a few trucks) and needs to visit farms, collecting waste and dropping it off at one or more waste sites in the area. There is a road system, represented by a network, that connects the farms, the waste removal company, and the waste drop-off sites. Edge weights in the network represent the length of each road. Different farms have different average rates of waste production, but it is not fully deterministic so there will be more waste on some days than others. Determine a good collection schedule and route for the waste removal company given the uncertainty in the network. There is a possibility that a farm might have no or very little waste.

Assumptions

- Each truck has a finite capacity for carrying waste.
- Each truck has a finite fuel supply. When it is running low on fuel, it needs to return to the company headquarters to refuel.
- The waste removal company wants to optimize for time and profit by driving between farms as little as possible and conserving fuel.
- A truck might drive to another farm, wait at the current farm (using time but not fuel), go to a waste drop-off site, or return to the company headquarters where it can refuel.

Note: It is not expected that you find the optimal strategy in this problem – just a fairly good strategy. This is a difficult optimization problem – essentially, the traveling salesman problem with some extra complications.

How to submit your deliverables

Submit a PDF report with your model description and results. You must also include a zip file with all code (in a Jupyter notebook) and data files needed to reproduce your results. Your instructor should be able to run your code and reproduce the same results you got.

PDF report

- Provide a PDF file containing a neatly typed report with numerical results, graphs, and text describing your work.
- Write this as if you are writing a report to a client or a supervisor. The purpose of the report is to explain your statistical models and present your results. Make it professional.
- Do not simply make a PDF out of your Python notebook. A client or supervisor will not read through pages of code to get to your explanation of the model and the results.
- A good report length is approximately 10 pages – much more or less than this means you are either not providing enough information, or you are not being concise enough.

Python notebook

- The purpose of the Python notebook is to show how you implemented your model and calculated your results.
- It has to be reproducible (as far as possible given that there is random variation in the simulation). Make sure your instructor can run your notebook from start to finish without any bugs or errors.
- Remember to include any data you used in the zip file you upload, so your code can be run to reproduce your results.
- Make your code readable – use good variable names, add code comments and docstrings, and organize your code using functions.

Where appropriate, identify any specific HCs that you apply in your work and write footnotes explaining your applications of those HCs.