Modeling and Simulation

CSE Written Qualifying Exam

Spring 2022

Instructions

- Please answer three of the following four questions. All questions are graded on a scale of 10. If you answer all four, all answers will be graded and the three lowest scores will be used in computing your total.
- Please write clearly and concisely, explain your reasoning, and show all work. Points will be awarded for clarity as well as correctness.

Problem 1

Problem setup: Consider a parallel discrete-event simulation consisting of logical processes (LPs) that interact exclusively by exchanging time-stamped messages. We regard the simulation as *correct* if it processes the events in a manner consistent with a sequential execution: a correct sequential execution would process those events in nondecreasing timestamp order, assuming unique timestamps. To attain correctness in the parallel case, the usual condition we try to enforce is the *local causality constraint*: each LP only processes events in nondecreasing timestamp order.

Suppose you are building a parallel discrete-event simulator but your customer requires *repeatable results* across the simulation runs. That is, if the customer runs your simulator twice—each time with the same external inputs (e.g., initial events), the same initial state, and the same number of both logical processes (LPs) and physical processes (e.g., compute nodes or processors)—then she expects the final results of the two runs to be the same.

Questions: Is enforcement of the local causality constraint sufficient to ensure repeatability? Why or why not? If you believe it can be sufficient, explain what assumptions are needed to justify that claim. If you do not believe it is sufficient, explain what else your simulator needs to do to make its runs repeatable. (You should always assume that the simulation is stochastic, meaning any event handlers may invoke pseudorandom number generators.) If you believe the answer depends on whether the synchronization scheme is conservative or optimistic, explain the two cases separately; if you believe it does not, explain why not.

Problem 2

Consider a parallel discrete-event simulation that implements an optimistic synchronization protocol, such as TimeWarp. A tricky issue is *error handling*, for errors like dividing by zero or executing an out-of-bounds array reference (e.g., reading or writing to an array location `x[i]` for an invalid `i` value). In particular, if an error occurs during the execution of an event handler, the program should **not** abort if that event is later rolled-back. (In such a case, the execution should appear as though the event *never* executed.) Explain what you might do to make an optimistic simulator more reliable, in the sense of reducing or eliminating the possibility of fatal program errors during events that are later rolled back. To help answer this question, you might consider different types of errors that could occur and give examples of strategies to handle those cases.

Problem 3

In this problem, you will consider a population of bacteria that produces a waste product that in high enough concentrations can be toxic to the bacterial population. The concentration of bacteria, represented by x, and the concentration of the waste product, represented by y, obey the following equations:

$$\frac{dx}{dt} = (a - by)x, \quad (1)$$

$$\frac{dy}{dt} = cx - dy. \quad (2)$$

Consider all four parameters (a, b, c, and d) to be positive.

a) Explain the model. What are the roles of the four parameters? What is the biological meaning of each of the terms on the right-hand sides of the equations?

b) Find all equilibria (fixed points) of the system and analyze their stability.

c) You should find in part (b) that a change in dynamics occurs for some condition. Explain the biological meaning of the condition.

Problem 4

The following model is being proposed to study the transmission of a disease.

$$\frac{dS}{dt} = b - ((1-f)\lambda_N + f\lambda_R)\, S - \mu S, \tag{3}$$

$$\frac{dI_N}{dt} = (1-f)\lambda_N S - (d+\mu) I_N, \tag{4}$$

$$\frac{dI_R}{dt} = f\lambda_R S - (d+\mu) I_R, \tag{5}$$

$$\frac{dP}{dt} = rP\left(1 - \frac{P}{K}\right) + \xi_N I_N + \xi_R I_R - \mu_P P. \tag{6}$$

The state variables have the following meanings.

S	Size of human population not infected with pathogen (susceptible)
I_N	Size of human population infected with non-resistant pathogen
I_R	Size of human population infected with resistant pathogen
P	Size of pathogen population

The parameters have the following meanings.

b	Birth rate of human population
f	Fraction of infecting pathogens that are resistant
λ_N	Infection rate of human population with non-resistant pathogens
λ_R	Infection rate of human population with resistant pathogens
μ	Natural human death rate
d	Human death rate from infection
ξ_N	Pathogen shedding rate by individuals infected with non-resistant pathogens
ξ_R	Pathogen shedding rate by individuals infected with resistant pathogens
Γ	Birth rate of pathogen
K	Carrying capacity of pathogen
μ_P	Death rate of pathogen

a) Describe in words what each equation means by discussing what each term in each equation indicates.

b) List at least 10 assumptions or choices that the model authors make. For each, discuss whether each assumption/choice made seems reasonable or is potentially problematic.

c) For four problematic assumptions/choices you identified in part (b), suggest a modification to the model that could counteract the problem and discuss how practical such a solution would be given the ease of obtaining relevant data for validation and/or determination of parameter values. (You may not address a problem by outright removing the problematic component!) The modification may be a verbal description; it does not have to be a mathematical expression.