Automation in Complex Systems – EIEN35

Exam Wednesday June 2, 2021

You may bring the course book and the reprints (defined in the course requirements), *but not* the solution to problems or your own solutions to simulation tasks and home works or other personal notes. A calculator is also permitted (memory cleared). You may answer in **Swedish** or in **English**.

Grading: There are 30 points all together. The following grades will apply:

Grade 3:	at least 15 points
Grade 4:	at least 20 points
Grade 5:	at least 25 points

Note that all answers should be complete and well motivated. Your line of thought should be easy to follow and hand calculations should be provided for all mathematical problems. Corrections will be completed not later than *Wednesday **June 23,** 2021.*

As this is a remote exam (due to the Corona pandemic) there is an **extra need** that you provide complete analytical/well formulated/well written solutions to all your answers. By taking this exam you hereby confirm that you have read all the information and instructions regarding the exam sent by email 2021-05-26 and you have agreed to follow all those instructions to the letter. If NOT, you should stop the examination NOW and come back for the next exam opportunity on August 25, 2021.

Also note that this written exam may be complemented by an individual oral exam before a final grade is issued to any student (as described in the emailed instructions).

Good Luck!

Problem 1 (4 points)
Numerical stiffness can be a major problem when simulating dynamic systems.

a) Describe the principles of how a stiff solver works. (1 p)
b) Discuss at least two situations when a stiff solver does not work very well. (1 p)
c) Motivate why an Euler forward approximation may produce numerically unstable solutions if the integration time step is set too large. (1 p)
d) A system is simulated using a 4^{th} order Runge-Kutta-solver. The progress forward in time is slow. A quick analysis shows eigenvalues of the system ranging in size by a factor 10^6. The overall system characteristics are not suited for a stiff solver. How can the situation be improved? (1 p)

Solution

a) All stiff solvers are implicit solvers (e.g. Gear) and take into account also future values x(n+1) when determining in which direction to move in the next integration step. Therefore at least one algebraic equation must be solved for each integration step, and this may require some kind of built in iterative solution method for non-linear systems (e.g. Newton-Raphson). The advantage is that implicit solvers can take long integration steps to compensate for the extra computational burden. Stiff solver are often based on the so called corrector-prediction method, where an explicit method, the predictor step, gives an *estimate* of the value of x(n+1). This value is then used in place of x(n+1) in the right hand side of the implicit method, the corrector step. The more sophisticated explicit AND implicit solvers used today are variable step-size solvers that adjust the integration step according to the selected tolerances and the system dynamics on-line.

b) A problem with stiff solvers is that they do not handle transient and random conditions well. Typical examples of such conditions are: poor initial conditions, noise, (highly) dynamic input, discrete events, mixture of continuous and discrete systems. All these situations will prevent the stiff solver from using its great advantage of taking long integration steps.

c) For example if we apply the Euler forward to a system $dx(t)/dt = -ax,\ a > 0,\ x(0) = 1$ we find

$$\frac{dx(t)}{dt} \approx \frac{x(t+h) - x(t)}{h} \Rightarrow x(t) \approx (1 - ah)x(t)$$

For this system to be stable, we must have

$$|1 - ah| \leq 1$$

which means that

$ah \leq 2$ and $ah \geq 0$

Thus, while Euler backward (implicit) has no upper limit on the step length (not shown here), Euler forward (explicit) soon runs into instability issues when the step length is increased. The same principle holds for any system we apply Euler forward to.

d) The difference among the eigenvalues shows that the dynamics of for the system is very different in terms of time scale. Consequently this is a stiff system and therefore the RK-solver (explicit) does not work well. Apparently a stiff solver is not an option due to other reasons. Two possible solutions exist: 1) buy a faster computer (not a good option) and 2) rewrite some of the equations as algebraic equation (either making the changes instantaneous for the fastest dynamics or using constant values for the slowest dynamics. Thereby the stiffness can be reduced and the explicit solver will work fine. Note that an iterative solver may be needed to internally solve the algebraic equations at each integration step. Also, this solution implies a slight approximation of the system, which must be acceptable to the user.

Problem 2 (6 points)

As an automation consultant you are supposed to help a manufacturing firm choose between three potential production systems. In all cases the arrival and production rates are considered to be stochastically (exponentially distributed) and all queuing disciplines are based on FIFO. The *total* arrival rate into the system is 10 per hour.

System 1: one machine with the production rate = 15 per hour.

System 2: three identical machines in parallel each with production rates = 5 per hour. One common queue for all machines.

System 3: three identical machines with *individual* queues and each machine with production rate = 7 per hour. An arrival dispatcher system that divides the arrivals equal among the three queues is assumed to exist.
No limitation on the buffer (queue) size has been defined by the manufacturer, so space is apparently not a major issue.
Analyse the three systems based on number of jobs in the system, the waiting time in the system and in the queue(s). Which system would you recommend? For the system you deem best also provide the probabilities of there being 0, 1, 2, 3, 4 items in the system (i.e. p_0, p_1, p_2, p_3, p_4) on an average.
Also briefly discuss (in general terms) the potential cost and complexity of the three different systems as well as the redundancy and availability of the systems.

Solution:
System 1 is a standard M/M/1 system, since we have no limitation on queue size. For this system the utilization factor = 10/15 = 0.6666. The number of jobs in the system is

$$L = \frac{\rho}{1-\rho} = \frac{0.6666}{1-0.6666} = 2 \text{ items}$$

The waiting time in the system is

$$W = \frac{1}{\mu-\lambda} = \frac{1}{15-10} = 0.2 \text{ hours}$$

The waiting time in the queue is

$$W_q = \frac{\rho}{\mu-\lambda} = \frac{0.6666}{15-10} = 0.1333 \text{ hours}$$

System 2 is a standard M/M/3 system, since we have no limitation on the common queue size, where

$L_q = \frac{m^m \rho^{m+1}}{m!(1-\rho)^2} p_0$ and $p_0 = \left[\frac{m^m \rho^{m+1}}{m!(1-\rho)} + \sum_{n=0}^{m} \frac{(m\rho)^n}{n!}\right]^{-1}$. The utilization rate is 0.6666 and p_0 is therefore equal to 0.1111. This gives L_q = 0.8888 items. The waiting time in the queue is

$$W_q = \frac{L_q}{\lambda} = \frac{0.8888}{10} = 0.0889 \text{ hours}$$

The waiting time in the system is

$$W = W_q + \frac{1}{\mu} = 0.0889 + \frac{1}{5} = 0.2889 \text{ hours}$$

The number of jobs in the system is

$$L = \lambda \cdot W = 10 \cdot 0.2889 = 2.889 \text{ items}$$

System 3 is made up of three parallel standard M/M/1 systems, since we have no limitation on queue size and each machine has its own queue. They will all behave in the same way. For one of the systems the utilization factor = 3.3333/7 = 0.4762. The number of jobs in that system is

$$L = \frac{\rho}{1-\rho} = \frac{0.4762}{1-0.4762} = 0.9091 \text{ items}$$

The waiting time in the system is

$$W = \frac{1}{\mu-\lambda} = \frac{1}{7-3.3333} = 0.2727 \text{ hours}$$

The waiting time in the queue is

$$W_q = \frac{\rho}{\mu - \lambda} = \frac{0.4762}{7 - 3.3333} = 0.1299 \text{ hours}$$

For the combination of the three subsystems in system 3, L = 3*0.9091 = 2.7273 items, W = 0.2727 hours and W_q = 0.1299 hours (the waiting times are the same for the subsystems as for the complete system since they operate independently of each other).

If we consider L then system 1 is the best. However, L does not make much difference in this case since we have no space limitation in the queues. In terms of waiting time in the queue, system 2 is best and systems 1 and 3 are almost identical. But for the overall production the most important variable is the waiting time in the system (queue time + production time) and here we see a clear advantage for system 1. So based on the information we have available, system 1 should be considered to be the best one. As system 1 is an M/M/1 system the resulting p vector is simply calculates as:

$$p_n = \rho^n (1 - \rho)$$

which gives us $\mathbf{p} = [p_0 \;\; p_1 \;\; p_2 \;\; p_3 \;\; p_4] = [\, 0.3333 \;\; 0.2222 \;\; 0.1481 \;\; 0.0988 \;\; 0.0658 \,]$.

With regard to cost and complexity the normal situation would be that a single large machine is cheaper than several small machines (with the same total production) since normally less material, power electronics and supply, local controllers are needed. Also the complexity with regard to the surrounding equipment is reduced if only one queue (and associated equipment for moving items from the queue to the machine) is required and everything is fed into one machine rather than several. However, the internal complexity and need for control etc. may be higher in the larger machine. With regard to redundancy and availability it is clear that the probability that something will be produced at all times is highest if there are three separate systems (individual queue and machine) and also higher for system 2 than for system 1. This is a drawback of using one efficient machine – if it breaks down then the production will be completely stopped. So if the demand that something is produced more or less at all times is very high then we may need to change our priorities somewhat with regard to if system 1 is really the best or not.

Problem 3 (2 points)

More and more renewable energy is being developed, mainly solar PV and wind energy. These energy sources become part of a distributed generation power system. Explain some differences between a conventional (centralized) power system and a decentralized distributed generation system with respect to:

a) Describe the power flow in the network in the two cases. (1 p)
b) How is the variability of the wind and solar compensated in a distributed system, so that the customers do not suffer? You can assume that the wind and solar are connected to a larger electric grid. (1 p)

Solution:

a) In the centralized system there is one direction of the power flow, from the top generators down to the customer. In the decentralized system the power flow will go in different directions, depending both on the loads and the present production.
b) In Sweden we use the hydropower to compensate. However, this will challenge the capacity of the transmission lines, in particular from northern Sweden to southern Sweden.

Problem 4 (2 points)
The Swedish power grid is, as partly described in the previous problem, a complex system. The generation of power can have very different degree of controllability, especially with an increasing amount of renewables. The consumption is mostly not possible to control but external factors, like the weather, have an impact. Import and export of power is controllable but with restrictions. Still we want to keep the voltage and the frequency close to nominal values.
Use the framework presented by T. Gillblad to describe as many types of complexity as you can find in the control of the power grid. Motivation is crucial.

Solution:
The structure of the power grid immediately tells us that there has to be **integrative complexity**. Very many nodes with generation and consumption all connected to the same grid and consequently have to work at the same voltage and the same frequency. Definitely a system where fairly simple components are interacting, thus creating complexity.
All the external factors that will affect the grid also creates **operational complexity**. Weather conditions, social events and individual behaviour all create a varying, and not easily predictable, demand for power. With the renewable generation in the form of wind turbines and PV local weather conditions that are very difficult to forecast will affect the production.
Finally one can also say that the grid operation contains **functional complexity**. To maintain the voltage and the frequency a sophisticated centralized semiautomatic control of the controllable production units (hydropower, nuclear etc.) must be implemented.

Problem 5 (2 points)
You are implementing a data collection system with very many distributed intelligent sensors. This should be an IIOT (Industrial Internet Of Things) application and should use 5G for the communication. With this said – what is defined and what is left to define according to the OSI framework? Motivate and briefly describe referring to the levels in OSI.

Solution:
The 5G network with IIOT will provide layer 1-3 (physical-data link-network). This means that you only have to provide valid IP-adresses (probably IP V6) and then make definition about the presentation layer (how you format the data in the transfer), the session layer (when and how you start a transfer) and the application layer (the meaning of the data and how you use it).

Problem 6 (3 points)
A process industry is having frequent problems due to production faults and disturbances, which are normally detected too late, causing periods of process shut-down. The existing SCADA system collects some 1000 different measurements on-line and the process manager realises that the existing traditional univariate monitoring systems need to be updated. The company has access to several years of good quality data in its data bases and has also kept excellent track of problems in the past in terms of written log books describing when different problems occurred, what they led to and what the probable causes of the problems were. It is your job (as an automation consultant) to lead the work of setting up a new monitoring system for detection and isolation of faults and also add diagnostic features. Discuss the principles of

how you would approach and solve this task (you do not need to explain all details but focus on the main important issues of such a system and its development and how it relates to the people working as operators).

Solution

There is no absolute correct answer but an example is given here. Based on the large number of measurements a multivariate monitoring tool (e.g. based on PCA) is a possible solution. Based on the historic data and the log books, periods when the process is working within the allowed 'normal' boundaries can be identified and used for creating a PCA model of the plant, which would drastically reduce the number of variables to be monitored. The model would then be used on old data together with the log books to properly calibrate warning and alarm limits for criteria such as SPE and T^2, but also to identify which principal components are usually changing and in what way when the most common types of process faults occur. Thereby 'scatter plots' of combinations of principal components could be used on-line both to detect process problems and also diagnose the most common causes at a very early stage. Potentially a knowledge based system could also be developed to assist the operator in the diagnostic work, since we have access to significant amounts of old data together with detailed log books. Good knowledge of the process is essential and also to package the whole system in an intuitive and easy-to-use tool based on a good HMI, thereby using graphics, sound, colour etc. to enhance the work of the operator. Good instructions for the operator of how different situations should be responded to are needed and the plant should continue to store data, maintain log books (preferably integrated in the monitoring system) for any future modifications of the system. The involvement of the operators during the development and test phase of the system is also essential. Training for operators also essential. Special efforts would be required to identify and diagnose new types of problems, which the process has historically not yet encountered and integrate those into the existing system. A strong commitment by the plant managers/owners is required since these types of monitoring and diagnostic systems are never really finished but need to updated and improved and tuned over time. (Some graphical descriptions of PCA principles and scatter plots with diagnostics could also be included in the response.)

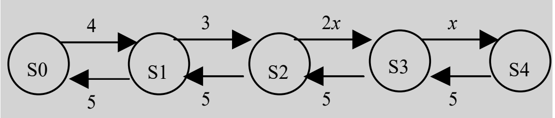
Problem 7 (3 points)

Consider a machine with a production rate of 5 jobs per hour, where the production rate is exponentially distributed. The arrival rate of jobs (exponentially distributed) is controlled by an ideal job dispatcher based on the following rules:

1. If the machine is empty then jobs are sent to the machine system with a rate of 4 jobs per hour;
2. If 1 job is present in the machine system then the job delivery is 3 jobs per hour;
3. If 2 jobs are present in the machine system then the job delivery is $2x$ jobs per hour;
4. If 3 jobs are present in the machine system then the job delivery is x jobs per hour;
5. If 4 jobs are present in the machine system then incoming jobs will be rejected and not be processed in the machine.

The manufacturer accepts a maximum rejection probability of 5%. Set up the state graph and calculate the required value of x to meet the demand of the manufacturer. Also provide the complete stationary probability vector for the limit value of x.

Solution*:*
This is a classical birth-death process with 5 states.



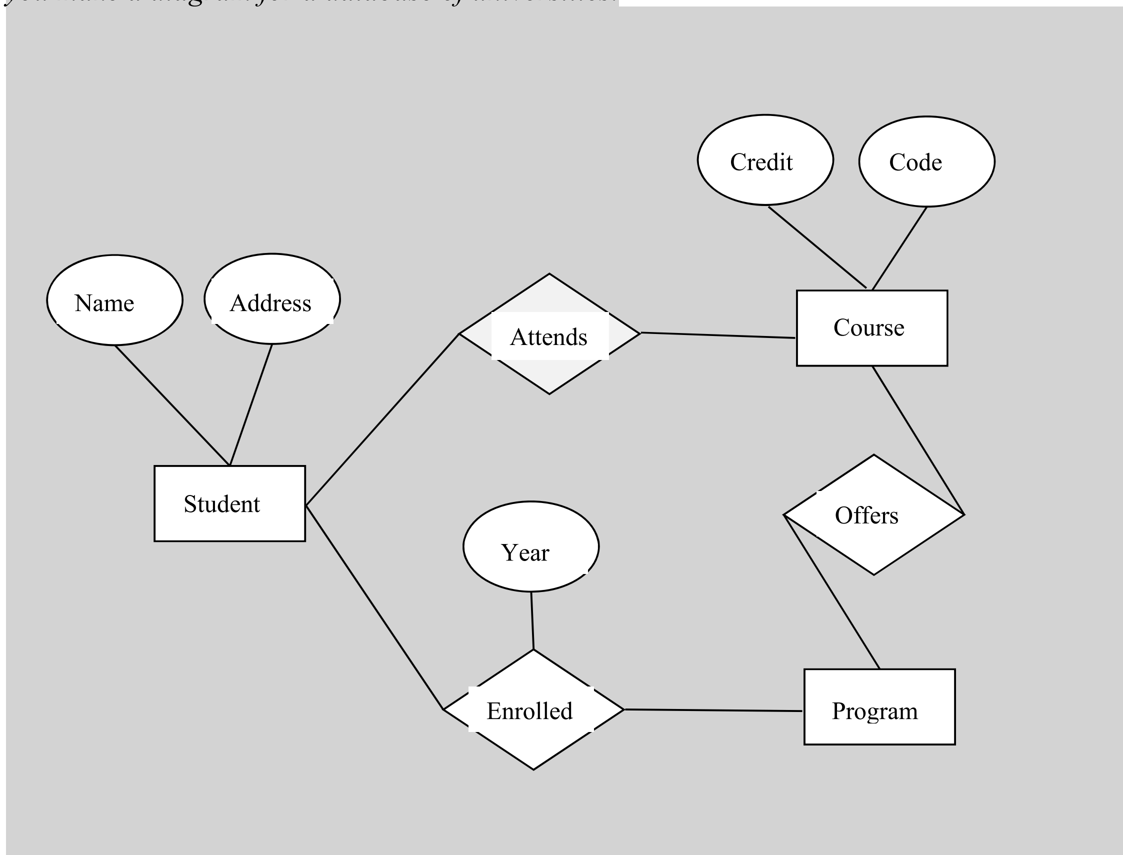
There can be 0, 1, 2, 3 or 4 jobs in the system, having the probability p0, p1, p2, p3, p4, respectively. The production rate is all the time $\mu = 5$. The arrival rates are:
s0 -> s1 = 4; s1 -> s2 = 3; s2 -> s3 = 2x; s3 -> s4 = x; s4 -> s5 = 0.
The following condition must hold: p0 + p1 + p2 + p3 + p4 = 1.
The general expression for this type of system is:
$p_{k+1} = \frac{\lambda_k}{\mu_{k+1}} p_k$, which gives us

p1 = 4/5*p0; p2 = 3/5*p1 = 12/25*p0; p3 = 2x/5*p2 = 24x/125*p0; p4 = x/5*p3 = $24x^2$/625*p0.
So the full probability vector is:
p = p0 * [1 0.8 0.48 0.192x $0.0384x^2$] and when we normalize this condition (since p0 + p1 + p2 + p3 + p4 = 1) we get
p = 1/(2.28+0.192x+$0.0384x^2$) * [1 0.8 0.48 0.192x $0.0384x^2$]. The rejection probability is equal to p4, i.e. $0.0384x^2/(2.28+0.192x+0.0384x^2) <= 0.05$. Solving this equation gives the roots $x_1 = 1.9042$ and $x_2 = -1.6411$. A negative delivery rate is not a scientifically feasible solution, i.e. $x <= 1.9042$
For the limit value of x, **p** ≈ [0.3591 0.2873 0.1724 0.1313 0.05]

Problem 8 (2 points)
Make a schema for a database, using the ER (Entity Relationship) -model, that describes the current basic education at LTH. (No history but the current situation.) The model should contain: Courses, Students and Programs. You should add some additional data like course credits and home address for students.

Solution

Entities: Student, Course, Program
Relations: Attends, Offers, Enrolled_at
Attributes: Name, Address; Code, Credits; Year
Note: If you add an entity like University or similar in the diagram it is wrong. Then you make a diagram for a database of universities.



Problem 9 (2 points)

a) If an unexperienced person designs an HMI-system for a continuous process it is a common error that only the "normal" control is supported. What other situations should also be taken care of? (1 p)

b) There are "funny" bikes built with two cogwheels between the handlebars and the front wheel. Since the wheel turns the opposite direction, compared to a normal bike, these bikes are very difficult to ride. Explain using the models from HMI why. (1 p)

Solution

a) Startup and shutdown should be a part of the design but are often forgotten. An even wider area is the operation in alarm situations. It is very easy to handle single alarms but to anticipate the combination of alarms in alarm chains where dependent functions give multiple alarms is very difficult. To help the operator find the root cause is a real challenge.

b) Since we have learned how to ride a bike so that it is treated on the skill based behaviour it is very difficult to avoid these fast reactions. We have to block these "spinal" reflexes and try to learn the "new" way first through "knowledge based behaviour" and then as "rule based behaviour". This is not an easy relearning process and emphasizes the recommendation to use previous experience in all HMI design.

Problem 10 (3 points)

a) In "plant wide control", according to Skogestad, the term "degree of freedom" is important. Describe, using your own words, the meaning of this term in this context. (1 p)

b) Give your own description of the different main types of "degrees of freedom" and explain each one of them with a simple example. (1 p)

c) Give an example when two actuators form only one dynamic degree of freedom. (1p)

Solution

a) A degree of freedom is typically an actuator of some kind. It is a way of changing the way the production works. It could result in changes in flow, temperature, pressure mix of ingredients but is something that has got an impact on the production process.

b) **Dynamic degree of freedom** includes most of the actuator signals with the exception of signals that are paired so that nothing happens if only one of them is activated. In that case the pair will form one dynamic degree of freedom.
Steady state degree of freedom should be something that has an impact over a longer time. E.g. a buffer tank can have one actuator to fill it and another actuator to empty it but the flow through the system is the only long term effect and thus the whole system inly represents one steady state DOF.
Optimization degree of freedom is the degree of freedom left when all active constraints are considered. In some processes the constraints make it impossible to vary some parameters since a certain objective just has to be fulfilled. E.g. the power grid has to continuously deliver the required power even if it would have been at a better price one hour later.

c) Two valves in series on one tube could very well be a sensible design to be able to change an instrument between them but for the process flow they form only one dynamic DOF.

Problem 11 (1 point)

Describe in general terms the advantage of using ISA88 (S88) for batch processing.

Solution

S88 defines a way of handling the equipment description and then recipe for a batch production separated. In this way the same production can be run on another equipment with minor changes and a lot of the control operations can be generated automatically.