Automation in Complex Systems – EIEN35

Exam Wednesday June 5, 2019

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 26,*** *2019.*

Good Luck!

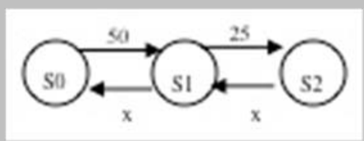
Problem 1 (3 points)

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

1) If the machine is empty then jobs are sent to the machine system with a rate of 50 jobs per hour;
2) If 1 job is present in the machine system then the job delivery is 25 jobs per hour;
3) If 2 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 10%. Set up the state graph and calculate the minimum possible value of *x* to meet the demand of the manufacturer. Also provide the complete stationary probability vector for the calculated value of *x*.

Solution:
This is a classical birth-death process with 3 states.



There can be 0, 1 or 2 jobs in the system, having the probability p_0, p_1, p_2, respectively. The production rate is all the time $\mu = x$. The arrival rates are:
s0 -> s1 = 50; s1 -> s2 = 25; s2 -> s3 = 0.
The following condition must hold: $p_0 + p_1 + p_2 = 1$.
The general expression for this type of system is:

$p_{k+1} = \frac{\lambda_k}{\mu_{k+1}} p_k$ which gives us

$p_1 = 50/x*p_0$; $p_2 = 25/x*p_1 = 1250/x^2*p_0$. So the full probability vector is:
p = p_0 * [1 50/x 1250/x^2] and when we normalize this condition (since $p_0 + p_1 + p_2 = 1$) we get
p = $x^2/(x^2+50x+1250)$ * [1 50/x 1250/x^2]. The maximum allowed rejection probability is equal to p_2 = 0.10, i.e. $1250/(x^2+50x+1250) = 0.10$. Solving this equation yields two solutions, -133.97 and 83.97. Obviously the production rate cannot be negative so 83.97 per hour is the minimum required production rate for the above system if the rejection probability should be 10% or less. The complete stationary probability vector for the limit value of *x* is consequently: **p** = [0.5641 0.3359 0.10].

Problem 2 (5 points)

Consider a system where four independent machine works in parallel, each with the average capacity 10 jobs per hour. The machine outputs are exponentially distributed. Likewise, jobs arrive to the machines stochastically (one common queue) with 30 jobs per hour at an average. Also the job arrival time is exponentially distributed. The service discipline is FIFO.

a) Assume that a maximum of 5 jobs is allowed in the queue.
   - Calculate the average waiting time in the queue (**1 p**)
   - Calculate the average waiting time in the system (**1 p**)
   - Calculate the average number of jobs in the system (**1 p**)
   - Calculate the rejection rate, in other words, the number of jobs per hour that are never admitted to the system (**1 p**)

b) You want to replace the system in (a) with an M/M/1 system (FIFO, infinite queue) that has the same average waiting time in the system. Calculate the necessary average capacity of the single machine to achieve this. (**1 p**)

Solution:

a) This is a traditional M/M/4/9 system. The utilization rate is 30/40 = 0.75. The probability vector is calculated as

$$p_0 = \left[\frac{m^m \rho^{m+1}\left(1-\rho^{K-m}\right)}{m!(1-\rho)} + \sum_{n=0}^{m} \frac{(m\rho)^n}{n!} \right]^{-1}$$

$$p_n = \frac{(m\rho)^n}{n!} p_0 \quad \text{(n=1,2,3,4)}$$, which gives (after some calculations)

$$p_n = \frac{m^m \rho^n}{m!} p_0 \quad \text{(n=5,6,7,8,9)}$$

p = [0.0415 0.1245 0.1867 0.1867 0.1401 0.1050 0.0788 0.0591 0.0443 0.0332].

$\bar{L}_q = \frac{m^m \rho^{m+1}}{m!(1-\rho)^2}\left[1-\rho^{K-m}-(1-\rho)(K-m)\rho^{K-m}\right]p_0$, which gives L_q = 0.7833 jobs. From here we calculate W_q as $W_q = \frac{\bar{L}_q}{\bar{\lambda}}$ where $\bar{\lambda} = \lambda(1-p_9) =$ 29.0029 and W_q = 0.0270 hours. The total number of jobs in the system is calculated as

$\overline{L} = \frac{\rho}{1-\rho} - \frac{(K+1)\rho^{K+1}}{1-\rho^{K+1}}$, which gives $L = 3.6836$ jobs

The average waiting time is given by Little's theorem where the average arrival rate into the system is 29.0029. Therefore $W = \frac{\overline{L}}{\overline{\lambda}} = 3.6836/29.0029 = 0.1270$ hours.

The rejection rate is $\lambda - \overline{\lambda} = 0.9971$ jobs per hour.

b) The system is now modified into a M/M/1 system. For such a system $W = \frac{1}{\mu - \lambda} = \frac{1}{\mu - 30} = 0.1270$ hours. Therefore the production rate of the new machine must be at least 37.87 per hour.

Problem 3 (4 points)

The numerical solution to a stiff system is sometimes hard to find.

a) Consider the simple first order system: $dx(t)/dt = -ax,\ a > 0,\ x(0) = 1$. Motivate, using Euler backward and forward as demonstration on this system, why implicit solvers are less sensitive to long integration steps compared to explicit solvers. (**2 p**)
b) Unfortunately, stiff solvers cannot always be used. Discuss some important situations, which are not handled well by stiff solvers. (**1 p**)
c) One of the most popular explicit solver is the 4th order Runge-Kutta method. Explain in words and in graphs in detail how this method calculates its next value x_{n+1} for a fixed *h*. (**1 p**)

Solution:

a) Euler backward:

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

For this system to be stable

$$\left| \frac{1}{1+ah} \right| \le 1$$

which means that

$ah \le -2$ or $ah \ge 0$

Euler forward:

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

For this system to be stable

$$|1-ah| \le 1$$

which means that

$ah \le 2$ and $ah \ge 0$

Thus, while Euler backward (implicit) has no upper limit on the step length, Euler forward (explicit) soon run into instability when the step length is increased.

b) A problem with stiff solvers is that they do not handle transient and random conditions well. Typical examples of these are: poor initial conditions, noise, (highly) dynamic input, discrete events, mixture of continuous and discrete systems. Stiff solvers can also get completely stuck.

c) The 4th order Runge-Kutta method is a straightforward explicit solver based on the equations:

$$x_{n+1} = x_n + h\left(\frac{f_1}{6} + \frac{f_2}{3} + \frac{f_3}{3} + \frac{f_4}{6}\right)$$

$$f_1 = f(x_n, t_n)$$

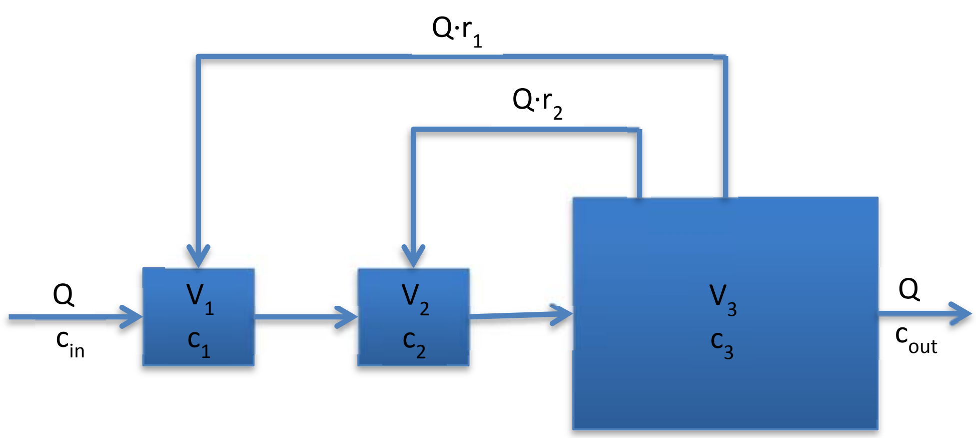
$$f_2 = f\left(x_n + \frac{h}{2} f_1, t_n + \frac{h}{2}\right)$$

$$f_3 = f\left(x_n + \frac{h}{2} f_2, t_n + \frac{h}{2}\right)$$

$$f_4 = f(x_n + hf_3, t_n + h)$$

, where f is the function value (the derivate) of our ordinary differential equation (linear or nonlinear), i.e. $dx/dt = f(x,u,t)$ if we assume a simple one-dimensional system, and h is the integration step. x_n represents the initial or current value of the state variable. First RK calculates the derivative f_1 (i.e. the direction of the state variable) at the current point. Then it moves forward half an integration step in the direction f_1 and calculates the derivative (f_2) at that point. It then goes back to the starting point and moves forward half an integration step in the direction f_2 and calculates a new derivative (f_3) at that point. Finally it moves from the starting point a full integration step forward in the direction of f_3 and there it calculates the derivative f_4. These four directions $f_{1..4}$ are then weighed together by factors (1/6, 1/3, 1/3, 1/6; note that the sum should equal 1) and the algorithm moves forward a full integration step in this direction and adds that value to the previous state variable value (x_n). Thereby it has arrived at a new value for the state variable (x_{n+1}) and then the entire procedure is repeated from that point to calculate x_{n+2} and so on.

Problem 4 (3 points)
You need to simulate a system as defined in the figure below but you only have access to an explicit solver. The input flow rate is Q and the input concentration is c_{in}. The system is completely mixed and volumes are constant. No reactions take place, there is only mixing of the compound c (dissolved in water) in the three tanks. Two recycle flows exist as defined in the figure by factors r_1 and r_2.



a) Set up the three ordinary differential equations that describe the dynamics of the state variables c_1, c_2 and c_3 (**1 p**)
b) Simulations are terribly slow because V_1 and $V_2 \ll V_3$ making the system very stiff. But you have no other solver. You decide to rewrite the system based on two

differential algebraic equations for tanks 1 and 2 and one ordinary differential equation for tank 3, accepting the simplification that changes in tanks 1 and 2 will be immediate by letting V_1 and V_2 go towards zero. Provide the algebraic equations for c_1 and c_2 (as functions only of c_{in}, c_3, r_1, r_2 and Q) and the resulting differential equation for c_3. Simplify the equations as much as possible. (**2 p**)

Solution:

a)
$$\begin{cases} \dfrac{dc_1}{dt} = \dfrac{Q}{V_1}\left(c_{in} + r_1 c_3 - (1+r_1)c_1\right) \\ \dfrac{dc_2}{dt} = \dfrac{Q}{V_2}\left((1+r_1)c_1 + r_2 c_3 - (1+r_1+r_2)c_2\right) \\ \dfrac{dc_3}{dt} = \dfrac{Q}{V_3}\left((1+r_1+r_2)c_2 - (1+r_1+r_2)c_3\right) \end{cases}$$

b) Using singular perturbation we move V_1 and V_2 to the left hand side and let them go towards zero. This yields the following two equations:

$$\begin{cases} Q\left(c_{in} + r_1 c_3 - (1+r_1)c_1\right) = 0 \\ Q\left((1+r_1)c_1 + r_2 c_3 - (1+r_1+r_2)c_2\right) = 0 \end{cases}$$

from the first equation we get an expression for c_1 as a function of c_{in}, c_3, r_1, which we can then use in the second equation to write c_2 as a function of c_{in}, c_3, r_1, r_2. After some rewriting we get: the simple equations

$$\begin{cases} c_1 = \dfrac{c_{in} + r_1 c_3}{1+r_1} \\ c_2 = \dfrac{c_{in} + (r_1+r_2)c_3}{1+r_1+r_2} \end{cases}$$

and finally we can write the differential equation for c_3, as expected after the simplifications, as:

$$\frac{dc_3}{dt} = \frac{Q}{V_3}(c_{in} - c_3)$$

As the equations are linear we can find direct analytical solutions for c_1 and c_2 and there is no need for any iterative solver for the internal algebraic equations.

Problem 5 (1 point)
Discuss the principle differences, advantages and disadvantages of univariate and multivariate monitoring methods.

Solution:
Univariate monitoring is based on the traditional way of checking every variable individually without making use of correlations between the measurement data. By using statistical methods various types of alarm and warning limits can still be defined and the methods are fairly straightforward and easy to grasp and implement. In multivariate monitoring we take advantage of how the different measurements impact one another and how they are correlated. Thereby much more information can be gathered from the data and still presented to an operator in a fairly simple way. However, the underlying mathematics and statistics used are much more complicated. Modern industrial processes involve many hundreds or thousands of

signals. Operators cannot handle information from all individual measurements and it may be difficult to define which are the most essential ones in order to cover all possible types of process disturbances. For most complex processes signals are coupled and not independent. But those synergetic effects cannot be detected by individual signal analysis used in univariate monitoring. Using multivarite monitoring algorithms the thousands of signals can often be reduced into only a few key components that describe the main mechanisms or variability in the process by using the fact that the signals are correlated and dependent.

Problem 6 (3 points)
Use the framework presented by T Gillblad to describe the type of complexity, and explain why it occurs, in the following examples:

a) Making the trains arrive according to the timetable in the national railway system.(**1 p**)
b) Keep a stable voltage and frequency in the national power grid. (Including renewable sources.) **(2 p)**

More than one type of complexity may be involved but the explanation should connect the type and the origin of the complexity.

Solution:

a) First, it is an obvious example of **integrative complexity**. To make one train start on time, travel along on free track to the destination, and arrive on time is not a complex task. However, when hundreds or thousands of trains share the same railway system the probability that something fails becomes very large. Since the trains share the same tracks, the same stations and it should be possible for the passengers to change trains the system is truly complex. To be able to handle this all involved components should be of high quality and the maintenance standard should be high. Redundancy (alternative tracks and extra trains) can help in many cases.
Secondly, we can also find **operational complexity** in the case of extreme weather like much snow or high temperature.
b) To produce the power in the same moment as it is consumed without control of the consumption is a clear example of **operational complexity**. Furthermore, when an error occurs, the consequences of this error (brown out/black out) could usually be avoided if the correct control actions took place. However, the information about the error travels only slightly faster than the consequences. This also makes it operationally complex.

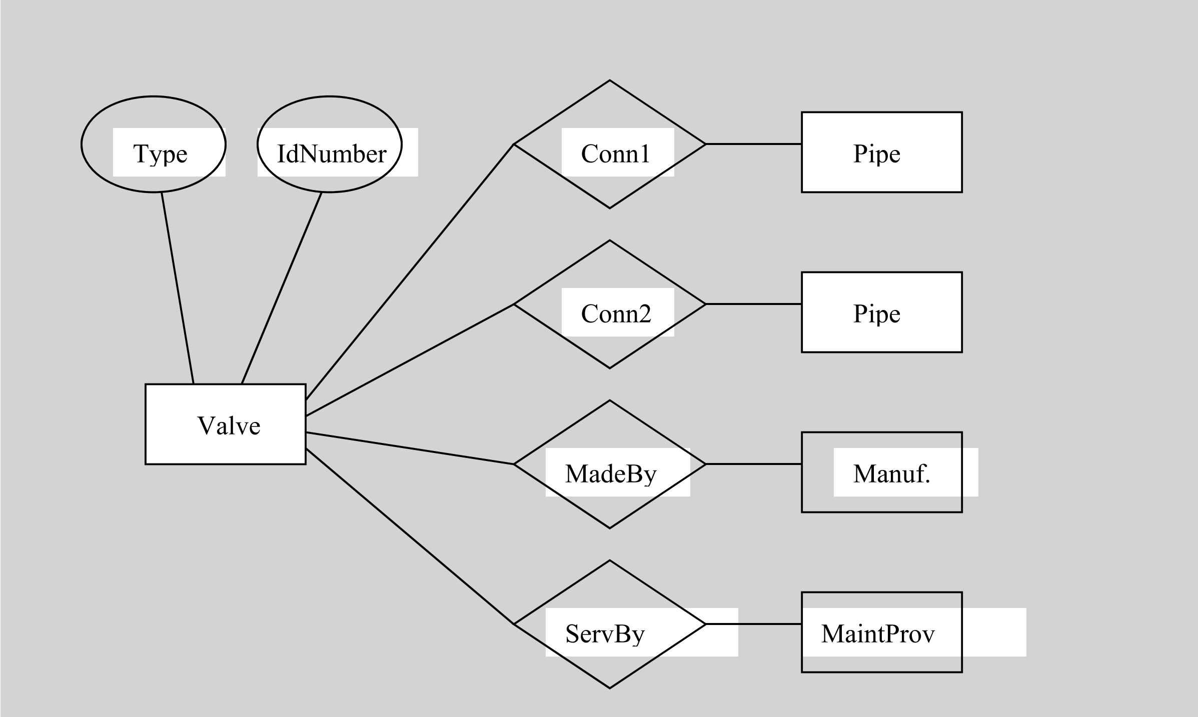
Problem 7 (3 points)

a) What is a "Data Warehouse" structure and when is it used? (**1 p**)
b) Explain the "Entity-Relationship model" in a graphical example with e.g. newspapers or movies as objects. (**1 p**)
c) What is a database "schema"? **(1 p)**

Solution

a) A Data Warehouse is a collection of data from different databases together with tools for easy access. A typical situation is a merge of two companies that have their own databases. To translate all data and put into one database is a very costly operation and it is often a better way to use a DW that extracts the data from the original databases.

b) Entities: Valve, Manufacturer, MaintenanceProvider, Pipe
Relations: MadeBy, ServiceBy, Connection1, Connection2
Valve attributes: Type, IdNumber



c) In the schema the database objects, their relations and their attributes are included. This could be done with e.g. the entity-relationship model.

Problem 8 (2 points)

a) OPC-servers are frequently used in Automation. Why and what do they do? (**1 p**)
b) A crucial part of HMI is alarm handling. Explain the difficulties around this and give some examples of tools that can be used to make a good solution. (**1 p**)

Solution

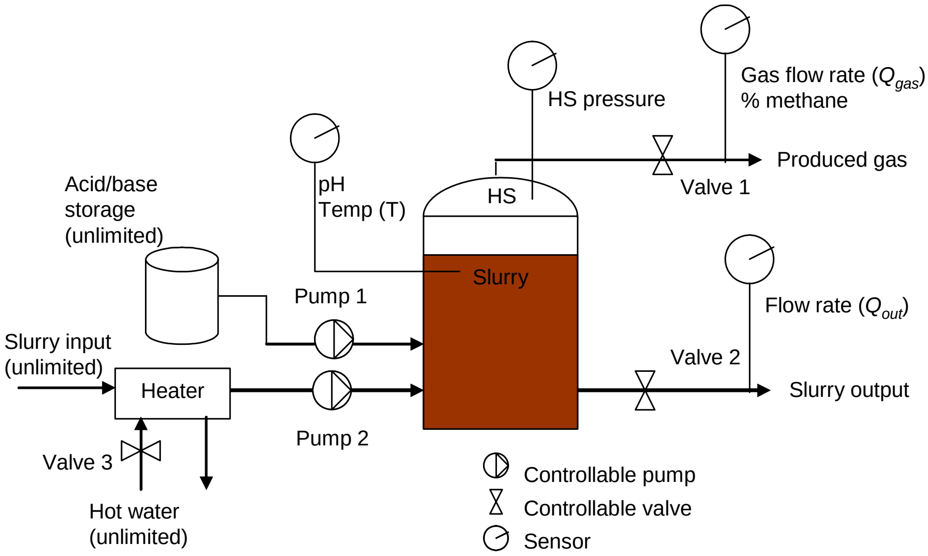
a) OPC-servers are used for communication in automation applications. Data exchange, alarm handling, historical data are also included in different parts of the standard for OPC-servers. They could be described as information hubs in configurations containing e.g. PLC:s, SCADA systems and HMI systems.
b) In most processes are many measurements somewhat related. This means that one error that gives an alarm usually also will trigger many other alarms. This connection could be both in the process variables and in the logical setup of the system. The result is however, an alarm flood that is difficult to handle. The challenge is to find the root cause. Structuring using priorities, alarm groups and summary/history-presentation can help but more advanced methods are available.

Problem 9 (2 points)
Industry 4.0, smart manufacturing and digitalization are terms that are used in the course material both by C Johnsson and T Antius. Together they describe the even more complex industrial future that seems very likely. Write 5-10 points that describe these expected changes both in the perspective of the producer and of the consumer. (Technology is of course the base but don’t forget the market consequences.)

Solution:
Some tools that will be used to a much higher functionality than today are IoT, large data volumes (big data) and AI.
The possibility to use the information about everything will affect both the production and the market for the products
The focus will probably not be on products but rather the outcome.
VR, AR and ML will change the role for the human and the expectations
Natural language and ML will make the HMI very different from today
The smart manufacturing requires a lot of new standardization
The result of smart manufacturing could however, be a more flexible structure compared to the one used today.
…..and many more aspects that can be found in the slides….

Problem 10 (4 points)
Consider the process in the figure. It is an anaerobic digestion process used to produce biogas for heaters and vehicles. The overall goal of the process is to produce as much biogas as possible with the required quality (percentage of methane). One problem is the changing nature of the raw material composition made of organic solids in slurry. The composition of changes results in changes in the biological reactions (BR) in the process. The process consists of a reactor with a continuous input and output of organic slurry and the mean retention time (MRT) (uppehållstid) is about 12 days. The gas produced raises to the surface of the slurry and is trapped in the gas filled head space (HS) above the slurry. The amount of slurry may change over time (input ≠ output) which implies that the HS changes in volume. The main mechanisms and the relations between important variables of the process are described below.

Biological reactions :	$BR = f\,(MRT, T, pH, composition)$
Gas pressure HS:	$P_{HS} = f\,(gas\ valve, BR)$
Gas flow rate:	$Q_{gas} = f\,(BR, P_{HS})$
pH in the reactor:	$pH = f\,(addition\ of\ acid/base, BR)$
Temperature:	$T = f\,(flow\ rate\ of\ heated\ water, ambient\ temperature, BR)$
Retention time:	$MRT = f\,(added\ slurry, removed\ slurry)$



Most of the variables can be varied within reasonable ranges. Reactor temperature may vary between 30 and 38°C, the MRT can range between 8 and 16 days without causing any process problems, the pH can be controlled between 6.5 to 8 and the pressure may range from atmospheric (1 bar) to 1 bar overpressure (2 bar). Liquid levels in the reactor and the storage tank are naturally limited by the tank geometry.

Based on the information above and using the definitions in Skogestad 2004:

a) What is the number of dynamic degrees of freedom (N_m)? **(1 p)**
b) What is the number of the steady state degrees of freedom (N_{ss})? (**1 p**)
c) What is the number of the unconstrained degrees of freedom (N_{opt})? (**1 p**)
d) What can be solved on Skogestad level 3 "Controlled variables"? (**1 p**)

In all answers the motivation and your assumptions are very important to get credit.

Solution

a) Nm=5. Pump 1, 2 + valve 1, 2 and 3.
b) N_{ss} = 4. In steady state, the liquid level must be constant. This means that pump 2 and valve 2 are dependent and must be balanced. The independent manipulated variables are: valve 3 + pump 1 + valve 1 + one of pump 2 and valve 2.
c) N_{opt} = 4. The unlimited input of slurry means that all steady state variables can be used for optimization
d) The temperature of the slurry might be a suitable variable. Depending on the process parameters, pH control could also be suitable. The pressure could probably also be controlled at this level with set points available for a supervisory controller.