Analytical Network and System Administration Workbook

Mark Burgess

August 8, 2008

What is the course about?

This course is about asking searching questions about systems and answering them as impartially as we can. We want to ask questions like the following:

- How to we describe systems conceptually, separating what is important from what is not important?
- How do we maximize performance of the system?
- How do we minimize risk of ‘loss’?
- How do we achieve stability and reliability?
- How do we repair damage?
- What subset of maintenance operations depends on the order in which we perform them?

To answer questions like these, we need to be more precise than we are used to being. We need to learn to think a new language: the language of precision thinking — mathematics.

Some of the problems are simply artificial, designed to discipline and train you into thinking analytically. Others are directly practical. *When you are answering these questions, think of how you would apply this knowledge, for instance in your final masters thesis.*

Write your solutions to these problems legibly by hand. Do not waste time typesetting the answers – spend your time thinking about them!

Some additional literature is recommended to you in answering these problems.

- *Performance by Design*, D.A. Menasce, V.A.F. Almeida and L.W. Dowdy, Prentice Hall, ISBN 0-13-090673-5.
- *Measurement Uncertainty (Method and Applications)*, R.H. Dieck, 3rd edition, ISBN 1-55617-759-X, Intrumentation, Systems and Automation Society (ISA).

Chapter 1

Philosophy of science

The aim of this week's exercises is to make you reflect on what the course will be about and what you hope to achieve from it. In addition, there are some exercises that review basic analytical skills.

Exercise 1 *Skim through the course book to get an idea of what this course is about. Read the introduction and the conclusions. Think of at least one question from networking or system administration that you would like to know the answer to, where you think it might be possible to calculate or learn the answer by analytical methods.*

Solution 1

Exercise 2 *Comment on the difference between the scientific method and the mathematical method (logic). Can we prove that something will never happen, or that something will always work:*

1. *Using the scientific method?*
2. *Using mathematics?*

What are the limitations of these approaches?

Solution 2

Exercise 3 *Comment on the ethics of using the scientific method in system administration. Is it ethical to investigate and analyse systems relentlessly to make decisions about system policy? Would it be just as good to 'cheat' by taking short cuts and making assumptions that are not supported by evidence? Discuss this.*

Solution 3

Exercise 4 *Consider the statement often made in connection with security: "Happy users are well-behaved users". What does this mean? How might you go about testing this hypothesis? How could you characterize happiness of users and well-behavedness. How might you design an experiment to test this? mean*

Solution 4

Exercise 5 *Write down and learn the upper and lower case Greek alphabets and the common names of the glyphs.*

Solution 5 *The Greek alphabet:*

Alpha	A	α
Beta	B	β
Gamma	Γ	γ
Delta	Δ	δ
Epsilon	E	ϵ
Zeta	Z	ζ
Eta	H	η
Theta	Θ	θ
Iota	I	ι
Kappa	K	κ
Lambda	Λ	λ
Mu	M	μ
Nu	N	ν
Xi	Ξ	ξ
Omicron	O	o
Pi	Π	π
Rho	P	ρ
Sigma	Σ	σ
Tau	T	τ
Upsilon	Υ	υ
Phi	Φ	ϕ
Chi	X	χ
Psi	Ψ	ψ
Omega	Ω	ω

Exercise 6 *Let the set x_i for $i = 1, 2, \ldots, 10$ be given by*

$$\{x\} = \{1, 4, 3, 2, 6, 5, 8, 9, 7, 10\} \tag{1.1}$$

1. *What is $\sum_{i=1}^{3} x_i$?*
2. *What is $\sum_{i=1}^{10} x_i$?*
3. *What is $\prod_{i=1}^{4} x_i$?*
4. *What is the mean value of x_i, $\overline{x}$? Explain the meaning of the mean value of a set of numbers.*
5. *What is the expectation value of x_i, $\langle x \rangle$? Explain the meaning of the expectation value of a distribution of numbers.*
6. *What is the difference between the expectation value and the mean value of a set of values? Given an example of usage where expectation value and mean are the same and an example where they are different. (This question is about explaining how an idea is used in two different ways, and it is really a philosophical question.)*

Solution 6

Exercise 7 *Sketch a system consisting of a web server, a DNS server and appropriate hardware, connected to the Internet. Suppose you wanted to study and describe the following; what variables or attributes would you want to measure?*

1. *Performance.*
2. *Architecture.*

3. *Utilization.*

In each case describe the variables you could measure or model, what values they can assume, and explain how they vary or change from place to place.

Solution 7

Exercise 8 *This problem is about dimensional analysis, i.e. how we relate different scales of measurement. Dimensional analysis is a basic skill in science and engineering, where measurements are made and relationships are expressed.*

1. *When FastSearch left Boston, all its computers had to be transported across the country to Sacramento. (Transportation is paid by weight.) Suppose we have N computers and W is the total weight to be transported. Clearly*

$$W \propto N \tag{1.2}$$

so we can write

$$W = kN + c, \tag{1.3}$$

for some constants k and c. If W is measured in kilogrammes, what are the units of k and c? Since Americans do not understand SI units, they have to convert W into pounds, using a formula $P = \alpha W$. What are the dimensions (units) of α?

2. *If A is measured in nanometres and B is measured in kilogrammes, what is the meaning of $A + B$?*

3. *Let A be the number of arrivals (packets) per second in a computer network. Suppose that the conversion factor κ is a constant measured in kilobytes per packet. Write down an expression for the number of kilobytes that arrive in time T seconds.*

4. *In wave theory one expresses waves in the form $A \sin(kx - \omega t)$. Let $\theta = kx - \omega t$, where θ is an angle measured in radians, x is distance measured in metres and t is time measured in seconds. What are the dimensions of k and ω? The wavelength λ is measured in metres per cycle and is related to k. On dimensional grounds, argue how λ and k must be related.*

5. *The frequency f is measured in Hertz (cycles per second). How many radians are there per wave cycle? Find a relationship between f and ω.*

6. *The speed of a wave c is measured in metres per second. Find a relationship between f, λ and c, using dimensional arguments.*

Solution 8

Exercise 9 *In a human-computer system human values are an important part of system design. Comment on whether you think human qualities can be quantified and dealt with using the scientific method.*

Solution 9

Chapter 2

Observations, data collection and uncertainty

The problems this week are aimed at developing the skill of being precise about what you say. Many misunderstandings and excuses for poor practice in system administration are the result of lack of adequate modelling. Mathematics helps us to formulate models in two ways:

- It is a precise language that is well known and standard.
- It expresses many concepts that are relevant to systems and contains tools that will help us to unravel things that are 'implicit' in our assumptions.

Exercise 10 *This problem is about using graphs to explain relationships. In order to estimate the average usage of a computer system a system administrator measures the time interval between each new request for data from disk modification times Δt and counts the frequency $N(\Delta t)$ with which a nearest whole number of seconds occurs.*

Δt	$N(\Delta t)$
1	*57364*
2	*12052*
3	*7005*
5	*4679*
10	*3050*
30	*795*
37	*531*
38	*509*
40	*552*

1. *What do you think these measurements tell us about the pattern of usage of the system?*
2. *Plot the data in the table:*
   (a) $N(\Delta t)$ *against* Δt.
   (b) $\log N(\Delta t)$ *against* $\log \Delta t$.
3. *The system administrator has heard that activity that is generated by human patterns follows an empirical law of the form*

$$N(\Delta t) \propto \Delta t^{-\alpha} \tag{2.1}$$

for some positive constant $\alpha > 0$, and that data stored by automated computer processes follows the form

$$N(\Delta t) \propto \mathrm{e}^{-\beta t}. \tag{2.2}$$

From your data plots suggest which of these two laws best fits the observed data and find the approximate value of α or β.

Why do we care which of the above is true?

Solution 10

Exercise 11 *This exercise is to remind you about the basic 'hands-on' skills of calculus. The great engineers of the Victorian age used these methods to build the industrial revolution. You can use them to revolutionize your understanding of systems.*

When answering these exercises, approach them as you would approach any problem in system administration: do not simply copy an answer from somewhere uncritically, but make sure you understand all the steps underpinning the result, i.e. create a HOW TO for solving these exercises! The expressions below are based on expressions you will meet in the course.

1. *Differentiate the function $\alpha(t) = c_1 t + c_2 + c_3 t^3$, with respect to t, where $c_i, i = 1, 2, 3$ are constants.*
2. *Differentiate the function $\beta(t) = \alpha(t)\mathrm{e}^{-\lambda t}$*
3. *Differentiate $q(t) = q_0\,(1 + \sin(2\pi t/P))$ wrt t. Sketch the function roughly by hand.*
4. *Differentiate the function $Q(t) = q_0\,(1 + \sin(2\pi t/P))\,\mathrm{e}^{-\lambda t}$ with respect to t. Sketch the function roughly by hand.*
5. *Differentiate $\rho(t) = \int_0^b q(t)(t' - t)dt'$ wrt t.*
6. *Differentiate $\rho(t) = \int_0^b q(t')(t' - t)dt'$ wrt t.*
7. *Differentiate $\log(\alpha t)$ wrt t.*
8. *Find the indefinite integral with respect to t of $\sin(\omega t)$*
9. *Find the indefinite integral with respect to t of $\cos(2\pi f t)$*
10. *Find the indefinite integral with respect to t of $\mathrm{e}^{-\lambda t}$*
11. *Find the indefinite integral with respect to t of $\mathrm{e}^{-\lambda t}\sin(\omega t)$*
12. *Find the indefinite integral with respect to t of $c_1 t + c_2 t^2 + c_3 t^3$, where $c_i, i = 1, 2, 3$ are constants.*
13. *Find the indefinite integral with respect to t of $\sum_{n=1}^{m} c_n t^n$ where m, c_n are constants.*
14. *Find the indefinite integral with respect to t of $\log(\alpha t)$*
15. *Evaluate $\int_0^1 \alpha t\, dt$.*
16. *Evaluate $\int_{-1}^1 \alpha t\, dt$.*
17. *Evaluate $\int_0^\pi \sin(\omega t) dt$ as far as possible.*
18. *Evaluate $\int_{-a}^{+a} t\cos(\omega t) dt$ exactly.*

19. *(Hard) Evaluate* $\int_{-\infty}^{+\infty} \mathrm{e}^{-at^2}\, dt$.

20. *Describe in words the fundamental theorem of calculus (hint: what is the physical interpretation of the derivative and of the integral?).*

Solution 11

Exercise 12 *This exercise is about developing your use of mathematics as a natural language for expressing ideas about systems. You need to know how to "say" (read/pronounce) these expressions in technical terms and in common terms, as they express behaviours in systems.*

1. *Express* $v(t) = \sin(\omega t)$ *in English words.*

2. *Express* $v'(t) = \frac{d}{dt} \sin(\omega t) = \omega \cos(\omega t)$ *in Engish words.*

3. *Express* $v'_0 = \frac{d}{dt} \sin(\omega t)\Big|_{t=0}$ *in English words.*

4. *What is the significance of an even function?*

5. *What is the significance an odd function?*

6. *What kind of function is the product of an even and an odd function?*

Solution 12

Exercise 13 *This exercise is about matters related to* functions. *We take as an example the use of functions of time and space to describe waves, since the description of wavelike phenomena is one obvious application. Some more practice at differentiation is part of the exercise.*

1. *How many radians are there in a single cycle?*

2. *Sketch by hand the function* $\psi(t) = A\sin(2\pi t)$ *for* $t \in [0, 1]$.

3. *Explain how this function can be considered to be a mapping.*

4. *What is the domain and range (co-domain) of this function mapping?*

5. *Suppose we drop the condition* $t \in [0, 1]$, *what are the domain and range of the function now?*

6. *Sketch (any way you like) the function* $\psi(x, t) = A\sin(kx - \omega t)$.

7. *The equation for a one dimensional wave* $\psi(x, t)$ *is written*

$$\frac{\partial^2 \psi(x,t)}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 \psi(x,t)}{\partial t^2}, \tag{2.3}$$

   where c *is the speed of the travelling wave. Show that the sine wave*

$$\psi(x, t) = A\sin(kx - \omega t) \tag{2.4}$$

   is a solution of the wave equation, provided that

$$k = \pm\omega/c. \tag{2.5}$$

Sketch this function for some value of k and ω. k is called the wavenumber *and ω is called the* angular frequency. *What do these quantities represent? Use your sketch to help you explain.*

The frequency *of a wave is defined by*

$$\omega = 2\pi f \tag{2.6}$$

and f is measured in Hertz (Hz) or 'cycles per second'. Draw one cycle of the wave on your sketch. The wavelength of the wave is defined by

$$\lambda = \frac{2\pi}{k}, \tag{2.7}$$

and is measured in metres (m). Draw one wavelength on your sketch. Using the condition for the wave solution above, show that

$$\lambda f = c. \tag{2.8}$$

8. *Suppose we take a general linear combination of waves, with different wave-numbers k and frequencies ω,*

$$\Psi(x,t) = \int dk d\omega \; c(k,\omega) \; \sin(kx - \omega t), \tag{2.9}$$

show that this combination of signals also is a solution of the wave equation, given the same relationship between frequency and wavenumber as before. Hence conclude that any signal shape that can be constructed by adding different frequencies together can only be transmitted at the speed of waves in the carrier medium.

Solution 13

Exercise 14 *Express in your own words the kinds of problems that calculus (differentiation and integration) is useful for. Comment on the relationship between calculus and the continuum approximation. How do these two go together?*

Solution 14

Exercise 15 *This exercise is to remind you about the meaning of maxima and minima of functions. Determine any maxima or minima in the following functions.*

1. $f(x) = x - 3$
2. $f(x) = (x-3)^2$
3. $f(x,y) = \sqrt{x^2 + y^2}$
4. $f(x,t) = A\sin(kx - \omega t)$
5. $f(x) = k^2 - x^2$.

Solution 15

Chapter 3

From the discrete to the continuous

Exercise 16 *(NEW) A system administrator's help-desk has a ticket handling system for user queries that classifies tasks in the following states of completion:*

1. *New problem ticket*
2. *Waiting to be assigned to a handler*
3. *Assigned to handler, waiting for response from handler*
4. *Reply sent to user, waiting for response from user*
5. *Ticket closed*

Show that this can be modelled as a finite state machine, by drawing a diagram with appropriate transitions. Explain your diagram.

Solution 16

Exercise 17 *What is meant by a finite state machine? Keyboards are configured to a wide range of specifications around the world; all others have additional symbols that have to be squeezed onboard. The simplest keyboard is the American one.*

On a Norwegian keyboard, when you press the ˆ key or ALT GR key followed by the ˜ sign, the tilde does not appear straight until you press another key. Suggest a way of modelling this keyboard map behaviour.

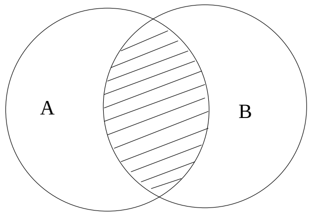
Solution 17

Exercise 18 *Draw a very simple non-deterministic finite state automaton that recognizes the TCP protocol, using the SYN, ACK, FIN flags as transition instructions between internal states. (It does not have to be elaborate, as long as it illusrates the main features.)*

The automaton should end up in an 'accept' state for valid protocol streams. Can you make this automaton detect incorrect streams? Could you think of a way in which this idea could be used for network intrusion detection?

Solution 18

Exercise 19 *This problem is about sets states and logic. You have to think about how different sets of values should be grouped together, named and identified. Consider the Venn diagram below: Let A and B be abstract sets to be defined below. Decide whether the following definitions for A, B lead to a picture that agrees with the figure.*



1.

$$\begin{aligned} A &= \{1, 2, 3, 4, 5, 6, 7\} \\ B &= \{3, 7, 19, 24\} \end{aligned} \tag{3.1}$$

2.

$$\begin{aligned} A &= \{644, 600, 400, 555\} \\ B &= \{755, 775, 555, 511\} \end{aligned} \tag{3.2}$$

3.

$$\begin{aligned} A &= \{Windows, Unix, Macintosh\} \\ B &= \{Debian, RedHat, SuSE\} \end{aligned} \tag{3.3}$$

4.

$$\begin{aligned} A &= \{0 \leq t \leq 1, 2 \leq t \leq 3\} \\ B &= \{t = 0.5, t = 0.7\} \end{aligned} \tag{3.4}$$

5.

$$\begin{aligned} A &= \{/*/passwd, /etc/*\} \\ B &= \{/etc/passwd, /etc/shadow, /tmp/passwd\} \end{aligned} \tag{3.5}$$

6.

$$\begin{aligned} A &= \{Windows, Linux, Solaris, MacIntosh\} \\ B &= \{Sparc, Intel, AMD, 68000, zSeries\} \end{aligned} \tag{3.6}$$

Solution 19

Exercise 20 *This exercise is about using sets to describe networks.*

1. *What is the set theoretical meaning of the Boolean expressions* **AND** *and* **OR** *? Draw these as Venn diagrams. Do Venn diagrams give a good representation of what the sets really look like, e.g. in part 4 of the previous exercise?*

2. *In configuration management we think of systems as belonging to a set of all possible system types. For instance, in cfengine systems are classified by the sets to which they belong. e.g.*

   - `Hr01` *is the set consisting of all times between 01:00:00 and 01:59:59.*
   - `linux` *means the set of all systems that run any kind of linux operating system.*

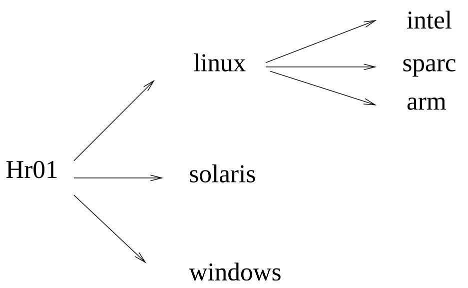
Operators `.` `&` `|` `!` *are provided to combine these using Boolean logic. The symbols* `.` *and* `&` *have the same meaning.*

3. *Explain with the help of a Venn diagram showing the machines in a network what is meant by the cfengine class expressions:*

```
  i)  linux|solaris::
 ii)  Hr01.crayos::
iii)  (Hr01|Hr03).(linux|solaris)::
```

4. *Compare the meaning of* `linux.Hr01` *and* $\{linux\} \cap \{Hr01\}$.

Some configuration management systems use hierarchical decompositions of the hosts in a network. Could we write the expressions above as tree-like hierarchies? e.g. as in the figure Compare these two methods critically.



Solution 20

Exercise 21 *This problem is about the use of probabilities. Probability can be used either descriptively or predictively. When making predictions based on past experience we must use careful judgement.*

1. *Suppose we make a survey of computers at a computing company and we count the number (frequency) of occurrences of three categories of machine:* $S = \{windows, linux, solaris\}$. *The following results are obtained:*

```
n(windows) = 25
n(linux)   = 16
n(solaris) = 3
```

   We call this kind of data statistics. *From these values, compute the probability that a computer is windows, linux or solaris at the company.*

2. *Probabilities derived from data are often used to predict the likelihood of finding a similar situation elsewhere, e.g. in intrusion detection. Comment on whether we can use the probabilities above to estimate the fractions of computers with these operating systems*

   (a) *At the same company in the future.*

   (b) *In Norway today.*

   (c) *Anywhere in the world today.*

3. *Suppose that the computers counted are all desktop machines. Given only the probabilities above, if someone tells you the size of the company will grow to three times the total number of employees in the next year, can you say how many windows, linux and solaris machines will be needed at the company next year?*

4. *The network engineers at the company are analyzing the traffic that passes in and out of the company to the Internet. From the packet headers they can see the origin and destination of each packet and can therefore determine which type of machine at the company generates which traffic. Using the frequency data above, what can you say about the average level of traffic from solaris, windows and linux if*

   (a) *All employees work on similar tasks.*

   (b) *If employees have differentiated tasks.*

Solution 21

Exercise 22 *(NEW) This problem is about using an approximation to a discrete process to answer a practical question. We use expectation values and integration to help us. This question addresses disk backups and many other issues in an abstract way*

The Snow Clearing Model (or Windshield Wiper Model) is a way of thinking about maintenance in systems. Imagine that snow is falling on the ground (many discrete snowflake events are arriving) or that rain is falling on your car windshield. A plough or wiper-blade clears the falling 'events' and then waits. When and how should you drive your snowplough to clear the snow? We shall show that there is a 'best' answer to this problem.

You can imagine that the snowflakes are like changes happening to the data on a disk, or they are like problems that are occuring in a system. We want to clear up these events by performing maintenance (clearing the snow), but the snow never stops falling. The snow falls mainly during the day, less at night (like users working) so we would like to decide when the best time to start clearing is (or when should a system administrator make a disk backup)?

To answer this, we formulate a 'risk function' ρ based in the idea of Mean Time To Repair. The risk here is the risk of loss due to a catatrophe. For instance, as long as the snow is not cleared, a driver could crash; if there are non-backed-up changes to a disk, then the risk of losing data is high. Since a catatrophe can occur at any time, the risk is proportional to the time between 'now' and the next maintenance cycle:

$$\rho(t_0, t_b) = \rho_0 \langle T \rangle \tag{3.7}$$

1. *Assume that maintainence (ploughing, backup) is performed once per day at time t_0 and takes time t_b to complete; also the intensity of events (snowflakes, file changes) is represented by a function $q(t)$ which is periodic (see fig. 3.1). Let the length of a dayly period be P; now show that we can use the function*

   $$q(t) = q_0 \left(1 + \sin\left(\frac{2\pi t}{P}\right)\right) \tag{3.8}$$

   to model a Probability for Required Maintenance $p(t)$ during the time, by normalizing. Is this always positive? Does it lie in the range $[0, 1]$?

2. *Suppose we can define a Time To Repair function $T(t)$ for events that occur at time t, express the expectation value $\langle T \rangle$ as an integral expression over the time period $[0, P]$*

3. *Suppose an event occurs in region I or in region III, what is the average time before it will be ploughed/backed up for each of these two regions?*

4. *If maintenance is quick, and t_b is fairly short, then we can assume that half the changes that occur in region II occur only after the plough has made its sweep (we miss them). These changes will have to wait a full period P to be swept again. Thus the Time To Repair in this region is the average of that in region I and P. What is it?*

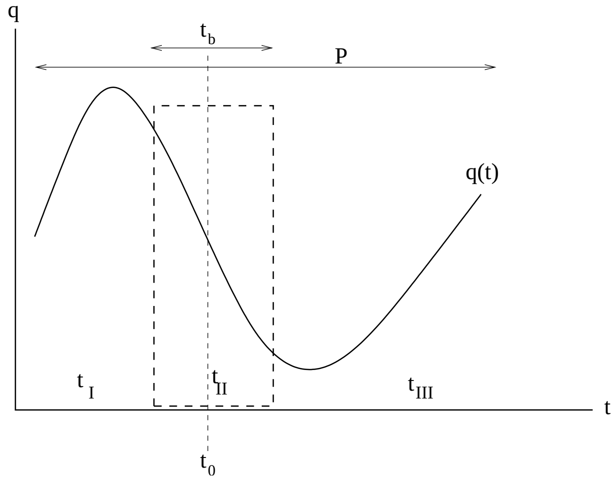


Figure 3.1: The snow clearing/windshield wiper model of maintenance. Three regions are identified: before, during and after maintenance.

5. *Show that the risk function $\rho(t)$ can be written*

$$\begin{aligned}\rho(t_0,t_b) &= \rho_0\int_0^{t_0-\frac{1}{2}t_b} q(t)(t_0-t)dt \\ &+ \rho_0\int_{t_0-\frac{1}{2}t_b}^{t_0+\frac{1}{2}t_b} q(t)(\frac{1}{2}P+t_0-t)dt \\ &+ \rho_0\int_{t_0+\frac{1}{2}t_b}^{P} q(t)(P+t_0-t)dt. \end{aligned} \tag{3.9}$$

6. *This function can be minimized with respect to t_0 to find the optimal time to start ploughing. The answer is $t_0 = P/2$. Comment on this answer. Is this the time you would intuitively pick for taking a disk backup? Explain your thoughts.*

Solution 22

Exercise 23 *(REVIEW) This problem is to remind you about matrix operations. Let*

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \tag{3.10}$$

$$B = \begin{pmatrix} 2 & 1 & 1 \\ 2 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix} \tag{3.11}$$

Find:

1. $A + B$
2. AB
3. BA

Solution 23

Exercise 24 *(REVIEW) This problem is to remind you about solutions to matrix equations.*

1. *Consider the matrix equation:*

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0. \tag{3.12}$$

   (a) *This is really two coupled scalar equations. Write these two equations separately.*

   (b) *Show that $x = y = 0$ is a solution of the matrix equation.*

   (c) *How can you determine quickly whether any other solutions exist?*

2. *Consider the matrix equation:*

$$\begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0. \tag{3.13}$$

   (a) *Show that $x = y = 0$ is a solution of the matrix equation.*

   (b) *How can you determine quickly whether any other solutions exist?*

3. *Consider the matrix equation:*

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}. \tag{3.14}$$

   (a) *Show that $x = y = 0$ is a solution of the matrix equation.*

   (b) *Find any other solutions. (Hint: recall the previous problem above.)*

Solution 24

Chapter 4

Configuration management

This exercise is about configuration management of computers and network devices. i.e. how we construct the details of a system and maintain them over time. It develops the abstract formulation of the preceding exercise so as to strip away the issues surrounding configuration management down to the essentials. By doing this, we can identify the fundamental issues.

Exercise 25 *This problem is about* operations *of the kind used in configuration management.*

Consider the problem of file permission attributes. To keep the problem simple, we shall consider only the attributes for the owner of a file f. Let the vector

$$\vec{P} = \begin{pmatrix} 1 \\ r(f) \\ w(f) \\ x(f) \end{pmatrix} \tag{4.1}$$

be formed from a constant value 1 followed by three Boolean functions that return 1 or 0 for the permissions for read, writing and execution of the file.

1. *Which of the operations below is represented by the following matrix operation?*

$$\vec{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ r \\ w \\ x \end{pmatrix} \tag{4.2}$$

   (a) *chmod u+r f*

   (b) *chmod u+w f*

   (c) *chmod u+x f*

Solution 25

Exercise 26 *In configuration management we have a map or specification of how a system is supposed to be configured. This map is called the* policy *of the system. The process of maintenance is about checking whether a system is still compliant with policy, and repairing it if it isn't. Random errors and changes creep into systems for a number of reasons. The* snow-clearing model *of system maintenance imagines that random changes to the system are like falling snow and they have to be swept away at regular intervals. The* immunity model *tries to construct dumb 'antibodies' (at set of operators $\hat{A}$) that, when applied persistently and without intelligence, lead to the snow-clearing model.*

1. *Let $\vec{\psi}$ be a vector of configuration parameters and let $\vec{\psi}_0$ be the state of the vector when it complies with policy.*

   The immunity model says that regardless of the current state of the system ψ, some number of iterations of an operator $\hat{A}$ will return the system to its policy state. Once the system has reached that state, $\hat{A}$ has no further effect.

$$\hat{A}^n\psi = \psi_0 \tag{4.3}$$
$$\hat{A}\psi_0 = \psi_0 \tag{4.4}$$

   This property is called convergence. *Show that an operator that satisfies*

$$\hat{O}^2 = \hat{O} \tag{4.5}$$

   satisfies the immunity property (i.e. it is convergent) for a state, provided it satisfies eqn. (4.3), and hence show that convergence is a stronger constraint than idempotence. This property is called idempotence.

2. *Show that the matrix chmod operator is both idempotent and that it satisfies the immunity property. Explain why this operator is both convergent and idempotent. Describe or characterize the state ψ_0 for the chmod operator.*

Solution 26

Exercise 27 *Let us extend the configuration model by allowing for the creation and deletion of configurable objects (e.g. files). We define the extended state vector for a file object by*

$$\vec{\psi}_f = \begin{pmatrix} 1 \\ \sigma \\ r \\ w \\ x \end{pmatrix} \tag{4.6}$$

where σ is a string that is the contents of the file and r, w, x are the permission attributes as before. If $\sigma = \epsilon$ the file is empty. If $\sigma = 0$, the file does not exist.

1. *Show that the file creation operator (like Unix* `creat` *or* `cat < /dev/null > file` *with start permissions R, W, X is $\hat{C}(\epsilon, R, W, X)$, where:*

$$\hat{C}_f(\sigma, r, w, x) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ \sigma & 0 & 0 & 0 & 0 \\ r & 0 & 0 & 0 & 0 \\ w & 0 & 0 & 0 & 0 \\ x & 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.7}$$

2. *Show that $\hat{C}_f$ is idempotent, i.e. $\hat{C}^2 = \hat{C}$.*

3. *Show that the permission setting operator (chmod) $\hat{P}$ is also idempotent:*

$$\hat{P}_f(r', w', x') = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ r' & 0 & 0 & 0 & 0 \\ w' & 0 & 0 & 0 & 0 \\ x' & 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.8}$$

4. *Show that* $\hat{C}_f \hat{P}_f = \hat{C}_f$ *and that* $\hat{P}_f \hat{C}_f \neq \hat{C}_f \hat{P}_f$, *i.e. that the operators do not commute.*

5. *What consequences could this have for the configuration of the system if we base configuration management on operations that do not commute?*

6. *What modification could we make to* $\hat{C}$ *in order to allow* $\hat{C}$ *and* $\hat{P}$ *to commute?*

Solution 27

Exercise 28 *Let us now extend the model above to include the editing of file contents. This is not a trivial operation like setting a file permission and it introduces new complexity mainly because of the way we think about editing files – not because there is a difference in principle.*

1. *Consider the file modification operator* $\hat{M}_f(\delta\sigma)$ *which appends a string* $\delta\sigma$ *to the end of the file* f, *i.e. such that*

$$\sigma \to \sigma + \delta\sigma, \tag{4.9}$$

where "+" means "join" or append. Show that $\hat{M}_f(\delta\sigma)$ *can be written*

$$\hat{M}_f(\delta\sigma) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ \delta\sigma & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.10}$$

and that no other file attributes are altered by this operation. Show that if $\hat{I}$ *is the identity matrix, then we can write (for convenience)*

$$\hat{M}_f(\alpha) = \hat{I} + \hat{\alpha}, \tag{4.11}$$

where α *is the matrix generator of an appendage* α. *Thus we see the relative change to the file attribute vector.*

2. *What happens if the file does not exist before applying this operation* $\hat{M}_f$*?*

3. *Is the operation* $\hat{M}_f$ *idempotent?*

4. *Does this operation commute with* $\hat{C}$ *and* $\hat{P}$*? Do we need to insist on any other constraints to make this true?*

5. *Do different modifications to the same file commute with one another?, i.e. what is:*

$$[\hat{M}_f(\alpha), \hat{M}_f(\beta)]? \tag{4.12}$$

6. *Do different modifications to different files commute with one another?*

$$[\hat{M}_f(\alpha), \hat{M}_{f'}(\beta)]? \tag{4.13}$$

7. *In cfengine we have the operator* `AppendIfNoSuchLine` *which can be interpreted as a conditionally constrained* $\hat{M}$ *(call it* $\hat{m}$*). This operator has the form*

$$\hat{m}_f = \begin{cases} \hat{I} + \hat{\delta\sigma} & \text{if } \delta\sigma \not\subset \sigma \\ \hat{I} & \text{if } \delta\sigma \subset \sigma. \end{cases} \tag{4.14}$$

where σ *is understood to be the current contents of the file prior to operation. Is this operation idempotent? Is it convergent?*

Solution 28

Exercise 29 *This exercise summarizes the foregoing exercises.*

1. *If you were writing a new tool for automatically configuring and maintaining devices with special properties what properties would you need to secure of this tool?*

2. *What is the difference between the properties of idempotence and convergence?*

3. *In the configuration management tool cfengine, operators are purposely made to be convergent wherever possible. The operators can also be made to commute to avoid ordering problems. Some authors have proposed an alternative to having many commuting operators by suggesting that any maintenance should begin by reinstalling the system from a fixed image and then applying configuration changes in a strict order to guarantee the outcome. They call this approach* congruence, *i.e. the correctness of policy configuration depends on the precise order in which the operations are performed from a known state.*

   - *Convergence:* $\hat{C}(\sigma), \hat{P}(r, w, x), \hat{m}(\delta\sigma), \ldots$ *(Many operators)*
   - *Congruence:* $\hat{T}(\sigma + \delta\sigma, r, w, x, ...)$ *(One big operator)*

   (a) *Are these alternatives both idempotent?*

   (b) *Is either one of these methods more reliable than the other?*

   (c) *Is either one more correct than the other?*

   (d) *Can both of these methods be used 'on the fly' while devices are in operation?*

   (e) *Given that regular maintenance is necessary, which of these approaches does the least violence to the system?*

   (f) *The Simple Network Management Protocol (SNMP) which is used for monitoring and configuring simple network devices has two operations: GET and PUT (like READ/WROTE or PEEK and POKE). Are these operations idempotent? Is there any problem with using SNMP to configure workstations and computers?*

Solution 29

Chapter 5

Simple systems

Exercise 30 *This question is about the fundamentals of systems: what constitutes a system and what makes one system better than another.*

1. *Explain the difference between a static and a dynamic system.*
2. *Describe the necessary components of a dynamical system.*
3. *What is meant by freedom and constraint? Identify the freedoms and constraints in the following systems:*
   (a) *A pendulum.*
   (b) *Web client-server communication.*
   (c) *A help-desk.*
   (d) *A configuration engine (e.g. cfengine).*
   (e) *An SNMP monitor (e.g. MRTG).*
4. *Explain the relationship between an algorithm and a protocol, and explain how freedoms and constraints enter into these.*

Solution 30

Exercise 31 *(NEW) In business modelling, one begins often with an activity chart, which is a crude finite state representation of a customer's behaviour within a system. This allows us to measure user behaviour as a pretext to capacity planning. Examine fig. 5.1*

1. *Write down the transition matrix for this set of states* $\{e, h, s, v, g, c, b, x\}$.

Solution 31

Exercise 32 *(NEW) This problem is about a technique called Linear Programming. It is a simple technique, in two dimensions that used to be taught to school children. It teaches you to use a pen and paper to solve simple constraints.*

A service offering business runs two different services, X *and* Y*, for its customers; these have service rates of* x *and* y *Gigabits per second. It leases a network connection with a maximum capacity of* C *Gigabits per second. The company pays* R *euros per second for its network connection, so it has to earn at least this much to make a profit. We want to find out how to choose (prioritize)* x *and* y *so that we can program a DiffServ router to maximize the profit for the company.*

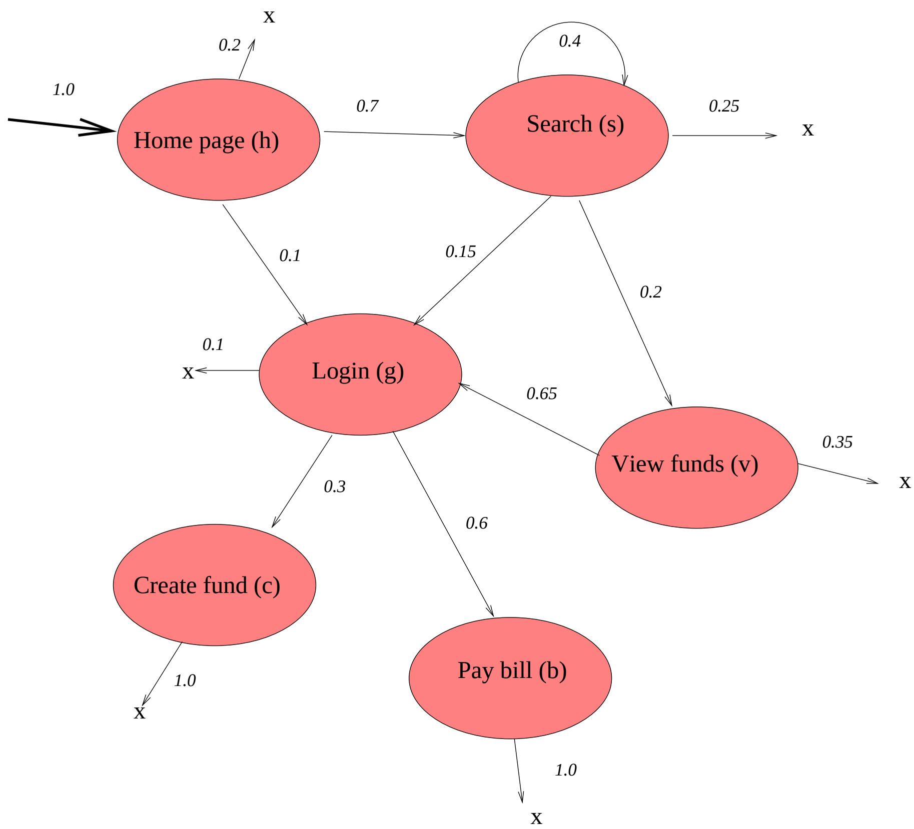


Figure 5.1: An activity diagram for customers arriving at an online bank's website. Transition probabilities between the different activities are shown on arrows. Each state has a link to the exit state (x), which is not drawn.

Consider fig. 5.2. We do this by taking the contraints we know and drawing them as lines and excluded regions on a diagram like this. For instance, we know that

$$x + y \leq C \tag{5.1}$$

since the service levels cannot be larger than the transport limit. To represent this on the disgram, we draw the line $x + y = C$ and shade the region above it to show that it is excluded.

1. *Draw the axes and exclude the region $x + y \leq C$ as explained above.*

2. *For dependency reasons, the company knows that the traffic to Y will be at least twice that to X, thus $y \geq 2x$. Plot the line $y = 2x$ on your diagram and shade the region under it.*

   The clear (unshaded) region in this area represents the company's "business plan" or operating zone. As long as x and y are in this area, or on its boundary, the constraints are obeyed.

3. *Let P_x, P_y be the price per completed transaction (euro/tr), let D_x, D_y be the data per transaction (Gb/tr). The, using dimensional analysis, show that the rates of earning/income of each machine (euros/sec) are $P_x x/D_x$ and $p_y y/D_y$.*

4. *The company knows that it can charge up to three times as much for X as for Y, i.e. $P_x \leq 3P_y$, so they want to maximize traffic stream x. Unfortuantely they can only*

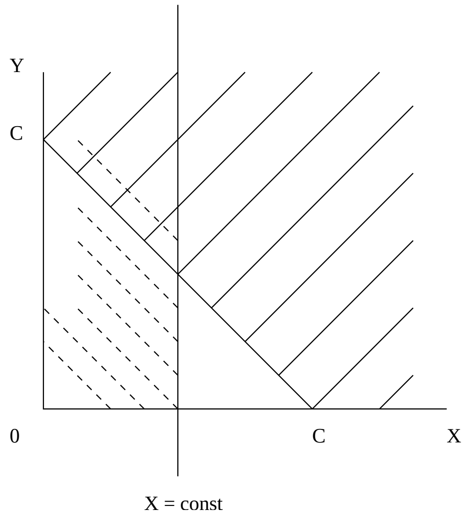


Figure 5.2: Linear programming uses a diagram like this.

limit the maximum level using the Diffserv router. If they limit y to allow more x in, they do not know that the demand for x will fill the empty capacity. They must choose the price and maximum values of the streams to maximize their profits. Can they calculate a best answer?

5. *What is the condition for the company to make a profit?*

Solution 32

Exercise 33 *This exercise is about system architectures or structural properties of systems.*

1. *Compare the structure of a StarLAN network with the architecture of a Linux Beowulf cluster. What are the similarities and differences? (Discuss this in terms of the main elements of a system.) When would you choose these architectures, i.e. what tasks are they best suited to?*
2. *What is meant by top-down and bottom-up design of a system? Comment on the use of these strategies for (a) system design and (b) system maintenance.*
3. *What is meant by hierarchical organization? Explain how the depth of a system hierarchy can affect the performance of a system.*
4. *Abstraction is used in system design to separate independent issues from one another. Explain how a hierarchy achieves such a separation. Is the hierarchy the only model of separation of independent issues?*
5. *What is meant by system normalization? Explain the purpose of normalization.*

Solution 33

Exercise 34 *This exercise is about system architectures or structural properties.*

1. *What is meant by a data structure? What is its relationship to systems?*
2. *Explain how the organization of a system can affect the efficiency with which different types of input are processed. Give an example of a system that is designed to handle a special kind of input.*

3. *How would one view a network design as a data structure? Explain how the use of an abstract data model can help in the design of a network.*

4. *Compare an old-fashioned shared-bus Ethernet architecture (or Hub) with a Star-LAN switched architecture and a wireless WLAN network. Assuming that a number of PCs are connected together by these two methods, comment on the freedoms and constraints in the two systems. Comment on the criteria you might use for determining which type of design would be best in a given situation.*

Solution 34

Chapter 6

Reading week

You should use this time to do some general reading in the book to enlarge your knowledge of the course materials and work through examples.

Chapter 7

Diagrams and graphs

Exercise 35 *(NEW) This exercise is about thinking of structural problems in terms of graphs. Graphs are a powerful way of analysing many qualitative and quantitative relationships.*

1. *A company has three clients for which it provides a service. Draw this relationship as a graph.*
2. *The same company decides to outsource part of its business to another company. Modify your diagram to show this.*
3. *After a while, the risk of relying on a single outsourcing company is found to be unacceptable and outsourcing is divided between two separate companies. Modify your graph to show this.*
4. *Comment on how you might represent load-balancing in a graph, using weighted links.*

Solution 35

Exercise 36 *Encryption keys for private communication between users or systems are of two kinds: (i) shared keys (128 bytes), or (ii) public-private key pairs (1024 bytes + 8 bytes).*

1. *In order to communicate privately with (i), any pair of users must share a unique key. Show that if there are N users who need to communicate potentially with every other user, a total of $N(N-1)/2$ keys is required.*
2. *To communicate privately with public keys, show that each user needs only 2 keys, i.e. the total number of keys is $2N$.*
3. *Calculate how many users (find N) such that it is worthwhile introducing public-private keys, i.e. before the number of shared keys is much larger than the number of public-private keys.*
4. *Calculate how many users (find N) such that it is worthwhile introducing public-private keys, i.e. before the amount of memory used to store the shared keys is much larger than that required to store the of public-private keys.*
5. *Comment on which of the two criteria above makes most sense.*

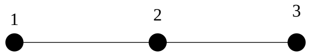
Solution 36

Exercise 37 *(NEW) Find the centrality eigenvector for the transition matrix in exercise 31.*

Solution 37

Exercise 38 *This exercise is about determining the relative importance of nodes in a network.*

1. *Consider the graph: Show that the graph has adjacency matrix*



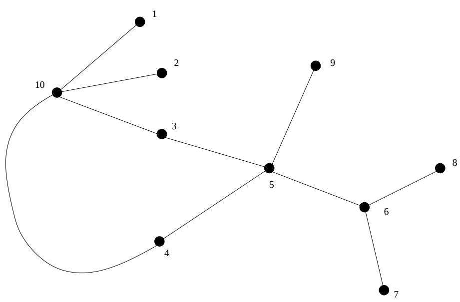
$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \tag{7.1}$$

2. *Find, by hand, the eigenvalues λ_i and eigenvectors ψ_i of this matrix, satisfying the secular equation*

$$A\vec{\psi}_i = \lambda_i \psi_i. \tag{7.2}$$

The principal eigenvalue is special; explain how the components of this vector characterize the graph.

3. *Consider now a metropolitan area network of routers in the figure: Find the adja-*



cency matrix of this network and find its principal eigenvector.

4. *Assuming that all of the routers are connected to local networks that are approximately as active as all the others, rank the nodes according to their importance as communication hubs in the network.*

5. *The remainder of this problem is about the interpretation of the eigenvector ranking. Explain how the local importance I_i of each node i is proportional to the number of nearest neighbours that interface with it:*

$$\begin{aligned} I_j &= \sum_{i=\text{neighbours of } j} 1 \\ &= A\vec{1}, \end{aligned} \tag{7.3}$$

where $\vec{1}^{\mathrm{T}} = (1, 1, 1, ..., 1)$. Hence show that a weighted sum of neighbours with weights α_i for the ith neighbour, i.e.

$$\vec{I} = A\vec{\alpha}. \tag{7.4}$$

Show that, if we weight the local importance of each node with the relative importance of its neighbours, i.e. $\vec{\alpha} \propto \vec{I}$, then we find the following self-consistent equation for $\vec{I}$:

$$A\vec{I} = \lambda\vec{I}, \tag{7.5}$$

for some constant of proportionality λ^{-1}. Hence one concludes that $\vec{I}$ is both the importance ranking over all the nodes and the solution of the eigenvector equation, i.e. that the two are equivalent.

6. *How useful do you think this ranking is? What assumptions lie behind it?*

7. *(Hard) The eigenvector equation has N solutions if there are N nodes. Can you see why the solution we want is the principal eigenvector?*

Solution 38

Chapter 8

A model of human-computer systems

Exercise 39 *An antarctic research station measures that the signal strength of its satellite link varies during the course of a day according to the empirical formula*

$$S(t) = S_0(1 + \sin(2\pi t/24)) \tag{8.1}$$

where time is measured in hours GMT and $t = 0$ is calibrated to 5:30 a.m GMT.

The system administrator guesses that the rate of transfer of data $R(t)$ via its satellite line varies with time in strict proportion with the signal strength: $R(t) \propto S(t)$.

1. *Does the assumption that data rate is proportional to signal strength sound reasonable to you? Explain your answer.*

2. *Based on his assumption, the system administrator scribbles down a formula for the data rate in* Giga-bytes per second.

$$R(t) = R_0(1 + \sin(2\pi t/24)). \tag{8.2}$$

If the maximum data rate is known to be 1 Giga-bit per second, find the value of R_0.

3. *To be certain of being able to download the evening's movie entertainment in time for its screening, the administrator needs to calculate how long the data transfer will take. He writes down the equation*

$$d(t) = \int_{t_s}^{t_f} R(t')dt' \tag{8.3}$$

for the amount of data $d(t)$ (in Giga-bytes) received as a function of the start time t_s and final time t_f, on the cailbrated hour-scale, given that he starts downloading at time t_s. Explain the thinking behind this formula and explain why the system administrator has made a simple mistake, which has to do with measuring units.

4. *The administrator corrects his mistake and starts with*

$$d(t) = 3600 \int_{t_s}^{t_f} R(t')dt' \tag{8.4}$$

Suppose that the movie must be downloaded by 18:00 GMT; by solving this equation for t_s, calculate how long it takes to download 4 Giga-bytes of data, and then 40 Giga-bytes of data. (Hint: be careful to convert times to the calibrated scale.)

5. *As more penguins start to run Linux, the total bandwidth available to the station has to be shared amongst mulitple customers and the station can only use one tenth of the total channel capacity. Recalculate the time taken for the download if the total channel capacity is reduced to i)* $1/10$ *and ii)* $1/100$ *of the total capacity. How would the answers be different if the film was to be downloaded for 21:00 GMT.*

Solution 39

Exercise 40 *A system administrator is trying to decide when she should take a backup of the system server. By measuring the number of client connections to the server over a number of weeks, she comes up with the following formula for the* average *number of connections to the server.*

$$N(t) = 10^4 \left(\frac{3}{2} + \sin\left(\frac{2\pi t}{24} + \phi \right) \right) \tag{8.5}$$

where ϕ *is a constant phase angle which she will use to fit the curve to her observations.*

1. *The number of connections is an integer, but the sine function is a continuous function. Explain why this makes sense for the average number of connections.*

2. *The maximum download activity is observed to occur in the middle of the day at 12:00 noon (*$t = 12$*). By differentiation, find the maxima and minima of* $N(t)$ *and determine the value of* ϕ *such that the maximum value of* $N(t)$ *is at 12 noon.*

3. *The system administrator decides to take a backup of the system when there is least activity. Suggest a time to take a backup, based on the formula.*

Solution 40

Exercise 41 *This exercise is about arbitrary decisions made in deciding policy. In modelling human-computer systems we are often required to make value judgements about different resources. In order to compare the different scales or 'currencies' we must relate them to one another. This allows us to make trade-offs and decisions about strategy. There are no right answer here; this problem is meant to illustrate how difficult it is to be precise about human-motivated policy.*

1. *What is meant by the expression 'time is money'? Can you formalize this expression mathematically?*

2. *Suppose that knowledge/experience is represented in a database in a binary tree data structure. This knowledge is used as a cache to avoid costly computation (e.g. a web browser, a DNS resolver or a numerical processor). Explain how you would go about describing the amount of CPU/network time saved as a function of the size of the cache.*

3. *Suggest a formalization of the expression 'security is the opposite of convenience'. Check that you answer makes sense. Criticize your expression, noting any problems or limitations.*

4. *The convenience of a having all hosts in a network configured identically is often used to express an amount of time saved in setting up the machines. Suggest an expression for the convenience associated with having all hosts identical. Does this "agree" with your expression about security above?*

5. *Is convenience for users the same as convenience for system administrators? Suggest a relationship between convenience for users and convenience for system administrators for some company of your choice. Can you think of an example of a company where the relationship is very different from your first example?*

Solution 41

Exercise 42 *This problem is about grading or ranking system qualities using arbitrary currencies as measures of integrity. This is a common practice in computer security.*

1. *In the Biba model of computer integrity one uses integrity levels as a currency for data and system security. If data or parts of a system are contaminated by contact with something with a lower integrity level then the integrity level of the data must be downgraded.*

   Consider a firewall model in which there is a "secure" region on the inside of the firewall and an "insecure" region outside. Users and systems need to access data on the inside of the firewall from outside. By assumption the integrity level outside is lower than that inside.

   (a) *If users or hosts outside the firewall can read information from the secure area, is the integrity of the information downgraded outside the firewall?*

   (b) *If users or hosts outside the firewall can write information to the secure area, is the integrity of the information downgraded inside the firewall?*

   (c) *Explain the usefulness and limitations of a firewall, especially when used in conjunction with a Virtual Private Network that tunnels through it?*

Solution 42

Chapter 9

Integrity, information and noise

Exercise 43 *This problem is about the technical meaning of informational entropy and how it applies to noisy measurements.*

1. *Explain what is meant by the uncertainty of an observation or measurement.*

2. *How is the Shannon entropy defined for a data stream with C possible measurement classes?*

3. *How do these classes of measurement relate the to information in a coded message?*

4. *Explain what is meant by a high entropy message and a low entropy message.*

5. *Suggest two different ways in which C symbols can be coded for transmission by optical fibre.*

6. *The usual model of signal noise is to assume that the true signal can be separated from a background of noise that we know little about. It is assumes that the sources of noise are independent, i.e. that they make orthogonal contributions to the uncertainty of signal measurement,*

$$\mathrm{Error}^2 = \Delta q_1^2 + \Delta q_2^2 + \ldots + \Delta q_C^2. \tag{9.1}$$

   Explain the reasoning behind this formula.

7. *One way to model this noise is to say that we wish to assume as little as possible about it, i.e. we have maximal uncertainty but we know that it has a limited extent. This leads us to a model of Gaussian noise.*

   Use the method of Lagrange multipliers to find the noisiest probability distribution p_i of symbol errors $\Delta q_i = q_i - \overline{q}_i$, subject to the condition that the noise falls within a bandwidth of approximately $\pm\sigma$ of the true signal $\overline{q}_i$. i.e. Maximize the Lagrangian function

$$L = -\sum_{i=1}^{C} p_i \ln p_i - \alpha \left(\sum_{i=1}^{C} p_i - 1 \right) - \beta \left(\sum_{i=1}^{C} p_i (q_i - q_i)^2/\sigma^2 - 1 \right) \tag{9.2}$$

   Hence show that the distribution that arises from many independent sources of noise is the Gaussian (normal) distribution.

8. *Consider the continuum limit of the measurement classes, i.e. $C \to \infty$, where the summation over i is replaced by an integral over a range of q. Use the argument in*

section 15.6 to derive the well-known Shannon formula for the capacity of a noisy channel:

$$C = B \log \left(1 + \frac{S}{N}\right). \tag{9.3}$$

where S is the signal power, N is the noise power and B is the bandwidth of the signal classes.

Solution 43

Exercise 44 *(Difficult) This exercise is about the interpretation of entropy to system integrity.*

1. *Show that the Shannon entropy H, multiplied by N, can be interpreted as the average length of the label that is needed to uniquely distinguish a message from all others of length N, within a fixed alphabet, i.e. that it .*

   In other words, this says that NH tells us the shortest number of symbols in which we can describe the exact message assuming one has an appropriate coding scheme that maps messages into strings of the specified alphabet. Hence it is the amount of non-redundant information in the message.

2. *In computer security, one often speaks about the entropy of an encryption key. What is meant by this terminology?*

3. *In what sense is system policy a message that is to be transmitted?*

4. *Suggest a simple alphabetic reclassification to compress the message:*

   ```
   BIGCATBIGCATBIGCAT
   ```

   by a factor of 3.

5. *The entropy (log base m) represents the uncertainty per message symbol, in units of 'digits belonging to an alphabet of m' symbols. Calculate the $\log_6$ entropy of the message above. What is the uncertainty per symbol of the message? Explain what this means.*

6. *If the entropy H can tell us about the compression of a message, what does this tell us about the last example? Take the log_6 entropy and multiply it by the length of the message above. Is the result shorter than the original message above? (Explain your result.)*

7. *Entropy assumes that the order of symbols in a string must be distinguished to capture the information in a message. Suppose we form an alphabet of C commuting, convergent operators. What is the shortest number of symbols that can faithfully transmit* any *message from this alphabet to an unconfigured computer?*

   What is the total number of different messages one can create from C symbols

   - *If order matters?*
   - *If order and repetition do not matter?*

8. *Consider a system in which there are eight possible policy rules that can be violated. The faults are called $\Delta_1, \Delta_2, ...\Delta_8$. Over time, a statistical observation tool (like cfenvd) observes that these faults occur with the following frequencies:*

$$p(\Delta) = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{64}, \frac{1}{64}, \frac{1}{64}, \frac{1}{64}\right)$$

in strings of the form:

$$\Delta_1\Delta_2\Delta_1\Delta_3...$$

The mean time before failure is $\langle T \rangle$ *(or average time between* Δ_i*).*

Consider a set of operations that counters each of these problems $\hat{O}_1, \hat{O}_2, ...$ *It is rarely possible to detect change and apply these until a certain time has elapsed, so imagine that several faults can have occurred in the time it takes to schedule maintenance (Mean Time To Repair or MTTR). Faults that occur most frequently should be performed most often, but multiple repairs can be eliminated when performing scheduled (not event driven) maintenance. This allows us to compress the work to be done by avoiding multiple operations of the same type, at the expense of the prolonged uncertainty.*

If T_m *is the time between maintenance checks, show that the average number of faults to be corrected is the uncertainty of the system over the time:*

$$\mathrm{Maintenance} - \mathrm{length} = H\langle T_m/T \rangle.$$

This is shorter than the full average fault number $N = \langle T_m/T \rangle$ *of full symbols. Explain why.*

Based on the previous questions, can these information-compressed maintenance operations be shorter than a maintenance procedure based on commuting, convergent operations?

Solution 44

Chapter 10

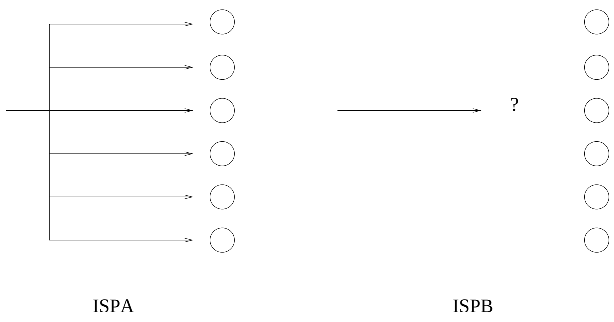
Arrivals and queues

Exercise 45 *Explain what is meant by a time-series. Explain what is meant by the Hurst exponent for a time-series. What does it show?*

Solution 45

Exercise 46 *This question is about task management and service handling. It concerns the efficiency of handling services by queueing, and the deployment of resources in parallel to cope with demand.*

1. *Explain what is meant by the term "arrival process". What is the most common model for such a process?*
2. *Consider the arrival of queries to an ISP help desk. The arrival process contains both long and short jobs in random order. How would you determine the type of the arrival process? What do you think is the best way to model such a process to discuss the efficiency of the queueing?*
3. *Two ISPs compete for customers by claiming to have an efficient help service. The two companies decide to organize their services differently (see figure). Both ISPs*



have six persons staffing their phone lines. ISPA decides to use a telephone menu to divide queries into different types in advance, so that a specialist can deal with each problem. ISPB decides that it is best to keep all lines open for all queries.

Use you intuition to argue which of these two methods is most efficient. Hint: what happens if there are just three calls in the queue all on the same specialist subject?

4. *ISPA and ISPB both receive an average of 0.9 calls per minute. The average service time at both companies is 5 minutes per call.*

   Assume the arrival processes are Poissonian in nature. Explain how you would compare the efficiency of these two queueing methods formally. Which system is superior?

5. *Calculate the utilizations of the help desks of ISPA and ISPB.*

6. *Suppose that arrivals follow a long-tailed distribution for inter-arrival times (non-Poissonian). What does this mean? How might it affect the results?*

7. *Compare the conclusions you have found in this exercise with the folk theorems for redundancy in chapter 18.*

Solution 46

Chapter 11

Workflow models

Exercise 47 *This exercise is about the scalability of system management. It relates to the models described in section 18.5-18.6.*

Suppose errors occur in N devices in the manner of a random arrival process that is equally likely to affect all devices (hosts, routers etc) in a network.

1. *If the average failure rate for the total network is I faults per second, what is the average failure rate* per device.

2. *Suppose that we centralize the administration of all the nodes so that a system administrator or automatic system (like an SNMP monitor) deals with all the errors that occur. Suppose that this centralized administrator has a work capacity of C repairs per second, how much work can be done on each host on average?*

3. *We call the behaviour of the expressions as a function of the number of devices N the scaling behaviour of the model. Explain what is "good" scaling behaviour and what is "bad" scaling behaviour in a maintenance model.*

4. *Explain the expression net failure rate of systems over time in eqn. (18.55). What does this expression really mean? What are its limitations?*

5. *Suppose the probability of the central controller being able to communicate with a device is less than 1, i.e. communications are unreliable. How does this affect the scalability of the model?*

6. *Suppose that every device in a network receives its policy from a central location, but is able to carry out repairs by itself. What is the scalability of this model?*

Solution 47

Exercise 48 *Consider the directory service LDAP.*

1. *What function does it perform?*

2. *Discuss this system from the perspective of the basic system principles. Mention redundancy, hierarchy, dependency.*

3. *What principles are at work in LDAP?*

4. *If you were to verify how effective this software is at performing its task, how would you test this by experiment?*

Solution 48

Exercise 49 *This problem relates maximization of resources to graphical methods of analysis-*

The network connectivity of N hosts can be written

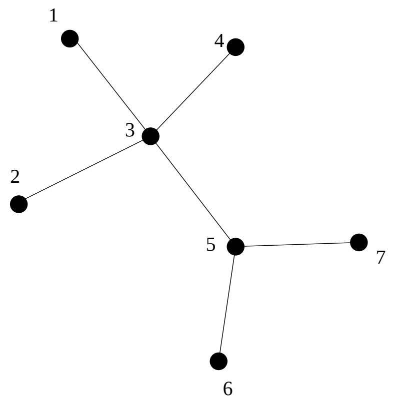
$$\chi = \frac{1}{N(N-1)} h^{\mathrm{T}} A h \tag{11.1}$$

has a maximum value when all hosts are connected and are available for communication. Consider a pervasive computing network in which the available channels of communication are described by the adjacency matrix A. Use the method of Lagrange multipliers to maximize χ with respect to h^{T} subject to the condition $h^{\mathrm{T}} h = H$, i.e. the total availability of resources in the network is fixed. Show that χ is maximized when

$$A\vec{h} = \lambda \vec{h}, \tag{11.2}$$

for some constant λ.

Use Octave, Mathematica or some other tool to find $\hat{h}$ and χ for the network: i.e.



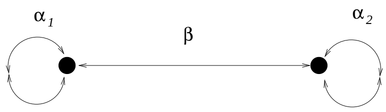
$$A = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}. \tag{11.3}$$

Which node in this network needs to be able to handle to greatest traffic load?

Solution 49

Exercise 50 *(Difficult) This problem is about algebraic modelling collaboration and dependency in systems. The aim here is to see how simple mathematical expressions reflect properties about systems and reveal features that are perhaps not obvious to us at the outset. It will also reveal some limitations: if we are too simplistic, then the results will not necessarily behave as we expect. The value of this kind of exercise is that it helps analysts and researchers to think and understand.*

1. *Show that, if we constrain two variables q_1 and q_2 so that their product is fixed, we create a* dependency *between the variables. Hint: plot $q_1 q_2 = k$, for various values of k on a plot of q_1 versus q_2.*



2. *Consider the following graph of two communicating systems. The arrows represent work-flow and the arrows indicate the direction of the flows. The real, positive adjacency matrix of this simple graph is*

$$A = \begin{pmatrix} \alpha_1 & \beta \\ \beta & \alpha_2 \end{pmatrix} \tag{11.4}$$

We define a workload vector $\vec{L} = \begin{pmatrix} L_1 \\ L_2 \end{pmatrix}$ *and the productivity (by analogy with connectivity) as* $P = \vec{L}^{\mathrm{T}} A \vec{L}$. *Write down the productivity and show that it is a scalar.*

3. *What is the minimum value of* P.

4. *What is the maximum value of* P?

5. *Use the method of stationary variation on the productivity to see what it tells you about the non-trivial solutions for* $\vec{L}$. *Show that non-trivial solutions exist only if* $\beta^2 = \alpha_1\alpha_2$. *Bearing in mind that the two nodes are independent of one another, what is the significance of this solution?*

6. *Let us now couple to two systems together formally, making them dependent on one another. Consider now the same system with the additional constraint* $L_1 L_2 \leq k$, *some some constant* k. *Use the Lagrange method with multiplier* λ *to maximize*

$$\mathcal{P} = \vec{L}^{\mathrm{T}} A \vec{L} - \lambda(L_1 L_2 - k) \tag{11.5}$$

and show that this leads to the coupled linear equations

$$A \begin{pmatrix} L_1 \\ L_2 \end{pmatrix} = \lambda \begin{pmatrix} L_2 \\ L_1 \end{pmatrix} \tag{11.6}$$

Find the values of λ *for which non-trivial solutions exist for* L_1, L_2. *Show that there are two solutions* $(L_1, L_2) \propto (\pm\alpha_2, \alpha_1)$.

7. *Show that the stationary value of the productivity is*

$$P \propto \beta L_1 L_2. \tag{11.7}$$

Suggest a reason why the result is proportional to β. *Comment on the symmetry of this answer with respect to* α_1 *and* α_2.

8. *Repeat the calculation for a graph in which the work-flow between the hosts travels only one way, i.e.*

$$A = \begin{pmatrix} \alpha_1 & \beta \\ 0 & \alpha_2 \end{pmatrix}. \tag{11.8}$$

Draw the graph of this matrix and comment on the values. What does it tell you? Does it make sense? The graph is now asymmetrical with respect to α_1 *and* α_2. *Find the stationary value of the productivity and determine which host drives the productivity? If you were to spend more money on one of the hosts in this setup, which one would you choose?*

Comment on this analysis. What does the mathematics show you? What are the limitations of the model.

Solution 50

Chapter 12

Decision theory

Exercise 51 *This problem is about the theory of rational decision making.*

1. *Explain in your own words what is meant by a mathematical game.*
2. *Explain what is meant by 'payoff' or 'utility'. What role do currency systems play in defining payoff?*
3. *Explain how a two-person game can be represented as a* pair *of matrices.*
4. *Examine the game in example 169. Explain, with a critical eye, how the values are arrived at. If you disagree with the values in this matrix, explain why and argue for your own values.*

Solution 51

Exercise 52

1. *Use the Gambit game software to analyze the model of software updating in example 169 in section 19.4. Look for the Nash equilibria of the game. If you have modified the game, in the previous exercise, work out the solutions of both versions and compare them.*

   Note, the Gambit software uses the term 'outcomes' for the payoff. You should choose the normal form for game modelling.
2. *Try modifying and extending the model to investigate how stable the equilibrium points are to changes in the precise values. Explain your thinking.*
3. *Evaluate and comment on the usefulness of the procedure you have just undertaken. How reliable is it in making decisions? How realistic is it as a model of rational decision making? How might you use it in practice?*

Solution 52

Exercise 53 *Think of a decision situation of your own that can be represented as a two person game and use the Gambit software to find a solution to the game. Interpret the results – you will be presenting this example to the class.*

Solution 53

Chapter 13

Game theory

Exercise 54 *These questions are to guide your reading. You do not need to submit them.*

1. *What does it mean to say that a game is 'zero-sum'?*
2. *Describe the so-called Prisoner's Dilemma game. Is it a zero-sum game? Explain your answer.*
3. *Mention some applications of the game.*

Solution 54