Software Evolution

Date: Tuesday 31st May 2011
Time: 14:00 - 15:30

Please answer any TWO questions from the FOUR questions provided.

Use a SEPARATE answerbook for EACH Section

This is a CLOSED book examination

The use of electronic calculators is NOT permitted
1. a) You have recently joined a team using agile development to create a Web application giving advice to customers of a major DIY store chain on which of the company’s products are useful in different DIY situations. The system must interact closely with the company’s existing product database and the software that implements the business rule relating to product sales and use. On your first day in the new job, you are asked to work with another new employee to familiarise yourself with this existing system and its interface. Your colleague confesses to nerves, because:

- He has never worked with someone else’s code before
- He has never worked on a large system before
- He has never worked with the technology used by the system before (language, database management system and operating system)
- He knows nothing about DIY.
- The small amount of documentation you can find turns out to be out-of-date and largely useless.

i) Which of your colleague’s concerns is/are likely to pose the biggest problem for him in attempting to quickly grasp the structure and behaviour of the legacy system?

What advice would you give to help your colleague work more effectively over the next few weeks, as he examines the code and begins to make decisions for the new agile development? For each question, be sure to justify your answer clearly. (5 marks)

ii) What code reading strategy would you advise your colleague to adopt in first familiarising himself with the legacy code? Justify your answer with reference to the specific scenario given above. (2 marks)

b) The code fragment shown overleaf is taken from the “Solitaire Collection for Android” app (licence: http://www.apache.org/licenses/LICENSE-2.0) by kmagic. This app implements several standard solitaire card games for use on smartphones running the Android operating system. This fragment shows part of the implementation of the class which models the decks of cards involved in the game.

Which of the idioms mentioned in the course unit is present in this fragment? Give the language-independent statement of each idiom you discover, and state clearly (e.g. using the line numbers provided) which parts of the code correspond to which components of the idiom. State clearly any assumptions you make about the rest of the code. (6 marks)

(Question 1b continues on the following page)
(Question 1b continues from the previous page)

```java
public class Deck {

    ...

    public void Shuffle() {
        int lastIdx = mCardCount - 1;
        int swapIdx;
        Card swapCard;
        Random rand = new Random();

        while (lastIdx > 1) {
            swapIdx = rand.nextInt(lastIdx);
            swapCard = mCard[swapIdx];
            mCard[swapIdx] = mCard[lastIdx];
            mCard[lastIdx] = swapCard;
            lastIdx--;
        }
    }
}
```

c) Draw a control flow graph for the following fragment of code from the Solitaire Collection for Android project.

(7 marks)

```java
int suits = mView.GetSettings().getInt("SpiderSuits", 4);
mDeck = new Deck(2, suits);
int i = 54;
while (i > 0) {
    for (int j = 0; j < 10 && i > 0; j++) {
        mCardAnchor[j].AddCard(mDeck.PopCard());
        mCardAnchor[j].SetHiddenCount(mCardAnchor[j].GetCount() - 1);
    }
}
while (!mDeck.Empty()) {
    mCardAnchor[10].AddCard(mDeck.PopCard());
}
ignoreEvents = false;
```

[PTO]
2. Government cutbacks have recently forced two established national charities to merge. Mothers Against Pwnage (M4P) is a charity devoted to helping computer games addicts both to beat their addiction and to deal with its consequences. It offers a telephone and internet Help Line, an outreach service that sends trained addiction therapists into people’s homes, and a specialist legal advice and aid service. DeLink is a smaller charity focussing on the problems of internet addiction. Its client base is growing at a much faster rate than M4P’s, and has just overtaken it in terms of size. At present, its main service is a telephone helpline but the charity is hoping to use the merger as a chance to offer further services of the kind currently offered by M4P.

The new charity will be called EndGame. It requires a new information system that combines the functionalities offered by the existing systems run by the charities, over their combined client bases. The new system should ideally also make it easier to add new functionalities and new types of client (e.g. Twitter addicts) so that the combined charity can take advantage of any specialist funding opportunities that arise. Changes in government regulations mean that one piece of new functionality is required urgently (within a matter of months). The legislation change is that now, in order to obtain the grants on which the charities rely, regular reports on the services provided (number of calls dealt with on help lines, number and duration of outreach visits, etc.) must be submitted to government on a quarterly basis. New software is needed to extract and submit these statistics. In addition, since the size of the grants now depends on the number of clients helped, it is vitally important that the new charity should be able to continue to operate all its services, without delays caused by problems in cutting over to new software.

As manager of the team responsible for the creation of the new computer system, you decide to adopt the Chicken Little approach to re-engineering the new system. You therefore ask your team to document the architectures of the two existing systems.

M4P’s computer system is an ageing dinosaur implemented in COBOL. Your team manage to distinguish the components and their relationships, producing the architecture diagram shown in appendix A. The database contains client data, plus records of all the charity’s activities.

DeLink has a newer system, but with a much more limited scope. It is built using C on top of a relational database, with Web-based clients. The architecture your team extract is shown in appendix B. Its database contains client details, plus logs of helpline calls.

a) As manager of the team responsible for this re-engineering project, you begin by assessing the two legacy architectures. How would you characterise each one in terms of its decomposability? Justify your answers briefly with reference to the specific features of the architectures. (3 marks)

(Question 2 continues on the following page)
b) Next, your team designs the target architecture for the new system. The result is shown in appendix C. The target architecture supports the same three service types currently offered by the charities, but now the software for each is able to distinguish between a greater range of client/addiction types. This requires the classification of all existing clients in order to create this new data. New functionality has also been added to gather the existing logs of activity and aggregate them into the statistics required by the government grant awarding bodies.

Based on these considerations, discuss the suitability of each of the three migration strategies (forward, reverse and general migration) for this application. State clearly which strategy you would adopt for this project, based on your analysis of the business case for each one. (7 marks)

c) You must next design a migration path for the system based on your selected migration strategy. The migration should aim to bring important functionality on-line fast, while allowing existing mission-critical functions to operate reliably and as usual.

- Indicate what changes will occur in each migration step. This will best be done by drawing a diagram of the architecture that results from the migration step, but partial diagrams showing just the elements that have changed, or even brief text descriptions, are also possible if short of time. Take care, however, to distinguish target from legacy components, and to show the exact placement and type (forward/reverse) of gateways.
- Justify your choice of component(s) to migrate in this step, in terms of the benefits, costs and risks of the step.
- State any assumptions you make about aspects of the system not covered by the description above.

(10 marks)
Section B

3. a) Each of the following scenarios describe difficulties that are experienced by an organisation, due to some underlying data quality problem in its information systems. In each case, state the kind of data quality problem (DQ dimension) that is most likely to be the root cause of the anomalous software behaviour (from the dimensions presented in lectures). Briefly justify each answer.

i) A class of university students have been informed that their coursework marks are available through the university’s student information system. However, when students try to access them, they discover that marks are only present for students on single honours programmes. No marks are shown for students on joint programmes. (2 marks)

ii) disorganised.com, a website offering last-minute holiday and flight deals, has recently experienced falling sales of holiday packages from one of the major tour companies. When former customers are contacted, they report that they had planned to book through disorganised.com, but found lower prices on a competitor website, which were not offered by disorganised.com until a week later. (2 marks)

iii) SoftStuff, a European software company, has recently contracted out some modules to EngCalc, a company based in the Far East specialising in software to perform engineering safety calculations. To get things started, SoftStuff imported its detailed development schedules for these modules into EngCalc’s project management system. The first deliverable was due on 7th May, but nothing has been supplied. On investigation, the EngCalc development team express surprise, since the schedules sent by SoftStuff gave a deadline in July. (2 marks)

iv) A government-funded project has been created to pool data from local intensive care units. The goal is to mine the data for possible early warning signs of various forms of heart condition. The data mining team are surprised, on viewing the data from one of the units, to discover blood test results from patients who had been deceased for several years when the test was performed. Other tests were performed on patients who were apparently older than 250 at the time of the test. (2 marks)

(Question 3b continues on the following page)
b) Mognor Entertainment Holdings (MEH) manage theatres and music venues across the south of England. Their ticket sales and reservations system (TSR) handles an average of 100 operations (sales, transfers and cancellations) every second during normal working hours (7.00am to 2.00am). Outside these hours, only online sales are allowed, leading to a reduced average load of 2 sales per second. The company will typically be managing 5000 events at any one time, with an average of 500 tickets to be sold per event.

A new events comparison website has recently opened and MEH is keen to be properly represented on it. The website (WhatsHot.com) provides a Web service interface for event managers to upload summaries of their events and ticket sales data. It then collates this information to provide personalised recommendations to users, based on what is popular and where bargains are to be had. MEH’s competitor organisations appear to be updating the site every 8 hours. MEH needs to match this to remain competitive.

Because of the mission critical nature of the heavily-loaded TSR, MEH management decides that a new database is needed, fed from TSR, to support the creation and upload of WhatsHot.com summaries. TSR is an old system, based on a DBMS that does not support triggering. However, update logs are created daily (2.00am-2.00am) and remain accessible for a week before being archived and removed. TSR already supports a feed to the management information system (MIS) with a window of 2.00am to 4.00am, and a feed to the marketing support system (MSS) with a window of 4.00am to 5.00am.

i) Draw an architecture diagram to show the new WhatsHot.com Upload application, in the context of the existing systems (i.e., show TSR, MIS, MSS and all other components implied by the above descriptions). (3 marks)

ii) State what strategy you would use for this feed: push, pull or hybrid. Justify your answer by referring to the specific features of the system, as described above. (4 marks)

iii) MEH management decides to upload to WhatsHot.com three times a day. Based on the information given about the loading on the system, suggest a schedule for the feeds needed to accomplish this (i.e., state at what time each feed should begin). (2 marks)

iv) What mechanism would you propose to capture the changed data, under the strategy you selected in your answer to b) ii), given the capabilities of the TSR system? (3 marks)

[PTO]
4.  a) When using a black-box element in the development of a new software system, it will need to be adapted to the environment in which it must operate. The two main options for this are bridges and wrappers. Describe how each of these operate and discuss the advantages and disadvantages of each. (5 marks)

b) Describe the characteristics that make the evolution of software systems containing black-box elements different to the evolution of software systems not containing black-box elements. (3 marks)

c) During the development of a new software system, it has been decided that money will be saved by using some existing commercial off-the-shelf products. For one element of the system that performs some analysis of data and returns the results of this analysis, three possibilities have been identified:

- choice A uses a command line to receive data and return its results,
- choice B uses a Java Swing-based GUI to receive its data and return the analysis results
- choice C uses HTML forms to capture its input and returns the results of its analysis via an HTML page.

Assuming that all three choices offer all of the functionality required, evaluate the advantages and disadvantages of using the available choices and, with reasons, identify the choice that you would use. (8 marks)

d) Open source software, by its very nature, includes access to the source code of the system. Evaluate, with reasons, whether open source elements in a software system should be treated as white-box or black-box components. Your answer must address not only initial integration issues, but also on-going maintenance cost. (4 marks)
Appendix A: Architecture of M4P Information System

The architecture diagram given below is for use in question 3, parts a), b) and c). It shows the three main applications for the Clients/Appointments Data files, owned by the M4P charity. These data files store details of past and present clients of the charity, plus a log of the actions that have been undertaken in connection with the client (called “appointments” in the M4P business model, even after they have taken place or failed to take place).
Appendix B: Architecture of DeLink’s Information System

The architecture diagram given below is for use in question 3, parts a), b) and c). It shows the internal structure of the main application owned by DeLink, operating over the Client-Actions DB. This database contains a full client history for past, present and referred clients, including details of all actions taken by the charity in respect of each individual. Actions include taking phone calls, referring clients to other service providers and closing a client’s record with the charity.
Appendix C: Target Architecture for EndGame Information System

The architecture diagram given below is for use in question 3, parts b) and c). It shows the desired architecture for the EndGame information system. Any migration plan should result in the system having this shape when the last migration step is completed.