DAT321/DIT847:
Software Quality

Welcome to the examination for the *Software Quality* course. The examination is intended to last for max **4 hours** and is intended to be **anonymous** (i.e., the teacher grading your exam will not know your name). Therefore, it is important that you follow the instructions (in the separate exam cover sheet) and **do NOT leave any information that would reveal your name on these pages.**

Each question has a number of points assigned shown in the square brackets. When the question is broken down into smaller sub-questions the part of the points for that specific sub-questions are also shown as following:

1. *[10 pts].*
    a. *[2 pts]*
    b. *[8 pts]*

The percentage of points and the corresponding grade is presented below (100 points in total):

% of points	DAT321	DIT847
[0, 50%)	U	U
[50%, 65%)	3	G
[65%, 85%)	4	G
[85%, 100%]	5	VG

Before handing in your exam, number and sort the sheets in task order. Write your **anonymous code** and page number on every page!

It is important that you write **clearly** so that the examiner can read you. If your handwriting is unreadable, then you will not get any points for that question. We will NOT assess grammar or spelling as long as your answer is readable, understandable and unambiguous.

The questions in this exam refer to the **ISO 25010:2011** that categorises internal and external software quality attributes into eight characteristics.

You are allowed to have the English ←→ Swedish dictionary during the examination. Simple calculators are also allowed, but NOT calculators in mobile phones.

Questions about the exam contact:
Richard Torkar, tel. 031 772 57 19, richard.torkar@cse.gu.se
Francisco Gomes, tel. 031 772 69 51, gomesf@chalmers.se

The exam review is scheduled for 2019-09-04, between 15:00 – 16:30 at Jupiter building, 4th floor, Room 482.

The questions in this exam are related to the following context. Note that the answers should, in turn, be justified based on i) the theory and terminology from software quality and ii) their connections to the elements in this context (team, tools, processes, system under test, etc.).

Context: Software Development for a Car

Description of the context:

You become a part of a team responsible for developing software components that will be part of a car. The components developed by your team will interact with mechanical parts of the car and will have significant roles in controlling essential parts of the vehicle. The project you are assigned to needs to develop a component that will use information from various sensors in the car to assist driving of the vehicle (by notifying drivers of objects around), as well as monitoring speed of the car and fuel consumption. Failures in the sensor, or even delays in transmitting data can lead to severe consequences to safety.

Your team is composed of software and mechanical engineers with wide experience in working with multi-disciplinary teams. Additionally, most of those engineers have worked with agile software development where changes are welcome, and testing activities are done often via automated tests. The company also invested in state of the practice tools for traceability and configuration management, such that all artefacts and documents produced by your team is under version control.

Even though there are automated tests for most of the code, your team is highly dependent to **two** existing legacy components. The first legacy component is a big monolithic piece of code that requires a lot of refactoring, but with a large number of automated tests available as well as documentation. The second legacy component has a Graphical User Interface (GUI) with calls to an API, where no documentation or tests are available.

For now, your development process has to use both legacy components, but the company's expectation is that the legacy components will be replaced by newer ones. Keep in mind that your components will be part of a larger system with distributed components and access to online applications hosted on the cloud.

1. **[20 pts]** Using the context above, answer the following questions:
   a. [5 pts] Two underlying characteristics related to software quality in this context are: safety and security. Describe the differences between safety and security. Use elements from the provided context (e.g., artefacts and stakeholders) to answer your question.
   b. [10 pts] Besides safety and security, describe two product quality's characteristics, **and** provide an example on why they are important for the component in our context (e.g., the component that communicates with the sensors in the car).
   c. [5 pts] Explain the differences between internal and external quality.

2. **[20 pts]** Using your knowledge on software quality measures, answer the following:
   a. [10 pts] McCabe and Henry and Kafura are two distinct complexity measures used in software quality. Are they used for internal or external quality? Justify your answer.
   b. [10 pts] What are the differences between both measures in terms of software complexity?

3. **[20 pts]** Using your knowledge on Verification and Validation (V&V).
   a. [10 pts] Considering the System, Integration and Unit levels of the V-model, explain how we can plan the testing activities in the Car project in the context described above. Your answer should contain, **for each level of testing**: i) what is being tested, ii) the purpose of the test, and iii) relevant dependencies/frameworks required for the test.

   b. [10 pts] Your colleague Alice suggests using exploratory testing to test both legacy components in your automotive project. Do you agree with Alice? Justify your answer pointing to information from the context of the project and describing the trade-offs in using exploratory testing.

4. **[20 pts]** Using your knowledge on Bayesian Data Analysis (BDA), answer:
   a. [10 pts] Define and distinguish the prior and posterior distributions in a BDA workflow.

   b. [5 pts] In the model definition below, what is each line? How many parameters are in the posterior distribution?
   $y_i \sim \text{Normal}(\mu, \sigma)$
   $\mu \sim \text{Normal}(0, 10)$
   $\sigma \sim \text{Uniform}(0, 10)$

   c. [5 pts] One of the advantages of Multivariate Linear Models (MLMs) is that they are supposed to strike a balance between under- and overfitting. Explain and contrast the two concepts (under- and overfitting) and why they are bad.

5. **[20 pts]** Among the documentation of your monolithic legacy component, one engineer has logged, for each version of the source code, the number of tests executed and the corresponding number of failures. An example of the dataset is shown below.

VersionID	Tests_Executed	N_Failures
1	500	45
2	420	38
3	450	35
…		

In order to make more informed decisions on maintenance of your legacy component, Alice suggests we use Bayesian Data Analysis and the data from three years maintenance to fit a linear regression predicting the number of failures given the number of executed tests and the probability of a test failing.

Write down the *mathematical model definition* for this regression using *any* variable names and priors of your choice. Defend your choice of priors and remember to state your assumptions regarding your model!