

Ecorrection

Place student sticker here

Note:

- During the attendance check a sticker containing a unique code will be put on this exam.
- This code contains a unique number that associates this exam with your registration number.
- This number is printed both next to the code and to the signature field in the attendance check list.

Advanced Topics of Software Engineering

Exam: IN2309 / Retake

Date: Thursday 21st April, 2022

Examiner: Prof. Dr. Florian Matthes

Time: 11:00 – 12:40

P 1 P 2 P 3 P 4 P 5 P 6 P 7 P 8

--	--	--	--	--	--	--	--

Working instructions

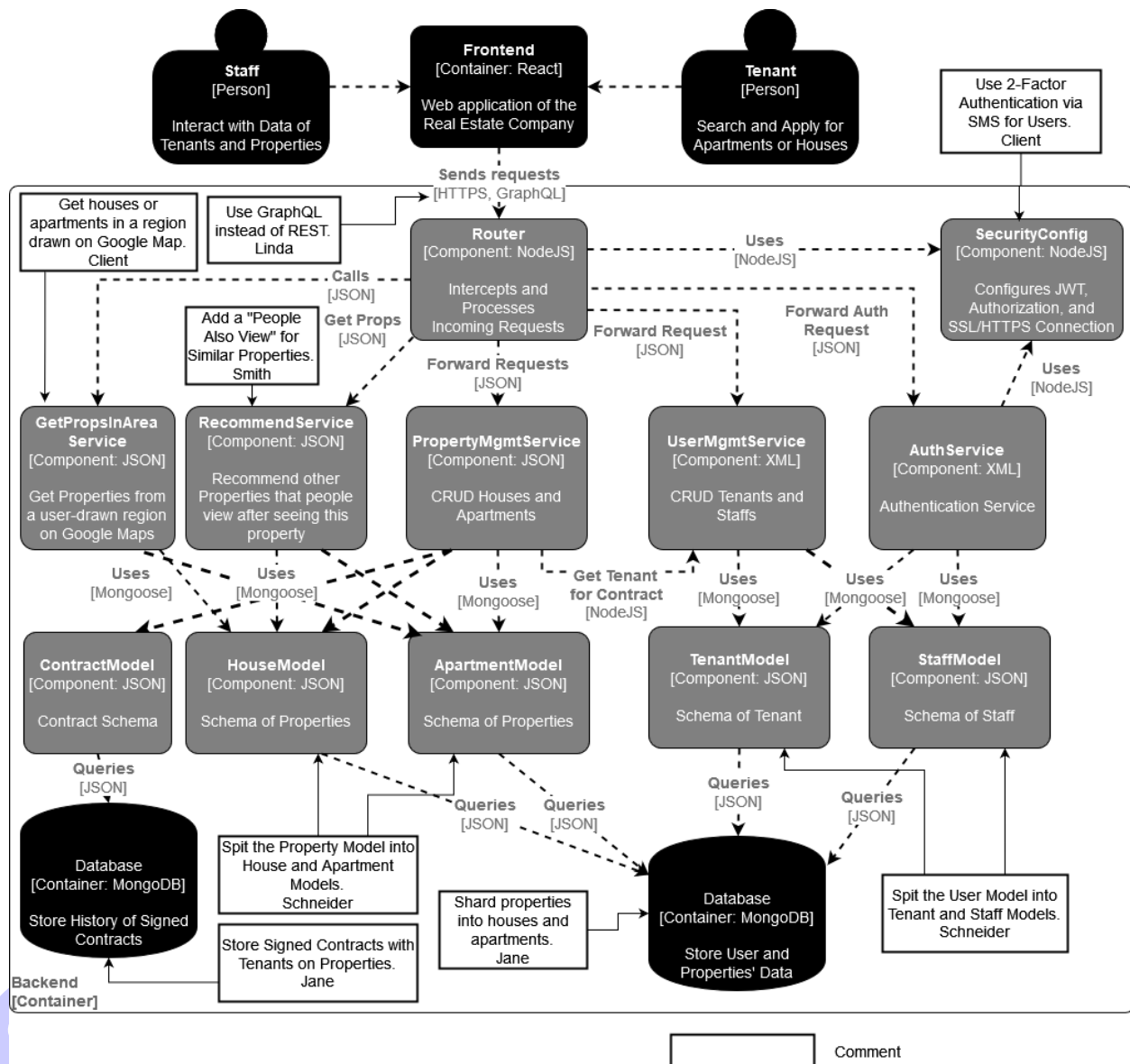
- This exam consists of **16 pages** with a total of **8 problems**.
Please make sure now that you received a complete copy of the exam.
- The total amount of achievable credits in this exam is 100 credits.
- Detaching pages from the exam is prohibited.
- Allowed resources:
 - one **non-programmable pocket calculator**
 - one **analog dictionary** English ↔ native language
- Subproblems marked by * can be solved without results of previous subproblems.
- **Answers are only accepted if the solution approach is documented.** Give a reason for each answer unless explicitly stated otherwise in the respective subproblem.
- Do not write with red or green colors nor use pencils.
- Physically turn off all electronic devices, put them into your bag and close the bag.

Left room from _____ to _____ / Early submission at _____

Problem 1 Anti-pattern (12 credits)

Jonas and his team is extending a web application to showcase apartments and houses of Semobilien, a small real estate company. Each month, Semobilien offers around 20 active real estates for rent and purchase in five areas: Munich city, Freising, Dachau, Eching and Ismaning. 50 percent of Semobilien's properties are in Munich City. The company requires their customers to register and login for applying to a house or apartment. Semobilien only aims to serve customers in the Munich and nearby districts, without expanding to another area.

Jonas and his teammates are designing the backend. Below is the most recent sketch of their backend architecture expressed as a C4 Component Diagram:



a) * What is the anti-pattern in the backend architecture design? Give two reasons to justify.

This is design by committee. (1 point)

2 Reasons (2 point each). The list is not exhaustive:

1. Unnecessarily complex: as the models for User and Properties do not have to be degeneralized.
2. Overly complex feature: find properties in a drawn region is not helpful with a handful of properties.
3. Over-engineered feature: Shard the properties database with around 20 active records each month increases complexity with insignificant improvement.
4. Make everyone happy: satisfy all the requests without justifying the necessity or complexity of the architecture
5. Gold Plating: Add features based on client interest of marketability because the feature is offered by another proprietary, like the Recommender system or find properties in drawn area
6. No experienced project architect to advise the team client on feasibility, complexity and necessity of features. One example from 2. or 3.

b) * Propose one solution to solve the anti-pattern. Briefly justify your answer

1 Point for a reasonable solution, 2 Points for explanation. The list is not exhaustive.

1. Post the meeting goal/ Time the meeting as people tends to contribute more quality when the meeting time is boxed
2. Follow the convergent, divergent and information sharing processes: generate ideas to use for the (future) design, collect the ideas and make a consensus, and exchange knowledge in the form of teaching, presentation, review or writing
3. Designate roles and make decision in a group smaller than six, as a group of more than five are less effective in creative tasks.

<input type="checkbox"/>	0
<input type="checkbox"/>	1
<input type="checkbox"/>	2
<input type="checkbox"/>	3

c) * Name and briefly explain **two** of the seven deadly sins in software engineering.

For each sin, 0.5 point for the right name, 0.5 point for explanation

Haste: make a decision hastily due to time pressure

Apathy: not care about the problem nor attempt a solution

Narrow-mindedness: refuse to use or not willing to learn a widely known solution

Sloth: making poor decision based on easy answer/lazy to invest in an efficient solution

Avarice: Addicted to complexity/Excessive modelling of details without abstraction

Ignorance: Fail to seek understanding

Pride: Not willing to adopt anything from outside

<input type="checkbox"/>	0
<input type="checkbox"/>	1
<input type="checkbox"/>	2

d) * Briefly explain what is a software quality.

1 Point for mentioning characteristics of software product, 0.5 Point for fulfilling requirements, and 0.5 Point for fulfilling expectations

Software quality refers to the **entire characteristics of a software product** that influence its ability to **fulfill specified requirements and stakeholder expectations**

(Software) Quality is the degree to which a set of inherent **characteristics fulfills requirements, meaning needs or expectations** that are stated, generally implied or obligatory

<input type="checkbox"/>	0
<input type="checkbox"/>	1
<input type="checkbox"/>	2

Problem 2 Dependency Structure Matrix (13 credits)

You are a project manager at a software company and for your next project, you are assembling a web development team. To better understand the relationship between the developers, you asked an intern to build a Dependency Structure Matrix (DSM) of all existing web developers in the company. After a week, she proposed the following matrix where weights indicate the number of projects in which one developer depends on the other:

	Alice	Bob	Carrie	Dennis
Alice	-		5	
Bob	1	-	1	2
Carrie	3		-	4
Dennis				-

a)* In how many projects does Carrie depend on other developers?

$$5 + 1 = 6$$

b)* Upon further examination, you realized that the DSM is not partitioned yet. Draw a partitioned version of the DSM such that it is in a block triangular form (use crosses (X) instead of exact weights to represent dependencies).

A DSM which has been rearranged so that all dependencies either fall below the diagonal or within groups, is said to be in block triangular form
 0 pt if anything besides Alice-Carrie dependency fall over the diagonal
 Dennis first row (2pts), Bob last row (2 pts)
 Alice and Carrie 2nd and 3rd row (order doesn't matter) (1pt each)
 Correct row, wrong box marked (-0.5pt)

	Dennis	Alice	Carrie	Bob
Dennis	-			
Alice		-	X	
Carrie	X	X	-	
Bob	X	X	X	-

c) After partitioning the DSM, justify whether it is in an **optimal form** (lower triangular) or not. If not, propose a way to achieve it.

2 points for correct answer (1pt) and justification (1pt), 2 points for a correct way to achieve it
 The current DSM is not optional since there is a cyclical dependency between Alice and Carrie. To remove any cycles, Alice and Carry can be regarded as a single composite entity (A-C).

If student showed the optimal form (A-C composite entity) in the previous question, then full points will be given for stating here that it is in optimal form.

d)* List **two** benefits of using a DSM as discussed in the lecture.

Each benefit 1 point
 The matrix representation scales better than box-and-line diagrams
 Graphs are more intuitive but when the numbers of nodes and edges grow, they become very complex
 Helps to better understand the information flows
 The partitioning algorithms provide an automatic mechanism for architectural discovery in a large codebase possibility to spot structural patterns at a glance
 Efficient cycle detection in DSM

Sample Solution

Correction Notes

Problem 3 Message-oriented Architectures (14 credits)

A world-famous sushi restaurant in New York City, U.S.A contacted you for getting a consultancy regarding their supply chain. Currently, the restaurant is ordering fresh fish from a market in Japan, and it takes up to a day for the fish to be delivered. The restaurant utilizes the following **software and hardware components** to track the location of their deliveries and do menu planning:

- Sensors: Continuously transmits data about the location of the deliveries
- Track-App: Provides a web interface for restaurant managers to track the location of the deliveries
- Whats-on-the-Menu-App: Provides available product information based on the location of the deliveries (e.g. "Tuna sushi will be on the menu if the deliveries pass the Pacific Ocean by 10 AM"). Updates app users daily with the new menu

a) * Your first task is to identify which message queue pattern should be used between which components and entities (e.g. app users). Justify your decisions.

1 pt for the correct message queue, 1 pt for correct justification

- Sensors - Track-App: Pub-Sub model should be used. Sensors publish location data and track-app consumes this data to provide a web interface about the location of deliveries to managers.
- Sensors - Whats-on-the-Menu-App: Pub-Sub model should be used. Sensors publish location data and whats-on-the-menu-app consumes this data to provide available product information to app users.
- Whats-on-the-Menu-App - App Users: Pub-Sub model should be used. The app will send push notifications to app users regarding the daily menu.

b) * After your initial inspection, you realized that a traditional message-oriented middleware (MOM) is used between the components. You know that with the current setup, communication between each component is handled separately and each publisher needs its own configuration. Propose a solution that would create a *hub and spoke* communication infrastructure for integrating the components (just the name is sufficient) and briefly explain **two** challenges of it.

Message broker (1 pts)

0.5 pt for challenge name, 1 pt for explanation

Scalability - with thousands of clients, connections, and streams of messages

Security - as the number of SSL certificates (for client authentication) increase, managing them becomes a challenge

Monitoring and debugging within a broker becomes expensive over time

Maintenance of message brokers – message broker cannot be updated without affecting the clients

c) *Briefly explain **two** motivations for using messaging.

1 Point for correct motivation, 1 Point for correct explanation

Remote Communication: message is a simple yet efficient protocol to communicate between two remote systems

Decoupling of systems: a system does not have to depend on a particular function or module of another system, it only sends the message and the execution is done accordingly on the other side.

Platform/application integration: enable coordination among different independent components such that they can execute their own tasks and simultaneously send requests to other applications

Asynchronous communication protocol: with messaging, an application can communicate with another one without blocking itself for the response. The sender can still continue working on their task

Solve the issue of throttling: avoid overloading requests on the receiver side with a message queue. (When the receiver is fully loaded, it will stop receiving messages. The sender can store its message in a queue and send them when the receiver is available)

Problem 4 Security (12 credits)

CriticalEater is an online delivery application exclusively for high-end restaurants. The application provides an interface for users to order food from available restaurants and pay instantly through the payment service. The application also has a forum section where users can write their reviews about restaurants. Recently, users started complaining about getting blackmailed by people who claim to have their credit card information. You have been hired by the company as a security consultant to analyze this issue and explore any vulnerabilities in the codebase.

After inspecting the codebase, you discovered the following issues:

- No input sanitization is done on the reviews entered into the forum section
- Upon login, user credentials (including the credit card data) are stored in a **non-Http-Only cookie** on the client-side

a)* Name the attack (and its type) which can be used to send the credentials of every user that visits the forum section, to a malicious server. Briefly explain how the attack can be executed.

The goal can be reached with a persistent (1 pt) XSS attack (1 pt).

A malicious user will submit a review which is a script that reads the cookies (which includes credentials) and sends them to a server with a post request. Cookies can be read on the client side since they are non-HTTP-only. (1pt) Every time a user visits the review section, this script will be rendered by the application on the client machine. Thus, the script will be able to access the cookies (credentials) and forward them to another server. (1 pt)

b) Briefly explain how this attack can be prevented to some extent and why it is not trivial to prevent it completely.

A persistent XSS attack can be prevented to some extent by filtering the user input for embedded Java(script) (2 pt). However, this is not trivial since the string "<script>" can be encoded in multiple ways. Simply searching for all encodings is not trivial. (1 pt)

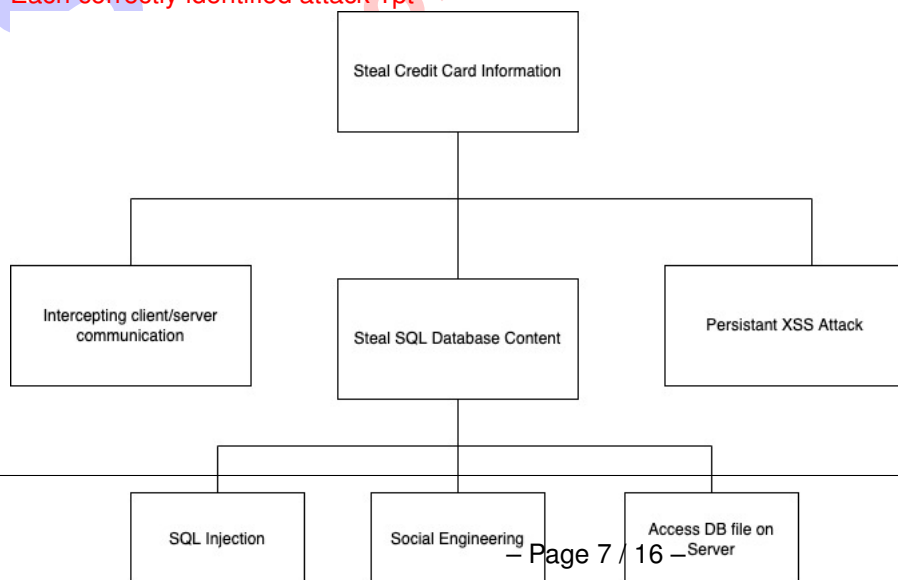
Using HTTP-Only cookies would help to expose less information about users, thus it would prevent credit card info or credentials getting stolen from cookies (1pt). However, this is not a measure against XSS.

c) Besides the attack you have discovered initially, you also found the following ways to steal the credit card information of users:

- By intercepting the client/server communication
- By stealing the SQL database content

You know that any of these three options (including the attack you have discovered initially) are sufficient enough to steal users' credit card information. Draw an attack tree including all these attacks and find **two** ways to steal the SQL database content (show them in the attack tree as well).

Each correctly identified attack 1pt



Problem 5 Blockchain (14 credits)

- 0 ☐
1 ☐
2 ☐
3 ☐
4 ☐
- a)* Blockchain is one of the most hyped technologies these days, and one of the reasons for that is the advantages it provides over centralized platforms. It is known that centralized platforms can suffer from *availability*. Briefly explain how blockchain-based architectures overcome this problem. Also, list **two** more drawbacks of centralized platforms as described in the lecture.

Since blockchain-based architectures are decentralized, there is no single point of failure. Thus, even if a computer from the network fails, the platform will continue to perform since other participants of the network are available. However, in centralized systems, the whole system can sometimes go down when a single server crashes. (2 pts)

1 pt each

manipulations

sufficient resources

fair conditions

trust

- 0 ☐
1 ☐
2 ☐
3 ☐
4 ☐
- b) * Proof-of-Work (PoW) is the protocol that has been utilized by some of the popular blockchains such as Bitcoin and Ethereum (1.0) for determining the next block creator. Explain how PoW makes it impractical for an attacker to successfully operate a 51% attack.

PoW requires the miners to solve a mathematical puzzle in order to get the right to propose the next block (2pts). This requires a significant hash power from the miner. Thus, in order to obtain the majority of the hashing power, the attacker has to be highly invested. This is already a barrier of entry for operating such an attack (1pt). Assuming the attacker has enough money to operate this attack since such an attack would destroy the trust in the system, the network will be destroyed and the attacker won't be able to profit anymore. (1pt)

- 0 ☐
1 ☐
2 ☐
3 ☐
4 ☐
- c)* Smart contracts cannot directly access external data available outside the blockchain (e.g., weather data, exchange rates). Name the characteristic property of smart contracts that ensures this and briefly explain why it is needed. Also, name the entity that allows smart contracts to access external data.

The deterministic execution property of smart contracts ensures that each node of the network can find the same result, given the same input for a contract function. Otherwise, reaching a consensus would not be possible since each node would find different results when validating a transaction. (1pt for the property name, 2 pts for explanation)

Smart contracts can access external data using oracles. (1pt)

- 0 ☐
1 ☐
2 ☐
- d)* Decentralized platforms such as blockchains are not the most fitting architecture for every use case. List **two** advantages of centralized platforms over decentralized platforms as described in the lecture.

Each 1 point

Efficiency

Scalability

Standardized components

Simpler and more mature technology

Less system complexity

Easier platform governance

Predictable, lower (even no) transaction costs

Problem 6 SCRUM (13 credits)

Steven and his team is developing a web application for ordering online fruits and vegetables for a grocery store. Currently, the shop owner requires the system to elegantly display the information of the shop and its goods, enable pick-and-go ordering of food in the store with credit or debit card payment, and the website is displayed in German or English.

However, in the future, the system can have other features that the owner have not decided yet, such as purchasing coupon, shop-to-home delivery, supporting various payment methods, or reserving seasonal food items.

Steven's team consists of five members: one UI Design expert, an experienced full stack developer, and three newly graduated Master's students who do not have remarkable practical experience in full-stack software development.

a) * Steven's team plans to use SCRUM for developing the web application. Is SCRUM applicable to the current scenario? Briefly explain **two** reasons to support your answer.

SCRUM is applicable to the context (1 Point)

3 points for each valid reason, 1 Point for mention only the reason without justification or illogical explanation

+ The requirements of the system is not fixed, so the team can apply SCRUM to iteratively and incrementally develop the current and future requirements

+ There are potential changes/risks in the future development, as the stakeholder has a vision, but not concrete expectations of the system's features and behaviors. Therefore, the team can experiment their design and get the feedback from the customers iteratively.

+ The team consists of unexperienced members, hence applying SCRUM gives the team a time to explore their capabilities, work ethics, strengths, weaknesses and potential domain of expertise.

0
1
2
3
4
5
6
7

* Briefly state **two** activities in each of the following SCRUM events:

b) Sprint Planning

Define a Sprint Goal that brings certain value to the stake holder Select items from Product Backlog to the Sprint Backlog

Refine items in the Sprint Backlog

Establish or Refine the Definition of Done

Plan the necessary work to create an increment

0
1
2

c) Daily Scrum

Inspect the progress of development

Identify or solve impediments occurred during the development

Adjust the plan to develop or deliver the Sprint increment

0
1
2



d) Sprint Retrospective

The Scrum Team discusses what went well during the Sprint, what problems it encountered, and how those problems were (not) solved
Identifies the most helpful changes to improve the team's effectiveness
Identifies potential risks to the development process

Sample Solution

Correction Notes

Problem 7 Microservice and Database (14 credits)

a) * Briefly explain **four** challenges of microservice development.

For each challenge, 0.5 Point for a correct description, 1 Point for a reasonable explanation. This list is not exhaustive, i.e., other reasonable challenges are also accepted

Additional complexity: Dev has to manage various aspects from inter-service communications to service orchestration

Testing distributed system: multiple services running simultaneously and potentially written in different languages creates challenges for the dev team has to implement seamless unit, integration, and system tests

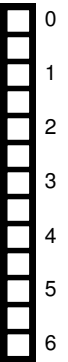
Distributed Transactions Management: how could one ensure that multiple transactions sent to different microservices are executed accurately and consistently. In case the transaction fails, the system has to know where the error occurs, how to handle the error, or report it to the user.

Service Partitioning: How to split the system into different services in an appropriate size

Security Protection: how to ensure the integrity of data flow within the microservice system, or secure the microservices with large attack surface and potential malicious insiders

CI/CD Pipeline Management: Choosing the best strategy to continuously integrate numerous microservices simultaneously requires experience and strategic vision. Otherwise, the CI/CD pipeline may run on an outdated version and endanger the correctness of the system

Service Orchestration: How to manage all microservices to work seamlessly given that they are distributed in different locations, environment, development languages, and configuration parameters



b) * Which problem is solved by the Service Discovery pattern?

How does the client (or consumer) of a service discovers the location (i.e. host and port) of a service instance (or provider)?

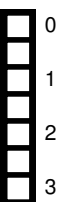


c) * Briefly explain what is a server-side service discovery and what is a client-side service discovery?

0.5 Point for how the service registry is reached from the client, 1 Point for the correct forwarding mechanism

Server-side service discovery: the client makes a request via a router (API Gateway or load balancer) that runs at a well known location. The router queries a service registry, which could be in the same router, to forward the requests to an available service instance.

Client-side service discovery: the client is registry-aware (i.e. it knows the location of the service registry). Therefore, it directly queries the service registry to retrieve a list of available instances of the requested service, and sends the request to a service instance





d) * Briefly describe **two** properties of document-oriented databases.

1 Point for each correct description
Collections do not enforce a schema
Collections can be created on demand
Document is a set of key-value pairs and can have dynamic schema
All documents within a schema belongs to a similar or related domain
Documents as independent units enable

- distributing the data accross multiple locations (Scalability)
- OR writing data sequentially off disk, improving Performance
- OR avoid costly migration as the database does not need to know the schema in advanced



e) * Given the following *students* Table:

ID	name	study_program	grade
1	Alice	Computer Science	1.7
2	John	Bioinformatics	1.7
3	Stein	Computer Science	3.0
4	Mary	Medicine	1.3

Transform the following query into Map-Reduce:

```
SELECT study_program, COUNT(*) FROM students GROUP_BY study_program
```

Note: Use function **size(array)** to get the number of elements in an array object.

map(key, record): // 0.5 Point
emit(record.study_program, 1) // 0.5 Point

reduce(iKey, iValues): // 0.5 Point
emit(iKey, size(iValues)); // 0.5 Point

Problem 8 Quiz (8 credits)

Evaluate the following statements (correct answers + 1 point, incorrect answers -1 point, you cannot get less than 0 points for this exercise):

Statement	True	False
The state modified by transactions is not stored on the blockchain.	X	
Safety ensures the protection of the system from external hazards.		X
Software should aim to achieve content coupling, not data coupling.		X
The non-repudiation security property ensures that it is impossible for a user to inappropriately deny a transaction or having sent a message.	X	
A high domain range ratio leads to a good testability.		X
Task parallelism focuses on distributing tasks across multiple cores.	X	
To develop a robust system, Autonomous Vehicles Architecture includes redundant functions and systems.	X	
In Function as a Service, the consumer has to allocate resources for handling triggered events.		X

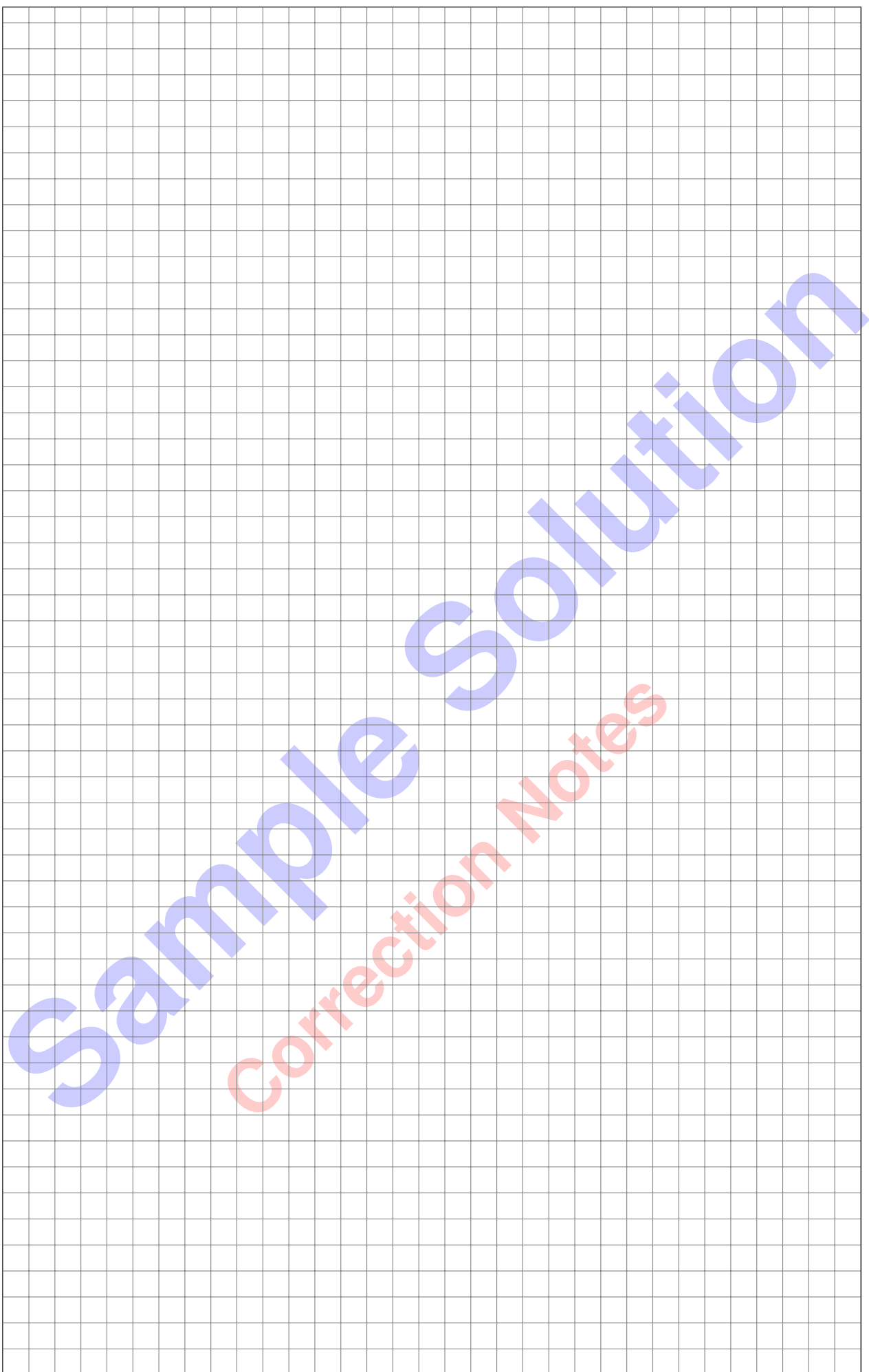
<input type="checkbox"/>	0
<input type="checkbox"/>	1
<input type="checkbox"/>	2
<input type="checkbox"/>	3
<input type="checkbox"/>	4
<input type="checkbox"/>	5
<input type="checkbox"/>	6
<input type="checkbox"/>	7
<input type="checkbox"/>	8

Sample Solution

Correction Notes

Additional space for solutions—clearly mark the (sub)problem your answers are related to and strike out invalid solutions.

The image shows a large rectangular area filled with a fine grid of squares, typical of graph paper. A diagonal watermark is overlaid across the grid. The text 'Sample Solution' is written in a large, blue, sans-serif font, slanted upwards from left to right. Below it, the text 'Correction Notes' is written in a smaller, red, sans-serif font, also slanted upwards from left to right.



Sample Solution

Correction Notes