Sample questions – part 2

Consider the following cosine similarity table. Which of the following threshold is more dependable in terms of TPR and FPR?
T = 0.50, 0.75, 0.95

	Exp1	Exp2	Exp3	Exp4
Act1	0,86	0,84	0,55	0,42
Act2	0,72	0,92	0,12	0,53
Act3	0,81	0,16	0,78	0,55
Act4	0,44	0,75	0,33	0,93

Consider the following function, proposed mutants and generated tests. Compute the mutation score of the test suite.

```
def calculate_final_price(base_price, discount, tax_rate, membership_status, loyalty_points):
    discounted_price = base_price * (1 - discount / 100)
loyalty_discount = loyalty_points * 0.05  # Each loyalty point gives a 5% discount
    discounted_price -= loyalty_discount
    final_price_with_tax = discounted_price * (1 + tax_rate / 100)
    if membership_status:
        final_price_with_tax *= 0.9  # Premium members get an additional 10% off
    if final_price_with_tax < 0:
        final_price_with_tax = 0
    return final_price_with_tax
```

discounted_price = base_price * (1 + discount / 100)

loyalty_discount = loyalty_points * 0.1

if not membership_status:

T1: assert calculate_final_price(100, 20, 10, True, 10) == 72.0
T2: assert calculate_final_price(200, 10, 0, False, 0) == 180.0
T3: assert calculate_final_price(100, 0, 0, False, 0) == 100.0

Consider the following requirements. Write a few-shot prompt to extract OOP Classes for a software architecture design.

REQS = “The platform allows customers to browse menus from various restaurants, place food orders, and track the delivery in real-time. It also includes features for reviewing and rating restaurants, as well as providing personalized recommendations based on the customer’s past orders. Additionally, platform managers can update restaurant menus, track orders, and manage customer feedback to ensure a smooth experience.”

Define a Router prompt for the LLAMA engine for the generation of the three parts of user stories based on the following requirements. The router prompt must have a default branch,

REQS = “The platform allows customers to browse menus from various restaurants, place food orders, and track the delivery in real-time. It also includes features for reviewing and rating restaurants, as well as providing personalized recommendations based on the customer’s past orders. Additionally, platform managers can update restaurant menus, track orders, and manage customer feedback to ensure a smooth experience.”

Which of the following is **not** a way to evaluate the correctness in generating code?

- a) functional correctness
- b) execution time
- c) execution success rate
- d) blind peer reviews
- e) developer feedback

In a Router Chain, what is the role of the 'Default Chain'?

A. To act as a fallback when no other subchain is selected.

B. To store previous results for future processing.

C. To route inputs to multiple subchains simultaneously.

D. To ensure inputs are processed in a sequential order.

What is the main function of the Reflection module in an agentic architecture?

A. To evaluate past decisions and improve future performance.

B. To manage and store intermediate outputs in short-term memory.

C. To dynamically route inputs to specialized subchains.

D. To decompose tasks into smaller, sequential subgoals.

E. To plan the most efficient strategy for task execution.

Few-Shot Prompting is particularly effective in which of the following scenarios?

A. When the task requires the model to autonomously explore various paths to a solution.

B. When there is sufficient prior training data, making examples redundant.

C. When the task demands clarity through structured reasoning without specific examples.

D. When a complex task needs to be demonstrated with multiple examples to guide model behavior.

E. When the output must follow a consistent tabular format for easy comprehension.

What is a potential limitation when using industrial datasets for training LLMs in software engineering (LLM4SE)?

A. They are often publicly available and lack restrictions on data usage.

B. They tend to focus primarily on high-level software concepts, neglecting specific code-based insights.

C. They contain sensitive and proprietary data, which could pose privacy and security concerns in the model's training.

D. They are mainly composed of graphical user interface (GUI) representations, making them less suitable for training on code-based tasks.

E. They primarily include bugs and patches but lack sufficient context on software development practices.