

Quiz 6 - Software Development Agents w/ Graham Neubig (10/14)

Total points 5/5

INSTRUCTIONS:

Each of these quizzes is completion based, however we encourage you to try your best for your own education! These quizzes are a great way to check that you are understanding the course material. You can attempt this quiz as many times as you wish. You only need to complete the quizzes if you wish to earn a completion certificate. More information at the bottom of the course website.

IMPORTANT:

In order to receive credit, use the same email address as the one used to sign up for the course. If you are not sure which email you used, just complete the sign up form again with your preferred email.

PROBLEMS?

If you have any technical difficulties about this quiz, please ask course staff in our LLM Agents Discord.

Email *

christopherbare@gmail.com

✓ What is *not* an effective method of file localization? *

1/1

- ☐ Creating a map of the repository structure and using it to guide the agent
- ☐ Offloading the task to the user, allowing experienced users to specify which files to use
- ☐ Using search tools to assist the agent in finding relevant files within the repository
- ☒ Manually scanning through every file in the project directory to find relevant files ✓



✓ **What does Pass@K measure? ***

1/1

- ☒ After generating K examples, whether there is at least one of them that satisfies the code specification ✓
- ☐ After generating K examples, calculate the average length of the generated outputs to assess their quality
- ☐ After generating K examples, evaluate their complexity using a predefined scoring rubric to determine how many are valid
- ☐ After generating K examples, count the number of unique outputs produced to measure the diversity of the generated responses

✓ **Why is data leakage a big problem when evaluating software agents? ***

1/1

- ☒ It can lead to overly optimistic performance metrics that do not reflect true generalization ✓
- ☐ It makes the code harder to read and maintain by increasing complexity
- ☐ It prevents agents to learn from a broader range of data, hurting their performance on all tasks
- ☐ It decreases the speed of training and evaluation processes, leading to slower model deployment



✓ **What is an example of how *not* to mitigate safety concerns of coding agents?** *1/1

- ☐ Review actions after they have been executed to detect and address any potential security issues or unauthorized activities
- ☒ Use automated code formatting tools to ensure that code generation is neatly organized and is easy to read in order to reduce the likelihood of human error ✓
- ☐ Restricting permissions and managing access tokens to ensure that users and applications only have the minimal level of access necessary
- ☐ Isolate the agent's execution environment to prevent potentially harmful actions from affecting the broader system

✓ **What is "code infilling"? *** 1/1

- ☐ A technique used to convert comments in code into fully functional methods
- ☐ The method of testing code by filling it with random values to check for errors
- ☒ The process of automatically generating the missing parts of code based on the surrounding context ✓
- ☐ The process of refactoring existing code to improve its readability and maintainability

This form was created inside of UC Berkeley. [Report Abuse](#)

Google Forms

