

RPM Project: Milestone 4 (Spring 2023)

First, make sure to read the [full project overview](#). It contains instructions for the project as a whole and getting started with the code. This page describes only what you should do for Milestone 4 within the broader context of that project overview.

For Milestone 4, your goal is to demonstrate that you have addressed all the problems covered by the final project. This should leave you the last part of the term to improve your agent, borrow and experiment with ideas from your classmates, and write your final, more formal report. 50% of your grade on Milestone 4 is earned by meeting the minimum performance requirement; 50% of your grade is earned by completing the milestone journal.

Performance Requirement

For Milestone 4, you will earn 2.5% of your Milestone grade for each Basic D and Basic E problem you get right up to a maximum of 25%. You will also earn 2.5% of your Milestone grade for each Test D and Test E problem you get right up to a maximum of 25%.

In other words: if your agent answers at least 10 Basic D+E and 10 Test D+E problems correctly, you earn full credit for the Performance Requirement of Milestone 4. Each problem fewer than that that your agent can answer results in a deduction of 2.5% from your Milestone grade.

Remember that when submitting answers to Gradescope, the answer choices are shuffled: the content of the correct answer to each problem will be the same every time, but it will be assigned a different number on each submission as a guard against hardcoding and overfitting.

Submission Instructions

To fulfill the Performance Requirement, follow the directions on the [full project overview](#) for submitting to Gradescope, and then submit your agent to the Milestone 4 assignment.

You may submit up to 40 times. The large majority of students do not need nearly that many submissions, so do not feel like you should use all 40; this cap is in place primarily

to prevent brute force methods for farming information about patterns in hidden test cases or submitting highly random agents hoping for a lucky submission. Note that Gradescope has no way for us to increase your individual number of submissions, so we cannot return submissions to you in the case of errors or other issues, but you should have more than enough submissions to handle errors if they arise.

This is an individual assignment. All work you submit should be your own. Make sure to cite any sources you reference or code you use (in accordance with the broader class policy on code reuse).

Grading Information

The Performance Requirement of this Milestone is graded out of 40 points. For this Milestone, in accordance with the [Project Overview](#), you must get at least 10 Basic D & E problems and 10 Test D & E problems correct to get full credit. For each problem fewer than 10 that your agent answers correctly in either set, you will lose 2 points. Additional correct answers in one set do not compensate for missed problems in the other; if you answered 11 Basic D & E problems correctly and 8 Test D & E problems correctly, for instance, your Performance score would be 36/40: 20 out of 20 for answering at least 10 Basic D & E problems, and 16 out of 20 for answering only 8 Test D & E problems.

Note that we do not differentiate which of the two sets (D and E) the ten correct problems come from. You could receive full credit by answering 10 D problems correctly, 10 E problems correctly, or any mixture of the two so long as the total adds to 10.

After submitting to the Milestone 4 assignment in Gradescope, make sure to select which submission you want to have count for your graded submission. By default, Gradescope will use your latest submission, but you may want to use an earlier one. After the deadline, this will be exported to Canvas to calculate your final Milestone grade.

Milestone Journal

In addition to submitting an agent to Gradescope, you will also submit a brief milestone journal to the Milestone 1 assignment in Canvas. Your Milestone Journal must be written in [JDF format](#). There is no maximum length; we expect most submissions to be around 5 pages, but you may write more if you would like. Writing the journal is intended to be a useful exercise for you first and foremost: it should let you externalize and formalize your ideas, it should let you get feedback from your classmates, and it should let your classmates learn from you. Your journal need should not include actual code; it should just include a description of your agent's approach.

Note that your Milestone 4 should be all original content; if there is content that you wrote for a previous Milestone that is still pertinent, you may refer back to that content again (including quoting yourself), and then go on to discuss what has changed or what is new (or, why the same content you wrote previously is still so applicable).

For example, you might write:

In Milestone 1, I wrote that my agent works by “calculating a percentage change in the number of black pixels between each pair of frames, and checking for mathematical patterns in the changing ratio of black pixels. Then, I checked each answer option to see if it maintained the observed mathematical pattern.” Then, in Milestone 3, I wrote that “For Set C, that continued to work, and in fact the patterns were easier to find because they were fit to 28 possible pairs instead of just 3.” However, for Sets D and E, that approach struggled because fewer problems are solvable by sequences of pixel ratio changes. So, I had to add an additional heuristic method...

If you need to quote large portions of your prior writing, you can use a blockquote, or include your prior Milestone in an appendix that you refer to. The important element is for TAs and classmates to be able to identify the new content.

For Milestone 4, you should answer the following questions:

- How does your agent currently function? Depending on the inner workings of your agent, there may be a lot of different ways to describe this. For example, does it select from multiple problem-solving approaches depending on what it sees in the problem? Does it perform shape recognition or direct pixel comparison? Does it generate a candidate solution and compare it to the options, or does it take each potential answer and assess its likelihood? You need not answer these specific questions, but they are examples of ways you might describe your agent's performance.
- How well does your agent currently perform? How many problems does it get right on the Set D+E problems?
- What problems does your agent perform well on? What problems (if any) does it struggle on? Why does it struggle?
- How efficient is your agent? Does it take a long time to run? Does it slow down significantly on certain kinds of problems?
- How do you plan to improve your agent's performance on these problems before the final project submission?
- What feedback would you hope to get from classmates about how your agent could do better? What challenges do you think could benefit from someone else's feedback?

Tip: Remember, we want to see how you put the content of this class into action when designing your agent. You don't need to use the principles and methods from

the lectures precisely, but we want to see your knowledge of the content reflected in your terminology and your reflection.

Submission Instructions

Complete your project journal [using JDF](#), then save your submission as a PDF. Journals should be submitted to the corresponding assignment submission page in [Canvas](#). You should submit a **single** PDF for this assignment. This PDF will be ported over to Peer Feedback for peer review by your classmates. If your assignment involves things (like videos, working prototypes, etc.) that cannot be provided in PDF, you should provide them separately (through OneDrive, Google Drive, Dropbox, etc.) and submit a PDF that links to or otherwise describes how to access that material.

This is an individual assignment. All work you submit should be your own. Make sure to cite any sources you reference, and use quotes and in-line citations to mark any direct quotes.

Grading Information

Your assignment will be graded on a 40-point scale coinciding with a rubric designed to mirror the question structure. Make sure to answer every question posted by the prompt. Pay special attention to **bolded** words and question marks in the question text. For further information on how the assignment is graded, see the rubric in Canvas.

Peer Review

After submission, your assignment will be ported to [Peer Feedback](#) for review by your classmates. Grading is *not* the primary function of this peer review process; the primary function is simply to give you the opportunity to read and comment on your classmates' ideas, and receive additional feedback on your own. All grades will come from the graders alone. See the course [participation policy](#) for full details about how points are awarded for completing peer reviews.