

Gradescope for Programming Assignments

August 12, 2022

Webinar Q&A

Question	Answer
Is there support for using my own docker container in the autograder?	You are able to use your own docker container for the autograder. Here is more information about its requirements and how you can set one up: https://gradescope-autograders.readthedocs.io/en/latest/manual_docker/
Is the python autograder library still being maintained? There hasn't been any recent activity and some pull requests have been waiting for 6+ months	Yes, the Python autograder library is still being maintained, and we have confirmed that all open pull requests have been reviewed.
Can we provide assignment instructions on the Gradescope platform?	There isn't currently an option to upload assignment instructions or a template file to a programming assignment, but this is on our roadmap.
Do you have updated docker base images? (I'm currently using your ubuntu 18 image)	Yes, you can select a newer version of Ubuntu from the base images version dropdown on the Configure Autograder

	<p>page. The current default base image is Ubuntu 22.04. You can see the list of base images which we have built here: https://hub.docker.com/r/gradescope/autograder-base/tags/</p> <p>Note that not all of the images there may appear in the Gradescope interface at the moment.</p>
<p>So Instructor uploads autograder.zip, gradescope builds the Docker image. How long does this image remain turned on for? When the student submits, does a copy of the autograder docker image spin up or does it run against the same docker container that the instructor generated?</p>	<p>To clarify the Docker terminology: the autograder setup process builds a Docker "image", which simply stores the filesystem after the result of downloading your zip file and running the setup script. After the setup script completes, the image is stored and the initial setup container shuts down.</p> <p>At run time, Gradescope starts a new "container" which starts from the stored Docker image and executes your autograder on the student's code. You can think of a container as a lightweight virtual machine, or essentially an isolated computing environment. Each student's submission will spin up a new Docker container instance.</p>
<p>Can you create an assignment that is a mixture of programming problems and "regular" written problems?</p>	<p>[answered live + following typed answer]</p> <p>Students can submit their written responses as PDF files (or any other file type) to a programming assignment, and you can then define questions to manually grade on the PDF on the Edit Outline page of the programming assignment. Programming assignments support uploading any number of files of any extension, so both a code file and PDF file(s) can be uploaded to the same assignment.</p>
<p>Is it possible to give line-by-line feedback comments if manual grading is turned off?</p>	<p>If you want to leave line-by-line feedback on students' code submissions, you need to</p>

	<p>turn on manual grading. If you don't want to award points with manual grading, you can set your manual grading question(s) to be worth 0 points and just use the manual grading interface for leaving in-line comments.</p>
<p>Can we compare to past semesters yet in the similarity checker? Or compare against Github repos / files?</p>	<p>The code similarity feature currently covers one assignment at a time. Expanding it across semesters is a feature that has been requested. You can see more details and vote it up on our roadmap: https://trello.com/c/8tjyQc4x/273-code-similarity-checks-against-a-wider-database-of-submissions</p>
<p>Does the programming grading work for R code?</p>	<p>Yes! You can build an autograder to grade an R coding assignment. You can read more about that and see a couple of R autograder examples here: https://gradescope-autograders.readthedocs.io/en/latest/resources/</p>
<p>Can we also grade several programming assignments inside a Jupyter notebook with Autograder at once?</p>	<p>Yes! You can find resources to help you build an autograder to grade a Jupyter notebook here: https://gradescope-autograders.readthedocs.io/en/latest/resources/ Specifically, Otter-Grader is a community built Jupyter Notebook autograder which is compatible with Gradescope.</p>
<p>For Python courses, especially Beginner level, many courses don't start by teaching Classes and methods (e.g., using "self" parameter) because they aren't required by the language. Can autograder work without classes and methods that use "self"?</p>	<p>The autograder does not require you to write it in a class format or have the student write classes. The student code can be in whatever format you want. You will need to build your autograder to understand your student's code (however you had them format it) and autograde the submission to give Gradescope the results.</p> <p>If the students are writing non-method</p>

	<p>functions (i.e. functions not using self) then it should still be straightforward to call the function with different inputs and check the returned values. For example, see the the “fibonacci” example in the webinar slides.</p> <p>If students are not writing code in functions at all, you may consider testing the behavior of the code in other ways, e.g. if their code communicates on standard input/standard output, you can run it in a subprocess and communicate with it that way. Some instructors have also used static analysis tools to parse the student’s code and check if variables exist and how they’re named, etc. In Python, you can also import a file to see any variables defined and their values by the end of execution.</p>
<p>Would be beneficial to be able to link/connect to enterprise deployments of github/gitlab.</p>	<p>We do support Gitlab and Github submissions, but this won’t work with Enterprise accounts at the moment. We can connect to institutional GitLab servers as long as they are accessible via the internet and not only via local networks.</p>
<p>The configurations and setup of autograders is a pain and it seems as if the system has been designed for IT folks rather than Instructors. This is especially true for Instructors teaching CS courses outside of computing departments. Does the team have a product vision on building GUI abstractions over the current setup for most common languages and simple use cases such as IO based testing? Or in the foreseeable future will the workflow remain the same as the status quo? I understand that the current workflow offers more flexibility and is powerful but from a usability perspective the barriers to</p>	<p>Thank you for the feedback. As you noted, the goal of the Gradescope autograder platform is to provide instructors the greatest amount of flexibility to support any programming language, assignment type (regardless of complexity), and code structure. Autograders can be as simple as just compiling student code and as complex as evaluating a student-written mobile application. Introducing a GUI to allow instructors to build autograders within the platform or an IDE for students to type code within Gradescope may limit this flexibility, but we are definitely considering ways to make code autograder</p>

<p>entry and setup are high.</p>	<p>setup more user-friendly. We hope to have more documentation, resources, user communities around autograder building, and webinars in the future to support instructors who are writing autograders with Gradescope.</p>
<p>Are there any plans to add Julia to your list of supported languages? I didn't see it on the list you showed.</p>	<p>Since our autograder platform gives you a generic docker image, you can set up pretty much whatever language you want. You should be able to install the Julia language when your autograder builds and then evaluate the student code and give the results back to Gradescope.</p>
<p>Looking at the outline page for a programming assignment, it appears there can only be one "Autograder" problem per programming assignment. Is there a way to have multiple different "Autograder" problems per assignment?</p>	<p>There is one autograder assignment because we only launch one run per submission. Within your autograder, you can have as many problems as you want and autograde all of the problems in the single autograder run. You would then return the results of all of the problems to Gradescope.</p>
<p>Can additional packages be added to the base Ubuntu image? For example LLVM's clang and updated version of GCC? My grading process requires specific tools and versions.</p>	<p>During the setup step of the autograder, you can install pretty much anything you want. In your case, you would need to make sure your `setup.sh` file contains the required steps to install LLVM and update GCC. Once built, the autograder will use those packages when it runs on a submission.</p> <p>Note that we now have multiple Ubuntu versions available as base images, so you may want to try different base images to see which ones have the package versions you want, or at least make it easier to install those versions.</p>
<p>Can students be given the option of what language to submit in?</p>	<p>The students can submit whatever is needed to the autograder. That could include a program written in different</p>

	<p>possible languages. Your autograder would have to figure out how to build/run their code to get the results back to Gradescope. You would need to include all relevant language tools and runtimes in your autograder image by installing them in your setup script.</p> <p>Depending on the set-up of your assignment, you could also add a test case that detects the language the student is writing their autograder in and fails the test if the language is not an expected language. You would just need to run the language detection as part of the autograder and give Gradescope the failure information.</p> <p>We would recommend that you tell your students that they must have a method to build their autograder and a method to run their autograder in a standard format so it is easier for your autograder to run the students' submissions.</p>
<p>In these examples, if the student's code doesn't define functions with the particular names that the autograder assumes, then it looks like the autograder would fail completely. Is there a recommended way to report "you didn't define function foo" as a single failed test case while still being able to autograde other functions that they did define? I can think of ways, but none particularly elegant.</p>	<p>You can detect if the students uploaded the required files. Additionally, depending on the language, you should be able to detect which functions/attributes they have defined. If you do not detect the required method, you can then print a pretty message to the students that you could not find that method instead of having the autograder crash. We recommend that you have error handlers around each of your autograder tests so buggy student code will not cause your autograder to crash.</p> <p>If students did not define a required function, one way to handle it for other tests is to mock/stub that function in other tests, or you can replace the student's</p>

	implementation with your own function for tests that aren't testing that function.
What about compile errors?	We recommend that you create a step in your autograder which will detect if the student submitted the required files and then compile them. If they failed to supply the needed files or you could not compile their code, your autograder should catch that and then return results indicating that it was unable to find the required files or compile their code. In such a case you may want to skip running tests.
What about segfaults or other runtime problems.	<p>We recommend that you structure your autograder in a way which has error handling around each test case so a buggy submission does not cause your entire autograder to fail. How you catch these errors depends on what language your autograder is written in.</p> <p>For languages where segfaults/memory safety may be an issue, it can be helpful to run each test as a separate subprocess. For higher-level languages, testing frameworks will often catch runtime exceptions and report them as test failures.</p>
Are there any gotchas he could mention about using Selenium during the Q&A?	<p>[answered live]</p> <p>At a high level, run Selenium and similar automation tools in headless mode since there is no GUI in an autograder environment. You should also ensure that the autograder does not need any manual input to proceed, it should be able to run completely unattended.</p>
The debug cycle for setting up an autograder is very slow. It would be really helpful if one could use docker locally to debug it, and push to gradescope using docker, with only the relevant parts of the	We recommend that you use GitHub or some other version control repository to push your autograder code to. You can then have the setup of the autograder pull the code and run the setup. Then, your

<p>gradescope image updated. It seems that every time one uploads a new zip file, the entire docker image is rebuilt from scratch at gradescope's end. It should be possible to avoid this constant rebuilding, using the features of docker. I think this is possible, but it would be nice if it was easy.</p>	<p>run_autograder (the file executed on a submission) can pull the latest changes of your autograder and run it. This way, to perform an update which doesn't require a rebuild, you would just have to push your changes to GitHub and then each run will pull the latest changes. We have an example of this here: https://gradescope-autograders.readthedocs.io/en/latest/git_pull/ and here is what I had written to do this: https://github.com/61c-teach/GradescopeDeploymentAutograder</p>
<p>Where would the autograder get a timestamp to determine whether the assignment was late?</p>	<p>The autograder gets a submission_metadata.yml file which contains some information about the submission such as the submission time. Your autograder can then use that data to adjust the scores it gives to students.</p>
<p>Regarding the compile error / runtime error answer, what does the Autograder environment do if this happens to a poorly written autograder?</p>	<p>If your autograder does not return a results.json file, we will output the standard out/error of your autograder in the Gradescope interface for you to see with an output message saying that your autograder did not produce results.</p>
<p>How do we contribute to the community resources?</p>	<p>To contribute to the Community Resources page, you can click "Edit on Github" at the top of the Community Resources page to submit a pull request with your autograder: https://gradescope-autograders.readthedocs.io/en/latest/resources/. Or, you can email your autograder example or Github link to our support team at help@gradescope.com.</p>
<p>Where is this documentation the speaker is using?</p>	<p>The gradescope_utils example can be found here: https://github.com/gradescope/gradescope_utils/blob/master/gradescope_utils/autograder_utils/decorators.py</p>

<p>Any security best practices available? As a student's code and the test case code both exist in the same process space, they probably have the ability to access the test scripts and other files that they should not be able to access.</p>	<p>[answered live]</p>
<p>Could you please send a link to the diff-based autograder example? Thanks.</p>	<p>Here is our documentation about the diff-based autograder (it has a link to example code as well): https://gradescope-autograders.readthedocs.io/en/latest/diff_general/</p>
<p>I can see how this could solve some issues, at the expense of slowing down an already slow grading process. But as you probably know, some updates, say, installing additional dependencies, can't be efficiently handled that way. I think it should be possible to use the features of docker to make this much less painful.</p>	<p>We found that using GitHub actually was super fast since it pulls the code almost instantly. If you have to perform updates to the setup, sadly our platform does not support updating a built image. Instead we rebuild the image from scratch. If you do not want to use the zip file upload, you can always build a manual docker container while you test and debug the autograder. You can then always upload your finished zip file later on if you do not want to continue to use a manual docker configuration. https://gradescope-autograders.readthedocs.io/en/latest/manual_docker/</p>

Getting Support

- help@gradescope.com
- <https://help.gradescope.com/>
- [Code Autograder Documentation](#)
- [Community Resources page](#)
- [Programming Assignments help center article and video](#)