DLUnit: A Unit Testing Framework for Simulated Digital Logic Circuits

Zachary Kurmas kurmasz@gvsu.edu Grand Valley State University Allendale, MI, USA

ABSTRACT

We present DLUnit, a unit test framework for simulated digital logic circuits. DLUnitprovides a quick and easy way to verify that a simulated digital logic circuit works as expected. It is based on JUnit in order to present a low learning curve to students. This testing ability benefits both instructors and students. It benefits instructors by automating the process of verifying the correctness of student submissions, thereby saving time. It benefits students by providing an easy (and familiar) way to test their assignments before submitting them, thereby helping them correct mistakes and misconceptions in a more timely manner. Furthermore, DLUnit can be used to provide students additional testing experience; therefore, it can serve as one component of a curriculum-wide emphasis on testing. DLUnit currently supports JLS and Logisim; but, is designed to be readily extended to work with other digital logic simulators.

CCS CONCEPTS

• **Applied computing** → *Computer-managed instruction*.

KEYWORDS

Digital logic, Testing, JUnit, JLS, Logisim

1 INTRODUCTION

We introduce DLUnit an improved unit testing framework for simulated digital logic circuits. DLUnit replaces JLSCircuitTester [7]. Both DLUnit and its predecessor provide (1) significant time savings for instructors when grading student submissions, and (2) a means for students to verify the correctness of their assignments before submitting them; however, DLUnit imporves upon JLSCircuitTe by using a JUnit-based interface that students find much easier to use, thereby reducing their cognitive load. Our students used JUnit in both CS 1 and CS 2 up until a recent switch from Java to Python. In contrast, JLSCircuitTester uses a custom syntax that most students found confusing and difficult to use. Although we didn't conduct a rigorous, formal evaluation, we expect that students will benefit from this reduced cognitive load (as compared to using JLSCircuitTester) because they will be able to redirect

FIE '23, October 18-21, 2023, College Station, TX

© 2023 Association for Computing Machinery.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...\$15.00

https://doi.org/10.1145/1122445.1122456

that time and mental energy elsewhere (whether it be to other topics in this course, other courses, or toward relaxation and better mental health).

1.1 Context

Automated grading frameworks, such as WebCAT [4] and Marmoset [13], have proven effective for both (1) helping instructors provide better, more timely feedback, and (2) helping students learn how to better test their code. In addition, support for automated grading is essential when using techniques like "second-chance" exams [6] and mastery grading [14] in large sections.¹ Over the past several years, these ideas have been incorporated into many tools including zyBooks, Courseara, Gradescope [12], Prairie Learn [9], Codio, TurningsCraft, and Vocareum. DLUnit can serve as the foundation for extensions that will allow these tools to determine the correctness of simulated digital logic circuits.

Second, the educational community is increasingly recognizing that testing is an integral part of any project. We should reinforce the importance of testing by requiring students to submit tests and/or a test plan for all computing projects. DLUnit supports this goal by providing a familiar, reasonably straightforward means for students to test their digital logic circuits as thoroughly as they would test their Java code.

Third, many students find testing frustrating and difficult. As with many skills, there is simply no substitute for practice. Therefore, we should expose our students to testing as much as is practical. DLUnit contributes to this "across-the-curriculum" exposure by providing a time-efficient means for students to gain additional testing experience.

Finally, our experience suggests that students learn better when they have a means of easily verifying the correctness of a project before submitting it. Students often "write off" points deducted from a project and never go back to learn the misunderstood topic. In contrast, most students hesitate to submit a project with known bugs. Therefore, they take the time to fix the bugs, and often learn the associated material in the process. We first noticed this trend when we introduced our original circuit testing tool [7]. Later, we noticed (but did not formally evaluate) similar improvements in (1) our CS 2 project submissions when we introduced unit testing, and (2) when we used MUnit to add unit testing to our assembly language assignments [8]. DLUnit provides the means for students to verify the correctness of their simulated digitial logic circuits before submitting them.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

^{1e}Second-chance" exams provide students two attempts at each exam, potentially doubling the exam grading workload. Similarly, students in courses that use mastery grading revise and resubmit each assignment until they have demonstrated mastery over the its learning objectives. In most cases, they are also allowed to attempt exam questions multiple times if necessary.

We could certainly achieve similar outcomes with respect to digital logic learning objectives by simply requiring students to resubmit assignments until they pass a suite of instructor-written tests; however, we believe that the additional experience writing tests is of particular importance to our students.

The vast majority of our Computer Science graduates begin their careers as software developers. Although computer organization is an important component of the Computer Science curriculum, our students' testing abilities have a much more immediate effect on their success as they begin their careers. Therefore, it is beneficial to include testing in as many courses as practical; but, in order to keep the course's workload reasonable, it is important that we provide our students tools like DLUnit to keep the workload resulting from the additional testing expectations as small as practical.

Our Computer Science curriculum includes only a single hardwarefocused course. This computer organization course covers basic digital logic, processor design, pipelining, and cache memory. Very few of our students work with hardware design after completing this course. Therefore, we have students use "drag-and-drop" simulators rather than requiring them to spend time and effort learning a hardware description language (HDL) and related tools which few will ever use again. Likewise, we based DLUnit on JUnit (with which our students have previous experience) as opposed to basing our design on (or even using) a professional tool such as a Quartus or Icarus.

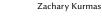
To the best of our knowledge, DLUnit is the only tool that uses a familiar unit test-based syntax for testing simulated digital logic circuits. Many digital logic simulators (JLS, LogiSim, TkGate, Digital [1], etc.) provide a command-line interface that allows users to initiate the simulation of a circut from the command line or a script. Instructors can use this feature to construct an ad-hoc testing system; but, such setups tend to be difficult for students to understand and use.

1.2 Contributions

Our primary goal in replacing JLSCircuitTester with DLUnit was to provide a better user experience for our students (i.e., students who are already familiar with JUnit, are taking only one hardware-related course, and are unlikely to directly design/program hardware after graduation).

JLSCircuitTester supported only the JLS digital logic simulator. In contrast, DLUnit supports both JLS and the more popular Logisim [3, 11]. (Although the original Logisim is no longer under development, several groups have forked the original tool and continued development. It appears that the most popular of these forks is Logisim Evolution [2].) Unlike JLSCircuitTester, DLUnit is designed to be extended to support additional simulators in the future. (The fact that DLUnit already supports two simulators means that the abstraction mechanism has already been used and tested.)

The remainder of this paper provides details of DLUnit and discusses our experiences using the tool.



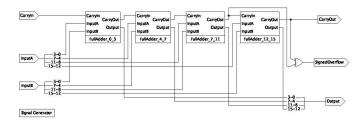


Figure 1: Sample circuit for testing

```
import static <witheld for anonymous review>
// Simple tests to help students get started
public class UAdderTest {
@Test
public void zero_zero_false() {
   setPinUnsigned("InputA", 0);
   setPinUnsigned("InputB", 0);
  setPin("CarryIn", false);
  run():
  assertEquals("Output", 0, readPinUnsigned("Output"));
  assertFalse("Carry Out", readPin("CarryOut"));
}
@Test
public void zero_one_false() {
  setPinUnsigned("InputA", 0);
   setPinUnsigned("InputB", 1);
  setPin("CarryIn", false);
  run();
  assertEquals("Output", 1, readPinUnsigned("Output"));
  assertFalse("Carry Out", readPin("CarryOut"));
1
// Place similar tests here....
```

Figure 2: Simple tests for adder

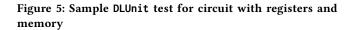
Figure 3: Running DLUnit

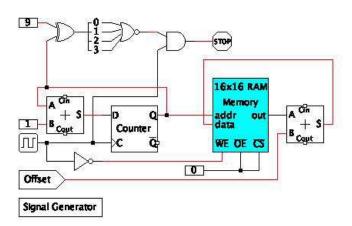
2 THE JAVA INTERFACE

DLUnit: A Unit Testing Framework for Simulated Digital Logic Circuits FIE '23, October 18-21, 2023, College Station, TX

```
import static <witheld for anonymous review>
// Use helper methods quickly write many tests
public class UAdderTest {
 // Helper method to calculate expected outputs
 // (Makes it easier to add tests.)
 protected void verify(long a, long b,
                        boolean carryIn) {
   long carryInAsInt = (carryIn ? 1 : 0);
   long exp = a + b + carryInAsInt;
   boolean expectedOverflow = exp > 65535;
   setPinUnsigned("InputA", a);
    setPinUnsigned("InputB", b);
   setPin("CarryIn", carryIn);
   run();
   // custom message to make it clearer
   // which test failed
   String message = "of " + a + " + " + b +
        with " + (carryIn ? "a " : " no ") +
       " carry in";
   // Output "wraps around" if overflow
   Assert.assertEquals("Output " + message,
       (exp % 65536), readPinUnsigned("Output"));
   Assert.assertEquals("CarryOut " + message,
      expectedOverflow, readPin("CarryOut"));
  }
  // With the helper method, we can focus on choosing
  // inputs without having to calculate outputs.
 @Test
 public void overflow1() {
   verify(65535, 1, false);
 }
 @Test
 public void overflow2() {
    verify(65535, 0, true);
  }
// Alternatively, we can automatically generate tests
// from a list of numbers. (Yes, I know this isn't
// the conventional way of generating tests; but,
// it's useful when grading student submissions.)
 @Test
 public void testMany() {
   long testIntegers[] = {0, 1, ..., 65535};
    int count = 0;
   for (long a : testIntegers) {
     for (long b : testIntegers) {
       verify(a, b, false);
       verify(a, b, true);
       count += 2;
     }
   }
 } // end testAll
```

Figure 4: More tests for a half adder







DLUnit is a Java interface to digital logic simulators (currently JLS and Logisim). It helps users write JUnit tests for their simulated digital logic circuits. In particular, DLUnit provides methods that

- configure the initial state of the circuit (i.e., place values in registers and/or RAM, if applicable),
- provide values for any input pins,
- instruct the chosen simulator to simulate the circuit under test, and
- query the final state of the circuit (registers, RAM, output pins, etc.)

Programmers then use JUnit's Assert methods to verify that the circuit's final state meets expectations.

Figure 1 shows an example circuit, a 16-bit adder. Figures 2 and 4 show example test suites. Users run a test suite from the command line by specifying the name of the file containing the circuit under test and the name of the class file containing the DLUnit tests (e.g., java -jar DLUnit.jar halfAdder.jls HalfAdderTest.class).

The setPin* methods specify the value for each input pin. The run method in each test starts the simulation of the circuit. The readPin* methods return the final values on the output pins. These values are then tested using JUnit's assert mechanism. Any errors are reported using JUnit's built-in error reporting functionality. Figure 3 shows (1) how to compile the tests, (2) the output from a passing test, and (3) the output when some test fails.

We test student submissions thoroughly. Typical software testing techniques, such as focusing on the corner cases, won't find hardware mistakes like a random missing or improperly connected wire. Because DLUnit tests are just Java/JUnit, instructors can easily write exhaustive tests. Our tests don't always follow typical unit tests patterns; but, they are effective for finding the types of mistakes our students tend to make.

Figure 4 shows part of one such test. The method verify tests the circuit using the parameters supplied (two integers and a Boolean representing the carry). It computes the expected output, simulates the circuit, then verifies that the circuit produced the expected output values.

DLUnit also provides similar methods to view and set registers and memory. Figure 5 shows a test for a circuit containing registers and memory (shown in Figure 6).²

2.1 Testing CPUs

In addition to simple combinatorial and sequential circuits, we also have our students implement the single-cycle datapath used as a running example in both the Patterson and Hennessy textbook and the Harris and Harris textbook [5, 10]. At a high level, using DLUnit to test a CPU entails writing a test class that (1) loads machine code into the circuit's memory, (2) simulates the circuit, then (3) confirms that registers and memory contain the expected values afterward. That class, SingleCycleCPUTest, is built into DLUnit. To further simplify the testing process, this CPU test class takes a MIPS assembly language file as input, then uses the MARS MIPS simulator [15] to (1) generate the machine code that is loaded into the circuit under test, and (2) execute the assembly code to determine the expected final register and memory values. By using MARS, the user need only provide the assembly code. She need not generate the machine code, nor must she determine the expected output (final register and memory state) "by hand".

Figure 7 shows the command for testing a CPU using DLUnit. This command is longer than we would like; but, the alternatives are (1) to provide separate jar files for circuit testing and CPU testing, or (2) provide a separate script to launch the CPU tester. Our experiences with JLSCircuitTester led us to conclude that the long command was preferable to expecting students to manage separate files/programs. (Instructors can still prepare scripts and shortcuts; students who aren't comfortable with such techniques still have the option of falling back on the straightforward, but long, command.)

We also provide a similar test class for testing the multi-cycle CPU discussed in the Harris and Harris text (and early editions of the Patterson and Hennessy text). Rather than having studens implement the finite-state-machine-based control unit discussed in the text, we have our students program a simple microcode.

3 OUR TYPICAL WORKFLOW

Although we do not formally present Test Driven Development (TDD) in this course, we strongly encourage students to follow a TDD-like workflow: We provide a few sample tests to help them learn the DLUnit API, ³ then encourage them to write a complete set of tests before implementing their circuit. Students submit their circuit when they are confident they have found and fixed all bugs. Students who submit buggy circuits are asked to write one or more failing tests (i.e., figure out why their test cases failed to detect their bugs), fix the bugs (i.e., get those failing cases to pass), and re-submit.

Our motivation for using this TDD-like workflow is to give students practice in devising test cases (thinking through all the requirements, corner cases, unusual situations, etc.). When a student's circuit passes all of her test cases, but fails one or more of the instructor's tests, that means the student's set of tests is incomplete. This situation provides an opportunity for the student to find the missing test case(s) and reflect on her process for writing tests, which will hopefully help her write a better set of tests for the next assignment.

We use GitHub Classroom⁴ and GitHub Actions⁵ to provide immediate feedback to students on the correctness of their circuits. GitHub Classroom is a service provided by GitHub that crates a separate git repository for each student/team. These repositories are initialized with instructions and, in most cases, some "starter circuits". (These "starter circuits" usually contain the input and output pins, so students are less likely to have an incorrect interface. Sometimes the repository also contains lower-level components that would be unnecessarily tedious to make the students build themselves – like the CPUs register file.)

These repositories also contain a GitHub Action configuration. When launched, the GitHub Action downloads the instructor's tests and uses DLUnit to run the tests against the circuit. These tests do not have descriptive names; so, although students can see which tests failed, they don't know what that test does. (The tests are stored on a web server with a "whitelist" that includes only the IP address of the virtual machines used by GitHub Actions. Thus, students can not easily access the tests. It is not impossible for the students to access the tests; but, the changes to the repo that would allow access should be detectable.)

Notice that the TDD-like workflow we use is optional. Instructors can simply provide a complete set of tests for students to use, then use a GitHub Action to verify that the submitted circuits pass the tests.

4 OUR EXPERIENCE

We began using DLUnit in our Computer Organization course several semesters ago. Our issues with JLSCircuitTester disappeared almost immediately. Only one or two students per semester have trouble setting up the tool; and, only a few students have trouble understanding its output. We still have several students a year ask for help writing tests. In contrast, when using JLSCircuitTester, well over half of the students would request help writing tests.

²Note: Logisim does not name its memory elements; therefore, (1) the name of a memory element is ignored when testing a Logisim circuit, and (2) DLUnit does not support the testing of Logisim circuits with more than one memory element per subcircuit (we have not found a good way to distinguish among them).

 $^{^{3}} https://kurmasgvsu.github.io/Software/DLUnit/dist/doc/index.html \\$

⁴https://classroom.github.com

⁵https://github.com/features/actions

DLUnit: A Unit Testing Framework for Simulated Digital Logic Circuits FIE '23, October 18-21, 2023, College Station, TX

java -jar DLUnit.jarmyCPU.jls builtin.SingleCycleCPUTest -param test1.a

Figure 7: command for testing a CPU using DLUnit

When using DLUnit, we are able to verify the correctness of the entire class's submitted circuits in a matter of minutes. We choose to review the students' circuits and comment on their design before assigning grades; but, in theory, instructors could use DLUnit to completely automate the grading. We use GitHub Classroom and GitHub Actions to completely automate the correctness checking of assignments: When students push changes to their main git branch, GitHub automatically runs DLUnit and reports any failures to the student. When GitHub reports a passing build, instructors then review circuits by hand in order to comment on the overall design.

Relatively few students need to submit an assignment more than once. During Fall 2019, we assigned three major projects. These projects were completed by 30 teams of two. Of the 90 submitted projects, 60% passed our tests after the initial submission. Another 27% passed after the second submission. 11% of submissions required 3 attempts; and 2% required 4. No team required more than four submissions to find and fix all bugs.

Needless to say, students did not always enjoy writing tests; but, survey results indicated that most recognized the value of the process. Students used a 5-point Likert scale to indicate their agreement with the following two statements:

- Using DLUnit improved my confidence in the correctness of my assembly code.
- Writing unit tests in this class was "busy work" that had little value.

As shown in Table 1: 61% of the 41 students agreed (responded with a 4 or 5) that using DLUnit improved their confidence in the correctness of their code. 41% of students disagreed that writing unit tests was "busy work". (25% agreed; and an additional 34% responded with a "neutral" 3.) The percentage of students explicitly disagreeing that writing unit tests is busy work was smaller than we expected, but not out of line with the "word on the street": We know from informal discussions with students in our program (not just in this course) that most don't enjoy writing tests. One student said "*It's kind of busy work, but works well for our testing*." We suspect that many students have a similar attitude.

The free-response section of the evaluation asked students "What suggestions do you have for improving DLUnit, and/or improving the way it is used in this course?"

The majority of responses fell into one of three categories: Six students offered complements. For example:

- DLUnit was super useful for validating my circuit solutions. After reading example test cases I built fairly thorough tests. Once I caught an error it let me find/fix my own problems.
- I thought it was great and had no issues

Three were not happy that DLUnit is a command-line tool:

• It certainly works, but I would rather have a simpler way of executing tests (not a fan of command line) ...

Four requested more directions on how and what to test:

- I feel that in some labs, more tests should have been provided to the student.
- I felt that writing the JUnit tests was pointless when we didn't get specifications on cases to test ...
- More suggestions on how one writes effective tests? It was sometimes hard for me to think of tests ...
- We need clearer instructions for writing the test. Some people have little experience writing unit tests.

We find the last set of comments somewhat disheartening because we have recently increased our focus on testing in our prerequisite CS 1 and CS 2 courses (partially in response to [7] and [?]). However, the difficulty students have designing a complete set of tests does reinforce our position that testing should be discussed and required in as many computing courses as practical.

We did not quantitatively compare DLUnit and JLSCircuitTester. Because it is based on JUnit, which is used in our CS 1 and CS 2 courses, DLUnit is clearly a better fit for our curriculum. In addition, our observations strongly suggest that our students are happier (and less confused) when using DLUnit. At this point, collecting quantitative data would require us to withhold DLUnit from a cohort of students, which we strongly suspect would not be in their best interest.

5 FUTURE WORK

Pedagogy: DLUnit helps instructors reinforce software testing in a course that does not traditionally contain a formal testing component. The effects of adding a testing component to a single course (e.g., Computer Organization) is probably small; but, is an important component of a more comprehensive "testing across the curriculum" approach that adds testing components to as many computing courses as practical. A comprehensive study of the effects of testing across the curriculum is beyond the scope of this Innovative Practice Report; but, we look forward to such a study in the future.

Tool Improvements: Currently, DLUnit only assesses a circuit's final state. For combinatorial circuits, this is when there are no further gate changes. For sequential circuits, DLUnit relies on a "halt" signal.⁶ We are considering adding an option that would allow users to check a sequential circuit's state after every clock tick.

Additional CPUs: Over the past several years, publishers have shifted away from MIPS and toward RISC-V and ARM. Upon finding a MARS-like tool for either RISC-V or ARM, we plan to build similar CPU-testing classes for them.

6 AVAILABILITY

The software, as well as full documentation and additional sample tests, is available at

http://www.blind.review

⁶JLS includes a "halt" gadget that can be activated. Sequential circuits in Logisim must have an output pin named "halt" that the circuit asserts when the simulation should terminate.

	Strongly				Strongly
Question	Disagree	Disagree	Neutral	Agree	Agree
Using DLUnit improved my confidence in	1	2	2	13	23
the correctness of my assembly code.					
Writing unit tests in this class was "busy	5	12	14	7	3
work" that had little value.					

Table 1: Responses to Likert questions

ACKNOWLEDGMENTS

Witheld for blind review.

REFERENCES

- [1] [n.d.]. Digital. https://github.com/hneemann/Digital.
- [2] [n.d.]. Logisim-evolution.
- [3] Carl Burch. 2002. Logisim: A Graphical System for Logic Circuit Design and Simulation. J. Educ. Resour. Comput. 2, 1 (March 2002), 5-16. https://doi.org/10.1145/545197.545199
- [4] Stephen H. Edwards. 2004. Using Software Testing to Move Students from Trialand-Error to Reflection-in-Action. In SIGCSE '04: Proceedings of the 35th SIGCSE Technical Symposium on Computer Science Education.
- [5] Sarah L. Harris and David Money Harris. 2015. Digital Design and Computer Architecture, ARM Edition. Morgan Kaufmann.
- [6] Geoffrey L. Herman, Zhouxiang Cai, Timothy Bretl, Craig Zilles, and Matthew West. 2020. Comparison of Grade Replacement and Weighted Averages for Second-Chance Exams. In Proceedings of the 2020 ACM Conference on International Computing Education Research (Virtual Event, New Zealand) (ICER '20). Association for Computing Machinery, New York, NY, USA, 56–66. https://doi.org/10.1145/3372782.3406260
- [7] Zachary Kurmas. 2008. Improving student performance using automated testing of simulated digital logic circuits. In *ITiCSE '08: Proceedings of the 13th annual conference on Innovation and technology in computer science education* (Madrid, Spain). ACM, New York, NY, USA, 265–270. https://doi.org/10.1145/1384271.1384342
- [8] Zachary Kurmas. 2017. MIPSUnit: A Unit Testing Framework for MIPS Assembly. In Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education (Seattle, Washington, USA) (SIGCSE '17). Association for Computing Machinery, New York, NY, USA, 351–355. https://doi.org/10.1145/3017680.3017747
- [9] Geoffrey L. Herman Matthew West and Craig Zilles. [n.d.].

- [10] David A. Patterson and John L. Hennessy. 2013. Computer Organization and Design: The Hardware/Software Interface, Fifth Edition. Morgan Kaufmann.
- [11] David A. Poplawski. 2007. A pedagogically targeted logic design and simulation tool. In WCAE '07: Proceedings of the 2007 workshop on Computer architecture education (San Diego, California). ACM, 1–7. https://doi.org/10.1145/1275633.1275635
- [12] Arjun Singh, Sergey Karayev, Kevin Gutowski, and Pieter Abbeel. 2017. Gradescope: A Fast, Flexible, and Fair System for Scalable Assessment of Handwritten Work (L@S '17). Association for Computing Machinery, New York, NY, USA, 81–88. https://doi.org/10.1145/3051457.3051466
- [13] Jaime Spacco and William Pugh. 2006. Helping Students Appreciate Testdriven Development (TDD). In Companion to the 21st ACM SIGPLAN Symposium on Object-oriented Programming Systems, Languages, and Applications (Portland, Oregon, USA) (OOPSLA '06). ACM, New York, NY, USA, 907–913. https://doi.org/10.1145/1176617.1176743
- [14] Ellen Spertus and Zachary Kurmas. 2021. Mastery-Based Learning in Undergraduate Computer Architecture. In ACM/IEEE Workshop on Computer Architecture Education, WCAE 2021, Raleigh, NC, USA, June 17, 2021. IEEE, 1–7. https://doi.org/10.1109/WCAE53984.2021.9707147
- [15] Kenneth Vollmar and Pete Sanderson. 2006. MARS: An Education-Oriented MIPS Assembly Language Simulator. In SIGCSE '06: Proceedings of the 37th SIGCSE Technical Symposium on Computer Science Education (Houston, Texas, USA). ACM, 239–243. https://doi.org/10.1145/1121341.1121415