(Board 52/Work in Progress) Datastorm: Using Data-Driven Competition to Improve Student Engagement in Computer Science

Dr. Ankunda Kiremire, Louisiana Tech University

Dr. Ankunda Kiremire is a lecturer of Computer Science and Cyber Engineering at Louisiana Tech University and serves as the program chair for its Computer Science Department. His research interests include Computer Science Education, Cyber Security, and Data Science.

Kevin A Cherry, Louisiana Tech University

WIP Datastorm - Using Data Driven Competition to Improve Student Engagement in Computer Science

Abstract

Studies have shown that Computer Science and other computing-based courses with traditional pedagogy suffer from low student engagement, low student retention, and a lack of real time evaluation tools for the instructors of these subjects. In this paper, the authors present an innovative solution to all three problems.

"Datastorm" is a series of data-driven competitions among teams of undergraduate students in which the student teams are tasked with imagining creative solutions to real data structure problems. The student teams must build and test their solutions using data from culturally relevant sources in a collaborative and time-restricted environment.

Pilot results show that the In Class Datastorm events increase student engagement, improve student retention, and provide a viable and alternative evaluation tool for computing instructors.

This paper and the work presented therein is relevant to instructors of computer science courses at the undergraduate and high school levels who are looking for ways to increase student engagement from current Computer Science students as well as interest in Computer Science as a potential education path for students outside of the field.

Future work includes expanding its implementation to more class sections, collecting data on these results, as well as identifying and cultivating relationships with local businesses and organizations to increase the quantity and scope of culturally relevant data that can be incorporated into the Datastorm challenges. We also plan to host annual full-day Datastorm events, which should provide visibility and outreach opportunities to other undergraduate students at our institution as well as highlight the relevance of the Computer Science program to the general public.

Introduction

Computer Science and computing based majors in general suffer from a variety of issues at the university level.

One of those issues is high drop out rates. The level of attrition in Computer Science is reported to be between 9.8% [1] and 28% [2]. This represents both a direct loss in terms of students not completing the major as well as an indirect loss in terms of students not encouraged to pursue it because of a perceived difficulty given its high withdrawal rates.

Figure 1 shows that we have a similar problem at Louisiana Tech University (LTU). The figure shows the total number of students who have taken the first 5 classes in our Computer Science curriculum over the last 5 years. While a direct comparison of attrition rates is difficult to determine, the figure shows that the number of students who take the Computer Science classes drops as students progress through the curriculum.

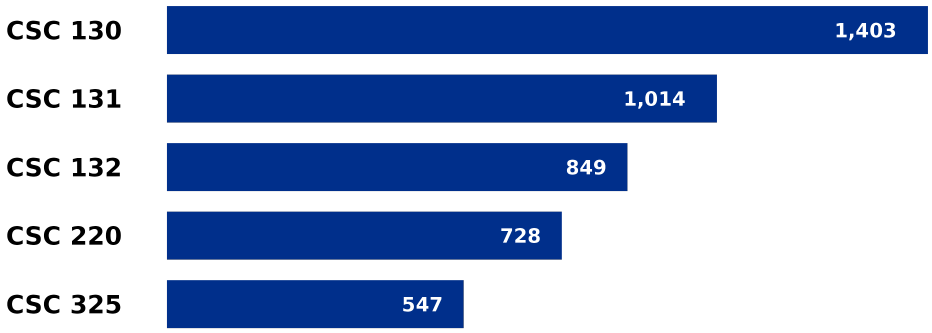


Figure 1: The size of LTU's first 5 Computer Science classes shows a significant drop in numbers as students progress through the curriculum.

Another issue with Computer Science and computing based majors in general is the antiquated nature in which the material is presented to the students. This problem manifests in a couple of ways. Firstly, the material taught in the classes is rarely culturally relevant. That is to say the material is typically taught in a manner that is removed from the social and cultural interests that make students passionate about their fields of study [3]. Secondly, Computer Science classes do not frequently offer a lot of opportunities for collaboration and real world application particularly in their freshmen and sophomore classes. As a result, some students become disillusioned about the nature of Computer Science work, assuming that the only kind of work that they will do will involve them writing code by themselves in a room.

The final issue with Computer Science and computing based majors in general is that the classes only typically employ a limited variety of student evaluation methods. While exams and programming assignments have value in showing an instructor and the student the level to which a concept from class has been grasped, there is always space to supplement the traditional educational tools [4]. Novel evaluation and educational tools would help narrow the gap between the time a concept is taught, and the time when the instructor and student are aware of how well that concept has been grasped. This would allow both instructors and students to take early corrective measures in or outside the classroom.

Datastorm is an innovative approach that allows us to tackle all three issues at LTU. The gaming component of Datastorm should improve student engagement in the classes and the major in general [5], [6]. Student engagement in the classes, and interest in the curriculum is also improved by the cultural relevance of the data employed in Datastorm events [3]. The Datastorm events also allow for students to engage in both collaborative and controlled competitive environments which allow them to hone the soft skills they'll need to navigate the work environment. The Datastorm system also provides a much needed supplementary educational evaluation tool for instructors of Computing based majors [4].

Methods

Datastorm involves pitting small teams of students in public competitions against each other to solve custom designed programming challenges in real time. These challenges involve the use of large data sets obtained from selected local businesses and vary in length, scope, and intensity. The challenges test the teams' ability to collaborate, design, and implement programming solutions to varying problems with scores given based on speed, accuracy, use of computational resources, and scalability of their solutions.

Data Sources

Typically, Computer Science courses are taught using hypothetical data. As an example, an instructor of a Computer Science Data Structures class is more likely to use random integers to demonstrate the merits/demerits of a specific data structure e.g. arrays in sorting. This data is easy to create and adjust and has value when an instructor wants students to pay more attention to the concept being covered than the data that that concept is being applied to. It is also very easy to re-use this data from year to year or even across different classes. However, there is value is using culturally relevant data – data from a real source in a field that students would be passionate about. As an example, a list of sales amounts from a local bakery in the same town as the college in question could also be used to cover the same material i.e. the merits/demerits of arrays in sorting. Culturally relevant data provides the students with a clear and recognized benefactor of their design decisions and shows them how the concepts that they are studying and applying could be applied to other interests and passions they have. Ideally, this would counteract the disillusionment they might feel from difficult concepts in those classes or difficult classes in the major/field.

System Design

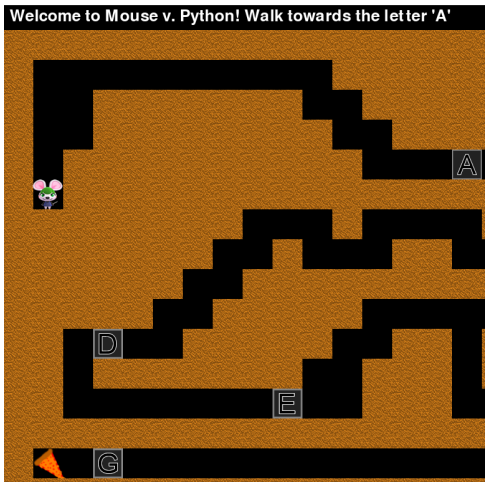
Datastorm is dependent on hardware and software components that work hand in hand to accomplish its goals.

On the hardware side, an institution requires either a local server or networked access to the server of another institution that is already implementing Datastorm events. The student teams also require either a computer or a mobile phone with internet access in order to access the server. While the system works with just a single computer/phone per team, multiple computers per team allow the teams to practice their collaboration and delegation tasks more effectively.

On the software side, the institution with the server requires the files necessary to host a Datastorm event. These include team score webpages where all the team scores are displayed and updated in real time, and Challenge webpages where student teams can access challenge files during the course of a Datastorm Challenge or event.

Both the team score and Challenge webpages are designed in a gamified manner to make user interaction with them more engaging. Figure 2 shows an example of both pages where the team is represented by a mouse character, and their progress through the challenges is represented by the mouse progressing through a maze. The faster and more successful the team is in solving their challenges, the faster the mouse representing their team progresses through the maze.

(a) The Team Score page shows the student teams' scores in real time



(b) The Challenge Page provides files and instructions while showing students' progress

Figure 2: Sample web pages hosted on server.

For the student portion of the software side, each computer should either allow for the running and writing of java files, or allow the user enough administrator privileges to install a java development kit (jdk) themselves.

In Class Challenges

Datastorm In Class Challenges are short challenges designed to span a single class period. These challenges typically cover a single topic. These events allow the instructors to evaluate student grasp of small but central concepts as early as possible with little prep requirement on both the student and instructor sides. For younger or less technically savvy students, there are In Class Challenges that will walk both students and instructors through the process of installation and set up of the entire system such that later In Class Challenges can focus on content coverage.

On top of allowing the faculty to identify any knowledge gaps in central concepts of the class, the brevity of these challenges allow for the faculty to identify and correct any set up or team management issues. They also allow the instructor to be as helpful as necessary especially during the first few In Class Challenges as students get used to the system and how to work with each other in that collaborative and yet competitive environment.

Datastorm Event

The final Datastorm event, which has not yet happened, will be a multi-hour event in which the student teams are presented with multiple challenges that span the entire class curriculum. The amount of challenges that the students will be presented with dictates that student teams will have to employ some task delegation in order to be successful. The challenges the students will face will include derivatives of the In Class Challenges as well as multiple novel comprehensive challenges. Team scores are displayed on a central screen for the duration of the event to show how each team is faring.

This event will also be open to the public which should allow students from other parts of the university to interact with the student teams as they tackle their challenges. This kind of visibility, as well as the competitive nature of the event should yield excitement and intrigue about the class specifically and the major in general for other prospective students.

To facilitate the competition and its visibility, participants will be provided with food, snacks, branded tshirts, and branded swag. The swag can also be given to teams that successfully complete challenges or win the entire event.

Student Surveys

To facilitate student feedback, we have collected voluntary survey results from 58 students. Part of this survey asks students to select which mode of teaching suits them between four different options that are used in their class i.e. traditional: listening to lectures and doing take home assignments, interactive: in-class coding or challenges with immediate feedback, visual: animations and visualizations showing coding concepts, and collaborative: team projects.

Results

The authors have designed eight sets of challenges for students learning either introductory programming or data structures in Python or Java. These challenge sets have been tested on college freshman and sophomore Computer Science and Cyber Engineering students at LTU over the last three years.

We have created and tested the system that allows for these challenges to be remotely launched over the last five years. This system allows students and instructors to participate remotely in In Class Challenges with only a single laptop/computer with a Java installation per student team. The system provides real-time feedback to both instructors and students about the efficacy and correctness of their solutions, allowing instructors to identify knowledge gaps in their classrooms in real time with very little set up overhead.

The feedback we have received from our student body has been overwhelmingly positive. A former student who was involved in a pilot version of the In Class Datastorm Challenges reported "*I thought the challenges were quite fun. [It] felt like we were working as an actual development team might [do] in the real world. We split up duties and helped each other when we finished our jobs. I thought it was fun working together and [it] felt great to get the challenges completed*."

Figure 3 also shows that out of the 58 students who have been polled, 88% expressed a preference for the interactive student experience captured by the Datastorm challenges. This kind of feedback has shown us that Datastorm is satisfying the student engagement need it was created to satisfy.

The feedback we have received from the faculty who have piloted the In Class Challenges has also been positive. They reported that the In Class Challenges allow them to "get a feel for where the students in the class are". The only administrative cost they mentioned is that they had to set aside full class periods for In Class Challenges that would have been used to cover more class

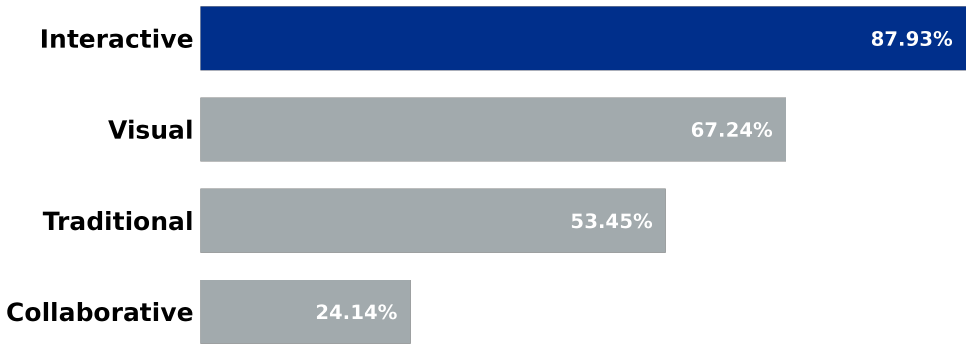


Figure 3: The vast majority of polled students expressed a preference for interactive challenges as a style of teaching/learning.

material. This feedback showed us that the In Class Challenges are a low cost resource for real time student evaluation.

Future Work

The authors plan on implementing the In Class Datastorm challenges across all sections of our program's sophomore Data Structures class initially, and then all our freshmen classes eventually.

We also plan on hosting our first day long Datastorm event in the near future. Our institution has successfully held a similar event called Cyberstorm [7] at least annually over the last 14 years. Cyberstorm has shown great success in increasing the visibility of both our institution's Cyber Engineering program as well as the Cybersecurity field of our Computer Science program. It has also served to increase student and community engagement in the field, and encourage students to pursue careers in these areas. We believe that the day long Datastorm event would have an even bigger impact because it would affect even more students at an earlier stage of their majors.

The authors have been able to identify and begin conversations with two local data sources that our student body is passionate about. The first of these is a female owned local bakery/cafe whose sale records can easily be incorporated into the challenges. The second is a K-12 mentoring program whose membership data can be anonymized and included in some challenges. The authors plan to identify and create partnerships with even more businesses and companies to broaden their data sources.

We are also in the process of brainstorming solutions to reduce the amount of class time that the Datastorm challenges take up that could be used to cover more material.

The authors are seeking feedback from the research community before taking the next steps outlined in this section.

References

[1] K. Flinders, "Computer science undergraduates most likely to drop out," August 2019, [Online; posted 1-August-2019]. [Online]. Available: https://www.computerweekly.com/news/252467745/Computer-science-undergraduates-most-likely-to-drop-out

[2] E. O'Brien, "Why are computer science drop-out rates so high?" February 2016, [Online; posted 21-February-2016]. [Online]. Available: http://trinitynews.ie/2016/02/why-are-computer-science-drop-out-rates-so-high/

[3] N. (IUSE:CUE), "Student-centered computing: Teaching computer science using culturally authentic practices," July 2022, [Online; retrieved 22-July-2022]. [Online]. Available: https://aaas-arise.org/2022/07/05/student-centered-computing-teaching-computer-science-using-culturally-authentic-practices/

[4] D. Hamilton, J. McKechnie, E. Edgerton, and C. Wilson, "Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design," *Journal of Computers in Education*, vol. 8, 07 2020.

[5] V. W. Imed Bouchrika, Nouzha Harrati and G. Wills, "Exploring the impact of gamification on student engagement and involvement with e-learning systems," *Interactive Learning Environments*, vol. 29, no. 8, pp. 1244–1257, 2021. [Online]. Available: https://doi.org/10.1080/10494820.2019.1623267

[6] P. Buckley and E. Doyle, "Gamification and student motivation," *Interactive Learning Environments*, vol. 24, no. 6, pp. 1162–1175, 2016. [Online]. Available: https://doi.org/10.1080/10494820.2014.964263

[7] B. Magazine. (2022) La tech computer science and cyber engineering juniors and seniors will hack away at 2022 cyber storm. [Online]. Available: shorturl.at/ezBNV