

Motivating Introductory Computing with Big Data and Blocks

Austin Cory Bart
Virginia Tech
2202 Kraft Drive
Blacksburg, VA, 24060, USA
acbart@vt.edu

ABSTRACT

My research is focused on motivating and scaffolding introductory computing experiences. Non-traditional computing students in courses such as Computational Thinking pose a unique challenge because of their limited motivation and skill, and the scale of such classes. To overcome these limitations, I propose a novel block-based programming environment leveraging mutual language translation to authentically transition students from the block-based code to a text-based programming language. The system will use sophisticated program analysis to guide students to success and report learner progress to instructors. The tool will synergize with a large collection of beginner-friendly big data sources in order to authentically contextualize the learning experience and thereby improve learner motivation. Data will be gathered on students' motivation and learning progress and instructors' perceptions of the success of the tool in order to answer my hypotheses about the efficacy of the this approach and the tools that are developed to support it.

Categories and Subject Descriptors

K.3.2 [Computers and Education]: Computers and Information Science Education—*Computer Science Education*.

Keywords

Computational Thinking, Big Data Science, Block-based Programming

1. PROGRAM CONTEXT

I am in the third year of a PhD program at Virginia Tech in Computer Science. I have completed all but two of the courses needed for my program. Additionally, I have just completed a 4-course certification in the Learning Sciences. My research is focused on motivating and scaffolding introductory computer science classes. I have established a major on-going research project to integrate Big Data Science into

introductory programming experiences as a motivating context (the CORGIS project). This project and its associated software scaffolding are now deeply integrated into a course on Computational Thinking meant for a diverse crowd of learners. Although preliminary results on the materials have begun to be gathered and published, much work remains to evaluate them and revise the associated technology.

2. CONTEXT AND MOTIVATION

Computational Thinking is increasingly considered a 21st century competency, and plays a growing role within universities' general education curriculum. Although the term itself is historically ill-defined, a given educator can operationalize it into well-defined learning objectives – typically, applying computer science concepts and computational techniques to frame problems and devise solutions in broad areas of study. At Virginia Tech, we have created a new course for non-major CS students on Computational Thinking. These students represent diverse majors: arts, sciences, agriculture, and many others. Non-major learners represent a challenge because they are not confident it will be a useful experience nor about their ability to succeed in the course. Complicating these problems is that enrollment must scale aggressively due to its critical role within general education requirements. The problem here is three-fold: How do we meaningfully motivate these learners, guide them to success, all while scaling the course offering to as large a population as possible? This work will provide new insights in how to successfully motivate and guide large numbers of non-major students through the introductory programming experience and also provide new technology to support the process.

3. BACKGROUND & RELATED WORK

The current dominant approach to motivating introductory students focuses primarily on games and animations as the primary context, such as the popular Scratch/Snap and Media Comp curriculums. Although these materials have proven successful for younger learners, they have been criticized for their inauthenticity for more mature audiences. An extensive intervention [2] failed to convince students that Media Computation was an inherently valuable skill.

Motivation is a complex construct, and there are many theories to explain engagement. The MUSIC Model of Academic Motivation [3] holistically explains motivation within academic contexts. The model identifies 5 components to motivation – eMpowerment (sense of agency), Usefulness (utility to their short-term and long-term interests), Success (self-efficacy), Interest (situationally and to their identity),

and Caring (of the stakeholders in the learning environment towards them). Students are motivated when one or more of these constructs is sufficiently activated. Since each learner and experience draws upon these components differently, it is crucial to appeal to as many as possible. This model identifies the limitation of relying on situationally interesting materials, instead of broadly satisfying learners' needs.

I suggest Big Data Science as an alternative context to motivate learners and convince them of the utility of the material while personalizing the experience and providing agency – every student can find data in their discipline and interests. This emerging context is growing in popularity [1]. Of course, motivation is only half the battle. Most of these students must overcome low self-efficacy and inexperience with the material. I propose using a new Block-based environment as our primary scaffold. To support the transfer from an inauthentic block-based editor to conventional text coding, we leverage the emerging area of Mutual Language Translation [4] to provide 1-1 isomorphism of blocks with the textual equivalent. This mechanism is coupled with state-of-the-art program analysis techniques to provide just-in-time feedback to the learner and feedback to instructors.

4. STATEMENT OF THESIS/PROBLEM

My primary thesis question is: How can we most effectively motivate a large quantity of introductory Computational Thinking students and guide them to success? My hypothesis is that a Big Data Science approach can be scaffolded with a block-based programming environment that supports learners and teachers in a blended learning experience. I will collect data to test this hypothesis by intervening directly in a new Computational Thinking class here at Virginia Tech using technology that I have created. We will gather self-reported motivational data, in-class observations of students, and fine-grained programming exercise progress. As the course materials are refined, I expect meaningful gains across motivation, ability, and the instructors' reported experiences. A number of secondary questions arises from Big Data Science as an introductory context. A major element of the course is the availability and range of datasets. How does the form and affordances of the data shape the learning experience, and how do we rapidly provide beginner-friendly datasets? These questions will be answered through the development and analysis of datasets.

5. RESEARCH GOALS & METHODS

To test my hypotheses, a number of new pieces of technology will need to be created. I am creating a new block-based programming environment with tools for Big Data Science and transitioning seamlessly to a serious programming experience. This environment will provide just-in-time feedback and report meaningful information to instructors. I am creating a massive collection of big data sources meant for introductory students, overcoming a number of difficult technical obstacles. Instructional materials will be created that take advantage of these tools to optimally instruct students. Motivation in the course will be primarily assessed through self-reported surveys based heavily on the MUSIC Model of Academic Motivation Inventory (MMAMI). This information will be supplemented with and compared against relevant measurements of student engagement (e.g., completion of activities, attendance in class, and course staff observations). Ability in the course will be primarily assessed

through the final semester project, designed as an authentic assessment experience: Students choose a dataset related to their major to answer questions computationally. Ability tracking will be supplemented by fine-grained student solutions to coding problems to evaluate student progress. As a major stakeholder in the course, the instructors' experience is also crucial to understanding the success of our approach. Course staff will be regularly interviewed about the automated tools used in managing a large-scale course. Students will also be surveyed to assess their perception of the courses size (e.g., whether feedback came timely enough, they had the in-class support they needed to succeed, they felt satisfied with the human interactions in the course).

6. DISSERTATION STATUS

Currently, I am writing my preliminary proposal, expecting to defend this summer before ICER. I have already published a SIGCSE paper and hosted two SIGCSE workshops on the novel technology used to integrate Big Data into introductory programming courses. My goal is to complete my dissertation work by Spring 2017, giving me approximately two years from now to implement interventions and gather data. I have helped to integrate the Big Data Science approach into a new Computational Thinking course, which has been offered twice now (first with 20 students, then with 40), enabling me to collect preliminary data (partially published at ITICSE 2015). This fall, the course will be offered to 80 students, and I expect to start gathering serious data to answer my research questions. Much of the technology I propose has been prototyped, including the block-based environment and a collection of 36 (and growing) data sources.

7. EXPECTED CONTRIBUTIONS

I anticipate that this approach to Computational Thinking will provide a useful new technique for introductory instructors. Adoption of the materials will vary widely, with some instructors taking the curriculum wholesale and some using pieces of it. The tools will be readily available and experience will indicate how they can be used most effectively.

8. REFERENCES

- [1] R. E. Anderson et al. Introductory programming meets the real world: Using real problems and data in cs1. In *Proceedings of the 45th ACM Technical Symposium on Computer Science Education, SIGCSE '14*, pages 465–466, New York, NY, USA, 2014. ACM.
- [2] M. Guzdial and A. E. Tew. Imagineering inauthentic legitimate peripheral participation: An instructional design approach for motivating computing education. In *Proceedings of the Second International Workshop on Computing Education Research, ICER '06*, pages 51–58, New York, NY, USA, 2006. ACM.
- [3] B. D. Jones. Motivating students to engage in learning: The MUSIC model of academic motivation. *International Journal of Teaching and Learning in Higher Education*, 21(2):272–285, 2009.
- [4] Y. Matsuzawa, T. Ohata, M. Sugiura, and S. Sakai. Language migration in non-cs introductory programming through mutual language translation environment. In *Proceedings of the 46th ACM Technical Symposium on Computer Science Education, SIGCSE '15*, pages 185–190, New York, NY, USA, 2015. ACM.

May 19, 2015

This letter is in support of Austin (Cory) Bart's proposal to participate in the ICER 2015 doctoral consortium program. I have been working with Cory since he entered the graduate program in Computer Science at Virginia Tech in Fall 2012, and am his co-advisor along with Eli Tilevich.

Since he was an undergraduate, Cory's focus has been on Computer Science Education as a research field. During the roughly three years that Cory has been in our program, he has done outstanding work on a series of related projects that seek to motivate students in a range of Computer Science courses through the use of authentic programming exercises. In recognition of his work and his potential, he has received an NSF Fellowship. He has already completed a certificate program from the Virginia Tech graduate school on learning sciences, and this training helps to inform the pedagogical side of his work.

Cory's research program is a good balance of technology and pedagogy. Technically, he has developed a series of tools that support the introduction of real-time and "big" data streams into introductory courses. This involves various levels of scaffolding (adjustable to the level of the course) to hide or expose details of the data streams. The tools collectively allow instructors to easily introduce more authentic projects into their courses, even to the level of allowing each student to use their own online data source from a wide variety of topics available in the support library. His software has been used in at least three courses that I am aware of. The overarching goal is to move the dial on motivation and ultimately retention of students, especially during their first experience with computing. So an important part of Cory's work is to study and evaluate the effects of the various interventions.

As you can see from this description, Cory is involved in the full gamut of CS Education research, from systems development to pedagogy to evaluation. His work has proved interesting enough to the community that he has been the lead presenter for his systems at SIGCSE workshops two years running, and has also presented his work at SPLASH-E.

Cory expects to graduate in two years time. He is now at a critical stage in his program where he has results to report, but still has enough time left to adjust his focus based on feedback from the ICER community. Thus, the ICER Doctoral Consortium this year will be an ideal experience for him. I enthusiastically endorse his proposal, and hope that he will be selected to participate.

Yours Sincerely,



Clifford A. Shaffer
Professor

AUSTIN BART

600 Appalachian Drive #12, Blacksburg, VA, 24060 | 302-384-2729 | acbart@vt.edu

EDUCATION

Virginia Tech, Blacksburg, VA

PhD in Computer Science

Expected 2017

Learning Sciences Certificate

Dissertation: "Authentic Introductory Computing with Big Data and Blocks: Motivation, Ability, and Scalability"

University of Delaware, Newark, DE

Honors Bachelor of Computer Science with Distinction

2012

Concentration: Educational Technology

Thesis: "Exploring the XO Laptop as a Platform for Encouraging Creative Writing by Children"

AWARDS

National Science Foundation Graduate Research Fellowship Program (NSFGRFP), *Virginia Tech*

2014 – 2017

Eugene DuPont Memorial Scholarship, *University of Delaware*

2008 – 2012

NSF Graduate Research Fellowship Program Honorable Mention, *Virginia Tech*

2013

SIGCSE'13 Graduate Student Research Competition – 2nd Place, *Virginia Tech*

2013

SIGCSE'15 Graduate Student Research Competition – 3rd Place, *Virginia Tech*

2015

Computer Science Department Outstanding Sophomore Award, *University of Delaware*

2009

Residence Life's Citizen of the Year, Upper Division, *University of Delaware*

2011

TEACHING EXPERIENCE

Virginia Tech, Blacksburg, VA

Associate Instructor – to Dr. Dennis Kafura in "Introduction to Computational Thinking"

2014-2015

Developed technical infrastructure, co-lectured, worked with students, grading

Volunteer Instructor – Governor's School for Agriculture "Introduction to Computer Science"

2014

Taught a 1-week class (1 hour per day) on Computer Science to 20 high school students.

University of Delaware, Newark, DE

Teaching Assistant – to Dr. Keith Decker in "Introduction to Computer Science I"

2011

Ran lab sessions, hosted office hours, and graded assignments for students

Lab Assistant – to Dr. Kathleen F. McCoy in "Introduction to Computer Science I"

2012

Assisted in running lab sessions (*Smaller role due to work on senior thesis*)

RELATED EXPERIENCE

Association for Women in Computing, Virginia Tech

Webmaster

Fall 2013 – Ongoing

Maintained and updated the AWC website, maintained and updated social media presence, organized and supported AWC events.

Computer Science Graduate Council, Virginia Tech

Webmaster

Fall 2013 – Fall 2014

Maintained and updated the Grad Council Website, administered grad council listserv, organized and supported graduate council events, introduced a new system for making graduate student pages indexable and searchable.

Honor Society, University of Delaware

Senior Fellow

Fall 2011 – Spring 2012

Planned and hosted a large number of community activities for Honors students. Also advised students on Honors degree progress.

Computer Science Department, University of Delaware

Undergraduate Research Assistant to Dr. Ben Carterette

Summer 2009 – Fall 2010

Studied evaluation and optimization of information retrieval systems.

Computer Science Department, University of Delaware

Undergraduate Research Assistant to Dr. Lori Pollock**Summer 2011**

Studied language patterns and naming conventions in a large corpuses of Java Programs.

PUBLICATIONS AND PAPERS

- D. Kafura, A. C. Bart, B. Chowdhury, "Design and Preliminary Results from a Computational Thinking Course". ITiCSE'15, Vilnius, Lithuania.* 2015
- A. C. Bart, E. Tilevich, T. Allevato, S. Hall, C. A. Shaffer, "Transforming Introductory Computer Science Projects via Real-Time Web Data", SIGCSE '14, Atlanta, Georgia.* 2014
- A. C. Bart, L. Pollock, "Exploring the XO Laptop as a Platform for Encouraging Creative Writing by Children", Honors BS Thesis, University of Delaware.* 2012
- A. C. Bart, E. M. Bart, Teaching Animal Science with Minecraft: AnimalScienceCraft. GSA Research Symposium at Virginia Tech, Blacksburg, VA.* 2015

POSTERS

- A. C. Bart, Situating Computational Thinking with Big Data: Pedagogy and Technology, SIGCSE 45th ACM technical symposium on Computer Science Education Graduate Research Poster Competition, Kansas City, MO.* 2015
- A. C. Bart, E. Tilevich, C. A. Shaffer, T. Allevato, S. Hall, Teaching Computational Thinking with Real-Time Data, Conference on Higher Education Pedagogy, Virginia Tech, Blacksburg, VA.* 2014
- A. C. Bart, L. Pollock, Wacky Writing: Enhancing the XO Laptop Platform to Motivate Creative Writing by Children, SIGCSE 44th ACM technical symposium on Computer Science Education Graduate Research Poster Competition, Denver, CO, March 2013* 2013
- A. C. Bart, R. Deaton, E. McGinnis, Lowering Development Barriers in Educational Game Design, Conference on Higher Education Pedagogy, Virginia Tech, Blacksburg, VA* 2013
- A. C. Bart, G. Sridhara, L. Pollock, V. Shanker, Reverse Engineering from Java Identifier Names: Conventions and a Grammar, Summer Scholars Poster Presentation, University of Delaware, Newark, DE.* 2011

WORKSHOPS AND CONFERENCE TALKS

- E. Tilevich, C. A. Shaffer, A. C. Bart. Creating Stimulating, Relevant, and Manageable Introductory Computer Science Projects that Utilize Real-Time, Large, Web-Based Datasets , SIGCSE'15, Kansas City, MO.* 2014
- E. Tilevich, C. A. Shaffer, A. C. Bart. Creating Stimulating, Relevant, and Manageable Introductory Computer Science Projects that Utilize Real-Time, Web-Based Datasets, SIGCSE'14, Atlanta, GA.* 2013
- A. C. Bart, J. Riddle, O. Saleem, B. Chowdhury, E. Tilevich, C. A. Shaffer, D. Kafura, Motivating Students with Big Data: CORGIS and MUSIC, Splash-E '14, Portland, Oregon.* 2014
- A. C. Bart, E. Tilevich, C. A. Shaffer, T. Allevato, S. Hall, Using Real-Time Web Data to Enrich Introductory Computer Science Projects, Splash-E '13, Indianapolis, Indiana.* 2013

BOOKS

- D. Kafura, A. C. Bart, B. Chowdhury, Computational Thinking, <http://think.cs.vt.edu/book>* 2014
- A. C. Bart, R. Deaton, E. McGinnis, Platipy Project, <http://www.platipy.org>* 2013

INTERESTS*Software Engineering, Gender Diversity, Introductory Computing Experiences, Motivation and Cognition*