LitterBox: A Linter for Scratch Programs

Gordon Fraser, Ute Heuer, Nina Körber, Florian Obermüller, Ewald Wasmeier

University of Passau

Passau, Germany

arXiv:2102.07440v1 [cs.SE] 15 Feb 2021

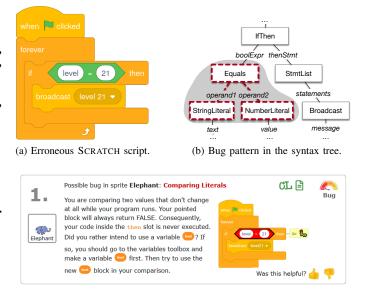
Abstract—Creating programs with block-based programming languages like SCRATCH is easy and fun. Block-based programs can nevertheless contain bugs, in particular when learners have misconceptions about programming. Even when they do not, SCRATCH code is often of low quality and contains code smells, further inhibiting understanding, reuse, and fun. To address this problem, in this paper we introduce LITTERBOX, a linter for SCRATCH programs. Given a program or its public project ID, LITTERBOX checks the program against patterns of known bugs and code smells. For each issue identified, LITTERBOX provides not only the location in the code, but also a helpful explanation of the underlying reason and possible misconceptions. Learners can access LITTERBOX through an easy to use web interface with visual information about the errors in the block-code, while for researchers LITTERBOX provides a general, open source, and extensible framework for static analysis of SCRATCH programs.

Index Terms-Scratch, bug patterns, code smells, linting

I. INTRODUCTION

The SCRATCH [7] block-based programming language is tremendously popular amongst teachers and programming novices. While it is easy to create games and animations using SCRATCH [10], programs may nevertheless contain bugs. This is particularly the case because SCRATCH programmers are usually young learners who might not know, or have misconceptions about, programming concepts. Even when programs suitably satisfy the behavior intended by the programmer, the quality of SCRATCH code has been reported to be low and riddled with code smells [6], [15]. This may have serious implications on learning outcomes, the general understandability of SCRATCH programs, and the overall enjoyment of programming. It can also severely inhibit teachers who may face the daunting task of debugging a potentially large and diverse set of student solutions at the same time.

In order to address this problem, we introduce LITTERBOX, a linter for SCRATCH programs. LITTERBOX is built on the observation that, even though there are abundant ways to produce bugs, many of them result from similar misconceptions and manifest in common *patterns* of bugs. As an example, consider the script shown in Figure 1a, which was written by a SCRATCH user with the intent to continuously check whether level 21 has been reached in a game, and if so to broadcast a corresponding message. The program contains a bug that is typical of learners who have not yet fully comprehended the concept of variables: Instead of comparing the value of the variable *level* with 21, the literal "level" is compared with the number 21. Obviously, this if-condition will never evaluate to



(c) Hint provided by LITTERBOX.

Fig. 1: Example bug pattern: Instead of comparing the variable *level* with the value 21, this script in the "Elephant" sprite compares the literal "level" with 21—a comparison that will never evaluate to true.

true. Such a comparison between literals can easily be described as a pattern on the abstract syntax tree of SCRATCH programs, and bugs matching this pattern occur frequently: In a recent study, we found 4,939 instances of this bug pattern in a dataset of 74,830 publicly shared SCRATCH projects [4].

LITTERBOX identifies bugs like this by statically analyzing the source code of SCRATCH programs, and applying a catalogue of finders for different bug patterns and code smells on the abstract syntax tree (Figure 1b) and control flow graph. For each issue identified, SCRATCH provides a visual summary of the relevant blocks together with an explanation of the issue, as well as possible underlying misconceptions. Figure 1c illustrates this for the *comparing literals* bug seen in Figure 1a.

LITTERBOX is developed as an open source Java project. On the command line, LITTERBOX can be used to check individual SCRATCH projects, collections of them, or for downloading projects from the SCRATCH website given their project IDs, and producing reports in configurable output formats. In addition to the command line interface, we provide a web interface which makes LITTERBOX easy to use for learners as well as for teachers having to assess multiple projects.

II. CODE QUALITY ISSUES IN SCRATCH

A. The SCRATCH Programming Language

SCRATCH [7] is popular amongst teachers and programming novices for several reasons: It is block-based, which means that programs are assembled by visually arranging blocks representing program statements and expressions. Blocks come in different shapes to visually demonstrate which combinations are valid. For example, boolean expressions are represented as pointed blocks, and other expressions in rounded blocks, and wherever such blocks can be inserted, there are holes in the appropriate shapes. Statements 'snap' together like Lego bricks, such that it is only possible to create programs of syntactically valid combinations of blocks. The blocks available are displayed in a toolbox, supporting recognition over recall. While SCRATCH blocks contain many standard programming language constructs such as conditions and loops, many statements represent high-level actions of the sprites interacting on a stage. This makes it easy to quickly arrange blocks in a way that results in fun and game-like programs.

While the tinkering-approach to programming is initially encouraging for learners, it has been observed that SCRATCH programmers tend to develop certain negative habits while coding [8]. Multiple studies have demonstrated that issues in code quality are prevalent in SCRATCH [1], [6], [11], [15] and have a negative impact on code understanding [5]. LITTERBOX therefore aims to identify recurring issues in SCRATCH code [4]. In the following, we discuss different categories of issues that occur in SCRATCH programs.

B. Syntax Errors in SCRATCH

Although the block-based nature of SCRATCH is intended to prevent syntax errors, it can only achieve this to a certain degree. Even though it is usually not possible to combine blocks in invalid ways, there are exceptions. For example, all 'reporter'-blocks have the same shape and so it is possible to use a reporter block for a costume name at a place where a color is expected. The notion of custom blocks gives rise to possible bugs, as SCRATCH does not validate the signature of these blocks; for example it is possible to have multiple parameters with the same name, or to use parameter blocks outside their scope. Many different types of syntax errors can arise from reuse and modification, for example when custom blocks are deleted, uses of their parameters may still remain in the program. Finally, incomplete programs can also give rise to bugs; for example, omitting a condition in a repeat until loop will lead to an infinite loop.

C. SCRATCH-specific Bugs

The program scenario of a stage with multiple sprites interacting on it gives rise to bugs specific to this setting. For example, programs may contain event handlers for scenerelated events (e.g., a change of the backdrop) without these events ever being triggered. The ability to use SCRATCH in a Logo-like mode, where sprites draw lines when moving, gives rise to various types of bugs when omitting to enable or disable the 'pen'-mode.

D. Patterns of General Bugs

Besides SCRATCH-specific bugs, there are countless ways for producing mistakes that may occur in any programming language. For example, it is possible to produce infinite recursions with custom blocks or message passing, there may be control flow anomalies, and there may be data flow issues such as lack of initialization. Common issues include omissions of **broadcast** or **clone** statements for which handlers are available, or vice versa, omissions of the handlers for events that are produced. There is no end to the creativity with which bugs can be produced in SCRATCH.

E. Code Smells

Even when SCRATCH programs do not contain bugs, the code may nevertheless be of mediocre quality. The concept of *code smells* describes different concrete quality problems in code which is not wrong per-se, but its quality issues reduce understandability and maintainability, and thus increase the chances that bugs are introduced later on. Common code smells known from regular programming languages often also apply to SCRATCH programs (e.g., spaghetti code, code duplication).

III. LITTERBOX: STATIC ANALYSIS FOR SCRATCH

A. Main Features of LITTERBOX

Bug finders. The central feature of LITTERBOX is to find issues in SCRATCH programs based on bug patterns and code smells. Table I provides a list of all issue finders implemented in LITTERBOX 1.5. Extending LITTERBOX with new issue finders is straightforward, so this list is growing continuously. **Code metrics.** Besides checking for bugs, LITTERBOX can extract different metrics on SCRATCH programs, such as the numbers of blocks, scripts, and sprites, or the overall weighted mean complexity of a program.

Code translation. LITTERBOX can translate SCRATCH programs to LeILA (Learners' Intermediate Language), the intermediate language used by the model checker BASTET [12], and to the SCRATCHBLOCKS¹ format.

Output formats. LITTERBOX can produce output in different formats: Besides basic information on the console, LITTERBOX can produce data files in CSV format, which is useful for researchers conducting analyses on datasets of SCRATCH programs. LITTERBOX produces a custom JSON report that contains detailed information about the issues found and their descriptions, which serves as a basis for displaying results to learners. LITTERBOX can export versions of the SCRATCH project in which all blocks associated with issues are annotated with comments explaining the issues, and giving hints on how the user could try to fix it.

Mining/downloading. LITTERBOX can analyze individual local files as well as folders containing multiple projects to check. Alternatively, LITTERBOX can also handle individual project IDs or lists of IDs, and will then download these projects from the SCRATCH servers before analyzing them.

Language support. LITTERBOX supports internationalization and currently provides output in English, German, and Spanish.

¹ http://scratchblocks.github.io, last accessed September 23, 2020.

TABLE I: Issue finders implemented in LITTERBOX 1.5.

Syntax Errors	
Ambiguous Custom Block Signature Ambiguous Parameter Name Call Without Definition Expression As Touching Or Color Illegal Parameter Refactor Missing Termination Condition Missing Wait-Until Condition Orphaned Parameter Parameter Out Of Scope	Several custom blocks have identical hats Custom block parameters with identical name Non existing custom block is called Reporter is used in color or object spot String parameter is used in bool condition Repeat until without condition Wait until without condition Parameter is not defined anymore Parameter outside custom block
SCRATCH-specific Bugs	
Missing Backdrop Switch Missing Erase All Missing Pen Down Missing Pen Up Missing Resource Stuttering Movement	Backdrop switch event never triggered Pen lines are not erased Pen is up but never down Pen is down but never up Used costume, sound or background is missing Annoying typematic delay
General Bugs	
Blocking If-Else Comparing Literals Custom Block With Forever Custom Block With Termination Delete Clone After Broadcast Endless Recursion Forever Inside Loop Inappropriate Hatblock Interrupted Loop Sensing Message Never Received Message Never Sent Missing Clone Call Missing Initialization Missing Loop Sensing No Working Scripts Position Equals Check Recursive Cloning Stop after Say	Terminates in both paths, code after if-else Strings/Numbers are compared directly Blocks after custom block never execute Blocks after custom block never execute Clone is deleted immediately after broadcast Custom block or script calls itself without termination Outer loop is never executed Greenflag handler in script with delete clone Block that takes time interrupts continuous sensing Broadcast does not trigger handler Broadcast for handler is never sent Answer is used without ask Clone event is never called Clone handler is never called Sprite is not initialized Condition is checked only a single time Only empty scripts and code lying around; no handler is connected to any blocks Positions are compared exactly Clones clone themselves without termination Script stopped immediately after say
Terminated Loop Type Error Variable As Literal	Loop is stopped during first iteration Incompatible blocks are compared Variable name instead of reporter
Code Smells	•
Busy Waiting Cloned Code Code Lying Around Double If Duplicate Sprite Duplicate Sprite Duplicate Script Empty Control Body Empty Custom Block Empty Script Empty Script Empty Sprite Long Script Message Naming Middle Man Multi Attribute Modification Nested Loops Same Variable Different Sprite Sequential Actions Sprite Naming Unnecessary If After Until Unnecessary Loop Unused Custom Block Unused Parameter Unused Variable	Constantly checking to stop script Code clones of types 1-3 Loose blocks without handler Consecutive if with same condition Two sprites are exact duplicates Two scripts in a sprite are exact duplicates C-block without sub stack Custom block without body Project without sprites Handler without body Sprite without scripts Script longer than 12 blocks Message with uncommunicative name Broadcast reception sends next broadcast; custom block calls next custom block Variable is changed multiple times in a row Loops without other blocks stacked in-between Same variable name in multiple sprites Sequence of repeated blocks instead of loop Sprite with uncommunicative name If checks same condition that terminated until Loop that runs never or one time Custom block is never called Parameter is defined but not used Variable is nittälized with different values in
varable initialization Race	scripts with same handler

B. Analysis Engine

LITTERBOX supports the analysis of projects in the most recent version of SCRATCH (3.0). SCRATCH projects are saved as .sb3 files, which are zip-archives containing the code of the project as a JSON file and all the assets such as sounds or costumes. LITTERBOX can process both, .sb3 and JSON files. A project is parsed to an abstract syntax tree (AST) in LITTERBOX, in which there is a distinct class for each type of block in SCRATCH. Stack blocks are represented as *Statements (Stmt)*, reporter blocks (including variables, lists and parameters of custom blocks) as *Expressions*. LITTERBOX also has wrapper classes to reflect the non typed structure of SCRATCH programs (e.g., with a wrapper a *String Expression* can be put into a *Stmt* expecting a *Number Expression*, for example when the **Username** block is used as parameter of the **move steps** block).

A visitor pattern is used to traverse this tree. So called *issue finders*, which check for bug patterns or code smells, are implemented as visitors to find idioms of SCRATCH blocks. Furthermore, LITTERBOX creates a control flow graph from the AST which is used in some of the finders, e.g., to report the missing initialization of variables or attributes of sprites.

IV. USING LITTERBOX

A. Command Line Usage

LITTERBOX can be accessed on the command line using an executable jar-file. An overview of all possible command line options can be displayed using the --help option:

\$ java -jar Litterbox-1.5.jar --help

The main modes of operation are --check to apply checks on a SCRATCH project, --stats to produce code metrics, and --leila to translate a program to LeILA. For example, to check a project stored on the local hard drive (as either .sb3 file or just the project.json file containing the code) with all default checkers on would use the following command:

\$ java -jar Litterbox-1.5.jar --check --path <filename>

Instead of a single file, one can also provide a folder, in which case LITTERBOX will check all SCRATCH projects contained in that folder. If the project is not available as a downloaded file yet, one can let LITTERBOX do this as well:

java -jar Litterbox-1.5.jar --check --projectid <id>

If no output directory is specified, LITTERBOX will place the file in a temporary directory.

There are command line options to select the output format. By default, LITTERBOX will only produce output on the console, but providing an output file name using --output=<filename> will produce a report. The format of this report is deduced from the filename (i.e., .csv or

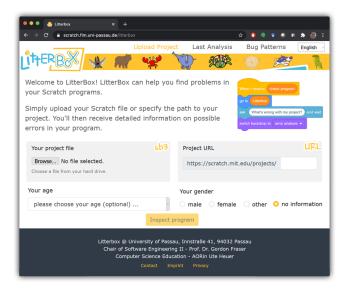


Fig. 2: Projects can be checked by uploading a file or entering the ID of a shared project.



Fig. 3: The result of the analysis is a list of issues and hints.

. json). Other options, such as whether to ignore unconnected blocks (i.e., scripts that have no hat-block), which checks to apply, and many other options using parameters can be displayed using the --help command line option.

B. The LITTERBOX Web Interface

For easier usage by learners LITTERBOX can also be accessed via a web interface. We provide a publicly accessible version of this at https://scratch-litterbox.org.

The web interface supports multiple languages (currently English, German, and Spanish). Users can upload a .sb3 file or simply provide the ID of a publicly shared SCRATCH project (see Figure 2). After the analysis is finished, a list of issues found is shown (Figure 3). For each issue, we display (1) the

```
public class MissingAsk extends AbstractIssueFinder {
private List<Answer> answerBlocks = new ArrayList<>():
private boolean askUsed = false:
@Override
public Set<Issue> check(Program program) {
   program.accept(this);
   // Add an issue for each answer block
   // if the program contains no ask-block
   if (!answerBlocks.isEmpty() && !askUsed)
      for (Answer answer : answerBlocks) {
         addIssue(answer, answer.getMetadata());
   }
   return issues;
00verride
public void visit(AskAndWait node) {
   askUsed = true;
00verride
public void visit(Answer node) {
   answerBlocks.add(node);
// ...
```

Fig. 4: Example bug finder to check if the Answer block is used without an AskAndWait block.

sprite which contains the affected script, (2) the block-code representation of the erroneous script, with the affected block highlighted in red with a worm attached to it, and (3) a textual description of the problem and possible remedies.

C. Extending LITTERBOX

Extending LITTERBOX with new bug patterns or smell checks is easy, as the internal structure allows extension by adding just a single class, a so called *finder*. As there is an abstract base class for finders which implements an AST visitor, the concrete finders have to override the existing visit methods only for the nodes that are relevant for the new check.

For convenience, the AST offers different levels of abstraction, for example, all reporter blocks implement one common interface Expression and all stack blocks implement Stmt, so that it is possible to handle all of them the same way without having to implement one visit method for each expression or stack block, respectively.

Figure 4 shows a simple bug finder which marks each usage of an **answer** block as an issue if the program contains no corresponding **ask** blocks. To achieve this, the finder simply needs to override the visit methods for the **AskAndWait** and **Answer** blocks to keep track of their usage, and afterwards can annotate all issues in the check method. LITTERBOX also makes more complex checks possible, using the control flow graph or support for classical data flow analysis.

Each finder needs to provide at least one hint message. It is possible to customize the hint messages to take specific aspects of the issue at hand into account, and provide more actionable help. Hint messages are internationalized through the use of resource bundles. Support for new languages can simply be added by including new resource bundles.

V. APPLICATIONS IN PROGRAMMING EDUCATION

A. Learners

Previous research has already shown that code smells in SCRATCH hamper the learning process of students [5]. LITTERBOX identifies not only code smells, but also bug patterns, which are more severe issues than smells, and thus likely have an even worse impact on learning.

LITTERBOX provides hints that consist of (1) an explanation of the issue focusing on meaningful context and (2) a short description of possible actions that might assist learners in solving the issue. This helps learners segmenting the complex debugging and refactoring process into manageable parts by locating potential issues one by one (Figure 3). It may also ease reflection on possible underlying misconceptions and foster resolving them by targeting inappropriate concepts for repair.

Deploying Eccles's expectancy-value theory [3] on students' debugging (and refactoring) processes yields the assumption that learners are more likely to keep on debugging if they expect to do well and they value debugging. As such, working with LITTERBOX learners might be more motivated to succeed.

B. Teachers

LITTERBOX can also be used by teachers to get an overview of potential bugs in their students' programs, in particular bugs that might be signs of misconceptions or missing concepts. It can thus support teachers in analyzing their learners' current knowledge and skills, facilitating instructional adjustments repeatedly. In particular, LITTERBOX endorses teachers assigning different tasks that are specifically set out to students differing in prior performance, thus alleviating individualized instruction.

LITTERBOX supports students when fixing their broken programs without needing extensive assistance of a human teacher while doing so. Teachers are then able to spend more time with learners in need of special support or extra attention. Note, however, that LITTERBOX finds only generic issues that are problematic in any SCRATCH program. Bugs that are specific to the task at hand require analysis with respect to a specification [12], [13].

VI. EMPIRICAL RESULTS

We investigated [4] a data set of 74,830 publicly shared SCRATCH projects (excluding remixes), and found a total of 109,951 instances of the 25 bug patterns implemented in LITTERBOX at the time of the study. Not every instance of a bug pattern may lead to visible failure of the program, and static analysis tools may in general produce false positives. We therefore manually inspected a stratified random sample of 250 bugs reported by LITTERBOX. In this sample, only 32 instances of bug patterns were revealed to be false positives. Overall, this demonstrates that bug patterns occur frequently in practice, and LITTERBOX helps to find them.

VII. RELATED WORK

The importance of analyzing SCRATCH programs was shown in previous studies [1], [6], [11], [15], which found code smells to occur frequently in practice. Hermans and Aivaloglou [5] determined empirically that code smells may inhibit learning and make it harder for pupils to extend SCRATCH programs with new functionalities. Techapalokul and Tilevich [15] also argue that code smells have negative effects on the SCRATCH community, for example as projects with code smells are less often remixed than those without smells.

The HAIRBALL [2] tool, which offers a basic Python API for analysing SCRATCH (version 2) programs, was instrumental in enabling this research on code smells, and other tools such as QUALITYHOUND [14] followed for statically finding code smells. LITTERBOX provides a generic program analysis framework for SCRATCH (version 3) programs. While it also incorporates basic code smells, it introduces the concept of bug patterns, the prevalence of which we demonstrated in an analysis of randomly sampled SCRATCH projects [4]. Whereas code smells identify low quality code that may lead to bugs in the future, patterns are indicative of bugs already present in the code. LITTERBOX offers various program analysis techniques, such as control- and data-flow analysis, to support the definition of further code smells, bug patterns, and also code metrics.

LITTERBOX also incorporates elaborate, multi-language hints to explain the issues found in the code. The web frontend in which these hints are displayed is inspired by DR. SCRATCH [9], which analyzes SCRATCH projects with respect to the computational thinking concepts applied.

VIII. CONCLUSIONS

SCRATCH is tremendously popular with learners, teachers, and researchers. There is, however, a need to provide support for all these user groups. This is what LITTERBOX aims to achieve: LITTERBOX is a static analysis tool that automatically detects bug patterns and code smells. A convenient web-interface and helpful explanations support learners and teachers alike, while the flexible command line interface and a modular design support teachers and researchers.

In this paper we have summarized the features of LITTER-BOX. By providing LITTERBOX as open source, we hope to foster research on analyzing SCRATCH programs as well as on SCRATCH-based programming education.

To try out LITTERBOX, visit our web site:

https://scratch-litterbox.org

The source code of LITTERBOX is available at:

https://github.com/se2p/LitterBox

ACKNOWLEDGEMENTS

This work is supported by DFG project FR 2955/3-1 "Testing, Debugging, and Repairing Blocks-based Programs" and BMBF project "primary::programming" as part of the Qualitätsoffensive Lehrerbildung. We would like to thank Christoph Frädrich, Miriam Münch, Gregorio Robles, Andreas Stahlbauer, and all other LITTERBOX contributors.

REFERENCES

- E. Aivaloglou and F. Hermans, "How kids code and how we know: An exploratory study on the Scratch repository," in *Proceedings of the ACM Conference on International Computing Education Research*, 2016, pp. 53–61.
- [2] B. Boe, C. Hill, M. Len, G. Dreschler, P. Conrad, and D. Franklin, "Hairball: Lint-inspired static analysis of scratch projects," 2013, pp. 215–220.
- [3] J. Eccles, "Who Am I and What Am I Going to Do With My Life? Personal and Collective Identities as Motivators of Action," *Educational Psychologist*, vol. 44, no. 2, pp. 78–89, 2009.
- [4] C. Frädrich, F. Obermüller, N. Körber, U. Heuer, and G. Fraser, "Common Bugs in Scratch Programs," in *Proceedings of the 2020 ACM Conference* on Innovation and Technology in Computer Science Education, ser. ITiCSE '20. ACM, 2020, pp. 89–95.
- [5] F. Hermans and E. Aivaloglou, "Do code smells hamper novice programming? A controlled experiment on Scratch programs," in 2016 IEEE 24th International Conference on Program Comprehension (ICPC), 2016, pp. 1–10.
- [6] F. Hermans, K. T. Stolee, and D. Hoepelman, "Smells in Block-Based Programming Languages," in 2016 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC). IEEE, 2016, pp. 68–72.
- [7] J. Maloney, M. Resnick, N. Rusk, B. Silverman, and E. Eastmond, "The Scratch Programming Language and Environment," ACM Transactions on Computing Education (TOCE), vol. 10, p. 16, 2010.
- [8] O. Meerbaum-Salant, M. Armoni, and M. Ben-Ari, "Habits of Program-

ming in Scratch," in *Proceedings of the Conference on Innovation and Technology in Computer Science Education*, 2011, pp. 168–172.

- [9] J. Moreno-León, G. Robles, and M. Román-González, "Dr. scratch: Automatic analysis of scratch projects to assess and foster computational thinking," *RED-Revista de Educación a Distancia*, 2015.
- [10] M. Resnick, J. Maloney, A. Monroy-Hernández, N. Rusk, E. Eastmond, K. Brennan, A. Millner, E. Rosenbaum, J. Silver, B. Silverman, and Y. Kafai, "Scratch: Programming for all," *Commun. ACM*, vol. 52, no. 11, p. 60–67, 2009.
- [11] G. Robles, J. Moreno-León, E. Aivaloglou, and F. Hermans, "Software clones in Scratch projects: On the presence of copy-and-paste in computational thinking learning," in 2017 IEEE 11th International Workshop on Software Clones (IWSC). IEEE, 2017, pp. 1–7.
- [12] A. Stahlbauer, C. Frädrich, and G. Fraser, "Verified from Scratch: Program Analysis for Learners' Programs," in *In Proceedings of the International Conference on Automated Software Engineering (ASE)*. IEEE, 2020.
- [13] A. Stahlbauer, M. Kreis, and G. Fraser, "Testing Scratch Programs Automatically," in *Proceedings of the 2019 27th ACM Joint Meeting* on European Software Engineering Conference and Symposium on the Foundations of Software Engineering, 2019, pp. 165–175.
- [14] P. Techapalokul and E. Tilevich, "Quality hound an online code smell analyzer for scratch programs," in 2017 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC), 2017, pp. 337–338.
- [15] —, "Understanding Recurring Quality Problems and Their Impact on Code Sharing in Block-Based Software," in 2017 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC). IEEE, 2017, pp. 43–51.