

MakeCode and CODAL: Intuitive and efficient embedded systems programming for education[☆]

James Devine^{a,*}, Joe Finney^a, Peli de Halleux^b, Michał Moskal^b, Thomas Ball^b, Steve Hodges^c

^a Lancaster University, UK

^b Microsoft, USA

^c Microsoft, UK

ARTICLE INFO

Keywords:

MakeCode
CODAL
BBC micro:bit
Embedded systems
Physical computing
Visual programming
Web-based programming

ABSTRACT

Historically, embedded systems development has been a specialist skill, requiring knowledge of low-level programming languages, complex compilation toolchains, and specialist hardware, firmware, device drivers and applications. However, it has now become commonplace for a broader range of non-specialists to engage in the making (design and development) of embedded systems - including educators to motivate and excite their students in the classroom. This diversity brings its own set of unique requirements, and the complexities of existing embedded systems development platforms introduce insurmountable barriers to entry.

In this paper we present the motivation, requirements, implementation, and evaluation of a new programming platform that enables novice users to create effective and efficient software for embedded systems. The platform has two major components: (1) Microsoft MakeCode (www.makecode.com), a web app that encapsulates an accessible IDE for microcontrollers; and (2) CODAL, an efficient component-oriented C++ runtime for microcontrollers. We show how MakeCode and CODAL combine to provide an accessible, cross-platform, installation-free, high level programming experience for embedded devices without sacrificing performance and efficiency.

1. Introduction

Recent years have witnessed expansive growth in the popularity and ubiquity of embedded systems. This growth can be primarily attributed to the emergence of new application domains ranging from wearables, to home automation, industrial automation, and smart grids – the phenomenon broadly referred to as the Internet of Things (IoT). As the IoT continues to grow, it has become more pervasive – far beyond the realm of domain experts and into the everyday lives of the public. This has led to growth in non-expert developers actively creating software for embedded systems. Small to medium sized businesses now create new products through *rapid prototyping* of embedded devices. Hobbyist *makers* create novel technical projects to inspire themselves and society. And now, more than ever before, *educators* are making extensive use of *physical computing* devices as a direct means to teach and inspire a generation of students – and to prepare them for a society where IoT will be the norm. These new developers all share a common characteristic: **they are not professional software developers** [11,16,28].

There are many examples of existing devices that aim to cater to non-expert developers such as the Arduino and Raspberry Pi [8,42], both of which can interface to external peripherals by connecting wires

to onboard GPIO. However, many barriers still exist for inexperienced users. The Arduino still requires programs to be written in the relatively unforgiving language of C/C++, and that these programs are written using a platform specific IDE installed on a personal computer. The Raspberry Pi gives access to higher level languages such as Python and Scratch [35] via a high performance CPU and full operating system. However, this comes at the cost of simplicity and transparency - requiring users to attach dedicated displays and keyboards, master the Linux operating system, and tolerate the non-realtime behaviours associated with an interpreted language on a general purpose OS.

United under the aim of further lowering the barrier to entry for non-expert developers, a consortium of industry and academic partners came together in 2015 to develop the BBC micro:bit – an embedded device designed specifically for education. One million of these devices were delivered to UK school children in 2016. The micro:bit is a highly capable IoT device containing a 32-bit ARM Cortex-M0 processor, integrated light level, temperature, acceleration and magnetic sensors, touch sensitive inputs, and USB and 2.4GHz radio communications.

Fig. 1 highlights two example educational projects based on the BBC micro:bit. The first is a data gathering experiment, where multiple sensor data was recorded to non-volatile memory as the

[☆] This paper is an extended version of the conference paper [15]

* Corresponding author.

E-mail addresses: j.devine@lancaster.ac.uk (J. Devine), j.finney@lancaster.ac.uk (J. Finney), jhalleux@microsoft.com (P. de Halleux), michal.moskal@microsoft.com (M. Moskal), tball@microsoft.com (T. Ball), steve.hodges@microsoft.com (S. Hodges).

<https://doi.org/10.1016/j.sysarc.2019.05.005>

Received 16 November 2018; Received in revised form 29 March 2019; Accepted 25 May 2019

Available online 28 May 2019

1383-7621/© 2019 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license. (<http://creativecommons.org/licenses/by/4.0/>)



Fig. 1. Example projects undertaken within education: Rishworth School sent a micro:bit to space [31] (left); The micro:bit placed in a custom-built rocket car for telemetry [32,33] (right).

device was launched into near space (32.5km altitude), along with an externally interfaced camera and GPS unit. The second highlights a live data telemetry application, where acceleration data was streamed in real-time via Bluetooth Low Energy to enable the profiling of chemical-rocket powered model vehicles. These projects are highly sophisticated and were undertaken by high school educators and their students.

In this paper we introduce an open-source¹ platform for embedded devices such as the micro:bit that enables the development of embedded applications by non-expert programmers. This platform consists of two major components: (1) Microsoft MakeCode (www.makecode.com), a web app providing a beginner IDE for embedded systems; and (2) CODAL (Component-Oriented Device Abstraction Layer), an efficient C++ runtime with high-level programming abstractions.

We contribute an approach to creating development tools for non-experts (exemplified through MakeCode and CODAL) that offers a simpler development experience using: multi-paradigm web-based drag-and-drop visual programming with in-browser compilation, device simulation, and driverless usb flash drive microcontroller programming. Despite providing such high-levels of abstraction, we also demonstrate up to 50x better performance than other state-of-the-art implementations, in some cases nearing the performance of native C++.

The remainder of the paper describes the design, implementation, and evaluation of MakeCode (Section 3), the CODAL C++ runtime (Section 4), and the UF2 bootloader (Section 5) which enables driver-less microcontroller programming. Sections 6 and 7 show that combined, these technologies enable simplified programming while maintaining a relatively high degree of temporal and spatial efficiency. MakeCode has been live since the fall of 2016, at first supporting just the micro:bit, but now supporting many more devices, most of which are based on CODAL. Section 8 discusses related work and Section 9 concludes the paper.

2. Design principles and architecture

In this section we present the major design challenges in bringing embedded systems to novice users and briefly describe the architecture of our solution.

2.1. Design principles

Enabling novice programmers to successfully develop embedded applications is a non-trivial task. Throughout our research we have established a number of design principles that are addressed through MakeCode and CODAL:

High Level Languages: Programming languages for microcontroller units (MCUs) have not kept pace with advances in hardware. Despite active research in the field, the C/C++ languages remain the standard for embedded systems: they provide a familiar imperative programming model, with compilers that produce highly efficient code, and enable low level access to hardware features when necessary. The

¹ MakeCode is open-source at <https://github.com/microsoft/pxt>; CODAL is open-source at <https://github.com/lancaster-university/codal>.

Table 1

Example microcontroller devices in relationship to a typical PC. Device abbreviations: Uno (Arduino Uno), micro:bit (BBC micro:bit), CPX (Adafruit Circuit Playground Express), PC (Personal Computer).

Device	RAM	Flash	Word Size	CPU Speed	CPU
Uno	2 kB	32 kB	8	16MHz	AVR
micro:bit	16 kB	256 kB	32	16MHz	ARMv6-M
CPX	32 kB	256 kB	32	48MHz	ARMv6-M
PC	16 GB	1 TB	64	3GHz	x86-64

Arduino project (www.arduino.cc), started in 2003, and ARM's Mbed platform (www.mbed.org) both rely heavily on a C/C++ programming model [6,39]. However, the limitations of using C/C++ as an application programming language for inexperienced developers are well understood [10]. To address this, higher level languages such as JavaScript, Python, and even visual programming languages are required.

Zero Installation Architecture: The development environment for existing embedded systems typically requires the installation of code editors, custom device drivers, compiler toolchains, and even additional programming hardware (such as a JLink programmer). For many, particularly in the field of education, this presents a high adoption barrier as in many schools, custom hardware and software is simply not permitted by policy and/or access to the necessary technical support is not present. An effective solution must therefore provide a fully transparent, platform agnostic, zero installation experience to developing embedded software.

Optimization for Code Efficiency: The projects of Fig. 1 are enabled by small, highly resource-limited programmable MCUs, which may have as little as a few kilobytes of RAM and FLASH memory. Table 1 compares the core capabilities of the class of MCU-based devices typically used in the education domain to a typical PC. Note that these devices have a proportionally large amount of processing power, relative to their storage. Consider the BBC micro:bit vs. a typical PC: the micro:bit has about 100 times less CPU power, but 10⁶ times less RAM, and 10⁶ times less storage. A language/runtime should therefore not only seek to provide high code density and spatial efficiency, but actively trade off temporal for spatial efficiency where possible.

Asynchronous and Concurrent Programming: It is already well understood that novice programmers benefit from event-based programming paradigms [29,30,41]. This is increasingly relevant for embedded systems due to the typically asynchronous nature of their hardware. MCUs still follow Moore's law, but this additional capacity is not typically invested in speeding up processors. Instead, more independent peripherals (such as Bluetooth/WiFi radio modules, audio inputs/output processors, etc.) are integrated onto the same package as the CPU as a system-on-chip. Such peripherals often operate independently of the main CPU. An effective language/runtime should directly support an asynchronous interaction model designed to cooperate with the independent nature of peripherals while remaining highly intuitive to the programmer.

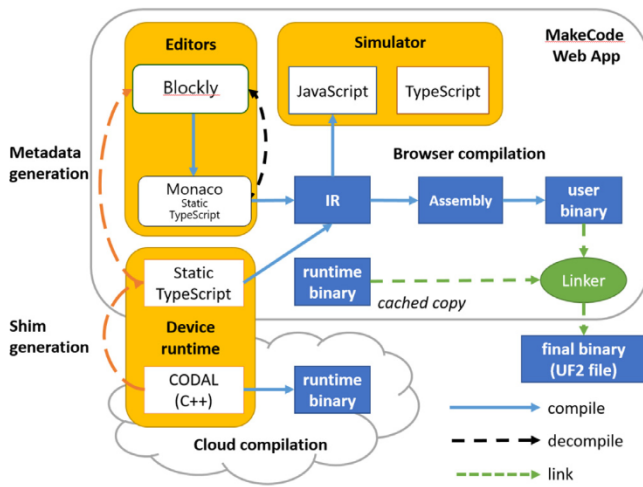


Fig. 2. The MakeCode and CODAL architecture.

Intuitive and Extensible APIs. Intuitive APIs and programming models are required to support novice users, yet it is equally important that these APIs remain complete enough to realise the ambitious projects that may be undertaken as students advance: simplification via the reduction of functionality is not a valid approach. *An effective solution must provide APIs that are consistent, easy to understand/use, and progressive, to address the growing capabilities of the programmer.*

2.2. Architecture

Fig. 2 illustrates the architecture of our platform. The MakeCode web app is the primary entry point for the end-user. MakeCode supports the simplified programming of MCUs via editors for visual blocks and the textual TypeScript² language. CODAL is a component-oriented, event-driven, fiber-based C++ runtime environment that bridges the semantic gap between the higher-level languages (such as TypeScript) and the hardware (bottom-left of the Figure). Enabling the flashing of the microcontroller is UF2, a new file format and bootloader for the simplified transfer of binaries to the device over USB (bottom-right).

MakeCode can be accessed from any modern web browser and cached locally for *entirely offline use*. The MakeCode web app incorporates the open-source Blockly³ and Monaco⁴ editors (upper-left), an in-browser device simulator (upper-right) for testing programs before transferring them to the physical device, as well as *in-browser compilation* of TypeScript to machine code and linking against the pre-compiled CODAL C++ runtime (lower-left).

MakeCode devices appear as USB pen drives when plugged into a computer, thanks to UF2. After a user has finished developing a program, the compiled binary is “downloaded” locally to the user’s computer (lower-right) and then transferred (flashed) to the MCU by a simple file copy operation. This works out-of-the-box on any OS with built-in support for USB pen drives (MacOS, Windows, Linux, ChromeOS).

3. MakeCode

The key technical contribution of MakeCode is to provide users with a basic environment to write programs for MCUs, enabling a simple progression from visual block-based programming to text-based programming in Static TypeScript (STS), while leveraging C++ on the backend for efficient use of MCU resources. MakeCode uses the open-source Blockly and Monaco editors to allow the user to code with visual blocks

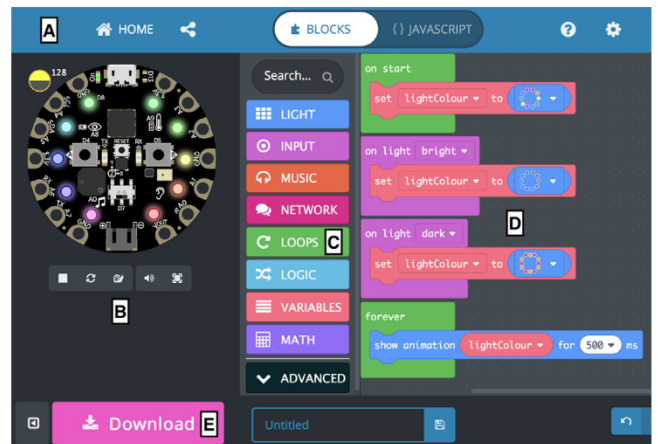


Fig. 3. Screen snapshot of the MakeCode web app.

or STS. The editing experience is parameterized by a fully-typed device runtime, which provides a set of categorized APIs to the end-user.

Fig. 3 is a screenshot of the MakeCode web app for Adafruit’s Circuit Playground Express (CPX) device.⁵ The web app has five sections: (A) the menu bar allows switching between the two editors; (B) the simulator shows the CPX board and provides feedback on user code executed in the browser; (C) the toolbox provides access to device-specific APIs and programming elements; (D) the programming canvas is where editing takes place; (E) the download button invokes the compiler, producing a binary executable.

The web app encapsulates all the components needed to deliver a programming experience for MCUs, free of the need for a C++ compiler for the compilation of user code. The web app is written in TypeScript and incorporates the TypeScript compiler and language service as well. The app is built from a MakeCode “target” which parameterizes the MakeCode framework for a particular device. The remaining subsections describe the essential components of the web app in Fig. 2.

3.1. Static TypeScript

TypeScript is a typed superset of JavaScript designed to enable JavaScript developers to take advantage of code completion, static checking and refactoring made possible by types. As a starting point, every JavaScript program is a TypeScript program. Types can be added gradually to JavaScript programs, supported by type inference. While TypeScript provides classes and interfaces with syntax like Java/C#, their semantics are quite different as they are based on JavaScript. Classes are simply syntactic sugar for creating objects that have code associated with them, but these objects are indeed JavaScript objects with all their dynamic semantics intact.

Static TypeScript (STS) is closely related to StrongScript [37], which extends TypeScript with a type constructor for *concrete* types, allowing the programmer to choose between untyped, optionally-typed, and concretely typed code. This provides traditional type soundness guarantees, as in Java and C#. STS can be seen to be StrongScript where every variable and expression has a concrete type. As in StrongScript, classes are nominally typed, which permits a more efficient and traditional property lookup for class instances. Currently, STS goes further than StrongScript by outlawing downcasts.

3.2. Device runtime and shim generation

A MakeCode target is defined, in part, by its device runtime; which is a combination of C++ and STS code, as shown in the lower-left of

² <https://www.typescriptlang.org>.

³ <https://github.com/google/blockly>.

⁴ <https://github.com/microsoft/monaco-editor>.

⁵ <https://makecode.adafruit.com>.

Fig. 2. All the target's C++ files are precompiled (by a C++ compiler in the cloud) into a single binary, which is stored in the cloud as well as in the HTML5 application cache. Additional runtime components may be authored in STS, which allows the device runtime to be extended without the use of C++, and permits components of the device runtime to be shared by both the device and simulator runtimes. The ability to author the device runtime in both STS and C++ is a unique aspect of MakeCode's design.

Whether runtime components are authored in C++ or STS, all runtime APIs are exposed as fully-typed TypeScript definitions. A fully-typed runtime improves the end-user experience by making it easier to discover APIs; it also enables the type inference provided by the TypeScript compiler to infer types for (unannotated) user programs.

MakeCode supports a simple foreign function interface from STS to C++ based on namespaces, enumerations, functions, and basic type mappings. MakeCode uses top-level namespaces to organize sets of related functions. Preceding a C++ namespace, enumeration, or function with a comment starting with `/// indicates that MakeCode should map the C++ construct to STS. Within the ///`

```
1 ///

```

Fig. 3(C) shows the toolbox of API and language categories, where the category `INPUT` corresponding to the namespace `input` can be seen (second from the top).

Mapping of functions and enumerations between C++ and STS is straightforward and performed automatically by MakeCode. For example, the following C++ function `onLightConditionChanged` in the namespace `input` wraps the more complex C++ needed to update the sensor and register the (Action) handler with the underlying CODAL runtime:

```
1 ///

```

MakeCode generates TypeScript declaration file (here called a shim file) to describe the TypeScript elements corresponding to C++ namespaces, enumerations and functions. Since the C++ function above is preceded by a `///`

```
1 ///

```

Since the `///block attribute, MakeCode creates a block (named "on light"), which can be seen in the upper-left of Fig. 3(D).`

To support the foreign function interface, MakeCode defines a mapping between C++ and STS types. Boolean and void have straightforward mappings from C++ to STS (`bool` → `boolean`, `void` → `void`). As JavaScript only supports number, which is a C++ double, MakeCode uses TypeScript's support for type aliases to name the various C++ integer types commonly used for MCU programming (`int32`, `uint32`, `int16`, `uint16`, `int8`, `uint8`). This is particularly useful for saving space on 8-bit architectures such as the AVR. MakeCode also includes reference-

counted C++ types for strings, lambdas (Action in C++, with up to three arguments and a return type) and collections, with mappings to STS.

MakeCode does not yet include a garbage collector, so advanced users who create cyclic data structures must be careful to break cycles to ensure complete deallocation.

3.3. Browser compilation

When a user requests a download of the compiled binary, MakeCode first invokes the TypeScript language service to perform type inference and type checking on the user's program, the device runtime written in STS, and the TypeScript declarations corresponding to the C++ device runtime. It then checks that the combined TypeScript program is within the STS subset through additional syntactic and type checks over the typed AST. Assuming all the above checks pass, MakeCode then performs tree shaking of the AST to remove unused functions. The reduced AST is then compiled to an intermediate representation (IR) that makes explicit labelled control flow among a sequence of instructions with conditional and unconditional jumps, heap cells, field accesses, store operations, and reference counting.

There are three backends for code generation from the IR. The first backend generates JavaScript, for execution against the simulator runtime. A second backend, parameterized by processor type, generates assembly code. Currently supported processors include ARM's Cortex class (Thumb instructions) and Atmel's ATmega class (AVR instructions). A separate assembler, also parameterized by an instruction encoder/decoder, generates machine code and resolves runtime references, producing a final binary executable. A third backend generates bytecode instructions. MakeCode can encode the resulting binary in several formats, including Intel's HEX format [22] and the UF2 format, documented in [Section 5](#).

The MakeCode compiler supports the STS language subset of TypeScript with two compilation strategies: untagged and tagged. Under the untagged strategy, a JavaScript number is interpreted as a C++ `int` by

default and the type system is used to statically distinguish primitive values from boxed values. As a result, the untagged strategy is not fully faithful to JavaScript semantics: there is no support for floating point, and the `null` and `undefined` values are represented by the default integer value of zero. The untagged strategy is used for the micro:bit and Arduino Uno targets.

In the tagged strategy, numbers are either tagged 31-bit signed integers, or if they do not fit, boxed doubles. Special constants like `false`, `null` and `undefined` are distinguished by specific values. The tagged execution strategy has the capability to fully support JavaScript semantics and is used by all ATSAM21 targets, including the CPX.

4. The CODAL runtime

CODAL is a lightweight, object-oriented, componentized C++ runtime for microcontrollers designed to provide an efficient abstraction

```

1  #include "CircuitPlayground.h"
2  CircuitPlayground cplay;
3
4  void onBright() { // user defined code }
5
6  int main() {
7      cplay.messageBus.listen(ID_LIGHT_SENSOR, LIGHT_THRESHOLD_HIGH,
8          onBright);
9  }

```

Fig. 4. Example of the CODAL MessageBus.

layer for higher level languages, such as JavaScript. CODAL has five key elements:

1. a *unified eventing subsystem* (common to all components) that provides a mechanism to map asynchronous hardware and software events to event handlers;
2. a *non-preemptive fiber scheduler* that enables concurrency while minimizing the need for resource locking primitives;
3. a *simple memory management system* based on reference counting to provide a basis for managed types;
4. a *stream processing framework* based on a composable, receiver-driven component model;
5. a *set of drivers*, that abstract microcontroller hardware components into higher level software components, each represented by a C++ class;
6. a *parameterized object model* composed from these components that represents a physical device.

There are discussed in detail below.

4.1. Message bus and events

CODAL offers a simple yet powerful model for handling hardware or user defined events. Events are modeled as a tuple of two integer values - specifying an *id* (namespace) and a *value*. Typically, an id correlates to a specific software component, which may be as simple as a button or something more complex as a wireless network interface. The value relates to a specific event that is unique within the id namespace. All events pass through the CODAL MessageBus. Application developers can then *listen* to events on this bus, by defining a C/C++ function to be invoked when an event is raised. Events can be raised at any time simply by creating an *Event* C++ object, which then invokes the event handlers of any registered listeners.

Continuing the example of detecting the brightness of a room used in [Section 3.2](#), a listen call to the MessageBus with a component ID of the light sensor and a threshold event is the underlying mechanism enabled by the runtime, as illustrated by the equivalent C++ code snippet in [Fig. 4](#).

Unlike simple function pointers, CODAL event handlers can be parameterized by the event listener to provide decoupling from the context of the code raising the event. The receiver of an event can choose to either receive an event in the context of the fiber that created it, or can be decoupled and executed via an Asynchronous Procedure Call (APC). The former provides performance, while the latter provides decoupling of low level code (that may be executing, say, in an interrupt context) from user code. Each event handler may also define a threading model, so they can be reentrant or run-to-completion depending upon the semantics required.

4.2. Fiber scheduler

CODAL provides a *non-preemptive* fiber scheduler with asynchronous semantics and a power efficient implementation. CODAL fibers can be created at any time but will only be descheduled as a result of an explicit call to `yield()`, `sleep()` or `wait_for_event()` on the MessageBus. The latter enables condition synchronization between fibers through a wait/notify

mechanism. A round-robin approach is used to schedule runnable fibers. If at any time all fibers are descheduled, the MCU hardware is placed into a power efficient sleep state.

The CODAL scheduler makes use of two novel mechanisms to optimize for MCU hardware. Firstly, CODAL adopts a stack paging approach to fiber management. MCUs do not support virtual memory and are heavily RAM constrained, but relatively cycle rich. Therefore, instead of overprovisioning stack memory for each fiber (which would waste valuable RAM), we instead dynamically allocate stack memory from heap space as necessary and copy the physical stack into this space at the point at which a fiber is descheduled (and similarly restored when a fiber is scheduled). This copy operation clearly incurs a small CPU overhead, but brings greater benefits of RAM efficiency - especially given that MCU stack sizes are typically quite small (~200 bytes is typical).

Secondly, the CODAL scheduler supports transparent APCs. Any function can be *invoked* as an APC. Conceptually, this is equivalent to calling the given function in its own fiber. However, the CODAL runtime provides a common-case transparent optimization for APCs we call *fork-on-block* - whereby a fiber will only be created at the point at which the given function attempts a blocking operation such as `sleep()` or `wait_for_event()`. Functions which do not block therefore do not incur all of the context switch overhead.

When invoking an APC, the scheduler snapshots the current processor context and stack pointer (but not the whole stack). If the scheduler is re-entered before the APC completes, a new fiber context is created at the point of descheduling, and placed on the appropriate wait queue. The previously stored context is then restored, and execution continues from the point at which the APC was first invoked. This mechanism provides potentially high RAM savings for the processing of MessageBus event handlers in particular.

CODAL's scheduling and eventing models are shared by both high and low level languages, and therefore handled uniformly. As a result, when a foreign function call is mapped to C++, that C++ function is capable of blocking the calling fiber without infringing on the concurrency model of the higher level language. This enables, for example, a C++ device driver to block a JavaScript program when awaiting data without changing the behavior of other JavaScript code acting asynchronously (as in [Fig. 3](#)).

4.3. Memory management

CODAL implements its own lightweight heap allocator, introducing reentrant versions of the `libc malloc` family of functions, permitting universal access to heap memory in user or interrupt code. The heap allocator is flexible and reconfigurable, allowing the specification of multiple heaps across memory and it is optimized for repeat allocations of memory blocks that are commonplace in embedded systems.

CODAL also makes use of simple managed types, built using C++ reference counting mechanisms. C++ classes are provided for common types such as strings, images, and data buffers. A generic base class is also provided for the creation of other managed types. This simple approach brings the benefits of greater memory safety for application code, but with the expense of suffering from the issues related to circular references. We take the view that such scenarios are rare in MCU

```

1  class DataSink
2  {
3      public:
4      virtual int pullRequest();
5  };
6
7  class DataSource
8  {
9      public:
10     virtual ManagedBuffer pull();
11     virtual void connect(DataSink &sink);
12 };

```

Fig. 5. CODAL stream interfaces.

```

1  CircuitPlayground cplay;
2  SAMD21DMAC dmacc;
3  SAMD21PDM microphone(cplay.io.microphoneData, cplay.io.microphoneClock,
4      dmacc, 10000);
5  LevelDetector level(microphone.output, 70, 30);
6
7  void onLoud(DeviceEvent)
8  {
9      cplay.serial.printf("LOUD\n");
10 }
11
12 int main()
13 {
14     microphone.enable();
15     cplay.messageBus.listen(DEVICE_ID_SYSTEM_LEVEL_DETECTOR,
16         LEVEL_THRESHOLD_HIGH, onLoud);
17 }

```

Fig. 6. CODAL stream processing example - a sound level detector.

applications, justifying this approach over a more complex garbage collection scheme and its overhead.

4.4. Streams

Whilst some embedded systems applications only process discrete data, many operate on data flows. Examples include digital signal processing (DSP), audio recording/playback, and gesture and voice recognition. CODAL defines a standardized mechanism to efficiently handle data streams that can be then exported to higher level languages as a data flow model. Following on from the Object Oriented approach adopted by CODAL, data streams are modelled using two well defined C++ abstract base classes - *DataSource* and *DataSink*, as defined in Fig. 5. These two simple interfaces contain the minimal set of functionality to enable simple, efficient, safe and extensible stream processing. As can be seen, all data is conveyed using a managed type (*ManagedBuffer*). This ensures that any device drivers that source streamed data or DSP software or application code that process it need not be burdened with the detail of memory management.

Any component capable of generating stream data (an *upstream* component) implements the *DataSource* interface, and likewise a consumer (*downstream* component) implements *DataSink*. A receiver driven approach to data streaming is adopted. *DataSinks* register with their upstream *DataSource* through an explicit call to the *connect* method. When data is subsequently generated by a *DataSource*, it then executes the *pullRequest()* method on its downstream component to indicate that data is available for processing, which can then be drained from the *DataSource* through its *pull()* method. This receiver driven approach allows for simple flow control to be implemented - placing the control of precisely when and how processing takes place in the hands of the application.

Components may be both a *DataSource* and a *DataSink*, allowing the creation of a data flow graph of components. The CODAL runtime contains reusable components that can be composed to undertake common operations. For example, analog to digital converter modules act

as *DataSources* and digital to analog converter modules as *DataSinks*. A *Mixer* component allows multiple *DataSources* to be aggregated into one *DataSink* (with clear applications for audio processing). A *LevelDetector* provides configurable smoothing and thresholding functionality with asynchronous outputs reported through the CODAL MessageBus. A *DataStream* provides user configurable first-in-first-out buffering functionality, and a *Synthesizer* frequency and phase shifted datastreams based on a user defined waveform template. For example, Fig. 6 illustrates an event based, direct-memory-access enabled, asynchronous sound level detector of a live data source from an attached microphone.

4.5. Device driver components

CODAL drivers abstract away the complexities of the underlying hardware into reusable, extensible, easy-to-use components. For every hardware component there is a corresponding software component that encapsulates its behavior in a C++ object. CODAL has three types of drivers:

1. A hardware agnostic abstract specification of a driver model (e.g. a Button, or an Accelerometer). This is provided as a C++ base class.
2. The concrete implementation of the abstract driver model, which is typically hardware specific. This is implemented as a subclass of a driver model, such as a LIS3DH accelerometer, as manufactured by ST Microelectronics.
3. A high level driver that relies only on the interfaces specified in a driver model (e.g. a gesture recognizer based on an Accelerometer model).

This approach brings the benefits of abstraction and reusability to CODAL, without losing the hardware specific benefits seen in flat abstraction models where every MCU is made to look the same, even though their capabilities are different (as in the Arduino and mbed APIs, for example).

Finally, we group together the components of a physical device to form a *device model*. This is a singleton C++ class that, through

```

1  class CircuitPlayground : public CodalDevice {
2      public:
3          MessageBus          messageBus;
4          CPlayTimer          timer;
5          SAMD21Serial        serial;
6          CircuitPlaygroundIO io;
7          Button              buttonA;
8          SAMD21I2C           i2c;
9          LIS3DH              accelerometer;
10         NonLinearAnalogSensor thermometer;
11         AnalogSensor        lightSensor;
12         ...

```

Fig. 7. Device model for the Adafruit CPX.

composition of device driver components, provides a configured representation of the capabilities of a device. Such a model allows: an elegant OO API for programming a device, and a static representation that forms an ideal target for the MakeCode linker to bind high level STS interfaces to low level optimized code.

An example device model for the CPX is shown in Fig. 7 for reference.

MakeCode is further supported by an annotated C++ library (*MakeCode wrappers*) defining the mapping from CODAL to TypeScript and Blockly. The use of MakeCode wrappers ensures that different MakeCode targets that use CODAL share a common TypeScript and block API vocabulary.⁶

5. UF2

The UF2 bootloader enables efficient, universal microcontroller programming through a USB pen drive interface — no drivers are required, as operating systems support pen drives out of the box. The UF2 bootloader builds upon the work of DAPLink [3], but offers a simpler implementation via a new flashing format, UF2.

DAPLink exposes a small virtual 512-byte block FAT file system (VFS), with an empty file allocation table and root directory. When the OS tries to read a block, DAPLink computes what should be there. During file system writes, DAPLink detects blocks of files in Intel HEX format [22], decodes them, and flashes the file’s contents into the target microcontroller’s memory. Other file system writes are ignored.

DAPLink implements many heuristics to deal with quirks of FAT file system implementations in various operating systems (order of writes, various meta-data files that are created and need to be ignored, etc.). However, we have found that the heuristics are fragile with respect to operating system changes, which are not infrequent.

Our new file format, UF2, consists of one or more 512-byte self-contained blocks (aligned to the block size of the VFS), removing the need for the complex heuristics. The blocks have magic numbers, the payload data to be written to flash, and the address where it should be written. Thus, on every 512-byte write via the USB controller, the bootloader can quickly and easily check if the block being written is part of a UF2 file (by comparing magic numbers) and if so, write it immediately in a streaming fashion.

For simplicity, the 512-byte UF2 blocks usually contain 256 bytes of payload data. While 50% density might seem low, the industry standard 16-byte-per-line HEX format has a density of around 35%. However, the files are small by modern computer standards (under 1000k) and we have not found the lower density to be a problem. On the MCU side, the bottleneck there is speed of flash erase, not the USB bus (which reaches 1000k/s).

The minimal implementation of the UF2 bootloader consumes just 1–2 kB of flash memory and less than 100 bytes of RAM, with some variability due to the microcontroller instruction set and USB hardware interfaces in use.

⁶ See <https://github.com/microsoft/pxt-common-packages>.

6. Systems evaluation

Our platform has been actively deployed for over a year, bringing the benefits of a safe programming environment for MCUs to hundreds of thousands of active users. In this section we provide a broad, quantitative evaluation of the cost at which these benefits are realized. We do this with several micro-benchmarks that give insight into the performance of MakeCode and CODAL across the Uno, micro:bit, and CPX devices. We break down results by layer (CODAL and MakeCode) to give an insight into how each performs.

6.1. Benchmarks, devices, and methodology

To analyze the performance of our solution, we have written a suite of programs to evaluate different aspects of MakeCode and CODAL on a representative selection of real hardware devices. Throughout, we use the C++ CODAL benchmarks as a baseline; the STS benchmarks show the overhead added by MakeCode. These programs were written in both C++ and STS, and evaluated on the three devices listed in Table 1: The micro:bit (Nordic nRF51 MCU), the CPX (Atmel ATSAM21 MCU), and the Uno (Atmel ATmega MCU).

The Uno is the simplest of these devices, consisting of an 8-bit processor running at 16 MHz, with only 2kB of RAM and 32kB of flash. The micro:bit has a 32-bit Cortex-M0 clocked at 16MHz, with 16kB RAM and 256kB of flash. The CPX is a 32-bit Cortex-M0+, which offers greater energy efficiency and performance; it clocks at 48 MHz, has 32kB of RAM and 256kB of flash. The Uno and micro:bit MakeCode targets use the untagged compilation strategy, while the CPX target uses the tagged strategy (see Section 3.3). The benchmarks are classified into two types, each with their own methodology:

1. *Performance Analysis*: Tests that capture time taken to perform a given operation. For these benchmarks, we toggle physical pins on the device at key points in the test code. We then measure the time to execute the operation, by using a calibrated oscilloscope observing these pins. This allows us to derive highly accurate real time measurements without biasing the experiment.
2. *Memory Analysis*: Tests that capture the RAM or FLASH footprint of a certain operation. A map of memory is logged before and after the execution of an operation, allowing us to compute the cost. A serial terminal captures the output of these tests.

Note that memory and performance analysis are done in separate runs to ensure logging does not affect time-related measurements.

6.2. Tight loop performance

To place the performance of MakeCode in context, we perform a comparative evaluation of MakeCode against two state-of-the-art solutions adopted by educators in the classroom, using native C++ as our baseline. The two points of comparison are MicroPython [20], an implementation of Python for MCUs, and Espruno [45], an implementation

Table 2

A comparison of execution speed between: native C++ with CODAL; MakeCode compiled to native machine code; MakeCode compiled to AVR VM; MicroPython; and Espruino. The first line lists the C++ time, while subsequent lines are slowdowns with respect to the C++ time. The 6.4x slowdown of MakeCode VM compared to native MakeCode on the Uno is compensated with 5x better code density.

	UNO	micro:bit	CPX
CODAL	171ms	102ms	31ms
MakeCode	2.4x	2.1x	7.3x
MakeCode VM	15.3x	–	–
MicroPython	–	101x	183x
Espruino	–	1139x	–

of JavaScript for MCUs. For the CPX, a fork of MicroPython known as “CircuitPython” was used. Both MicroPython and Espruino use virtual machine (VM) approaches.

To give an indicative general case execution time cost of each solution, we created a simple program that counts from 0 to 100,000 in a tight loop in each solutions’ respective language; the results are shown in Table 2. On AVR we count to 25,000 (to fit within a 16 bit int) and scale up the results.

For MicroPython and Espruino on the micro:bit, the run is *two or more orders of magnitude slower* than a native CODAL program. MakeCode performs only 2x slower. The slowdown reflects the simple code generator of our STS compiler. It should be noted that MakeCode for the CPX uses the tagged approach, which allows for seamless runtime switching to floating point numbers, resulting in a further 3x slowdown. For both devices, we can observe that MakeCode outperforms both the VM-based solutions of MicroPython and Espruino by at least an order of magnitude.

MicroPython and similar environments cannot run on the Uno due to flash and RAM size limitations. We also ran into these limitations, and as a result, developed two compilation modes for AVR. One compiles STS to AVR machine code, and the other (MakeCode VM) generates density-optimized byte code for a tiny (~500 bytes of code) interpreter. The native strategy achieves code density of about 60.8 bytes per statement, which translates into space for 150 lines of STS user code. The VM achieves 12.3 bytes per statement allowing for about 800 lines. For comparison, the ARM Thumb code generator used in other targets achieves 37.5 bytes per statement, but due to the larger flash sizes we did not run into space issues.

6.3. Context switch performance

To evaluate the performance of CODAL’s scheduler we conducted a test that created two fibers, continuously swapped context, and measured the time taken to complete a context switch. We performed this test in both STS and C++ and the resulting profiles can be seen in Fig. 8, which breaks the context switch down into three phases: (1) CODAL, the time it takes to perform a context switch in CODAL; (2) Stack, the time taken to page out the MakeCode stack; and (3) MakeCode, the overhead added by MakeCode.

From these results, we observe that context switches generally take tens of microseconds. The cost of CODAL’s stack paging approach can also be a significant, but not dominant cost. The cost of stack paging would of course grow with stack depth. Fig. 9 profiles the time a context switch takes with an increasing stack size across all three devices in CODAL. This is similar to the previous test, except we placed bytes (in powers of 2) on the stack of each fiber, starting from 64 and finishing at 1024. The difference in gradients, and ranges of values can be put down to device capability. For instance, the Uno has an 8-bit word size, which means more instructions are required to copy the stack, this results in a steeper gradient than the other two devices. The vertical band indicates

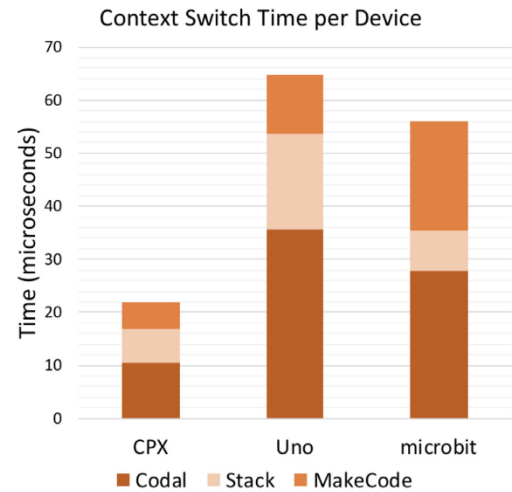


Fig. 8. Base context switch profiles per device.

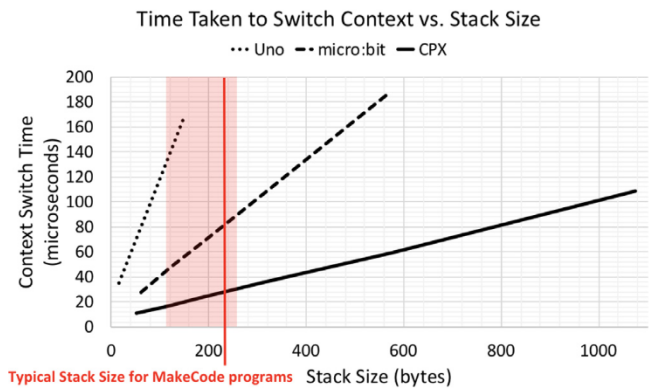


Fig. 9. Time taken to perform a context switch against stack size.

typical stack sizes for MakeCode programs based on a representative set of examples.

6.4. Performance of asynchronous operations

To gauge the cost of asynchronous operations in CODAL, we tested three commonly used code paths, designed to determine the efficiency of CODAL’s *fork-on-block* Asynchronous Procedure Call (APC) mechanism that underpins all event handlers in MakeCode and CODAL. We measured the RAM and processor cost of: (1) creating a fiber; (2) handling a non-blocking APC call; and (3) handling a blocking APC call. We used the CPX for this experiment.

Non-blocking APC calls, the best case, have a small overhead of 32 bytes of RAM and 4.01 microseconds of processing time. Blocking APC calls, the worst case, incur a large overhead of 204 bytes of RAM and 32.4 microseconds of processor time. Creating a fiber costs 136 bytes of RAM and 35.4 microseconds of processing time. These results highlight the performance gains of the opportunistic fork-on-block mechanism over a naive approach that would execute every event handler in a separate fiber.

6.5. Flash memory usage

MCUs make use of internal non-volatile FLASH memory to store program code. Table 3 shows the per device flash consumption of each software library used in the final MakeCode binary. To obtain these numbers, we analyzed the final map file produced after compilation. The ordering of the table aligns with the composition of the software layer: MakeCode builds on CODAL which builds on the C++ standard

Table 3
Flash consumption of a MakeCode binary (kB).

	CPX	micro:bit	Uno
MakeCode	20.46	12.14	7.79
CODAL	29.85	34.35	13.7
Supporting Libraries	14.99	24.28	–
C++ Standard Library	43.14	24	1.03

Table 4
Static RAM consumption of a MakeCode binary (kB).

	CPX	micro:bit	Uno
MakeCode	0.612	1.069	0.074
CODAL	0.369	0.214	0.156
Supporting Libraries	0.312	0.923	–
C++ Standard Library	0.161	0.149	0.074

library and supporting libraries. MakeCode and CODAL consume 108 kB of flash, whereas CircuitPython consumes 201 kB, MicroPython consumes 228 kB, and Espruino consumes 142 kB of flash. This means that users can write sizeable applications in MakeCode, without the worry of running out of flash memory.

From the bottom up, the profile of the standard library changes dramatically for each device: The Uno has a very lightweight standard library; the micro:bit uses 64-bit integer operations (for timers) which requires extra standard library functions; and the CPX requires software floating point operations pulling in more standard library functions.

The size of CODAL and MakeCode scales linearly with the amount of functionality a device has, due to the component oriented nature of CODAL and transitively MakeCode. For instance, the Uno has few onboard components when compared to the CPX and micro:bit. The modular composition of CODAL allows us to support multiple devices with a variety of feature sets, while maintaining the same API at the MakeCode layer.

6.6. RAM memory usage

Table 4 shows the per device RAM consumption of each software library used in the final MakeCode binary. To obtain these numbers, we analyzed the final map file produced after compilation. At runtime, MakeCode dynamically allocates additional memory: 1.56 kB for the CPX, 560 bytes for the micro:bit, and 644 bytes for the Uno. We also can see that in all cases, the RAM consumption of MakeCode and CODAL is well within the RAM available of each device.

MakeCode and CODAL consume a small amount of resources in comparison: CircuitPython (a derivative of MicroPython) consumes 12.8 kB, MicroPython consumes 9.5 kB, and Espruino consumes 5.3 kB of RAM. On the micro:bit, the Bluetooth stack requires 8 kB of RAM to operate. Due to MicroPython's RAM consumption this means that Bluetooth is inoperable. Comparatively, Espruino does enable the Bluetooth stack, but users have just ~300 bytes available for their programs due to the overhead incurred.

6.7. Compiling static typescript

During compilation, the entire STS program (including the STS runtime) is passed to the TypeScript (TS) language service for parsing. Then, only the remaining part of the program (after code shaking) is compiled to native code. On a modern laptop, using Node.js, TS parsing and analysis takes about 0.1ms per statement, and MakeCode compilation to native code takes about 1ms per statement. While the TS compiler has been optimized for speed, MakeCode's native compilation process has not. For example, the CPX TS pass is dominated by compilation of the device runtime and takes about 100ms, whereas the MakeCode pass typically only includes a small user program and a small bit of the runtime,

resulting in less than 100ms. Thus, compilation times are under 200ms for typical user programs of 100 lines or less.

6.8. Extensibility

Adding a new device in CODAL is trivial once a MCU has been ported. The porting of a MCU is where we observe the largest development overhead, as low-level implementations of drivers for I2C, Serial, and SPI may have to be re-written. Due to CODAL's abstraction model, once low-level drivers have been implemented, drivers for higher level components like Accelerometers (which depend on high-level interfaces for low-level drivers) can be immediately adopted if hardware is present. A similar technique is used in MakeCode for simulators.

7. User evaluation

In order to determine the overall usability of MakeCode and CODAL for end users, we conducted a workshop to evaluate the experience of novice programmers when using MakeCode and one of the supported physical computing devices, the BBC micro:bit. We wanted to determine how easily they engaged with the platform, and identify the challenges they faced, specifically seeking answers to the following questions:

1. *Visual Programming Language*: How do novice users engage with the block-based paradigm when applied to embedded computing? Do novice users find it easier to create software using synchronous or asynchronous paradigms (Fig. 10)?
2. *Programming Editor*: Does the design of the editor introduce any additional barriers for physical computing? Does the presence of an interactive hardware simulator affect the success of participants?
3. *Physical Device*: How do novice users engage with the physical device during the development cycle?

7.1. Demographics

To help answer these questions, we ran a workshop with 50 participants. Participants were over the age of 16, and encouraged to work in a group size that was comfortable for them (commonly groups of 2), with a leader completing the programming tasks. All were volunteers, recruited through a public engagement event, groups were self-forming and there was an even gender split. In total there were 30 groups, three groups requested to be omitted, and 2 groups failed to fill out the questionnaire correctly; only valid groups are considered for analysis. Of the lead programmers from each group, the majority (91%) had no or very limited programming experience. One lead programmer had significant prior experience.

7.2. Methodology

7.2.1. Workshop description

The workshop took 45 minutes. Participants were introduced to the MakeCode environment in a short 5 minute presentation, which explained the Block language, and guided them through the creation of a simple example program that scrolled text on the display, flashing the final program to the micro:bit. Participants were then asked to complete three physical computing programming tasks in an instrumented MakeCode environment, and an optional questionnaire aiming to capture qualitative information. We also made a record of any questions that were asked during the session. Participants were randomly allocated to one of two versions of MakeCode to mitigate any layout bias that may influence the blocks that the users chose. Version A had all asynchronous (Fig. 10) blocks moved to the top of the toolbox, whereas version B had all asynchronous blocks moved to the bottom of the toolbox.

The programming tasks began simply, and advanced to controlling a physical remote control vehicle using the radio interface for the final task. The participants were given the following instructions:

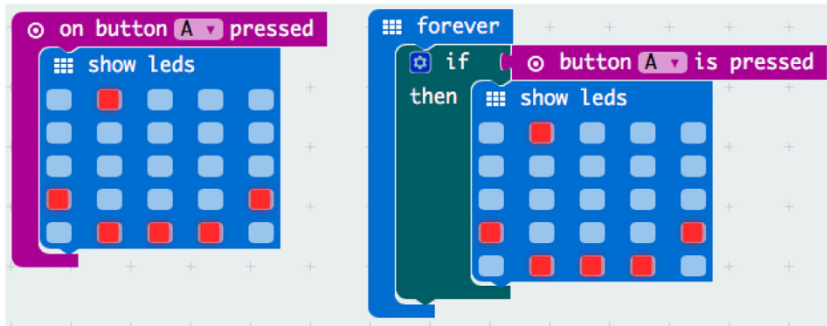


Fig. 10. Equivalent operations using asynchronous (left) or synchronous (right) paradigms.

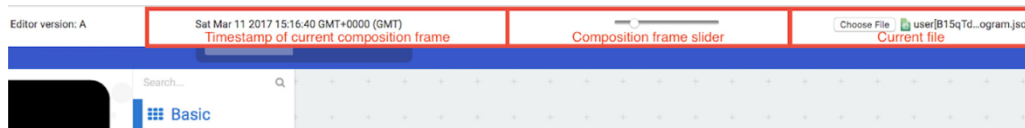


Fig. 11. MakeCode with the program reconstruction plugin (top). The slider will move through frames of program composition in the loaded log file.

1. Write a program that displays your name when you click button A. (Hint look for a block in the editor that mentions buttons)
2. Write a program that shows a winking smiley when you shake the device. (Hint look for a block in the editor that mentions shake)
3. Write a program to drive this remote controlled car using Radio. (Hint to get the car to move you need to send numbers using micro:bit radio in group number 52.)

The first task aimed to determine the paradigm groups implicitly picked, and could be solved either synchronously or asynchronously (Fig. 10). Participants were given *no assistance* and a strict time limit of 11 minutes for the first task, to gauge how well users engaged with the environment with minimal introduction. For the latter two tasks, limited assistance was offered by the observers.

7.2.2. MakeCode instrumentation

Our primary data source was obtained from a custom version of MakeCode that had increased logging, allowing us to capture mouse movements and clicks, simulator interactions, and program composition over time. Tracking of mouse movements was restricted to once every 25 milliseconds, and program composition was stored on every change offering greater insights into how groups composed their programs.

To analyze program composition over time, we developed a new plugin for MakeCode that enables step-by-step reconstruction of tracked programs. MakeCode is built on Blockly [18], which offers a simple XML export capability for portability; we exploit this feature for program composition tracking. Upon every program change, we captured the entire program XML generated by Blockly.

The plugin (Fig. 11) has a simple interface, showing date and time information for the current composition frame, a slider to move through the tracked frames of composition, and the currently selected file. We used this tool to step through user programs and extract commonalities between groups, gathering information on how far they progressed, whether they completed the first task in the allotted time, and the composition of their programs including discarded and unused blocks.

Finally, analytics were added to MakeCode to capture engagements with the simulator, allowing us to determine whether heavy usage of the simulator correlated with participant progression.

7.3. Overall results

The following sections present the key findings grouped by research question. They combine data from our three collection sources: MakeCode analytics, qualitative data from questionnaires, and in-workshop questions.

For context, 51% of groups completed the first task within the time limit and *without assistance*, 70% of our groups completed the second task, and 30% completed the final radio task involving the remote controlled car. 22% completed none of the tasks, of these, two groups did not attempt any tasks. We analyzed those who did not attempt any tasks, and from their logs we determined that one group curated programs that were not relevant to the tasks, and the other did not make any attempt to engage with the environment. All data is used to provide an indication of how the environment enabled groups to progress, and is not a reflection on learning.

7.3.1. Results: visual programming language

Users Preferred Asynchronous Blocks. Asynchronous blocks are event-driven blocks that execute when a condition is met, rather than a synchronous, polling-based, approach (Fig. 10). Of the 51% that completed the first task within the time limit, 43% used solely asynchronous blocks to solve the task, 50% began with a synchronous approach but completed the task using asynchronous blocks, just 7% completed the task using synchronous blocks only, and *no* participants attempted the task asynchronously and finished synchronously. The participant who solved the task synchronously completed the second task asynchronously, and utilised asynchronous button blocks to send commands to the remote control car in the final task. The version of the editor had no statistical significance on either approach.

Blocks are Not Intuitive for All Users. Although 51% of groups managed to complete the first task unsupervised and within the time limit, two thirds of groups experienced difficulties when using the block language. Fig. 12, right, shows the two most prolific examples we saw.

In Blocks, capped blocks are a way of protecting the user from creating invalid programming constructs or combinations of blocks that could produce a confusing result. Groups commonly tried to connect these capped blocks to other capped blocks, placing conditionals directly into capped-blocks rather than logic or variable assignment blocks. With a largely inexperienced audience who had a minimal introduction to the environment, this shows that the simplicity of a “jigsaw” model was not immediately intuitive to all of our groups. Feedback captured by observers reinforces this view, “Why can’t I plug ‘on shake’ (capped block) into my forever block (capped)?”, and “Where do I plug this (condition) block into?”. In contrast, one participant “Thought the blocks interconnect made it easy to see where things could be placed within the program”.

7.3.2. Results: programming editor

Program Validation was Confusing. When a block is not connected to any other blocks, or a duplicate event-based block handler is added,

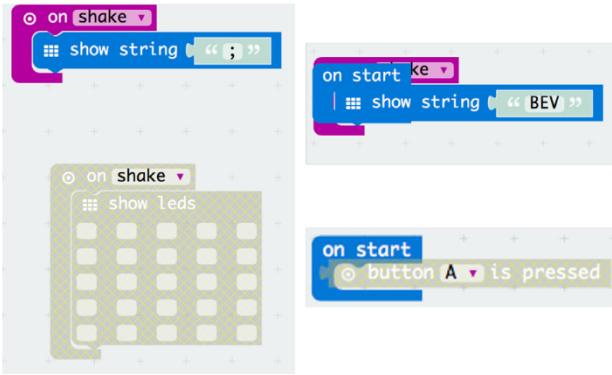


Fig. 12. Program Validation: (left) If there are duplicate event-based blocks registered by the user, one is greyed out; in this case the bottom *on shake* event handler duplicates the functionality of the top handler. (right, top) Groups tried to connect capped (no connection dimple at the top) blocks to other capped blocks; (right, bottom) and conditionals to capped blocks.

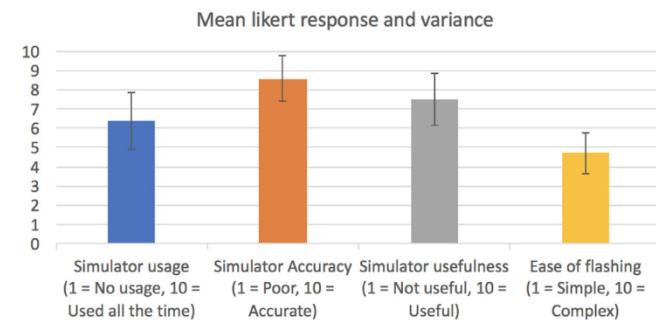


Fig. 13. Likert ratings captured from our questionnaire.

the block is greyed out (Fig. 12, left). Given that 75% of groups dragged blocks onto the canvas before connecting them to a valid block programming construct, this feature caused confusion for groups. Two groups asked why their blocks were greyed out as soon as they dragged them onto the canvas. From looking at program composition, we see that four groups ran into similar issues. The effect of this feature is exhibited in Fig. 12 (left), where perfectly valid code is displayed as invalid causing confusion when the program does not behave as expected.

The Simulator Improved Participant Progression. Responses to our questionnaire indicate (based on the mean rating in Fig. 13) that participants-groups found simulations accurate and useful. Groups optionally could give feedback on their experiences with the simulator: “It was good, I didn’t use it much as I had the micro:bit and so wanted to see the real effect. Helped to see if code was ok...”. Two groups found the simulator constraining for the third task involving the remote control car, commenting that there was “No ability to simulate the radio task”. One participant had real difficulties understanding the simulator.

Based on the responses to the questionnaire (Fig. 13; Simulator Usage), we see that groups were unsure of how often they used the simulator. We therefore used the MakeCode analytics to understand how many groups engaged with the simulator—where an engagement is defined as a click of any button: Button A, Button B, or the Shake button. The average number of engagements for all groups was $\mu = 50.5$ with a variance of $\sigma^2 = 22.4$. To understand the source of this high variance, we grouped participant-groups by their highest task completed and calculated the ratio of simulations to compiles (Fig. 14). One third of groups who advanced no further than task one compiled far more than they simulated, with the ratio inverted for those who advanced further than task one. This indicates that those who used the simulator advanced further than those who did not. We attribute the lack of testing via the simulator to incomplete assimilation with the environment.

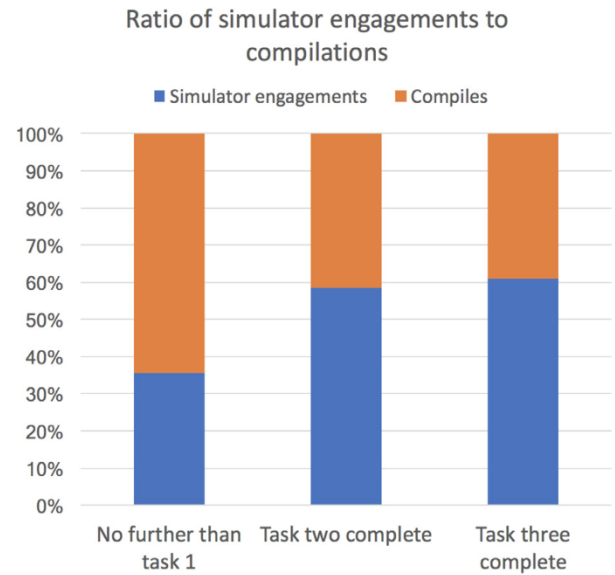


Fig. 14. Simulations vs. compilations.

7.3.3. Results: physical device

Flashing the Device was Challenging. Questionnaire data indicates that the flashing process was moderately difficult (Fig. 13; Ease of Flashing). Qualitative data describes a similar story with participant-groups stating that it “Took a while to understand the steps that had to be used” and “Once shown the process it was straightforward but not obvious without being given a hint”. Workshop assistants recorded 11 different questions from groups throughout the workshop, with one group asking “I hit download but my program is not running on my micro:bit, why?”. The flashing process involves three steps: (1) in-browser compilation, a user hits compile in the MakeCode editor; (2) the compiled program is downloaded to the PC; and (3) the user saves or “drags” the compiled program onto the micro:bit, which appears as a USB flash drive. This process was clearly not simple and intuitive for our participants.

7.4. Discussion and implications

In this section we reflect on our results, providing implications for the design of web-based physical computing. We begin by reflecting on the visual programming language, following with recommendations for improvements to the web-based programming editor generalisable to web-based editors and other visual programming environments. Finally, we reflect on how the move to the web impacts the user experience of flashing the device.

7.4.1. Visual programming language

Use Asynchronous Blocks more Widely. Given the demographics, it’s interesting to note that the majority completed the tasks using asynchronous, event-based blocks. 93% of groups used event-based blocks to complete the first task instead of their synchronous equivalents. A possible reason for this is that the participants relied heavily on the hints given alongside the tasks, resulting in them picking the first block they saw, rather than the most intuitive block. However, we can confidently say that inexperienced programmers found the asynchronous blocks the simplest way to solve the task, reinforcing previous observations [30,41]. We recommend that any visual programming environment for physical computing offer asynchronous blocks to increase the accessibility of physical computing for inexperienced users.

7.4.2. Programming editor

Program Validation Requires Explanation. Greying out disconnected blocks caused confusion. 75% of groups chose to build programs by dragging blocks from the toolbox to the canvas, dropping them, and finally connecting them to the desired block. Using the canvas as an

intermediary step caused confusion as the block was greyed out while in the disconnected state. Further, when two identical handlers are dragged onto the canvas one is greyed out—another form of program validation. One can imagine the confusion this could cause an inexperienced programmer: why does the environment allow many ‘forever’ blocks but only one ‘on shake’ block? Why are blocks greyed out whilst I am composing a program? Conversely, in Scratch, multiple event handlers can be registered and executed sequentially, allowing users to question why programs behave unexpectedly [29]. Each approach ultimately achieves the same goal, further user studies are required to determine which approach is better. In the case of MakeCode, a simple contextual explanation as to why a block was greyed out would sufficiently reduce user confusion and better scaffold their experience.

Hardware Simulators Should be Included in all Environments. Simulating the hardware was a positive experience for groups. We recommend that future programming editors include the capability to simulate hardware, reducing the need to flash the device and offering easy program validation. We found that the groups who used the simulator more progressed further. There was one clear area for improvement from our workshop: many projects involve external components or devices, the inability to simulate these rendered the simulator useless when groups were attempting the third task (the radio controlled car). This has wider implications on other physical computing communities beyond the micro:bit, including the Arduino, where Circuit Labs, a circuit prototyping solution, already simulates off-board components.

8. Related work

8.1. Novice programming environments

Arduino [39] is an environment for programming microcontrollers, aimed at novices. However, its C++ based APIs introduces barriers for novice programmers [10]. Scratch [36] is a widely adopted, event-based visual programming environment designed to introduce novice programmers to computer science concepts. Extensions enable the programming of physical devices with Scratch. However, devices require constant tethered connections to operate, restricting potential projects [16]. ArduBlock [2] brings visual programming to the Arduino, but it lacks the event-based blocks Scratch users are familiar with.

There are well-known benefits to using such visual programming languages: a simpler programming vocabulary, lower cognitive load, and easy assembly, making program composition less error prone [9]. Prior work in this area focuses on evaluating the learning of participants using Scratch, rather than the usability of the language or ecosystem. Furthermore, prior studies are longitudinal, evaluating the language over a number of sessions, hence, participants are experienced with the language by the end of the study and do not capture the learning curve involved [24,46].

Event-based programming is present in Scratch, and it has been noted that it is used preferentially by novice users [29,30,41]. In addition, Fraser provides valuable reflections on learnings from the development of Blockly [19]. We observe the same preference towards event based blocks, and offer reflections on different design decisions new environments are taking using Blockly.

With the environments above, additional software must be installed—this creates barriers for novice users in restrictive environments [14]. MakeCode and CODAL require *no installation* to support a diverse user base and support *event-based higher-level languages* to help beginners get a head start in the world of the microcontroller.

8.2. Hardware simulation

Some programming environments for physical computing include hardware simulators for use without a device in order to increase adoption and allow for easy debugging and testing before deploying to a real-

world device [12,13,21,34,38]; circuit labs⁷ even simulates off-board components. However, simulations in existing literature involve hardware register simulation rather than real-time, visual simulation of the device, resulting in a gap in our understanding of how interaction-level simulators are used [23,25].

8.3. Virtual machine-based languages

Recently, virtual machines supporting most of the semantics of higher level languages like JavaScript, Java, and Python, have been ported to 32-bit microcontrollers by maker communities [16]. Examples include: MicroPython [20], CircuitPython, and Espruino [45]. These VMs consume a large amount of RAM and flash memory, and run significantly slower than native languages.

The research community has worked to bring higher level languages to microcontrollers [26,40,44]. Rather than running a full-featured VM, others enable higher level languages to run efficiently by stripping out advanced language features, in favour of efficient, native execution [43]. Comparing these solutions to our solution is challenging due to a misalignment in evaluation metrics and microcontrollers. For example, the PICOBIt uses an 8-bit MCU, and evaluates the cost of a VM, without the cost of a runtime environment. Simply accounting for a 32-bit MCU in this case, results in factor of 4 multiplication of most metrics.

Our approach bears most similarity to [43], where we compile higher level languages to an *optimized, event-driven C++ runtime* (CODAL).

8.4. Embedded runtime environments

Arduino [39] is an example of a simple platform where the developer uses high-level APIs to control hardware; there is no scheduler and memory management is discouraged, with a heavy emphasis on the use of global variables.

TinyOS [27], Contiki [17], RIOT OS [7], Mynewt [4], mbed OS [6], and Zephyr [5] are RTOS solutions known widely in the systems community. The majority focus on the networking features of sensor based devices and commonly adopt a preemptive scheduling model, which leads to competition over resources resolved using locks and condition synchronization primitives. Contiki has a cooperative scheduler but uses proto-threads to store thread context—local variables are not allowed as the context of the stack is not stored.

Although the platforms above are widely used by C/C++ developers, none of these existing solutions align well with the programming paradigms seen in higher level languages. CODAL *bridges the semantic gap between the higher level language and the microcontroller*, offering appropriate abstractions and higher level primitives written natively in C++.

8.5. Flashing microcontrollers

There are two common ways to transfer a program to the flash of a microcontroller: for embedded developers, a specialized debugger chip; for hobbyists, a custom serial protocol [1]. Both approaches require operating system drivers. ARM’s mbed platform provides DAPLink [3], firmware that presents itself to an external computer as a USB pen drive. DAPLink exposes a virtual file system that caters for normal file system behaviour and handles the decoding of Intel HEX files [22]—the firmware consumes 66 kB of flash and 13 kB RAM. UF2 contributes a new file format that *greatly simplifies* the virtual file system approach, reducing complexity of the firmware and code size.

9. Conclusion

We have presented MakeCode: a *no installation*, web-based programming environment, that supports novice programmers with *block-*

⁷ <https://circuits.io/lab>.

based and text-based higher-level languages, and compiles programs in the browser. So as to not compromise the spatial efficiency of the microcontroller, we created CODAL: a C++ runtime that *bridges the semantic gap* between higher level languages in MakeCode and C++. To transfer programs compiled by MakeCode to the microcontroller without the installation of any drivers, we created UF2: a new bootloader and file format that enables the *simplified, driverless* programming of microcontrollers. Combined, our approach to running higher level languages on microcontrollers is up to 50x more performant compared to other approaches. Further, by using modern tooling, and higher level languages, our approach lowers the barrier to entry for microcontroller programming.

We have also presented a user evaluation of MakeCode and CODAL via a structured workshop under controlled conditions. This evaluation showed that novice users preferred, and were much more successful, when using simplified asynchronous programming constructs than they were when using synchronous constructs. It also showed that novice programmers typically benefit from a device simulator, but also revealed that this simulator needs to model the whole system, not just a microcontroller. Finally, it highlighted that the simplified, universal programming experience offered by UF2 still creates a barrier to entry for users, and further work is required to make this process fully transparent. These conclusions further reinforce the informal observations we have seen from the user community, where hundreds of thousands of MakeCode programs are written every month (at the time of writing).

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank the members of the MakeCode team for their many contributions to the success of the platform: Abhijith Chatra, Sam El-Husseini, Caitlin Hennessy, Guillaume Jenkins, Richard Knoll, Galen Nickel, Jacqueline Russell, and Kesavan Shanmugam.

Appendix A. Artifact appendix

Submission and reviewing guidelines and methodology: <http://cTuning.org/ae/submission.html>.

A.1. Abstract

This artifact allows others to reproduce the results seen in this paper for MakeCode and CODAL, using the BBC micro:bit. The artifact contains an offline build environment for CODAL and MakeCode, allowing evaluators to test and build programs locally. In addition, we also provide espruino and micropython virtual machines to further increase repeatability of our results. Evaluators should download the virtual machine containing all pre-requisite tools, and use an oscilloscope to observe wave forms (used for timing) generated by the micro:bit, and a serial terminal to observe results reported from the micro:bit over serial.

A.2. Artifact check-list (meta-information)

- **Program:** MakeCode & CODAL
- **Compilation:** arm-none-eabi-gcc
- **Binary:** espruino, and micropython binaries included; others compiled during testing
- **Run-time environment:** CODAL
- **Hardware:** BBC micro:bit
- **Output:** Waveforms, and serial output
- **Publicly available?:** Yes

- **Artifacts publicly available?:** Yes
- **Artifacts functional?:** Yes
- **Artifacts reusable?:** Yes
- **Results validated?:** Yes

A.3. Description

A.3.1. How delivered

The artifact is available hosted on GitHub: <https://lancaster-university.github.io/lctes-artefact-evaluation/> Alternately, the latest release is available for download: <https://doi.org/10.5281/zenodo.1242627> Finally, a virtual machine, based on debian, containing all the required software to reproduce our results is available here: <https://doi.org/10.5281/zenodo.1242605>.

A.3.2. Hardware dependencies

- A BBC micro:bit
- An oscilloscope
- A computer capable of running a virtual machine

A.3.3. Software dependencies

- A virtual machine obtained from the URL above.
- A serial terminal.

A.4. Installation

Use virtual box to install the image located at: <https://drive.google.com/open?id=1nxiorz6NRqjen89G59RCOEMklqAyaUv7> and the VirtualBox extension pack: <https://www.virtualbox.org/wiki/Downloads>.

A.5. Experiment workflow

Tests generally follow the following sequence of steps:

1. Perform small program modifications.
2. Compile the program.
3. Transfer program to the micro:bit (flashing).
4. Observe either a waveform generated by the micro:bit using an oscilloscope, or serial output from the micro:bit using a serial program.

A.6. Evaluation and expected result

We expect the results to be the same as those reported in the paper. The observed waveforms may differ in time due to different compilers, oscilloscopes, and oscilloscope calibration.

A.7. Experiment customization

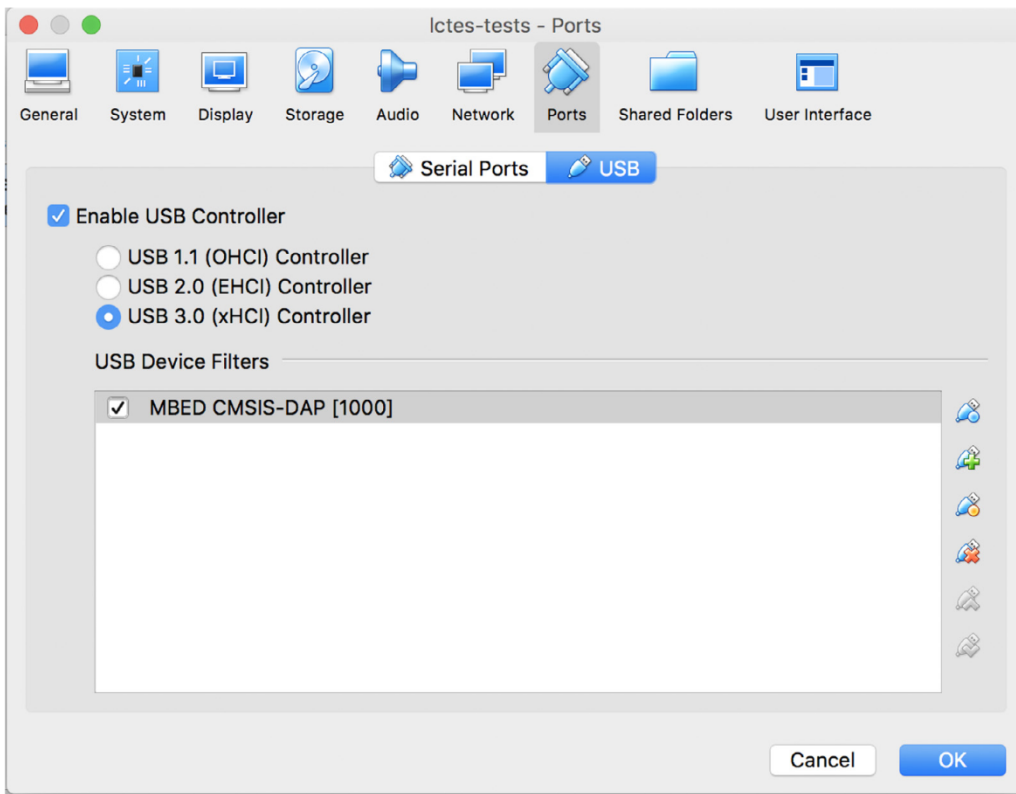
All tests provided have a clear set of corresponding instructions that evaluators should follow to observe the same results. Any steps involving customisation have been minimised.

A.8. Notes

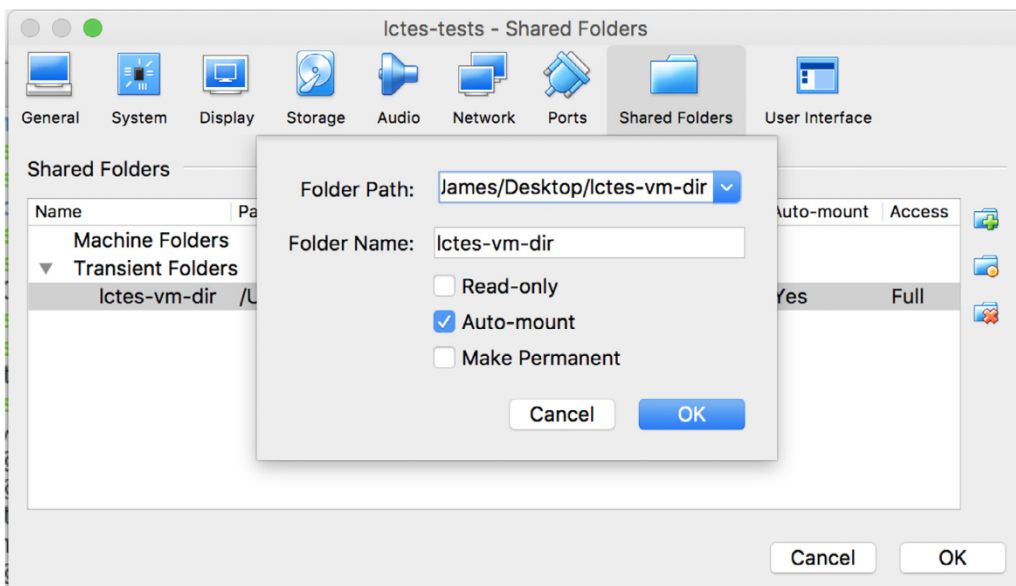
The virtual machine contains a folder named ‘evaluators’ which is placed in the home directory of the lctes user. The username for the virtual machine is: *lctes* and the password is: *lctes2018*. To become super user, type *su* in a terminal, and enter the same password (*lctes2018*).

Once logged in, and in the ‘evaluators’ directory, you can view the tests as markdown files in the ‘docs’ directory. Alternately, these markdown documents can also be viewed on the web by running ‘mkdocs serve’ in the evaluators folder, or browsing to: <https://lancaster-university.github.io/lctes-artefact-evaluation/> Which is a pre-built, and hosted version produced from the same source.

We recommend that you add the micro:bit usb device using the machine settings tab in virtual box as shown in the image below:



We also have a convenience script for mounting a shared folder between the host and the vm. Simply create a shared folder named 'lctes-vm-dir' and run 'sh mount.sh' (contained in evaluators) as a super user to mount the shared folder to vb-share (also contained in evaluators). Shared folder creation in VirtualBox is pictured below:



Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.sysarc.2019.05.005.

References

- [1] Avrdude - avr downloader/uploader, 2010, (<https://www.nongnu.org/avrdude/>) (Accessed on 02/22/2018).
- [2] Ardublock — a graphical programming language for arduino, 2012, (<http://blog.ardublock.com/>) (Accessed on 02/22/2018).
- [3] Github - armmbed/daplink, 2013, (<https://github.com/ARMMbed/DAPLink>) (Accessed on 02/22/2018).

- [4] Apache mynewt, 2015, (<https://mynewt.apache.org/>) (Accessed on 11/16/2017).
- [5] Home - zephyr project, 2017, (<https://www.zephyrproject.org/>) (Accessed on 11/16/2017).
- [6] ARM, The arm mbed iot device platform (2017).
- [7] E. Baccelli, O. Hahm, M. Gunes, M. Wahlsch, T.C. Schmidt, Riot os: towards an os for the internet of things, in: Computer Communications Workshops (INFOCOM WKSHPs), 2013 IEEE Conference on, IEEE, 2013, pp. 79–80.
- [8] M. Banzai, M. Shiloh, Getting Started with Arduino: The Open Source Electronics Prototyping Platform, Maker Media, Inc., 2014.
- [9] D. Bau, J. Gray, C. Kelleher, J. Sheldon, F. Turbak, Learnable programming: blocks and beyond, *Commun. ACM* 60 (6) (2017) 72–80.
- [10] P. Blikstein, Gears of our childhood: constructionist toolkits, robotics, and physical computing, past and future, in: Proceedings of the 12th International Conference on Interaction Design and Children, ACM, 2013, pp. 173–182.
- [11] R.F. Bruce, J.D. Brock, S.L. Reiser, Make space for the pi, in: SoutheastCon 2015, IEEE, 2015, pp. 1–6.
- [12] Calliope, Calliope — home, (<https://calliope.cc/en>). (Accessed on 01/03/2018).
- [13] CodeBug, Codebughome, (<https://www.codebug.org.uk/>) (Accessed on 01/03/2018).
- [14] M.J. Cox, K. Cox, C. Preston, What factors support or prevent teachers from using ict in their classrooms? (2000).
- [15] J. Devine, J. Finney, P. de Halleux, M. Moskal, T. Ball, S. Hodges, MakeCode and CODAL: intuitive and efficient embedded systems programming for education, in: Proceedings of the 19th ACM SIGPLAN/SIGBED International Conference on Languages, Compilers, and Tools for Embedded Systems, ACM, 2018, pp. 19–30.
- [16] D. Dougherty, The maker movement, *innovations* 7 (3) (2012) 11–14.
- [17] A. Dunkels, R. Quattlebaum, F. Österlind, G. Oikonomou, M. Alvira, N. Tsiates, O. Schmidt, Kontiki: The open source os for the internet of things, Retrieved October 13(2012) 2015.
- [18] N. Fraser, et al., Blockly: A Visual Programming Editor, Google, Place, 2013.
- [19] N. Fraser, Ten things we've learned from blockly, in: Blocks and Beyond Workshop (Blocks and Beyond), 2015 IEEE, IEEE, 2015, pp. 49–50.
- [20] D. George, MicroPython (2017).
- [21] P. Inc., Program with blocklyprop — parallax inc, (<https://www.parallax.com/product/program-blocklyprop>) (Accessed on 01/03/2018).
- [22] Intel, Hexadecimal object file format specification (1988).
- [23] Y. Kato, Splash: a visual programming environment for arduino to accelerate physical computing experiences, in: Creating Connecting and Collaborating through Computing (C5), 2010 Eighth International Conference on, IEEE, 2010, pp. 3–10.
- [24] B. Kaučič, T. Asič, Improving introductory programming with scratch? in: MIPRO, 2011 Proceedings of the 34th International Convention, IEEE, 2011, pp. 1095–1100.
- [25] C. Kelleher, R. Pausch, Lowering the barriers to programming: a taxonomy of programming environments and languages for novice programmers, *ACM Comput. Surv. (CSUR)* 37 (2) (2005) 83–137.
- [26] J. Koshy, R. Pandey, Vmstar: synthesizing scalable runtime environments for sensor networks, in: Proceedings of the 3rd International Conference on Embedded Networked Sensor Systems, ACM, 2005, pp. 243–254.
- [27] P. Levis, S. Madden, J. Polastre, R. Szewczyk, K. Whitehouse, A. Woo, D. Gay, J. Hill, M. Welsh, E. Brewer, et al., Tinyos: an operating system for sensor networks, *Ambient Intell.* 35 (2005) 115–148.
- [28] M. Maksimović, V. Vujović, N. Davidović, V. Milošević, B. Perišić, Raspberry pi as internet of things hardware: performances and constraints, *Des. Issues* 3 (2014) 8.
- [29] J. Maloney, M. Resnick, N. Rusk, B. Silverman, E. Eastmond, The scratch programming language and environment, *ACM Trans. Comput. Educ. (TOCE)* 10 (4) (2010) 16.
- [30] J.H. Maloney, K. Peppler, Y. Kafai, M. Resnick, N. Rusk, Programming by Choice: Urban Youth Learning Programming with Scratch, 40, *ACM*, 2008.
- [31] The micro:bit Educational Foundation, micro:bit : blast off: School launches bbc micro:bit into space!, 2016a, (<https://www.microbit.co.uk/rishworth-space>) (Accessed on 05/13/2017).
- [32] The micro:bit Educational Foundation, micro:bit : use the bbc micro:bit in the bloodhound rocket car contest, 2016b, (<https://www.microbit.co.uk/bloodhound-rocket-car>) (Accessed on 05/15/2017).
- [33] Microsoft, 20160705_microsoft_bloodhound_0409 - microsoft news centre uk, 2016, (https://news.microsoft.com/en-gb/2016/07/07/26785/20160705_microsoft_bloodhound_0409) (Accessed on 05/15/2017).
- [34] OzoBlockly, Welcome to ozoBlockly, (<http://ozoblockly.com/>) (Accessed on 01/03/2018).
- [35] M. Resnick, J. Maloney, A. Monroy-Hernández, N. Rusk, E. Eastmond, K. Brennan, A. Millner, E. Rosenbaum, J. Silver, B. Silverman, et al., Scratch: programming for all, *Commun. ACM* 52 (11) (2009) 60–67.
- [36] M. Resnick, J. Maloney, A. Monroy-Hernández, N. Rusk, E. Eastmond, K. Brennan, A. Millner, E. Rosenbaum, J.S. Silver, B. Silverman, Y.B. Kafai, Scratch: programming for all, *Commun. ACM* 52 (11) (2009) 60–67.
- [37] G. Richards, F.Z. Nardelli, J. Vitek, Concrete types for typescript, in: 29th European Conference on Object-Oriented Programming, ECOOP 2015, 2015, pp. 76–100, doi:10.4230/LIPIcs.ECOOP.2015.76.
- [38] O. Roberta, Open roberta lab, (<https://lab.open-roberta.org/>) (Accessed on 01/03/2018).
- [39] C.R. Severance, Massimo banzi: building arduino, *IEEE Comput.* 47 (1) (2014) 11–12.
- [40] V. St-Amour, M. Feeley, Picobit: a compact scheme system for microcontrollers, in: International Symposium on Implementation and Application of Functional Languages, Springer, 2009, pp. 1–17.
- [41] F. Turbak, M. Sherman, F. Martin, D. Wolber, S.C. Pokress, Events-first programming in app inventor, *J. Comput. Sci. Colleges* 29 (6) (2014) 81–89.
- [42] E. Upton, G. Halfacree, Raspberry Pi User Guide, John Wiley & Sons, 2014.
- [43] A. Varma, S.S. Bhattacharyya, Java-through-c compilation: an enabling technology for java in embedded systems, in: Design, Automation and Test in Europe Conference and Exhibition, 2004. Proceedings, 3, IEEE, 2004, pp. 161–166.
- [44] B. Vaugon, P. Wang, E. Chailloux, Programming microcontrollers in ocaml: the ocaptic project, in: International Symposium on Practical Aspects of Declarative Languages, Springer, 2015, pp. 132–148.
- [45] G. Williams, Making Things Smart: Easy Embedded JavaScript Programming for Making Everyday Objects into Intelligent Machines, Maker Media, 2017.
- [46] A. Wilson, D.C. Moffat, Evaluating scratch to introduce younger schoolchildren to programming, in: Proceedings of the 22nd Annual Psychology of Programming Interest Group (Universidad Carlos III de Madrid, Leganés, Spain, 2010).



James Devine is a PhD student working to receive a doctorate in Computer Science from Lancaster University. James has worked extensively on the CODAL runtime and support for the BBC micro:bit forming the basis of his PhD: investigating accessible tools and techniques for contemporary embedded systems. James' PhD research themes also extend into the world of the Internet of Things, where he aims to simplify this traditionally complex area so it can be more easily used by non-experts.



Joe Finney is a senior lecturer in the School of Computing and Communications at Lancaster University. His research interests include networked mobile systems, support for lightweight embedded systems, and novel mobile applications. Recently, he has also worked as a founding partner of the BBC micro:bit initiative, where worked with colleagues at Microsoft, ARM and the open source community to develop the underlying software for the BBC micro:bit - an inexpensive, lightweight computer designed to inspire the next generation of technologists.



Peli de Halleux is an engineer in the Research In Software Engineering group at Microsoft Research. He is working MakeCode; and previously worked on Touch Develop, Code Hunt, Pex.



Michał Moskal works at Microsoft Research in Redmond on programming languages and related topics. In the past Michał has worked on software verification (including VCC — a state of the art formal verifier for C programs) and automated theorem proving. In the past decade Michał has worked on programming experiences for beginners. In 2010, with Nikolai Tilmann, he started Touch Develop for programming phone on a phone, later adopted a programming environment for the very popular BBC micro:bit. In 2016, with Peli de Halleux and Tom Ball, Michał started Programming Experience Toolkit (PXT), which has grown into Microsoft MakeCode — a platform for creating domain-specific programming experiences for beginners, especially in education.



Thomas (Tom) Ball is a Principal Researcher at Microsoft. In 1999, He initiated the influential SLAM software model-checking project with Sriram Rajamani, which led to the creation of the Static Driver Verifier tool for finding defects in Windows device drivers. Tom is a 2011 ACM Fellow for 'contributions to software analysis and defect detection'. As a manager, he nurtured research areas such as automated theorem proving, program testing/verification and empirical software engineering. Tom's current focus is the Microsoft MakeCode platform for programming with physical computers.



Steve Hodges is a Principal Researcher at Microsoft where he combines device-related research insights with emerging technologies to create new hardware concepts, tools and technologies. By seeding adoption of these beyond the lab he ultimately aims to demonstrate new ways in which technology can empower individuals, organizations and communities. He is particularly interested in physical computing for computer science education. He led the team that developed the .NET Gadgeteer electronic device prototyping system which enjoyed success in schools, and instigated the BBC micro:bit hardware advisory committee which oversees development of micro:bit hardware.