Simplifying the Programming of Microcontroller-based Devices

Microsoft MakeCode and Lancester University Teams

Anonymous Author(s)

Abstract

Text of abstract ....

Keywords keyword1, keyword2, keyword3

1 Introduction

Over the last decade, microcontrollers, the workhorses of embedded systems, have become central to efforts in making and education. For example, the Arduino project (www.arduino.cc), started in 2003, created a printed circuit board (the Uno) based on the 8-bit Atmel AVR microcontroller unit that makes most of its I/O pins available via headers. Hardware modules (shields) may be connected to the main board to extend its capability. The Arduino ecosystem, based on an open hardware design, has grown tremendously in the past 15 years, with the support of companies such as Adafruit Industries (www.adafruit.com) and Sparkfun Electronics (www.sparkfun.com), to name a few.

In contrast, what has not changed much is the way microcontrollers are programmed, which is with the C/C++ programming language (as well assembly). This is not a huge surprise, given the low-level nature of microcontroller programming, where direct access to the hardware is the order of the day. There generally is no operating system running on such boards, as they have very little RAM (2K for the Uno, for example) and lack memory protection hardware. What is more surprising about the Arduino platform is that:

- it encourages the programmer to use polling to interact with sensors, which leads to monolithic sequential programs;
- its IDE lacks any code "intellisense" or common interactive features of modern IDEs;
- it loads code onto the microcontroller using 1980s era bootloader technology.

As a result, it is not simple to get started with Arduino-based systems, of which there are many. On the other hand, on the web we find many excellent environments for introductory programming. Visual block-based editors such as Scratch (https://scratch.mit.edu/) and Blockly (https://developers.google.com/blockly/) allow the creation of programs without the possibility of syntax errors. HTML, CSS and JavaScript allow a complete programming experience to be delivered as an interactive web app, including editing with intellisense, code execution and debugging. (While the Arduino IDE recently has been ported to the web, it lacks many of the above features and requires a web connection to a server which runs a C/C++ compile tool chain to compile user code.) The programming models associated with these environments are event-based, freeing the programmer from the need to poll.

We have created a new programming platform that bridges the worlds of the microcontroller and the web app. The major goals of the platform are to: (1) make it simple to program microcontrollers using an interactive web app that also functions offline; (2) allow a user's compiled program to be easily installed on a microcontroller; (3) support the addition of new of software/hardware components to a microcontroller. The platform is defined by four new technologies:

- *MakeCode* is a new web app (www.makecode.com) that supports both visual block-based programming and text-based programming using TypeScript (www.typescriptlang.org), a gradually-typed superset of JavaScript. The web app also converts between the two program representations. The web app supports in-browser execution via a device simulator, as well as compilation to machine code and linking against a pre-compiled C++ runtime. No C/C++ compiler is invoked to compile user code. The result of compilation is a binary file that is "downloaded" from the web app to the user's computer.
- *Static TypeScript* is a statically-typed subset of TypeScript for fast execution on low-memory devices and a simple model for linking against pre-compiled C++; Static TypeScript also can be used to write safe device driver code.
- *CODAL*, the Component-oriented Device Abstraction Layer, is a new C++ library that maps each hardware component of a device to one or more software components that communicate over a message bus and schedule event handlers to run non-preemptively on fibers.
- *USB Flashing Format* (UF2) is a new file format designed for flashing microcontrollers over the Mass Storage Class (removable USB pen drive) protocol. This new file format greatly speeds the installation of user programs and is robust to difference in operating systems.

These advances enable beginners to get started programming microcontrollers from any modern web browser, and enable hardware vendors to innovate and safely add new

PL'17, January 01–03, 2017, New York, NY, USA
2017.

components to the mix using Static TypeScript, and its foreign function interface to C++. Once the web app has been loaded, all the above functionality works offline (i.e., if the host machine loses its connection to the internet). All of the above components are open source under the MIT/Apache licenses, as detailed below.

Platform targets can be seen at www.makecode.com, where the MakeCode web app for a variety of boards is available, including the micro:bit (a Nordic nRF51822 microcontroller with Cortex-M0 processor, 16K RAM), Adafruit's Circuit Playground Express (CPX: an Atmel SAMD21 microcontroller with Cortex-M0 processor, 16K RAM), and the Arduino Uno (Uno: an Atmel ATmega328 microcontroller with AVR processor, 2K RAM).

We encourage the reader to choose a board and experiment with programming it, using the simulator to explore many of each board's features, to appreciate the qualitative aspects of the platform: its simplicity and ease of use. In this paper, we will evaluate quantitative aspects of the platform: compilation speed, code size, and runtime performance. In particular, we evaluate:

- the compile time of Static TypeScript compile/link of user code (to machine code) with respect to the GCC-based C/C++ toolchain, as well as the size of the resulting executable;
- the time to load code onto a microcontroller using UF2, compared to standard bootloaders;
- The performance of a set of small benchmarks, written in both Static TypeScript and C++, compiled with the MakeCode and GCC-based toolchains, as well as the performance of device drivers written in Static TypeScript compared to their C++ counterparts.

[evaluate with respect to the popular Arduino toolset, for boards with 8-bit (AVR) and 32-bit (Cortex-M0) microcontrollers. Summary of evaluation]

The platform's components are open source on GitHub. The MakeCode framework is at https://github.com/microsoft/pxt. (PXT is previous codename of MakeCode). MakeCode targets for the three previously mentioned boards are at microsoft/pxt-microbit, microsoft/pxt-adafruit, and microsoft/pxt-arduino-uno. The latter two targets make use of a common set of MakeCode libraries (packages) at microsoft/pxt-common-packages, Many other MakeCode packages, developed by Microsoft and hardware partners *[details later]. A few examples: XYZ.*

The rest of this paper is organized as follows. Section 2 presents the design and implementation of the MakeCode framework. Section 3 describes Static TypeScript and section 4 presents the CODAL C++ runtime. Section 5: USB Flashing Format; Section 6: Evaluation; Section 7: Related Work; Section 8: Conclusion and Future Directions.

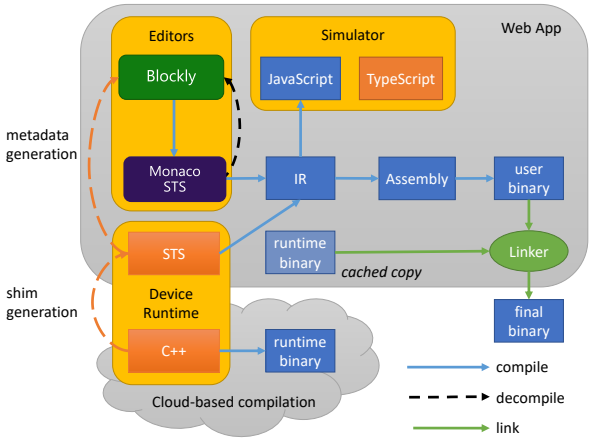

Figure 1. MakeCode Architecture

2 MakeCode: Design and Implementation

The MakeCode web app encapsulates all the components needed to deliver a programming experience for microcontroller-based devices, free of the need for a C++ compiler for the compilation of user code, as shown in Figure 1. The web app is written in TypeScript. The web app also incorporates the TypeScript compiler and language service (www.typescriptlang.org) which are used in several ways, as detailed below. The remaining subsections describe the parts of Figure 1.

2.1 Device Runtime and Shim Generation

A MakeCode target is defined, in part, by its device runtime, which can be a combination of C++ and Static TypeScript (STS) code. The C++ runtime for the target microcontroller is precompiled and stored in the cloud. The C++ runtime binary is cached in the HTML5 application cache (with other assets) so that the web app can function when the browser is offline. Additional runtime components may be authored in STS, which allows the device runtime to be updated without the need for any C++ programming, and permits components of the device runtime to be shared by both the device and simulator runtimes. We will describe later how elements of the C++ runtime are exposed to STS (via automated shim generation). 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 (for runtime components written in TypeScript) or declarations (for runtime components written in C++). A full-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 code.

MakeCode supports a simple foreign function interface from TypeScript to C++ based on namespaces, enumerations,

functions, and basic type mappings. MakeCode uses top-level namespaces (in both C++ and TypeScript) to organize sets of related functions. Preceding a C++ namespace, enumeration, or function with //% indicates that MakeCode should map the C++ construct to TypeScript. Within the //% comment, attributes are used to define the visual appearance for that language construct, such as for the led namespace in the micro:bit target:

```
//% color=#5C2D91 weight=97 icon="\uf205"
namespace led {
...
```

Here is the C++ file defining the micro:bit's led namespace and its functions: pxt-microbit/libs/core/led.cpp.

Mapping of functions and enumerations between C++ and TypeScript is straightforward. Here's an example of the C++ function plot in the led namespace that wraps a more complex function call of the underlying device runtime to set/plot an LED in the micro:bit display:

```
//% blockId=device_plot block="plot|x %x|y %y"
//% x.min=0 x.max=4 y.min=0 y.max=4
void plot(int x, int y) {
    uBit.display.image.setPixelValue(x, y, 1);
}
```

We'll describe the attribute definitions in the //% comment in the next section. MakeCode uses a TypeScript declaration file to describe the TypeScript elements corresponding to C++ namespaces, enumerations and functions. We call such files shim files. Since the C++ plot function is preceded by a //% comment, MakeCode adds the following TypeScript declaration to the shim file (shims.d.ts) and copies over the attribute definitions in the comment. MakeCode also adds an attribute definition mapping the TypeScript shim to its C++ function (shim=led::plot):

```
//% blockId=device_plot block="plot|x %x|y %y
//% x.min=0 x.max=4 y.min=0 y.max=4 shim=led::plot
function plot(x: number, y: number): void;
```

Here is the shim file generated from C++ micro:bit device runtime: pxt-microbit/libs/core/shims.d.ts.

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

MakeCode supports both untagged and tagged integers (described more later). Under the untagged strategy, a JavaScript number is interpreted as a C++ int by default, while the tagged strategy allows both interpretation as integer and double (with an optimization to convert from integer to double, as required). MakeCode includes reference counted C++ types for strings (StringData*) and parameterless lambdas (Action).

These C++ types are mapped to the TypeScript types string and ()=>void, respectively. MakeCode allows a set of C++ functions with the same first parameter (of type Foo) to be exposed as a TypeScript interface Foo as follows: this set of C++ functions must be grouped inside a namespace of the name FooMethods. See, for example, how a C++ buffer abstraction is exposed: pxt-microbit/libs/core/buffer.cpp.

You can find the resulting TypeScript Buffer interface in the shim file for the micro:bit (already referenced above).

2.2 Block Metadata Generation

Both C++ and TypeScript APIs can be specially annotated (minimally via *//% block*) so that the MakeCode compiler generates the needed Blockly metadata to expose an API as a visual block. So, to expose the previously encountered plot function as a visual block (as well as a TypeScript function), one simply needs:

```
//% block
void plot(int x, int y) { . . . }
```

Additional attribute definitions can provide text descriptions for the block, project function parameters (thus simplifying the API available via Blockly), and describe other visual/functional characteristics of the block. MakeCode uses the types of function parameters to select appropriate Blockly widgets. For example, an enumeration is represented by a dropdown menu in blocks. For more information on the block-specific annotations, see https://makecode.com/defining-blocks. MakeCode's support for Blockly means that for the common case, the target developer doesn't need to know anything about the Blockly framework. For more sophisticated needs, one can directly access the Blockly framework.

2.3 Editors and Code Conversion

MakeCode uses the Blockly (https://github.com/google/blockly) and Monaco (https://github.com/Microsoft/monaco-editor) editors to allow the user to code with visual blocks or TypeScript. The editing experience is parameterized by the full-typed device runtime, which provides a set of categorized APIs to the end-user, based on namespaces, as previously described. These APIs are visible in both editors via a toolbox to the immediate left of the programming area. The Blockly and Monaco toolboxes show the same set of APIs, to help in transition from coding with blocks to coding with JavaScript. More advanced TypeScript APIs can be discovered in Monaco via code intellisense.

The Blockly program representation is compiled to Static TypeScript in a syntax-directed manner (see https://github.com/Microsoft/pxt/tree/master/pxtblocks). A key issue is

the need for type inference on the Blockly representation, as variables generally are defined and used without being declared in Blockly. MakeCode uses a simple unification-based type inference to assign a unique type to each variable. In the future, we expect to use TypeScript's type inference instead and eliminate the need for separate type inference over the Blockly representation. TypeScript supports programming constructs that are not available in Blockly, such as classes. Such constructs are converted into grey uneditable blocks in Blockly, with the construct's program text intact. This means MakeCode always can decompile a TypeScript program to Blockly and then recover the program text of the grey blocks when converting from Blockly back to TypeScript (see https://github.com/Microsoft/pxt/blob/master/pxtcompiler/emitter/decompiler.ts).

2.4 Compilation Pipeline

MakeCode first invokes the TypeScript language service to perform type inference and type checking on the user's program, using the TypeScript declaration files for the device runtime. It then checks that the user's program is within the STS subset through additional syntactic and type checks over the adorned abstract syntax tree (AST) produced by the language service (detailed in Section XYZ). Assuming all the above checks pass, MakeCode then performs tree shaking and compilation of the AST of the user code and device runtime 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 two backends for code generation from the IR. The first backend simply generates JavaScript, for execution against the simulator runtime. The other backend generates assembler, parameterized by a processor description. 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. Finally, a linker completes the compile chain by resolving references to the driver runtime in the user's program, producing a binary executable. The compiler chain can be found at https://github.com/Microsoft/pxt/tree/master/pxtcompiler/emitter and https://github.com/Microsoft/pxt/tree/master/pxtlib/emitter.

2.4.1 Asynchronous Functions

A key part of the compilation process is to allow users to call async functions (identified through the *//% async* annotation) as if they were regular (blocking) functions. This is done by compiling each function so that it can be suspended (at the return of a call) and later resumed (at the same point). The default behavior at a suspension point is to immediately resume execution. If a call is to an async function then the default behavior is overridden by the compiler, which suspends execution of the current function. Upon completion of the async function call, the current function then is resumed. This feature greatly simplifies the JavaScript-based programming model which relies on promises to achieve asynchronous execution, breaking up sequential (blocking) code into a series of event handlers. The JavaScript-based simulator runtime uses promises to achieve asynchronous execution in a single-threaded context, but these promises are hidden from the end user. The CODAL C++ device runtime supports fibers with the ability to pause, so for compilation to a device, the compiler simply emits a call to pause at a suspension point.

2.4.2 Untagged and Tagged Strategies

The MakeCode compiler supports the Static TypeScript language subset described in Section X, with two compilation strategies: untagged and tagged. *TODO* Untagged compilation strategy No support for doubles. 32-bit integers (signed only). Use static type system to distinguish between integers and references. No runtime support. Not fully compatible with JavaScript semantics. Null = Undefined = 0 = default integer value. Support for different integer sizes using type aliases (this is used for AVR). Tagged compilation strategy In the tagged strategy, numbers are either tagged 31-bit signed integers, or if they do not fit, boxed doubles. Code generation to check tag bit. Special constants like false, null and undefined are given special values and can be distinguished. Fully compatible with JavaScript semantics (number, null, undefined). Can accommodate a richer type system.

2.5 Simulator

A MakeCode target can provide an alternate TypeScript implementation for each API in the device runtime, for use in the device simulator. As this code runs in the web browser (not on the actual device) and manipulates the DOM, the developer is free to use all of TypeScript/JavaScript's features. (As an aside, MakeCode also support "simulator-only" targets that have no associated device; in these cases, the "device runtime" is defined solely by the simulator APIs.) The simulator allows the user to experience the basic functions of the device in the browser and to test their code before deploying it to the actual device. The simulator has proxy widgets for sensors such as accelerometer (mouse motion), temperature and light, allowing the user to control the sensor's value. The simulator only provides basic functionality and is far from a complete device emulation. For example, it is not possible for the user to simultaneously modify two inputs to the simulated device, while it is possible with the actual device (i.e., shaking it to change the accelerometer reading while pushing one of the device's buttons).

MakeCode provides various components to make device-based simulators easier to build: board, parts, wiring, etc.

2.6 Packages and Custom Editors

Packages are MakeCode's dynamic/static library mechanism for extending a target (by adding new code/data to the device and simulator runtimes, as well as accompanying documentation). The following package extends the micro:bit target so that the micro:bit can drive a NeoPixel strip of RGB LEDs: https://github.com/Microsoft/pxt-neopixel. To see how this package is surfaced to the end-user, visit http://makecode.microbit.org/ and select the "Add Package" option from the gear menu; you will see the package "neopixel" listed in the available options. If you click on it, a new block category named "Neopixel" will be added to the editor. In this scenario, PXT dynamically loads the (white listed) neopixel package directly from GitHub, compiles it and incorporates it into the web app. Packages also can be bundled with a web app (the analog of static linking).

Hardware partners already have started to create MakeCode packages for the micro:bit. Seeed Studio (https://www.seeedstudio.com/) has created packages to add its Grove components to a micro:bit. Grove components are accessed via the I2C serial protocol, supported by the micro:bit device runtime. All micro:bit packages for the Grove components are authored in Static TypeScript (gesture, ultrasonic-ranger, 4-digital-display, two-led-matrix). These packages can be found under GitHub user "Tinkertanker", prefixed with "pxt-". Sparkfun has created MakeCode packages for its micro:bit shields (GitHub user "sparkfun").

3 Static TypeScript

TypeScript is a typed superset of JavaScript designed to enable JavaScript developers to take advantage of code intellisense, static checking and refactoring made possible by types []. You can edit and run simple TypeScript programs at http://www.typescriptlang.org/play. As a starting point, every JavaScript program is a TypeScript program. Types can be added gradually to such programs, supported by type inference. In TypeScript, the Any type represents any JavaScript value with no constraints. Type inference may assign the Any type to expressions for which no more specific type can be inferred.

In TypeScript, object types are used to describe dictionaries, functions, arrays, as well as class instances. Object types also describe objects that take on multiple of the above roles, as is common in JavaScript. Object types are related to one another by structural subtyping []. Interface declarations name object types; classes add implementation and the ability to statically inherit implementation from super classes. TypeScript also supports generics, as well as union and intersection types.

While TypeScript provides programming abstractions (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 JavaScript objects with all their dynamic semantics intact. This is not surprising, as TypeScript was designed to accommodate the dynamic nature of JavaScript and programming patterns familiar to the JavaScript programmer.

TypeScript has enough types to allow us to approach it from the viewpoint of the microcontroller programmer, who is familiar with the static (though unsound) type systems of C and C++. Our realization is that TypeScript contains a statically-typed sound subset (Static TypeScript: STS, for short) that closely resembles Java and C# in its semantics. STS arises from TypeScript by:

- *eliminating* the Any type from the type system, as well as JavaScript constructs that only can be typed with Any;
- *partitioning* the space of object types into functions, records, constructor functions, and arrays, with no casts possible between the partitions (and eliminating all other object types) - structural subtyping is used to relate records to each other, as in TypeScript;
- typing of classes *nominally* rather than structurally, as in TypeScript, and using traditional function (contravariant in the argument type, covariant in the return type) subtyping for functions and methods;
- *restricting* casts between records and classes: a class can be cast to a record, but a record cannot be cast to a class.

STS is a syntactic subset of TypeScript – there is no change to the syntax of TypeScript. *[TODO: also, we disllow update of method property]* In terms of type checking, STS also is a subset of TypeScript. Simply put, given a STS program P, for every type compatibility or subtype relation $R(P)$ of TypeScript, the corresponding relation $R'(P)$ in STS is a subset of relation R. Every STS program is a TypeScript program. *[TODO: semantics preserved by erasure for programs written entirely in STS]* STS guarantees the absence of runtime type errors, including downcasts. The major class of runtime errors still possible is dereference of null/undefined values. STS can be supported by a simple runtime model, like that for C++, using classic vtables and interface tables, rather than the prototype-based runtime model used for full TypeScript/JavaScript.

3.1 Eliminating Types: Any, Union, Intersection

By default, the TypeScript compiler assigns the Any type to an expression/declaration for which it is unable to infer a more precise type. STS uses the *–noImplicitAny* option to direct the TypeScript compiler to raise an error whenever it makes an (implicit) assignment of the Any type. STS also checks for an explicit use of the Any type in the program and raises an error. Many JavaScript constructs can only be typed with the Any type, including: prototype lookup, the eval statement, the with statement, a this reference outside of class context, an index access on non-array object, and

reflection on Function/Object objects. Therefore, these constructs are not in STS. STS also excludes TypeScriptâĂŹs union and intersection types.

3.2 Partitioning of Object Types

Key to TypeScript are the definitions of object types and interfaces. Per the TypeScript specification [**?**]:

- "object types are composed from properties, call signatures, construct signatures, and index signatures, collectively called members";
- "interfaces provide the ability to name and parameterize object types and to compose existing named object types into new ones.".

Object types and the interface declarations that describe them enable the typing of a JavaScript object that plays multiple roles (as a function, dictionary, etc.). For example, here is an interface that describes an object that is a function from number to number and also has property (foo) which is a string:

```
interface Bar {
    (x: number): number; // call signature member
    foo: string;         // property member
}
```

STS partitions the space of object types as follows:

1. a *record* type has only member properties and optionally, a string index signature;
2. a *function* type has exactly one member, a call signature;
3. a *constructor function* type: has at least one constructor signature and no other signature kind;
4. an *array* type has a numeric index signature and no other signatures;
5. an *other* type: object type not covered by the previous four categories.

STS excludes all "other" types, leaving us with a set of classic types XYZ.

[Interfaces describe all of the above, help to access class and record uniformly, as they support arbitrary subsetting of properties.

3.3 Nominal Typing of Classes

TypeScript introduces two (object) types for each class declaration: a constructor function type by which instances of a class are created (using new) and static properties accessed; a class type describing the instances of the class (records, as defined above). [but the class type contains a link back to the class declaration; STS treatment of classes using a nominal interpretation, names and extends for subtyping]

3.4 Compilation of STS

Classic vtable for classes (prefix); classic interface table for classes and records

4 The CODAL Runtime

TBALL comment: the main focus on CODAL here is the support for event-based programming with concurrrent handlers, providing the bridge between scripting languages such as JavaScript and the world of microcontrollers.

For reasons of efficiency and ease of access to hardware, code for microcontrollers is typically written in C/C++. However, this still leaves the programming model undefined. The Arduino programming model is based on polling: an Arduino "sketch" is a program template that consists of a setup procedure, for initialization of data structures, and a loop procedure; programmers implement the body of setup and loop, where they explicitly poll the state of the microcontrollerâĂŹs pins. The Arduino model emphasizes sequential composition, which is not well-suited to reactive systems where event arrival is unpredictable, and events must be handled quickly to achieve responsiveness.

The component-oriented device abstraction layer (CODAL) presents a programming model consisting of a set of components that abstract away microcontroller details (such as pins and whether one polls or uses interrupts) and communicate via events using a publish/subscribe mechanism. Components can schedule event handlers to run concurrently using a non-preemptive fiber scheduler. Each software component abstracts a hardware feature, sensor, or connected device. *TODO: how CODAL handles responsiveness?*

The CODAL runtime provides higher-level abstractions to the compiler developer than the Arduino runtime, making it easy to compile event-based mechanisms, such as found in scripting languages like JavaScript, to the world of the microcontroller. [Greater flexibility, learning opportunities. We will show it is competitive in terms of memory usage and performance.]

4.1 Components

TODO - CodalComponent: software model of device driver (physical and logical); components requiring periodic invocation

4.2 Events and the Message Bus

A message bus enables a simple but powerful eventing model. Every runtime component has its own namespace for events which can be raised at any time. Users may register event handlers for any number of such events. Moreover, event handlers are transparently run in their own fiber - providing decoupling of user code from system code, device drivers and interrupts. Fibers may also block awaiting events—a simple yet effective condition synchronisation mechanism.

4.3 Fiber Scheduler

A fiber scheduler enables a degree of managed concurrency on the device. Many high level languages require asynchronous execution for responsiveness. Support for these languages can be greatly simplified through the presence of a scheduler. The scheduler is, by design, non-preemptive to minimize race conditions and interleaving of user code. This is also informed by recent studies of concurrency in visual programming languages [**?**]. A further benefit of running a scheduler is that when no user code needs to be executed, the device can safely be put into a power efficient sleep mode, transparently to the user.

4.4 CodalDevice

Putting it all together

5 USB Flashing Format

The way that code gets from a host computer onto a microcontroller is deeply rooted in 1980's technologies - serial wires, obscure protocols, and text-based file formats with limited line length. Depending on exact circumstances, one must install serial USB drivers, select the right port and parameters, and then use a native application to access the microcontroller, as the browser is not allowed to. As the advance of maker content in educational curricula continues, this complicated flashing process presents one of the major obstacles to adoption in schools, where installing any software is usually the domain of IT administrators.

One solution would be to rely on emerging standards, like WebUSB and WebBluetooth to transfer a program from the browser to the microcontroller. These standards are however still in their infancy, and it may take even longer before they are deployed in schools.

Another solution was pioneered by ARM with its DAPLink firmware, where a USB-capable microcontroller presents itself to an external computer as a USB pen drive. No special drivers need to be installed, as operating systems support pen drives out of the box, typically formatted using FAT. The USB pen drive protocol (Mass Storage Class or MSC) is a block-level protocol (generally 512 bytes) with no concept of files. DAPLink exposes a virtual FAT file system, which is very small and never changes due to OS writes - it has an informational text file and an HTML file with a redirect to the online editing environment. Otherwise, the FAT and the root directory table are empty. When the OS tries to read a block, the DAPLink computes what should be there dynamically, based on compiled-in contents of the info and HTML files.

On file system writes, DAPLink detects the HEX format, and flashes it into the target microcontroller's memory, over debug wires. Let's consider the problem that DAPLink must solve. It sees a 512 byte block of data to be written at a particular block index on the device. It needs to decide if it's part of the file being flashed and if so, extract the data and write it to the target microcontroller. Furthermore, when the OS writes a HEX file, DAPLink needs to discard writes to the FAT or directory table, as well as writes of the meta-data files. It may need to deal with out-of-order writes. All these details mean that DAPLink is quite complex and sometimes needs to be updated when a new OS release changes the way in which it handles FAT. This also is the reason that some MSC bootloaders for various chips only support given operating systems under some specific conditions.

The task would be simplified then, if every 512 block of the file being flashed was easy to distinguish from meta-data or other random files, and easy to act on, independent of other blocks. The HEX file format doesn't give us these properties - the 512 byte boundary can be in the middle of a line in the HEX file, and even if we have an entire line, every line only contains the last 16 bits of the address where to flash, with the upper 16 supplied only when they change.

For these reasons, we designed the USB Flashing Format (UF2). It consists of 512 byte blocks, where each block contains:

- magic numbers at the beginning and end (to heuristically distinguish it from any other data the OS writes);
- the address in the target chip flash where the payload should be written;
- the payload data (up to 476 bytes)

6 Evaluation

6.1 Implementation

6.2 Benchmarks

6.3 Performance evaluation

7 Related Work

8 Conclusion

A Appendix

Text of appendix …