Why I Love D: An Undergrad Experience

Erich Gubler
Why I Love D: An Undergrad Experience

Erich Gubler
Objectives

- Recount the discovery process of D features during my project development

- What helped me really love D?

- Expose some cool D stuff!
CS 4450: Where It Began

- Chuck Allison got me into D while teaching me CS 4450!
- At the same time, I was implementing a virtual machine in C++ for CS 4380.
- I decided to port and continue development in D so I could get a better grade in CS 4450's D assignments. :)
- I stayed with D through implementing a compiler later!
Porting Over to D

- No more *.h and *.cpp
- "I don't have to deal with pointers anymore?"
- Pointers are like an unsheathed knife in your pants pocket...
- No more namespace operator!
- 'Twas a snap! Mostly **stripping** stuff.
Enums

Assembly

Add R1, R2
Sub R3, R4
LDR R5, DATA
LDR R5, R6

Assembler

Bytecode

0x0E | 0x01 | 0x02
0x0F | 0x03 | 0x04
0x0B | 0x05 | 0x??
0x0A | 0x06 | 0x??
String to enum and back!

- How would you implement it?

**C++**

- Switch/if-else statements for efficiency
- Map for ease on the mind

```cpp
CapstoneVM::Register getRegister(const std::string &name)
{
    // Singleton pattern goodness
    static std::map<std::string, CapstoneVM::Register> registerMap;

    if(registerMap.empty())
    {
        registerMap["R0"] = CapstoneVM::R0;
        registerMap["R1"] = CapstoneVM::R1;
        registerMap["R2"] = CapstoneVM::R2;
        registerMap["R3"] = CapstoneVM::R3;
        registerMap["R4"] = CapstoneVM::R4;
        registerMap["R5"] = CapstoneVM::R5;
        registerMap["R6"] = CapstoneVM::R6;
        registerMap["R7"] = CapstoneVM::R7;
        registerMap["PC"] = CapstoneVM::PC;
    }

    std::map<std::string, CapstoneVM::Register>::iterator it = registerMap.find(name);
    return ((it != registerMap.end()) ? registerMap[name] : CapstoneVM::INVALID_Register);
}
```

**D**

- Ported:

```d
Register toRegister(string name)
{
    try
    {
        return !Register(name);
    }
    catch {}
    return Register.INVALID_REGISTER;
}
```

...it's like taking a bath, isn't it?
Switch/if-else statements for efficiency
Map for ease on the mind

```cpp
CapstoneVM::Register getRegister(const std::string &name) {

    // Singleton-pattern goodness
    static std::map<std::string, CapstoneVM::Register> registerMap;

    if (registerMap.empty()) {
        registerMap["R0"] = CapstoneVM::R0;
        registerMap["R1"] = CapstoneVM::R1;
        registerMap["R2"] = CapstoneVM::R2;
        registerMap["R3"] = CapstoneVM::R3;
        registerMap["R4"] = CapstoneVM::R4;
        registerMap["R5"] = CapstoneVM::R5;
        registerMap["R6"] = CapstoneVM::R6;
        registerMap["R7"] = CapstoneVM::R7;
        registerMap["PC"] = CapstoneVM::PC;
    }

    std::map<std::string, CapstoneVM::Register>::iterator it = registerMap.find(name);
    return ((it != registerMap.end()) ? registerMap[name] : CapstoneVM::INVALID_REGISTER);
}
```
String to enum and back!

- How would you implement it?

C++

- Switch/if-else statements for efficiency
- Map for ease on the mind

```cpp
CapstoneVM::Register getRegister(const std::string &name) {
    // Singleton-pattern goodness
    static std::map<std::string, CapstoneVM::Register> registerMap;
    if (!registerMap.empty()) {
        registerMap["R0"] = CapstoneVM::R0;
        registerMap["R1"] = CapstoneVM::R1;
        registerMap["R2"] = CapstoneVM::R2;
        registerMap["R3"] = CapstoneVM::R3;
        registerMap["R4"] = CapstoneVM::R4;
        registerMap["R5"] = CapstoneVM::R5;
        registerMap["R6"] = CapstoneVM::R6;
        registerMap["R7"] = CapstoneVM::R7;
        registerMap["PC"] = CapstoneVM::PC;
    }
    std::map<std::string, CapstoneVM::Register>::iterator it = registerMap.find(name);
    return (it != registerMap.end()) ? registerMap[name] : CapstoneVM::INVALID_REGISTER;
}
```

Ported:

```d
Register toRegister(string name)
{
    try
    {
        return to!Register(name);
    }
    catch {}
    return Register.INVALID_REGISTER;
}
```

...it's like taking a bath, isn't it?
Can use std.conv's "to" function because enums are smart in D

```d
string s = //...
auto r = s.to!Register;
r = Register.R1;
s = r.to!string;
```
Ported:

```cpp
Register toRegister(string name) {
    try {
        return to!Register(name);
    } catch {}
    return Register.INVALID_REGISTER;
}
```

...it's like taking a bath, isn't it?
C++

- Switch/if-else statements for efficiency
- Map for ease on the mind

```cpp
CapstoneVM::Register getRegister(const std::string &name)
{
    // Singleton-pattern goodness
    static std::map<std::string, CapstoneVM::Register> registerMap;
    if (registerMap.empty())
    {
        registerMap["R0"] = CapstoneVM::R0;
        registerMap["R1"] = CapstoneVM::R1;
        registerMap["R2"] = CapstoneVM::R2;
        registerMap["R3"] = CapstoneVM::R3;
        registerMap["R4"] = CapstoneVM::R4;
        registerMap["R5"] = CapstoneVM::R5;
        registerMap["R6"] = CapstoneVM::R6;
        registerMap["R7"] = CapstoneVM::R7;
        registerMap["PC"] = CapstoneVM::PC;
    }
    std::map<std::string, CapstoneVM::Register>::iterator it = registerMap.find(name);
    return (it != registerMap.end()) ? registerMap[name] : CapstoneVM::INVALID_REGISTER;
}
```

D

Ported:

```d
Register toRegister(string name)
{
    try
    {
        return to!Register(name);
    } catch {}
    return Register.INVALID_REGISTER;
}
```

...it's like taking a bath, isn't it?
LDR  R5,  DATA

Another example: load pointer in register

STR  R6,  RESULT

LDR  R5,  R6

Super easy because D removed the obstacles.
public enum Opcode
{
    INVALID_OPCODE = -1,

    // ...

    STR,
    LDR,
    STB,
    LDB,

    // ...

    // Indirect load/stores
    STR_INDIRECT,
    LDR_INDIRECT,
    STB_INDIRECT,
    LDB_INDIRECT,
}

// Instruction
case Opcode:
    , Opcode,
    , Opcode,
    , Opcode
{
    auto x = ...
    if (x == 0)
    {
    }
    else
    {
    }
    break;
// Instructions that need some top-down parsing to decide what they need
      
    case Opcode.COMAT:
      
    } else {
      
    oToUse = (o.toString() + "_INDIRECT").toString().to!Opcodes;
      
    p = new RegisterRegisterParser(argumentString);

      } break;

    This is pretty neat!

eye because D removed the obstacles.
Super easy because D removed the obstacles.
The D std: Raising the Bar

- I've never been so rewarded by just reading through the docs!
- Because of this, it's tempting to say to newbies, "RTD or get out" (but we shouldn't)
- Of special note:
  - `std.getopt`
  - `std.csv`
  - `std.json`
  - `std.database` (?)

Why are these so nice?
- Reducing the reinvention of the wheel increases productivity
- Employers and clients are willing to listen if you've got it done, not necessarily if you've got things done "right"
- This is what happens with application development
- Generic database
  - You can optimize later if you need to.
std.getopt

This needs exposition! Some comments heard:

- "...that's standardized?"
- "That's cheating!"
- "No way..."

As fast as prototyping

Modular getopt calls

Where was this in C++?

- After using D's getopt, I realized from not having something nice like it in C++, so I wrote it.
- https://github.com/EricDongGable/getopt
Using the original GNU getopt

```c
// Command-line argument parsing
int argFlag;
while ((argFlag = getopt(argc, argv, "i:o:")) != -1)
{
    switch (argFlag)
    {
    case 'i':
        // ...
        break;
    case 'o':
        // ...
        break;
    case '?':
        if (optopt == 'o' || optopt == 'i')
            return Error(ERRORCODE_BADARGS, "Option -%c requires an argument.\n", optopt);
        return 1;
    default:
        goto EXIT;
    }
}
EXIT;
```
As fast as prototyping

```cpp
import std.stdio;
import std.getopt;

int main(string[] args)
{
    // Opt variables
    string myString;
    bool verboseFlag = false;

    auto result = getopt(args
        , "s|myString", &myString // -s or --myString
        , "verboseFlag", &verboseFlag); // --verboseFlag=[...]

    writeln("myString: ", myString);
    writeln("verboseFlag: ", verboseFlag);

    writeln("args left: ", args);

    // Print help
    if(result.helpWanted)
        defaultGetoptPrinter("Help:", result.options);

    return 0;
}
```
Modular getopt calls

- IO

```
// Grab a single output file and any number of input files from command args
2000xResult parseArgs(String[] args)
{
    try
    {
        getopt(args,
                "output", m_output)
        config.parseThrough();
    }
    catch (GetOptException e) {
        // ... }

    if (input.length == 1)
        arg.length = 1;
    return io;
}
```

getopt call

Return data structure

- Logging with channels

```
./compiler --channel=tg
```

Getopt makes this all work

Reuse and factorization of getopt options = win!

- Putting it all together

```
void main(String[] args)
{
    auto logging = CompilerLogger.parseArgs(args);
    auto io = parseFilesystem(args);

    if (logging.result.helpRequested || io.result.helpRequested)
    {
        import std::getopt; defaultGetoptPrinter;
        defaultGetoptPrinter("Usage: ", args[0] = "source file", logging.result.options = io.result.options);
    }

    // Validate IO
    if (io.inputs.length == 0)
        returnError("No input file specified");
    else if (io.inputs.length > 1)
        returnError("Multiple input files not supported");

    Call opt parsing modules

Use built-in help

Misc. validation

EPIC WIN

This is the face of an epic win

CanItBeSaturdayNow.com

Prezi
```
/// Grabs a single output file and any number of input files from command args
I0OptResult parseIOFiles(ref string[] args)
{
    I0OptResult io;
    try
    {
        getopt(args,
            "outfile|o",
            &io.output,
            config.passThrough
        );
    }
    catch(GetOptException e){ /* ... */ }
    io.inputs = args[1..$];
    args.length = 1;
    return io;
}
```
• Logging with channels

```
./compiler --channel=tg
```

```c
class ChannelLogger(ChannelEnum, PresetEnum)
{
    // ...

    static ChannelLoggerResult parseArgs(ref string[] args)
    {
        // ...
        try
        {
            results.result = getopt(args, "channel\|c", &rawChannels,
                                    "preset\|p", &rawPreset,
                                    config.passThrough);
        }
        catch(GetOptException e){ /* ... */ }
    }
}
```

`Getopt` makes this all work

```
enum CompilerLogChannel
{
    TOKENIZING = 't', // Prints output from the tokenizer
    GRAMMAR_STEP = 'g', // Prints output as the grammar is walked
    SEMANTIC_ANALYSIS = 'e', // Prints output from semantic analysis
    SYMBOL_RESOLUTION = 'y', // Prints output from the symbol table as it is built
    ICODE_GENERATION = 'i', // Prints icode as it is generated
}
```

```c
import std.traits : EnumMembers;
enum compilerChannelPreset
{
    QUIT = new CompilerLogChannel[0],
    VERBOSE = new CompilerLogChannel[0],
    DEBUG = [ EnumMembers!CompilerLogChannel ],
}
```

```c
import capstone.utils : ChannelLogger;
alias CompilerLogger = ChannelLogger!CompilerLogChannel, CompilerChannelPreset;
```

Reuse and factorization of `getopt` options = win!
logging with channel

./compiler --channel=tg

enum CompilerLogChannel
{
    TOKENIZING = 't', // Prints output from the tokenizer
    GRAMMAR_STEP = 'g', // Prints output as the grammar steps
    SEMANTIC_ANALYSIS = 'e', // Prints output from semantic analysis
    SYMBOL_RESOLUTION = 'y', // Prints output from the symbol resolution
    ICODE_GENERATION = 'i', // Prints icode as it is generated
}

import std.traits : EnumMembers;
enum CompilerChannelPreset
{
    QUIET = new CompilerLogChannel[0],
    VERBOSE = new CompilerLogChannel[0],
    DEBUG = [ EnumMembers!CompilerLogChannel ]
}

import capstone.util : ChannelLogger;
alias CompilerLogger = ChannelLogger!(CompilerLogChannel,
class ChannelLogger(ChannelEnum, PresetEnum)
{
    // ...

    static ChannelLoggerResult parseArgs(ref string[] args)
    {
        // ...
        try
        {
            results.result = getopt(args, "channel|c", &rawChannels,
                                    "preset|p", &rawPreset,
                                    config.passThrough);
        }
        catch(GetOptException e){ /* ... */ }
    }
}
enum CompilerLogChannel {
    TOKENIZING = 't', /// Prints output from the tokenizer
    GRAMMAR_STEP = 'g', /// Prints output as the grammar is walked
    SEMANTIC_ANALYSIS = 'e', /// Prints output from semantic analysis
    SYMBOL_RESOLUTION = 'y', /// Prints output from the symbol table as it is built
    ICODE_GENERATION = 'i', /// Prints icode as it is generated
}

import std.traits : EnumMembers;
enum CompilerChannelPreset {
    QUIET = new CompilerLogChannel[0],
    VERBOSE = new CompilerLogChannel[0],
    DEBUG = [ EnumMembers!CompilerLogChannel ]
}

import capstone.utils : ChannelLogger;
alias CompilerLogger = ChannelLogger!(CompilerLogChannel, CompilerChannelPreset);
• Putting it all together

```java
void main(string[] args)
{
    auto logging = CompilerLogger.parseArgs(args);
    auto io = parseIOFiles(args);

    if(logging.result.helpWanted || io.result.helpWanted)
    {
        import std.getopt : defaultGetoptPrinter;
        defaultGetoptPrinter("Usage: " ~ args[0] ~ " <source file>", logging.result.options ~ io.result.options);
    }

    // Validate IO args
    if(io.inputs.length == 0)
        returnError!(ErrorLevel.BAD_ARGS)("no input file specified");
    else if(io.inputs.length > 1)
        returnError!(ErrorLevel.BAD_ARGS)("multiple input files not supported");
}
```
Reuse and factorization of getopt options = win!

EPIC WIN
This is the face of an epic win
Where was this in C++?

- After using D's getopt, I nerdraged from not having something nice like it in C++...so I wrote it. :)

- [https://github.com/ErichDonGubler/getopt](https://github.com/ErichDonGubler/getopt)

```cpp
GetOpt::GetOptResultAndArgs results;
try
{
    results = GetOpt::getopt(argc, argv,
                              "l|length", &length,
                              "file|f", &data,
                              "verbose|v+", &verbosity,
                              "color|c|colour", &color,
                              "pretty|p", &usePrettyOutput);
}
catch(GetOpt::GetOptException e)
{
    cerr << e.what() << endl;
    return 1;
}
```
GetOpt::GetOptResultAndArgs results;
try
{
    results = GetOpt::getopt(argc, argv, "l|length", &length,
                              "file|f", &data,
                              "verbose|v+", &verbosity,
                              "color|c|colour", &color,
                              "pretty|p", &usePrettyOutput);
}
catch(GetOpt::GetOptException e)
{
    cerr << e.what() << endl;
    return 1;
}
Why are these so nice?

- Reducing the reinvention of the wheel increases productivity.

- Employers and clients are willing to listen if you've got it done, not necessarily if you've got things done "right"...
  - This is what happens with application development!

- Sensible defaults
- You can optimize later if you need to.
The Machine Shop

*Quote from Walter:* When I was in London for the 2010 ACCU (when the volcano stranded me there), I took a chance to tour the Belfast cruiser sitting in the Thames. One interesting aspect of it was the ship's machine shop.

It was full of carefully selected machine tools. It was pretty clear to me that an expert machinist could quickly and accurately make or repair about anything that broke on that ship.

Sure, you can make do with fewer, more general purpose machines. But it'll take you considerably longer, and the result won't be as good. For example, I've used electric drills for years. I was
the Thames. One interesting aspect of it was the ship's machine shop.

It was full of carefully selected machine tools. It was pretty clear to me that an expert machinist could quickly and accurately make or repair about anything that broke on that ship.

Sure, you can make do with fewer, more general purpose machines. But it'll take you considerably longer, and the result won't be as good. For example, I've used electric drills for years. I was never able to get it to drill a hole perfectly perpendicular. I finally got a drill press, and problem solved. Not only is it far more accurate, it's much faster when you've got a lot of holes to drill.

I prefer to view D as a fully equipped machine shop with the right tools for the right job. Yes, it
the Thames. One interesting aspect of it was the ship's machine shop.

It was full of carefully selected machine tools. It was pretty clear to me that an expert machinist could quickly and accurately make or repair about anything that broke on that ship.

Sure, you can make do with fewer, more general purpose machines. But it'll take you considerably longer, and the result won't be as good. For example, I've used electric drills for years. I was never able to get it to drill a hole perfectly perpendicular. I finally got a drill press, and problem solved. Not only is it far more accurate, it's much faster when you've got a lot of holes to drill.

I prefer to view D as a fully equipped machine shop with the right tools for the right job. Yes, it
Sure, you can make do with fewer, more general purpose machines. But it'll take you considerably longer, and the result won't be as good. For example, I've used electric drills for years. I was never able to get it to drill a hole perfectly perpendicular. I finally got a drill press, and problem solved. Not only is it far more accurate, it's much faster when you've got a lot of holes to drill.

I prefer to view D as a fully equipped machine shop with the right tools for the right job. Yes, it will take longer to master it than a simpler language. But we're professionals, we program all day. The investment of time to master it is trivial next to the career productivity improvement.
The Machine Shop

Quote from Walter:
When I was in London for the 2010 ACCU (when the volcano stranded me there), I took a chance to tour the Belfast cruiser sitting in the Thames. One interesting aspect of it was the ship’s machine shop.

It was full of carefully selected machine tools. It was pretty clear to me that an expert machinist could quickly and accurately make or repair anything that broke on that ship.

Sure, you can make do with fewer, more general purpose machines. But it'll take you considerably longer, and the result won’t be as good. For example, I’ve used electric drills for years. I was never able to get it to drill a hole perfectly perpendicular. I finally got a drill press, and problem solved. Not only is it far more accurate, it’s much faster when you’ve got a lot of holes to drill.

I prefer to view D as a fully equipped machine shop with the right tools for the right job. Yes, it will take longer to master it than a simpler language. But we’re professionals, we program all day. The investment of time to master it is trivial next to the career productivity improvement.
D tools: rdmd

- Lots of awesome tools for D I've seen!
- rdmd is an excellent example
- No need for big compiler commands
- Why is rdmd cool?
Command: rdmd -of "compiler.exe" compiler_main.d [args]
Unit tests in a large project

```cpp
enum Register {
    INVALID_REGISTER = -1,
    // Data registers
    R0, R1, R2, R3, R4, R5, R6, R7,
    PC, // Program Counter
    SL, // Stack Limit
    SP, // Stack Pointer
    FP, // Frame Pointer
    SB, // Stack Base
}

tunittest
{
    writeln("Hello, world!");
    import std.conv : to;
    assert("R0".to!Register == Register.R0);
    assert(Register.FP.to!string == "FP");
}
```

- I want to test just this module.
- But package root is the parent directory!

```
rmdm -I.. -unittest -main vm.d
```

- I have no excuse but to test!

Another hack with contract programming

- At one point during development, I became convinced that a bug was in one of my harder-to-debug modules because of how it was called.
- Because of logging, it was easy to track where I was, but I didn't want to insert a ton of code...

```cpp
class Logger {
    // Log function calls
    // ... but which one is the problem?!

    public:
    static void log(string msg);
    void print(string msg);
}
```
Another hack with contract programming

- At one point during development, I became convinced that a bug was in one of my harder-to-debug modules because of how it was called.
- Because of logging, it was easy to track where I was, but I didn't want to insert a ton of code...

```java
class ICodeGenerator : SimpleLogger
{

    // TONS of functions here
    // ...but which one is the problem? D:

    invariant
    {
        import std.stdio : writeln;
        writeln("Bloop!");
    }
}
```
class ICodeGenerator : SimpleLogger
{

    // TONS of functions here
    // ...but which one is the problem? D:

    invariant
    {
    import std_stdio : writeln;
        writeln("Bloop!");
    }
}
Last Opinions

- We, as the D community, shouldn't be afraid to break stuff so it can make progress.
  - Make transition easy as possible
  - But give priority to language development
- Making D easy to get into will be essential for D's adoption into mainstream.
Conclusion

- D removes obstacles from your path, and that makes development \textit{fast}.
- D's standard library sets a new standard, much like its predecessors.
  - The fully-equipped machine shop!
- D's willingness to pioneer (which may involve breaking occasionally) will be what makes it better in the long run.

\textbf{That's why I love D. :)}
Why I Love D: An Undergrad Experience

Erich Gubler
Credits

• Quote from Walter about the Belfast:
  • http://forum.dlang.org/thread/fbdeybmxbpbgxflmvny@forum.dlang.org?page=3#post-ki59uk:24fjc:241:40digitalmars.com

• Images of the HMS Belfast:

• Countdown image
  • http://www.iab.net/media/image/countdown1.jpg

• D Logo: