

Instituto Superior de Engenharia de Lisboa

Electronics, Telecommunications and Computers Engineering Department

BrightScript Debugger

Celso de Almeida Fernandes

(Graduated in Computer Science and Computer Engineering)

Dissertation for Master's Degree in Computer Science and Computer Engineering

Advisors: Eng. Paulo Pereira

Eng. Pedro Pereira

Jury:

President: Eng. Manuel Barata

# Abstract

BrigthScript is a programming language based on JavaScript and Visual Basic, created by Roku. Roku is a company who produces and commercializes devices that are used to watch movies and television. These devices are usually connected to a TV and are commonly known as TV boxes. BrightScript is the language on which applications for their boxes are developed.

After some analysis, we could not found many development tools and the existing ones are not functional enough. Roku provides an Eclipse plugin and the boxes expose a telnet port for basic debugging. The Eclipse plugin only makes syntax validation and exports application code to the box. There are several open source plugins for the most popular text editors, that perform syntax highlighting.

The goal of this project is therefore to implement an integrated tool for application development thereby simplifying the development process, that includes debugging. This tool supports syntax validation, code compilation, intellisense and graphical debug interaction.

Visual Studio is an Integrated Development Environment created by Microsoft and it is the main tool for the development of Windows applications. The tool is materialized as a Visual Studio plugin for the BrightScript language. This plugin uses the language services provided by Visual Studio SDK.

The tool could be enhanced with a box simulator that would be used to run the applications on the development machine, thereby further improving the development cycle; a goal identified as future work.

The idea for this project was a consequence of the author’s involvement on SkyStore Roku App development, for Sky UK Limited Company.

# Index

[Abstract i](#_Toc466406991)

[1. Index ii](#_Toc466406992)

[2. Introduction 1](#_Toc466406993)

[2.1. Compilers Theory 1](#_Toc466406994)

[2.2. Code Generation tools 1](#_Toc466406995)

[2.3. Implementation 2](#_Toc466406996)

[3. Compilers Theory 4](#_Toc466406997)

[3.1. Lexical Analysis 5](#_Toc466406998)

[3.2. Parsing 6](#_Toc466406999)

[3.3. Abstract Syntax 8](#_Toc466407000)

[4. Code Generation tools 9](#_Toc466407001)

[4.1. GPlex 9](#_Toc466407002)

[4.2. Gppg 10](#_Toc466407003)

[4.3. Irony 10](#_Toc466407004)

[4.4. JavaCC 11](#_Toc466407005)

[4.5. SableCC 11](#_Toc466407006)

[4.6. Conclusion 11](#_Toc466407007)

[5. Compiler 12](#_Toc466407008)

[5.1. Lexer 12](#_Toc466407009)

[5.2. Parser 13](#_Toc466407010)

[6. Debugger 15](#_Toc466407011)

[6.1. Telnet module 16](#_Toc466407012)

[6.2. Deploy 20](#_Toc466407013)

[6.3. Remote 22](#_Toc466407014)

[7. Visual Studio Plugin 24](#_Toc466407015)

[7.1. Project Type 24](#_Toc466407016)

[7.2. Builder/Deploy 26](#_Toc466407017)

[7.3. Editor Extension 26](#_Toc466407018)

[7.4. Debugger integration 28](#_Toc466407019)

[8. Conclusion 29](#_Toc466407020)

[9. References 30](#_Toc466407021)

[10. Figures Index 32](#_Toc466407022)

# Introduction

The project is divided into three stages. The first stage focuses the study off compilers theory. The second stage is about the investigation of existing tools for generating compilers code, how they work and their benefits. The third stage focuses the plugin implementation.

## Compilers Theory

In the first stage it was used the set of Compilers Theory videos [1], by Alex Aiken, a Professor of Computer Science in Stanford University and the book entitled Modern Compiler Implementation in Java [2]. These two sources have a very similar approach of compilers theory, suggesting a modular implementation.

BrightScript [3] is an interpreted language and therefore there is no need to implement all the customary compilation steps. In particular, the tool only has to contain the Lexer and Parser steps. If, on the other hand, we implement the simulator, then we need to implement all the compilation steps.

Regarding the two compilation steps to be implemented in the tool, the Lexer reads the code file and generates a list of tokens (also known as tokenizer) and the Parser receives the list of tokens, performs syntax validation and generates the corresponding abstract syntax tree.

## Code Generation tools

In the second stage we analyzed the use of tools to generate compiler code.

Two of the most useful abstractions used in modern compilers are context-free grammars, for parsing, and regular expressions, for lexical analysis. To make the best use of these abstractions it is helpful to have special tools, such as Yacc [4], which converts a grammar into a parsing program, and Lex [5], which converts a declarative speciﬁcation into a lexical-analysis program. According to this we analyzed the following tools:

* GPlex [6]
* Gppg [7]
* Irony [8]
* JavaCC [9]
* SableCC [10]

## Implementation

The third stage comprises the plugin implementation. The plugin is divided in three components, illustrated in Figure 1.



Figure 1 - Component diagram

The BrightScript Compiler is the generated code for syntax validation, therefore preventing compilation errors from occurring when deploying the code to the box.

The Debugger manages the connection with the box, using the telnet and HTTP ports. The box exposes an HTTP port to emulate remote inputs, a web page to deploy the apps and a telnet port to receive box output and send debug commands.

The Visual Studio Plugin is based on some samples, developed by Microsoft, Python Tools [11], Visual Studio Extension for Lua [12], and Visual Studio MI Debug Engine [13]. Python Tools is an extension for Visual Studio that adds Python support; Visual Studio Extension for Lua is a basic implementation for the Lua language; Visual Studio MI Debug Engine is a debugger implementation how supports the gdb Machine Interface ("MI") specification such as GDB, LLDB, and CLRDBG.

The plugin will use the compiler for syntax highlighting, syntax analysis and intellisense generation and uses the Debugger to interact with the box.

# Compilers Theory

A compiler is a computer program that processes source code, written in a specific language, and generates machine code that computers can run.

A compiler is a very complex software. To simplify its understanding and implementation, it’s modularized. Figure 2 shows compiler’s components.



Figure 2 - Phases of a compiler, and interfaces between them, extracted from [2]

Each of the above models corresponds to a compiler phase, described in table 1.

|  |  |
| --- | --- |
| **Phase** | **Description** |
| Lex | Breaks the source ﬁle into individual words or tokens. |
| Parse | Analyzes the phrase structure of the program. |
| Semantic Actions | Builds a piece of abstract syntax tree corresponding to each phrase. |
| Semantic Analysis | Determines what each phrase means, relates uses of variables to their deﬁnitions, checks types of expressions, requests translation of each phrase. |
| Frame Layout | Places variables, function-parameters, etc. into activation records (stack frames) in a machine-dependent way. |
| Translate | Produces intermediate representation trees (IR trees), a notation that is not tied to any particular source language or target-machine architecture. |
| Canonicalize | Hoists side effects out of expressions, and cleans up conditional branches, for the convenience of the next phases. |
| Instruction Selection | Groups the IR-tree nodes into clumps that correspond to the actions of target-machine instructions. |
| Control Flow Analysis | Analyzes the sequence of instructions into a control ﬂow graph that shows all the possible ﬂows of control the program might follow when it executes. |
| Dataﬂow Analysis | Gathers information about the ﬂow of information through variables of the program; for example, liveliness analysis calculates the places where each program variable holds a still-needed value (is live). |
| Register Allocation | Chooses a register to hold each of the variables and temporary values used by the program; variables not live at the same time can share the same register. |
| Code Emission | Replaces the temporary names in each machine instruction with machine registers. |

Table 1 - Compiler phases, extracted from [2]

## Lexical Analysis

The lexical analyzer takes a stream of characters and produces a stream of names, keywords, and punctuation marks. It discards white space and comments between the tokens. The parser doesn’t have to handle possible white space and comments at every possible point.

A lexical token is a sequence of characters that is a grammar unit of a programming language. A programming language classiﬁes lexical tokens into a ﬁnite set of token types.

A language is a set of strings and a string is a ﬁnite sequence of symbols. The symbols themselves are taken from a ﬁnite alphabet.

To specify the language alphabet, we use regular expressions. Each regular expression represents a set of strings.

Regular expressions can specify lexical tokens, but we need an algorithm that can be implemented as a computer program.

For this we use ﬁnite automata. A ﬁnite automaton has a ﬁnite set of states, edges lead from one state to another, and each edge is labeled with a symbol. One state is the start state, and certain of the states are distinguished as ﬁnal states. Figure 3 shows a finite automaton sample.



Figure 3 - Combined ﬁnite automaton, extracted from [2]

Lexical analyzer needs to find the longest matches, using the finite automata we just need to remember the last finite state. Each match corresponds to a token, after find a token the automaton reinitializes.

As result of lexical analysis phase, the lexical analyzer generates a list of tokens, to be consumed by the Parser.

## Parsing

The parser makes the syntax analysis, validates the order ofwords and the way they are put together to form phrases, clauses, or sentences.

For syntax analysis we use context-free grammars. Grammars deﬁne syntactic structure declaratively. We need something more powerful than ﬁnite automata to parse languages described by grammars.

A language is a set of strings, each string is a ﬁnite sequence of symbols taken from a ﬁnite alphabet. For parsing, the strings are source programs, the symbols are lexical tokens, and the alphabet is the set of token-types returned by the lexical analyzer. A context-free grammar describes a language. A grammar has a set of productions of the form

symbol → symbol symbol ···symbol

where there are zero or more symbols on the right-hand side. Each symbol could beterminal, meaning that it is a token from the alphabet of strings in the language, or non-terminal, meaning that it appears on the left-hand side of some production. No token can ever appear on the left-hand side of a production. Finally, one of the non-terminals is distinguished as the start symbol of the grammar.

To show that this sentence is in the language of the grammar, we can perform a derivation: Start with the start symbol, then repeatedly replace any nonterminal by one of its right-hand sides. There are many different derivations of the same sentence. A leftmost derivation is one in which the leftmost nonterminal symbol is always the one expanded, in a rightmost derivation, the rightmost nonterminal is always the next to be expanded. Figure 4 shows a state table sample for a grammar.

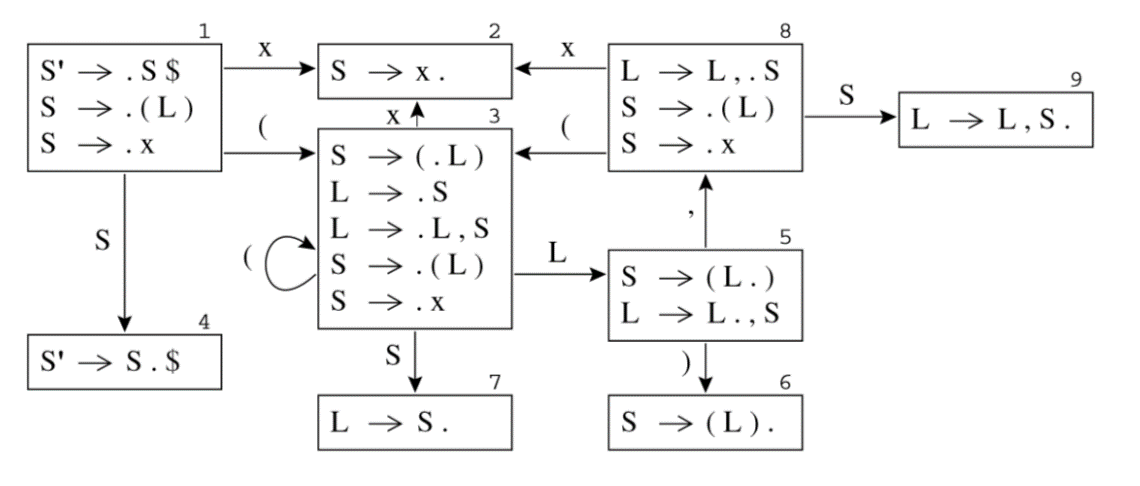


Figure 4 - LR(0) states for Grammar, extracted from [2]

A compiler must do more than recognize whether a sentence belongs to the language of a grammar, it must do something useful with that sentence. The semantic actions of a parser can do useful things with the phrases that are parsed. In a recursive-descent parser, semantic action code is interspersed with the control ﬂow of the parsing actions. In a parser implementation, semantic actions are fragments of code attached to grammar productions that generates the syntax tree.

The Abstract Syntax Tree (AST) is the result of syntax analysis phase.

## Abstract Syntax

Abstract syntax trees are data structures widely used in compilers, due to their property of representing the structure of program code. An AST is usually the result of the syntax analysis phase of a compiler. It often serves as an intermediate representation of the program through several stages that the compiler requires, and has a strong impact on the final output of the compiler. Figure 5 shows a syntax tree sample.



Figure 5 - Abstract Syntax Tree Sample, extracted from [2]

# Code Generation tools

Code generation tools are used to perform simple tasks that can be automatic. Deterministic Finite Automaton (DFA) construction is a mechanical task, so it makes sense to have an automatic lexical-analyzer generator to translate regular expressions into a DFA. The task of constructing a parser can be automated too.

## GPlex

Gardens Point LEX (GPlex) generates scanners based on ﬁnite state automata. The generated automata have the number of states minimized by default, and have a large number of options for table compression. The default compression scheme is chosen depending on the input alphabet cardinality, and almost always gives a reasonable result. However, a large number of options are available for the user to tune the behavior if necessary. The tool implements many of the FLEX extensions, including such things as start state stacks. The generated scanners are designed to interface cleanly with bottom-up parsers generated by Gardens Point Parser Generator (gppg).

GPlex is a scanner generator which accepts a “LEX-like” speciﬁcation, and produces a C# output ﬁle. The tool does not attempt to implement the whole of the POSIX speciﬁcation for LEX, however the program moves beyond LEX in some areas, such as support for Unicode.

The scanners produce by GPlex are thread safe, in that all scanner state is carried with in the scanner instance. The variables that are global in traditional LEX are instance variables of the scanner object. Most are accessed through properties which expose only a getter.

There are two main ways in which GPlex is used. In the most common case the scanner implements or extends certain types that are deﬁned by the parser on whose behalf it works. Scanners may also be produced that are independent of any parser, and perform pattern matching on character streams. In this “stand-alone” case the GPlex tool inserts the required super type deﬁnitions into the scanner source ﬁle.

The code of the scanner derives from three sources. There is invariant code which deﬁnes the class structure of the scanner, the machinery of the pattern recognition engine, and the decoding and buffering of the input stream. These parts are deﬁned in a “frame” ﬁle and a “buffers” ﬁle each of which is an embedded resource of the GPlex executable. The tables which deﬁne the ﬁnite state machine that performs pattern recognition, and the semantic actions that are invoked when each pattern is recognized are interleaved with the code of the frame ﬁle. These tables are created by GPlex from the user-speciﬁed “\*.lex” input ﬁle. Finally, user-speciﬁed code may be embedded in the input ﬁle. All such code is inserted in the main scanner class deﬁnition. Since the generated scanner class is declared partial it is also possible for the user to specify code for the scanner class in a C# ﬁle separate from the LEX speciﬁcation.

## Gppg

Gardens Point Parser Generator (Gppg) is a parser generator that produces parsers written in the C# language. Gppg generates bottom-up parsers. The generated parsers recognize languages that are LALR(1), with the traditional Yacc disambiguation’s. There are a number of extensions of the traditional input language that are necessary for correctness of the generated C# output ﬁles.

A particular feature of the tool is the optional generation of an html report ﬁle that allows easy navigation of the ﬁnite state automaton that recognizes the viable preﬁxes of the speciﬁed language. The report shows the production items, look ahead symbols and actions for each state of the automaton. It also optionally shows an example of a shortest input, and shortest FSA-path reaching each state.

Gppg parser generator accepts a “Yacc-like” speciﬁcation, and produces a C# output ﬁle. Both the parser generator and the runtime components are implemented entirely in C#.

## Irony

Irony is a development kit for implementing languages on .NET platform. Unlike most existing Yacc/Lex-style solutions, Irony does not employ any scanner or parser code generation from grammar specifications written in a specialized meta-language. In Irony, the target language grammar is coded directly in C# using operator overloading to express grammar constructs. Irony's scanner and parser modules use the grammar encoded as C# class to control the parsing process.

## JavaCC

Java Compiler Compiler (JavaCC) is the most popular parser generator for use with Java applications. A parser generator is a tool that reads a grammar specification and converts it to a Java program that can recognize matches to the grammar. In addition to the parser generator itself, JavaCC provides other standard capabilities related to parser generation such as tree building and parsing actions.

## SableCC

SableCC is a parser generator, which generates fully featured object-oriented frameworks for building compilers, interpreters and other text parsers. In particular, generated frameworks include intuitive strictly-typed abstract syntax trees and tree walkers. SableCC also keeps a clean separation between machine-generated code and user-written code which leads to a shorter development cycle.

## Conclusion

After analyzing the presented tools and the Visual Studio SDK, we get the following conclusions:

* The Visual Studio plugins only use .Net code, the Parser and the Lexer should be in C# or Visual Studio code.
* JavaCC only generates Java Lexer’s and Java Parser’s.
* SableCC can’t generate pure .Net code, only Java in C# style.
* Irony doesn’t use Yacc or Lex common definitions, the definition can’t be used in other implementations.
* GPlex generates a C# Lexer based on Lex definition.
* Gppg generates a C# Parser based on Yacc definition.

The GPlex and Gppg matches all requirements and are designed to work together. According to this we chose GPlex and Gppg.

# Compiler

The Compiler is composed of two components: The Lexer (also called Scanner) and the Lexer. The purpose of the compiler is to process code files, generating compilation errors, if they exist, and build the abstract syntax tree for intellisense functionality. Figure 6 shows components diagram of the compiler.



Figure 6 - Compiler Components

## Lexer

The lexical analyzer makes lexical validation and generates tokens for the Parser. This validation consists in verifying if the code is according to the lexical definition for this language.

The lexical analyzer is a state machine that tries to find the longest tokens. It reads character by character changing to the possible states. When it founds a token, removes the string form the source and generates the token.

The Lexer is generated by GPlex, who generates a C# file with the state machine implementation. GPlex reads the definition file (\*.lex) with the regular expressions and generates the Finite State Automata (FSA) tables. Figure 7 shows a FSA sample.



Figure 7 - FSA - Finite State Automata, extracted from [2]

The Visual Studio SDK uses the Lexer for two different purposes, the first is for syntax highlighting and the second is to generate tokens for Parser. The syntax highlighting needs less tokens then the parser, according to this we implement two Lexers one for each purpose.

## Parser

The Parser analyzes the grammatical structure of the language. It validates the sentences, if the tokens are in the right order. This analysis allows the Parser to generate the abstract syntax tree (AST), which is its output.

The grammatical structure is defined by a set of rules, named context-free grammars, defined in the Yacc file. These rules define the tokens’ order.

The Parser is generated by Gppg, which generates a C# file with the implementation of the Parser. The generated Parser implements the Shift-Reduce algorithm and generates the AST by calling actions on reduce. Figure 8 shows a shift-reduce sample table.



Figure 8 - Shift reduce table, extracted from [2]

Gppg allows to define actions written in C#, this actions returns the AST nodes and builds the AST using these nodes.

# Debugger

The Debugger is the tool that performs the deployment of the app to the box, and manages the interface between the telnet port and Visual Studio.

The debugger has three main components, namely: Telnet, Deploy and Remote, as depicted in figure 9.

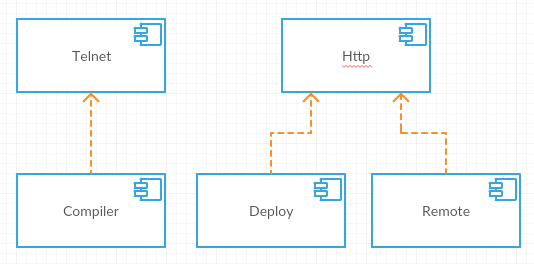


Figure 9 - Debuger diagram

The Telnet module uses a Compiler to parse the telnet output and generates debug context. The Deploy module generates the deployment package and uploads it to the box, using HTTP. The Remote module uses HTTP to emulate remote input.

The debugger will be integrated in Visual Studio Plugin, but should be independent of the Visual Studio SDK, for reuse purposes. To implement and test the functionalities we created. Figure 10 shows the debugger application.



Figure 10 - Debugger Application

The app is implemented according to the debugger diagram, depicted in Figure 9.

## Telnet module

The Telnet module has two functionalities: to show the box output, and send it debug commands.

The box output is done using print or stop keywords in the source code files. The print keyword writes to the output, similarly to the printf function from C’s standard library. The stop keyword makes puts the box in debug mode. In this mode, the box accepts debug commands, which are described in table 2.

|  |  |
| --- | --- |
| **Command** | **Description** |
| **bsc** | Print current BrightScript Component instances |
| **bscs** | Print a summary of BrightScript component instance counts by component type. |
| **brkd** | Toggle whether BrightScript should break into the debugger after non-fatal diagnostic messages. |
| **bt** | Print backtrace of call function context frames |
| **classes** | Print Brightscript Component classes |
| **cont**or**c** | Continue Script Execution |
| **down** or **d** | Move down the function context chain one |
| **exit** | Exit shell |
| **gc** | Run garbage collector |
| **help** | Print the list of debugger commands |
| **last** | Print the last line that executed |
| **list** | List current function |
| **next** | Print the next line to execute |
| **print, p,**or**?** | Print a variable or expression |
| **step, s, or t** | Step one program statement |
| **over** | Step over function |
| **out** | Step out of a function |
| **up** or **u** | Move up the function context chain one |
| **var** | Print local variables and their types/values |
| Any Brightscript statement | Execute an arbitrary Brightscript statement |

Table 2 - Debug commands, extracted from [3]

Using the debugger, it’s possible to get the current variables’ value and get information about the call stack.

The implementation of the component uses a socket to connect to the telnet port and a compiler that compiles the debugger output, to generate the debugger context. Figure 11 shows the Telnet component diagram.



Figure 11 - Telnet component diagram

The compiler recognizes the call stack and local variables patterns and parses them. This information is shown on Call Stack and Local Variables windows. The output of the component is shown on Console window. Figure 12 shows the compiler’s output windows.



Figure 12 - Compiler output windows

The output visualizer uses the socket output to show the text and send debug commands. Figure 13 show telnet output visualizer.



Figure 13 - Output visualizer

The tool bar contains icons that correspond to debug commands and makes more easy to use the debugger. Figure 14 shows app tool bar.



Figure 14 - UI Commands

## Deploy

The deploy process consists in generate a zip file with the code files and upload it to the box, using http port.

The process is configurable and should allow to:

* Select the folders to send
* Inject break points
* Remove specific parts of code
* Parameterize package, injecting code
* Execute unit tests
* Generate/Edit manifest file

The deploy needs to implement four steps:

1. Copy files to specific folder
2. Automatic edition, for configurations
3. Zip generation
4. Upload zip to the box

The process uses a several configurations, to support this was created a page for manage deploy configurations. Figure 15 shows the configuration window.



Figure 15 - Configurations Window

The upload step needs the developer user and password of the box to access http port. The optimize configuration will add an extra task that removes comments, empty lines and extra spaces. The includes section is the list of sub folders to include in the zip. The exclude section is the list of sub folders to delete (for unit tests purposes). The extra configs section is for use in replaces. The replaces is to replace code on files (allows to inject build configurations, replacing code).

SkyStore App is deployed using Coffee Script, to maintain compatibility with SkyStore and other Roku apps that is deployed using make files or other command line tools, was created a graphical component that allows to use Cygwin console. Figure 16 shows the Cygwin console window.

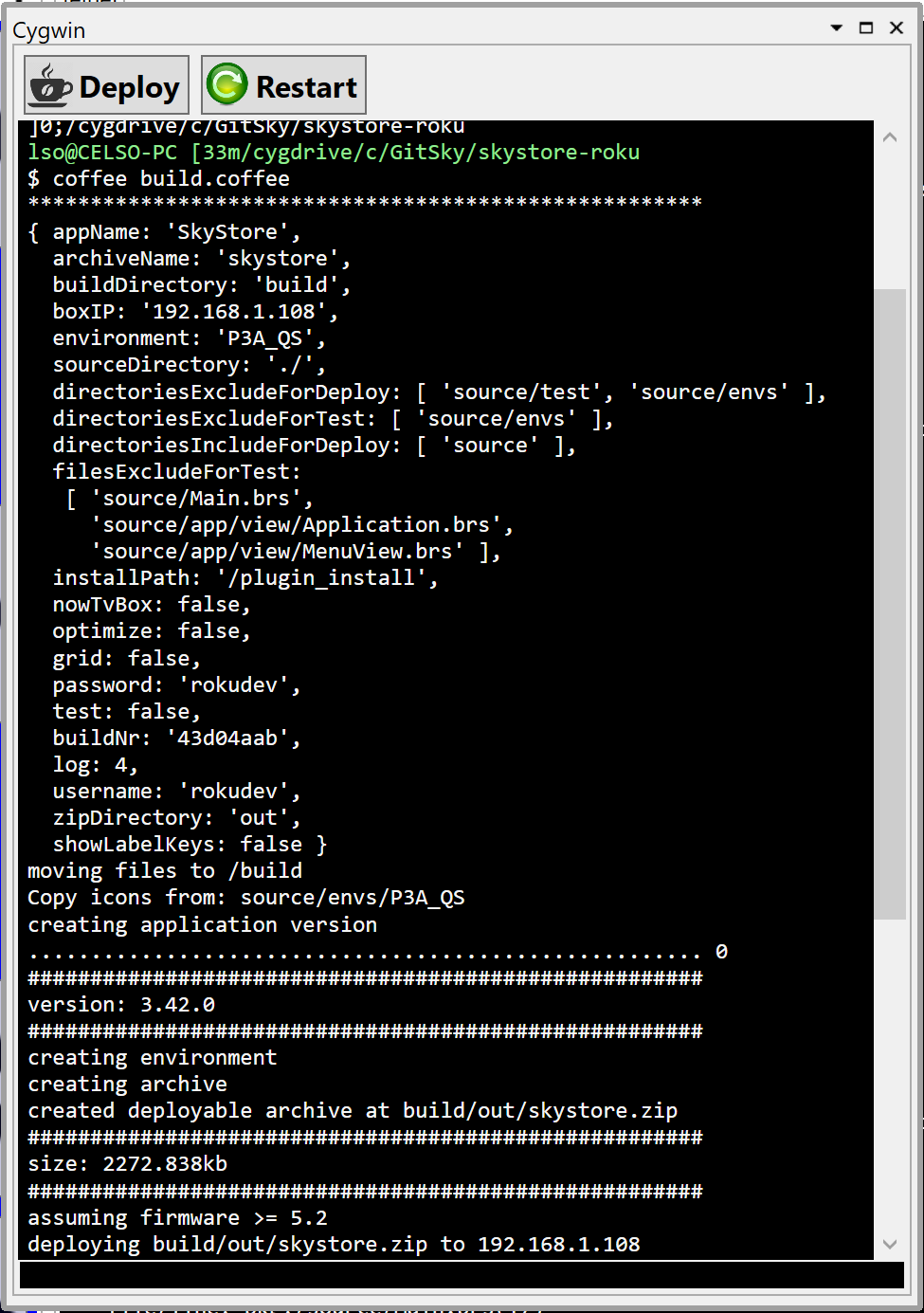


Figure 16 - Cygwin console

## Remote

Roku boxes accept remote commands over http, using this feacture we build a remote component, that allows to control the box form the PC.

The remote has the corresponding buttons and a text box that allows to send text to the box. Figure 17 shows remote window.



Figure 17 - Remote

# Visual Studio Plugin

The Visual Studio has four different components, the project type, the builder/deploy, the editor extensions and the debugger.

Visual Studio has several ways of extend his behavior. We are using “MefComponent”, “VsPackage”, “ProjectTemplate” and “ItemTemplate”.

The “MefComponent” uses Mef [13], MEF is a library for creating lightweight, extensible applications. It allows Visual Studio to discover and use extensions with no configuration required.

The “*VsPackage*” uses an implementation of “*Package*” to register extensions.

The “*ProjectTemplate*” exposes project templates to be used on project dialogs.

The “*ItemTemplate*” exposes templates to be used on new item dialog.

## Project Type

The project type provides the templates to create the BrightScript project and for create the code files. The project type will provide the settings to show on UI and to be used by the other components.

In project type implementation we used VSProjectSystem [14], this base implementation has the most common code to create project types.

The project template defines the base “*bsproj*” that needs to be registered on plugin to appears on new project dialog. Figure 18 shows the project dialog window.



Figure 18 - Visual Studio project dialog

The project template should have the base code files to be added on project creation. We could define a project factory to inject code on templates. Figure 19 shows project structure, after being created.



Figure 19 - Project created

Using VSProjectSystem [14], we need to define all item templates and register them, for the code files appear in solution explorer.

The item templates are base code files, that appears on new item dialog and has the base code structure. Figure 20 shows new item dialog.



Figure 20 - New item dialog

Figure 21 shows create file, after being created.



Figure 21 - Base code file

## Builder/Deploy

The builder/deploy component implement’s the following tasks:

|  |  |
| --- | --- |
| **Task** | **Description** |
| CompileTask | Uses the Compiler to validate the code before deploy. |
| CopyToOutputTask | Copy files to output path |
| GitVersionTask | Generates a output parameter width last tag from Git repository |
| ManifestTask | Edits manifest file replacing app name and version |
| ReplacesTask | Replace specific file contents. Used to pass parameters for app |
| OptimizeTask | Remove following code:   * Empty lines * Comments * Duplicated spaces * Code beteween ‘%--‘ and ‘--%’ tags |
| ArchiveTask | Generates the zip file to be deployed |
| DeployTask | Uploads the zip file to the box |

Table 3 – MSBuild tasks

All of these tasks, except CompilerTask, uses the implementations created for debugger, to deploy the apps.

## Editor Extension

The editor extensions provide syntax highlighting, compiler errors and intellisense.

The syntax highlighting is supported by the simpler Lexer, that generates tokens for syntax highlighting. To expose fuctionality we need to create and export a class that implements “*ITaggerProvider”* interface. Figure 22 shows syntax highlighting on editor.



Figure 22 - Syntax highlighting

The compiler errors are supported by the Parser who compiles the code and generate the compiling errors. The compiling errors is shown on editor as underlining mark and are listed on error window.

The editor errors are presented using a “*ITaggerProvider”* for “*ErrorTag”*. The Error tagger will expose the compiler errors. Figure 23 shows error underline.



Figure 23 - Editor error

The error window uses an implementation of “*ErrorListProvider” class* to show the generated errors. Figure 24 shows the error on error window.



Figure 24 - Error window

The intellisense uses the Parser to generate the AST and needs to implement an instance of “*ICompletionSourceProvider*” interface to generate the list of suggestions. We need to implement “*IVsTextViewCreationListener*” to register the command handler to show the list. Figure 25 shows intellisense popup.



Figure 25 - intellisense

Visual Studio generates events on every file change, this would make the code to be compile in different points for the same code. To avoid this, we cache the result of the compilation.

## Debugger integration

The debugger was based on Visual Studio MI Debug Engine [13] and use the implementation developed for Debugger App, that manage the connection with the box and shows debug output. The debugger output and debugger commands will be integrated with Visual Studio debugger extension.

To implement a debugger, we need to implement the following steps:

* Implement Visual Studio SDK debugger interfaces
  + Debug Engine
  + Debug Thread
  + Debug Stack Frame
  + Debug Document Context
  + Debug Events
  + Interop Structures
* Register Debug Engine
* Create Debug Launcher

Debug engine is the interaction between the Visual Studio and our debugger. This class exposes methods for Launch, Resume, Suspend, Step, and Terminate debugger.

Debug Thread represents a thread on debugging app, exposes thread id, thread name, and current stack list.

Debug Stack Frame represents a frame of stack list.

Debug Document Context represents the document and line of specific stack frame.

Debug Events implements the debugger events, to be send to Visual Studio. Using these events, we can pass current state of debugger to Visual Studio. For example, when program hits a break point.

Interop Structures used to pass data on events.

Debugger engine needs to be registered in Windows register to be used, this is implemented using a registration attribute, used on Visual Studio package.

Visual Studio needs a Launcher implementation to specify witch Debug Engine to create.

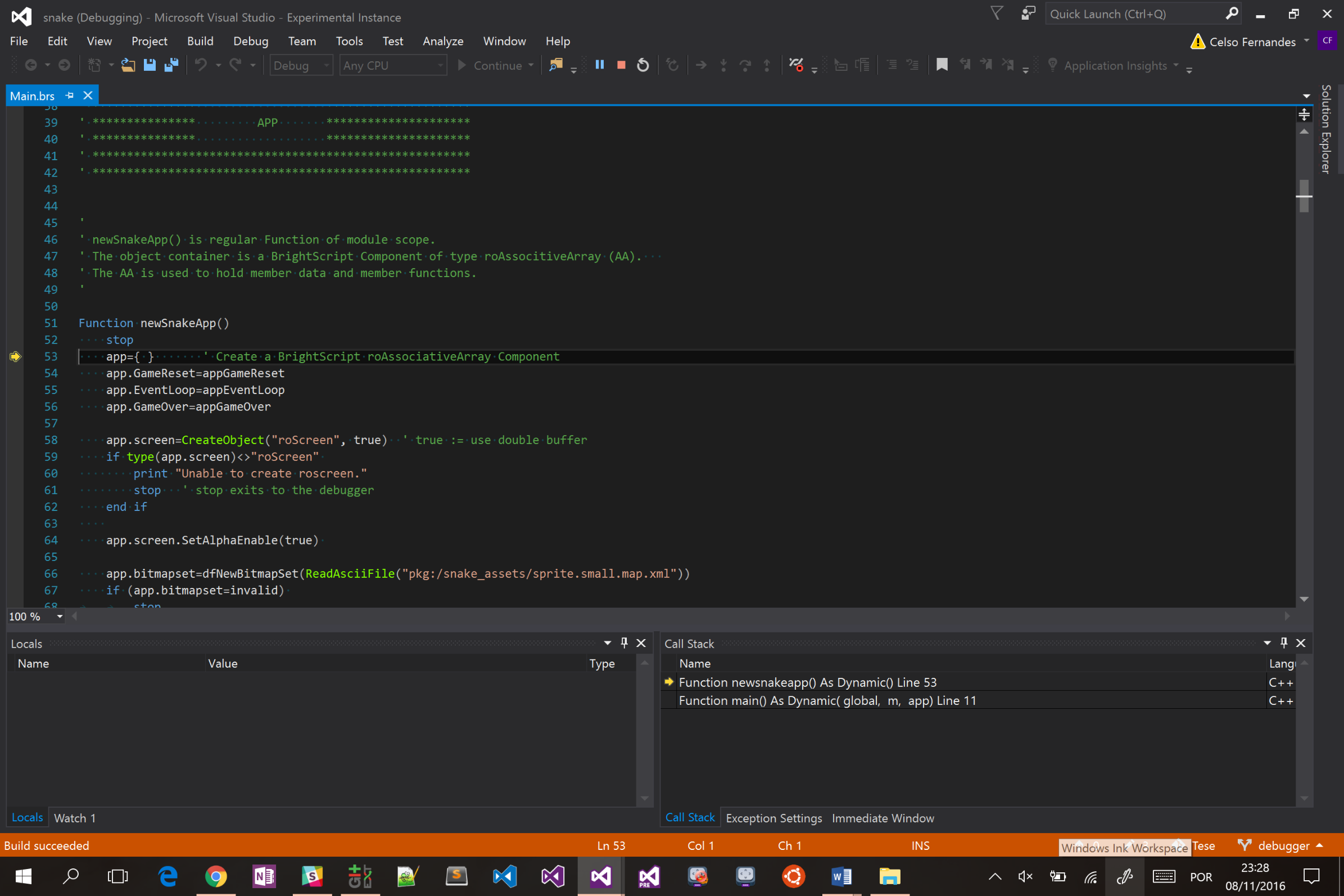


Figure 26 - Debugging BrightScript

# Conclusion

When the project starts we need to collect all the knowledge of theory and how to make IDEs extension. It was a long research to get all knowledge and it needs to be refresh along the implementation. When we join all the pieces it was very gratefully to see the result.

The Visual Studio plugin isn’t yet usable, but the debugger app is a very useful tool on day to day work and it makes much more easy to debug the apps. The all project will be much more useful.

We identify the following future work tasks:

* Implement Simulator
* Integrate Scene Graph framework
  + Scene Graph Parser
  + Debug additional threads

# References

|  |  |
| --- | --- |
| [1] | A. Aiken, "Compilers Theory," [Online]. Available: https://www.youtube.com/playlist?list=PLLH73N9cB21VSVEX1aSRlNTufaLK1dTAI. |
| [2] | A. W. Appel, Modern Compiler Implementation in Java, Cambrige University Press, 2002. |
| [3] | "BrightScript," [Online]. Available: https://sdkdocs.roku.com/display/sdkdoc/Roku+SDK+Documentation. |
| [4] | [Online]. Available: http://dinosaur.compilertools.net/#yacc. |
| [5] | [Online]. Available: http://dinosaur.compilertools.net/#lex. |
| [6] | "GPlex," [Online]. Available: http://gplex.codeplex.com/. |
| [7] | "Gppg," [Online]. Available: https://gppg.codeplex.com/. |
| [8] | [Online]. Available: https://irony.codeplex.com/. |
| [9] | [Online]. Available: https://javacc.java.net/. |
| [10] | [Online]. Available: http://www.sablecc.org/. |
| [11] | "Python Tools," [Online]. Available: https://github.com/Microsoft/PTVS. |
| [12] | "Visual Studio Extension for Lua," [Online]. Available: https://github.com/Microsoft/VSLua. |
| [13] | Microsoft, "Visual Studio MI Debug Engine," [Online]. Available: https://github.com/Microsoft/MIEngine. |
| [14] | "MEF - Managed Extensibility Framework," [Online]. Available: https://msdn.microsoft.com/en-us/library/dd460648(v=vs.110).aspx. |
| [15] | "VSProjectSystem," [Online]. Available: https://github.com/Microsoft/VSProjectSystem/. |
| [16] | "PowerShell Tools," [Online]. Available: https://github.com/adamdriscoll/poshtools. |

# Figures Index

[Figure 1 - Component diagram 2](#_Toc466411223)

[Figure 2 - Phases of a compiler, and interfaces between them, extracted from [2] 4](#_Toc466411224)

[Figure 3 - Combined ﬁnite automaton, extracted from [2] 6](#_Toc466411225)

[Figure 4 - LR(0) states for Grammar, extracted from [2] 7](#_Toc466411226)

[Figure 5 - Abstract Syntax Tree Sample, extracted from [2] 8](#_Toc466411227)

[Figure 6 - Compiler Components 12](#_Toc466411228)

[Figure 7 - FSA - Finite State Automata, extracted from [2] 13](#_Toc466411229)

[Figure 8 - Shift reduce table, extracted from [2] 14](#_Toc466411230)

[Figure 9 - Debuger diagram 15](#_Toc466411231)

[Figure 10 - Debugger Application 16](#_Toc466411232)

[Figure 11 - Telnet component diagram 18](#_Toc466411233)

[Figure 12 - Compiler output windows 19](#_Toc466411234)

[Figure 13 - Output visualizer 19](#_Toc466411235)

[Figure 14 - UI Commands 20](#_Toc466411236)

[Figure 15 - Configurations Window 21](#_Toc466411237)

[Figure 16 - Cygwin console 22](#_Toc466411238)

[Figure 17 - Remote 23](#_Toc466411239)

[Figure 18 - Visual Studio project dialog 25](#_Toc466411240)

[Figure 19 - Project created 25](#_Toc466411241)

[Figure 20 - New item dialog 26](#_Toc466411242)

[Figure 21 - Base code file 26](#_Toc466411243)

[Figure 22 - Syntax highlighting 27](#_Toc466411244)

[Figure 23 - Editor error 28](#_Toc466411245)

[Figure 24 - Error window 28](#_Toc466411246)

[Figure 25 - intellisense 28](#_Toc466411247)

[Figure 26 - Debugging BrightScript 30](#_Toc466411248)