

Cesium : A New Language

William B. Hart

November 23, 2012

Contents

1	Introduction	1
2	The Cesium interpreter	2
3	Code comments	2
4	Literals	2
4.1	Integer literals	2
4.2	Floating point literals	3
4.3	Character literals	3
4.4	String literals	3
5	Expressions	4
5.1	Tuples	4

1 Introduction

Cesium is a new programming language designed primarily for efficiency, easy interface with C libraries and with a straightforward high-level syntax. It uses the LLVM Jit to give very performant code, even when run in interactive mode. The language is polymorphic, statically typed (with type inference) and imperative in style, with some functional features.

The language fills a niche between Python and C. Python is not terribly performant, but has a very simple and flexible syntax. C is fiddly and low level, statically compiled and not interactive, but very fast.

Cesium is inspired by Julia, but is more suited to computer algebra, rather than numerical tasks. Of course Cesium can be used as a general purpose language, but design decisions are influenced by the desire to make it useful to a certain class of mathematical/computational people.

2 The Cesium interpreter

The Cesium interpreter is started by typing `./cs` at the command line.

When the interpreter starts, a prompt is displayed ready to accept Cesium code.

If you type an expression and press enter, Cesium immediately executes the expression and a return value is displayed. If there is no return value, `none` is printed.

Note that end-of-lines are significant in Cesium. In many cases they can be elided by addition of certain keywords. However, in general, statements must finish with an end-of-line.

3 Code comments

Cesium allows you to add comments to your code, which are ignored by the parser. For example, `/* this is a comment */`.

Comments are allowed anywhere that a space is allowed.

```
if 1 < 2 /* check if 1 is less than 2 */
  3 /* if it is, return 3 */
else
  4 /* if not, return 4 */
end /* end of if */
```

4 Literals

We first describe how to enter various types of constant literals in Cesium.

4.1 Integer literals

Here are some examples of signed and unsigned integers in Cesium:

```
1234
12u8
109i8
278009804u
234567i32
```

Integers with no suffix are signed and either 32 or 64 bits, depending on whether a 32 or 64 bit binary of Cesium is being used.

Integers which are unsigned are given the suffix `u`.

The other suffixes, e.g. `i8`, `u32`, etc., specify whether the number is signed (with a `i`) or unsigned (with a `u`) and how many bits the type is. The valid numbers of bits are 8, 16, 32, 64.

For example, the literal `109i8` is the integer 109 stored as a signed 8 bit value.

The suffixes are necessary in Cesium because it uses type inference to infer the type of a value from the value itself. In the case of numbers this is not possible without the extra annotation.

In practice, only unannotated types are required unless one has some specific bit fiddling to do or if one is interfacing with a C program which makes use of specific types.

4.2 Floating point literals

There are two types of floating point values in Cesium, `float` and `double` values.

Floating point values must have a decimal point and have an optional exponent. The exponents begin with an optional sign, followed by an integer.

Without a suffix, floating point values are considered to be double precision floating point values. However, if they have the suffix `f` they are considered to be single precision floats.

Here are some example floating point values:

```
1.0
2.1f
123.4
1.3e+12
1.4E-11f
```

For example, `1.4E-11f` is the single precision floating point value 1.4×10^{-11} .

4.3 Character literals

Cesium has a `char` type for single ASCII characters. To represent a character one simply wraps it in single quote marks.

```
'c'
's'
','
','
```

Some special characters are also available. These characters all begin with a backslash:

```
'\n' /* new line */
'\r' /* carriage return */
'\t' /* tab */
'\0' /* null character */
'\' ' /* backslash */
'\\" /* double quotes */
'\'' /* single quotes */
```

4.4 String literals

One can represent strings in Cesium by putting them between double quotation marks:

```
"this is a string"
"123"
"%51"
```

To include a double quotation mark in a string, it must be escaped with a backslash:

```
"this \"string\" contains double quotes"
```

The type of a string value in Cesium is `string`.

5 Expressions

When an expression is entered at the Cesium prompt, it immediately evaluates the expression and prints the result. For example, typing the expression `1 + 1` at the prompt results in the value `2`.

We now discuss the various ways of constructing expressions in Cesium.

5.1 Tuples

A tuple of values can be entered by separating the individual expressions by commas and wrapping in parentheses. Note that a tuple with only one entry must have a comma before the final parenthesis to distinguish it from a set of parentheses around an expression which are acting as a grouping rather than as a tuple constructor.

Here are some examples of tuples expressions:

```
(1, 2, 3)
(0,)
(3, 'c', "string", (2.3, 1.1f, 3u8))
```

A tuple can be thought of as an anonymous data type in Cesium. It can contain inhomogeneous data, i.e. the types of the expressions in the tuple do not need to be the same. This is different to arrays, which must have homogeneous data. However, unlike arrays, tuples cannot be accessed by index. To extract values from a tuple, one needs to use tuple unpacking, described later.

The type of a tuple value is given by a tuple of types, e.g. the tuple `(1, 'c', "string")` has type `(int, char, string)`.