

# Ruby Source Code Notes

Nicholas Yang

June 5, 2018

## Contents

<b>1</b>	<b>Intro</b>	<b>1</b>
<b>2</b>	<b>Globals</b>	<b>1</b>
<b>3</b>	<b>Helper Functions/Macros</b>	<b>1</b>
<b>4</b>	<b>Types</b>	<b>2</b>
<b>5</b>	<b>Lexer</b>	<b>2</b>
<b>6</b>	<b>Terminals</b>	<b>3</b>
<b>7</b>	<b>Grammar Rules</b>	<b>3</b>
<b>8</b>	<b>Sources</b>	<b>3</b>

## 1 Intro

These are my notes for the Ruby source code. I hope to document my attempts to understand the Ruby code.

## 2 Globals

- `Qundef` — (?) A magic number that represents undefined

## 3 Helper Functions/Macros

- `has_delayed_token` — For Ruby parser, always false

## 4 Types

- `parser_params` — Oh boy, this is a complicated one. Consists of a massive struct that contains:
  - A `rb_imemo_tmpbuf_t` pointer called `heap` (so I'm guessing this is just the heap/free list for the parser?).
  - A `YYSTYPE` pointer called `lval`
  - A struct called `lex`
  - Two `stack_type`, i.e. `VALUE`'s called `cond_stack` and `cmdarg_stack`
  - A bunch of `int`'s:
    - \* `tokidx`
    - \* `toksiz`
    - \* `tokline`
    - \* `heredoc_end`
    - \* `heredoc_indent`
    - \* `heredoc_line_indent`
    - \* `line_count`
    - \* `ruby_sourceline`
- `VALUE` — `uintptr_t`
- `yyltokentype`
- `YYSTYPE` —
- `rb_imemo_tmpbuf_t` — A linked list with a payload consisting of two `file:///Users/nicholas/cs_projects/gsoc_ruby_types/~VALUE~`'s: `flags` and `reserved` and a `VALUE` pointer called `...~ptr~` which apparently is a `malloc`'d buffer. Also has a `size_t` called `cnt` which is the size of that buffer. Finally has a pointer to another `rb_imemo_tmpbuf_struct` called `next`.

## 5 Lexer

- `yylex` — A wrapper for `parser_yylex` with some extra options. Initializes a `yyltokentype`. Assigns `p`, a struct of `file:///Users/nicholas/cs_projects/gsoc_ruby_types/~parser_params~` an `lval`

(with type `YYSTYPE`) and gives `lval` a `val` of `Qundef`. (I guess initializing `p` with an empty `lval`?) Then calls `parser_yylex` with `p` and assigns the result to the `yytokentype` variable. If `has_delayed_token` (never in Ruby parser) then call `dispatch_delayed_token`. Else if `parser\yylex` returns 0 then call `dispatch_scan_event`, which, uh, in Ruby parser does nothing (well, it casts 0 to void).

- `parser_yylex` —

## 6 Terminals

- — `alias`
- — Global variable
- — `$+`, `$&`, `$‘`, `$\`

## 7 Grammar Rules

- `program` — Just the top level rule. Alias for `top_compstmt`
- `top_compstmt` — Again just `top_stmts` with an `opt_terms` add on.
- — Either nothing, one single `file:///Users/nicholas/cs_projects/gsoc_ruby_types/~top_stmt~` or a
- 

## 8 Sources

- <https://whitequark.org/blog/2013/04/01/ruby-hacking-guide-ch-11-finite-state-lexer/>
- <https://ruby-hacking-guide.github.io/>