**Tokens**

There are several tokens in HAWK: identifiers, keywords, constants, string literals, operators, and patterns. Whitespace and comments are ignored except as separators of tokens.

**Comments**

Use /\* to begin a comment, and terminate it with \*/. Comments can extend across multiple lines.

**Identifiers**

An identifier is a sequence of letters, digits, dashes, and underscores. Uppercase and lowercase letters are different. Identifiers may have any length, and are separated from other tokens by whitespace. An identifier must have at least one letter or underscore, and must begin with either of the two.

**Keywords**

The following identifiers are keywords reserved for particular use—they are not for other use:

begin double else end if int return String this while

**Constants**

There three types of constants, described in detail below.

**Integer Constants**

An integer constant consists of a sequence of digits that does not begin with 0. All integers in HAWK are taken to base 10. Negative integer constants consist of a sequence of digits prefixed by a dash.

**Double Constants**

A double constant consists of an integer part, decimal point, fractional part, and optional exponential part, which consists of an integer prefixed by an e. Either the integer or fractional part, but not both, may be missing. Both integer and fractional parts are themselves integers, separated by the decimal point. The decimal point must be present.

**String Literals**

A string literal is a sequence of characters surround by double quotes. HAWK contains several escape sequences which can be used as characters within string literals:

newline “\n”, tab “\t”, backslash “\\”, single quote “\’”, double quote “\””

String literals are immutable, and thus cannot be altered. Any operations performed a string literal will not affect the original literal but instead generate a new string.

**Table Literals**

A table literal is a comma-separated sequence delimited by curly braces. The comma-separated sequence can take two forms. (1) A sequence of values of any type—there can be different types in this sequence. This generates a table where the values are keyed in sequential order from the integer 0 to *sequence length*-1. (2) A sequence of key-value pairs written in the form *key* : *val*. As before, values can vary within a table. Keys are restricted to integer or string types, and a table can be keyed by either integers or by strings, but not both.

**Patterns**

HAWK uses two types of patterns: regular expressions (regex) and CSS selectors.

**Regular Expressions**

Regular expressions are delimited by a pair of backslashes, in the form \*regex*\. For more information on what forms regular expressions can take, refer to Section IV-B.

**CSS Selectors**

CSS selectors are delimited by a pair of at signs (@), in the form @*selector*@. For more information on the different types of CSS selectors, please visit https://developer.mozilla.org/en-US/docs/Web/CSS/Reference#Selectors.