Formal Languages and Compiler Design

https://github.com/samzirbo/Formal-Languages-and-Compiler-Design

The **Data Structure** used for **PIF** is the Python build in list that stores a tuple consisting of: token name and position/-1.

The regex used for the identifiers is: "**^[a-z][a-z0-9]*$**". It matches any sequences that start with a letter (any case - used re.IGNORECASE) and is followed by letters or digits of indefinite length. The ^ and $ is used to make sure that only strings of this pattern are matched, and not just if parts of it match the regex. The * expresses the fact that elements in the list can appear multiple times or not at all.

The regex used for the integer constants is: "**^(0|\-?[1-9][0-9]*)$**". It matches 0 or any sequence of single digits (1-9) or multiple digits that must start with a non zero digit. The sign is optional and only - is used. The ^ and $ is used to make sure that only strings of this pattern are matched, and not just if parts of it match the regex.

The regex used for string constants is: "**^\"[a-z]+\"$**". It matches any combinations of letters (any case - used re.IGNORECASE) enclosed in quotation marks. The + expresses the fact that elements in the list appears at least once.