

Expression Tokenizer

©2025, Khang Vu

Algorithm - `_get_number`

1. **Set** accumulator to numerical value of digit.
2. Advance to the next character.
3. **While** current character is a digit **do**
4. Multiply accumulator by 10 and add numerical value of current digit.
5. Advance to the next character.
6. **Return** an **Integer** token.

Algorithm - `_get_identifier`

1. **Set** an identifier string to nil.
2. **Do**
3. Append the current character to the identifier string
4. Advance to the next character
5. **While** the current character is alphabetic **do**
6. **If** the identifier is in the identifier map **then**,
7. **Return** the associated token from the identifier map
8. **Else**
9. Add the identifier to the variable map.
10. **Return** a **Variable** token.

Algorithm - `tokenize`

1. **While** the current character is a space or tab **do**
2. Advance to the next character.
3. **If** the current character is an end-of-string **then**,
4. **Return** an **end-of-expression** token.
5. **If** current character is a digit then,
6. `_get_number()`
7. **If** current character is an equal sign **then**,
8. Advance to the next character.
9. **Return** an **Assignment** token.
10. **If** current character is an open-parenthesis **then**,
11. Advance to the next character.
12. **Return** an **Opening-Parenthesis** token.
13. **If** current character is a closed-parenthesis **then**,
14. Advance to the next character.

15. **Return** a **Closing-Parenthesis** token.
16. **If** current character is an asterisk **then**,
17. Advance to the next character.
18. **Return** a **Multiplication** token.
19. **If** current character is a virgule¹ **then**,
20. Advance to the next character.
21. **Return** a **Division** token.
22. **If** current character is an exclamation mark **then**,
23. Advance to the next character.
24. **If** the previous token is a **Closing-Parenthesis**, **Operand** or **Postfix-Operator** **then**,
25. **Return** a **Factorial** token.
26. **Else**
27. **Exception** "**Factorial must follow Expression**".
28. **If** current character is a plus-sign **then**,
29. Advance to the next character.
30. **If** the previous token is a **Closing-Parenthesis**, **Operand** or **Postfix-Operator** **then**,
31. **Return** an **Addition** token.
32. **Else**
33. **Return** an **Identity** token.
34. **If** current character is a dash **then**,
35. Advance to the next character.
36. **If** the previous token is a **Closing-Parenthesis**, **Operand** or **Postfix-Operator** **then**,
37. **Return** a Subtraction token.
38. **Else**
39. **Return** a **Negation** token.
40. **If** current character is an alphabetic character **then**,
41. _get_identifier()

¹ A virgule: The forward slash '/' character