

20174677전상호
20154093이정혁

code 설명문서

1. 사용언어 및 빌드 OS
python3을 이용해 제작. window 환경에서 개발 후 window exe빌드, ubuntu를 활용해 linux execution 파일 빌드.
2. input 처리
일단 토큰의 종류를 크게 3가지로 분류.
 - a) 특수문자로 시작하는 토큰(A_O, B_O, LPAREN, 등등...)
 - b) 문자로 시작하는 토큰(bool, type, string, keyword, identifier)
 - c) 숫자 토큰(int, float ...)들어온 character를 아스키코드로 변환 후 범위에 따라 만들어 놓은 함수로 이동.
함수 내에서는 들어온 문자와 인덱스를 가지고 해당 토큰을 판단.
들어온 인덱스 다음 인덱스부터 탐색하여 조건에 어긋날때까지 판단.
그 다음 탐색을 위한 index를 반환.
판단된 토큰은 token, value리스트에 저장됨.

예외처리

- A) -부호가 들어왔을 때 int나 float의 부호인지, arithmetic operator(A_O)인지 구분하지 못함.
해결방법으로는 -문자를 새로운 undefined_token과 같이 새로운 토큰을 생성하여 syntax analyzer에서 정의를 해주는 방법이 있고, lexical analyzer에서 앞 뒤 토큰을 확인하여 판단하는 방법이 있을 것으로 생각, lexical analyzer를 만드는 과제이므로 후자의 방법을 선택함.
-부호 뒤에 whitespace 없이 숫자가 오는 경우는 int 나 float으로 판단, 나머지 경우는 모두 AO로 판단
- B) float 숫자의 정의에 대한 예외처리.
숫자를 읽어오는 중 "." character가 들어오게 되면 dot이라는 bool문을 true로 설정. 이 dot을 가지고 float형인지 int형인지 판단하고, float형의 숫자의 정의를 만족하는지 판단. "."character가 2개 이상 있거나, 0이 아닌 다른 수로 끝나면 error를 출력.