

# Database

## Project 1-1: SQL Parser

컴퓨터공학부  
2017-18538  
황선영

### 1. 핵심 모듈과 알고리즘

Python parsing module인 Lark을 이용하여 grammar.lark에 EBNF 형식으로 parsing grammar를 정의한 후 Lark 함수를 써서 parser를 instance화 하였다.

grammar.lark 파일을 열어 Lark함수에 custom한 transformer와 함께 parameter로 넣어 parser를 생성한다. 그 후 input query를 받아 생성한 parser에 넣어주면 query를 분석하여 어떤 request인지 출력하거나 프로그램을 끝낸다.

### 2. 구현 내용

전체적인 알고리즘은 위에 기술한 것과 같고, 구체적인 구현 내용은 다음과 같다.

```
while True: #while loop for continued input query
    try:
        print("DB_2017-18538> ", end='')
        input_str = ""
        query = ""
        exit = "exit;"
        while True: #append input query until ';' is come
            temp = input().rstrip()
            if temp.endswith(";"):
                input_str += " " + temp
                break
            else:
                input_str += " " + temp

        if input_str.find(exit) != -1: #split input query for middle of 'exit'
            query = input_str.split(exit)[0] + exit
        else:
            query = input_str

        result = sql_parser.parse(query) #parse input query using parser
        if query.find(exit) != -1: #break if 'exit' query exist
            break

    except Exception as e: #except for catching exception
        print("DB_2017-18538> Syntax error")
```

먼저 input query를 끊임없이 받기 위해 while loop를 선언하고, 그 내부에서는 다시 while loop를 선언하여 엔터를 치더라도 input을 계속 받아 한 string 변수에 append하도록 하였다. 그 후 exit query가 중간에 있다면 exit 이후의 query는 버리고, 그렇지 않다면 그대로 유지한다. 그 후 lark으로 만든 parser를 이용해 parsing을 하는데, 이 때 try except 문을 선언하여 exception이 발생할 시

Syntax error임을 나타내게 하였다. Exception이 없는 경우 MyTransformer에서 선언한 함수들이 동작하고 exit query가 있다면 outer while loop를 break하여 프로그램을 끝낸다.

### 3. 가정한 것

Query의 끝에는 항상 semicolon이 존재하고, query의 중간에 개행 문자가 나타나더라도 query를 종료하지 않는다고 가정하였다. 여러 input이 들어왔을 때 중간에 exit이 있다면 exit 앞의 query까지만 수행하고 프로그램을 끝낸다고 가정하였다.

### 4. 구현하지 못한 내용

exit query가 대문자인 경우 제대로 request 받지 못하는 issue가 있다. MyTransformer class에서 exit 함수를 선언하여 이를 해결하고자 했지만 exit 함수를 선언해도 함수가 실행되지 않아 이 방법은 실패하고, while 문 안에서 적절한 함수를 사용하여 해결하려 했으나 방법을 찾지 못했다. exit과 semicolon 사이에 공백이 있을 때 이를 exit query로 인식하지 못하는 issue도 있다.

### 5. 느낀 점

스스로 간단한 DBMS를 구현해 보는 프로젝트를 진행하며 SQL에 대한 이해를 더 깊게 할 수 있을 것 같다. 이번 project에서는 parser를 직접 구현해봄으로써 query를 분석하는 방식에 대해 이해하였다.