

编译原理课程实践第九次作业

语法分析1-化学分子式解析

张昊 1927405160

[概述](#)

[编程环境说明](#)

[相关代码解释](#)

[运行](#)

[测试](#)

[运行](#)

概述

使用 Python3 与 PLY 模块编写程序，计算化学分子式中元素的数目（假定输入的化学式中不包含括号）。

使用如下文法：

```
species_list -> species_list species | species
species -> SYMBOL | SYMBOL COUNT
SYMBOL -> C[laroudsemf]?|Os?|N[eaibdpos]?|S[icernbmg]?|P[drmtboau]?
|H[efogas]?|A[lrsgutcm]|B[eraik]?|Dy|E[urs]|F[erm]?|G[aed]|I[nr]?|Kr?
|L[iaur]|M[gnodt]|R[buhenaf]|T[icebmah]|U|V|W|Xe|Yb?|Z[nr]
COUNT -> \d+
```

编程环境说明

- 语言：Python3
- 依赖：PLY, unittest
- 文件编码：UTF-8
- 测试环境：Python 3.8.10

相关代码解释

在语法分析的过程中，为不同的节点创建了一系列的数据结构：

```
class Atom:
    def __init__(self, symbol: str, count: int):
        self.symbol = symbol
        self.count = count

    def __str__(self):
        return f'Atom({self.symbol}, {self.count})'

class Formula:
    def __init__(self):
        self.atom_count = 0
        self.atom_list = []

    def __add__(self, other: Atom):
        self.atom_count += other.count
        self.atom_list.append(other)
        return self

    def __str__(self):
        return f'Formula(atom_count={self.atom_count}, atom=[{"",
".join(map(str, self.atom_list))}])'
```

并赋值给p[0]:

```
def p_species_list_species_list_species(p):
    """species_list : species_list species"""
    p[0] = p[1] + p[2]

def p_species_list_species(p):
    """species_list : species"""
    p[0] = Formula() + p[1]

def p_species_sym_cnt(p):
    """species : SYMBOL COUNT"""
    p[0] = Atom(p[1], int(p[2]))

def p_species_sym(p):
    """species : SYMBOL"""
```

```
p[0] = Atom(p[1], 1)
```

解析结束后将返回一个 `Formula` 对象（不成功为 `None`），可通过访问 `atom_count` 属性来获取该化学式的原子数。

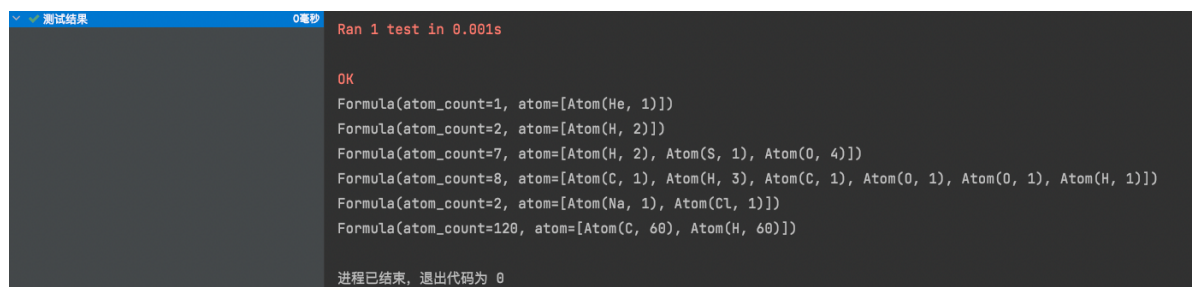
运行

测试

测试用文件为：`formula_test.py`。使用单元测试包 `unittest` 对如下测试用例进行测试：

```
atom_count("He") == 1
atom_count("H2") == 2
atom_count("H2SO4") == 7
atom_count("CH3COOH") == 8
atom_count("NaCl") == 2
atom_count("C60H60") == 120
```

测试结果如下：



```
测试结果 0毫秒 Ran 1 test in 0.001s

OK
Formula(atom_count=1, atom=[Atom(He, 1)])
Formula(atom_count=2, atom=[Atom(H, 2)])
Formula(atom_count=7, atom=[Atom(H, 2), Atom(S, 1), Atom(O, 4)])
Formula(atom_count=8, atom=[Atom(C, 1), Atom(H, 3), Atom(C, 1), Atom(O, 1), Atom(O, 1), Atom(H, 1)])
Formula(atom_count=2, atom=[Atom(Na, 1), Atom(Cl, 1)])
Formula(atom_count=120, atom=[Atom(C, 60), Atom(H, 60)])

进程已结束，退出代码为 0
```

运行

项目源文件为：`formula.py`。

程序从标准输入读取一行字符串，计算该化学式的原子数，运行方法如下：

```
$ python3 formula.py
```

样例1输入：

```
CH3COOH
```

样例1输出：

8

```
/Users/holger/codes/chemical_formula/venv/bin/python /Users/holger/codes/chemical_formula/formula.py  
请输入化学式: CH3COOH  
该化学式的原子数为: 8
```

样例2输入:

C60H60

样例2输出:

120

```
/Users/holger/codes/chemical_formula/venv/bin/python /Users/holger/codes/chemical_formula/formula.py  
请输入化学式: C60H60  
该化学式的原子数为: 120
```