

编译原理实践第 9 次课

(语法分析 1-化学分子式解析)

1. 请继续熟悉 PLY 基本编写规范

仔细阅读 PLY 使用手册

2. 熟悉 yacc_example 中分析四则运算的程序

3. 编写程序，计算化学分子式中元素的数目，并完成以下测试：

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

参考语法

```
species_list : species_list species
species_list : species
species : SYMBOL
species : SYMBOL COUNT
```

提交代码与 **readme.pdf** (**readme.pdf** 里需要有对于代码的使用，以及给定“CH3COOH”、“C60H60”等例子，其运行结果截图)

附录

(1) 元素周期表

1

H

hydrogen

[1.0078, 1.0082]

2

He

helium

4.0026

3

Li

lithium

[6.938, 6.957]

4

Be

beryllium

9.0122

11

Na

sodium

[22.989, 22.990]

12

Mg

magnesium

[24.304, 24.307]

19

K

potassium

[39.098, 39.098]

20

Ca

calcium

[40.078, 40.078]

21

Sc

scandium

44.956

22

Ti

titanium

[47.867, 47.867]

23

V

vanadium

50.942

24

Cr

chromium

51.996

25

Mn

manganese

54.938

26

Fe

iron

[55.845, 55.845]

27

Co

cobalt

58.933

28

Ni

nickel

58.693

29

Cu

copper

[63.546, 63.546]

30

Zn

zinc

[65.38, 65.38]

31

Ga

gallium

69.723

32

Ge

germanium

72.630

33

As

arsenic

[74.922, 74.922]

34

Se

selenium

[78.971, 78.971]

35

Br

bromine

[79.904, 79.907]

36

Kr

krypton

83.798

37

Rb

rubidium

85.468

38

Sr

strontium

87.62

39

Y

yttrium

88.906

40

Zr

zirconium

91.224

41

Nb

niobium

92.906

42

Mo

molybdenum

95.94

43

Tc

technetium

[98.906, 98.906]

44

Ru

ruthenium

101.07

45

Rh

rhodium

102.91

46

Pd

palladium

106.42

47

Ag

silver

107.87

48

Cd

cadmium

112.41

49

In

indium

114.82

50

Sn

tin

118.71

51

Sb

antimony

121.76

52

Te

tellurium

[127.6, 127.6]

53

I

iodine

126.90

54

Xe

xenon

131.29

55

Cs

cesium

[132.91, 132.91]

56

Ba

barium

137.33

57-71

lanthanoids

72

Hf

hafnium

[178.49, 178.49]

73

Ta

tantalum

180.95

74

W

tungsten

183.84

75

Re

rhenium

186.21

76

Os

osmium

[190.23, 190.23]

77

Ir

iridium

192.22

78

Pt

platinum

195.08

79

Au

gold

196.97

80

Hg

mercury

200.59

81

Tl

thallium

[204.38, 204.38]

82

Pb

lead

207.2

83

Bi

bismuth

208.98

84

Po

polonium

[209, 209]

85

At

astatine

[210, 210]

86

Rn

radon

[222, 222]

87

Fr

francium

88

Ra

radium

89-103

actinoids

104

Rf

rutherfordium

105

Db

dubnium

106

Sg

seaborgium

107

Bh

bohrium

108

Hs

hassium

109

Mt

meitnerium

110

Ds

darmstadtium

111

Rg

roentgenium

112

Cn

copernicium

113

Nh

nihonium

114

Fl

flerovium

115

Mc

moscovium

116

Lv

livermorium

117

Ts

tennessine

118

Og

oganesson

57

La

lanthanum

138.91

58

Ce

cerium

140.12

59

Pr

praseodymium

140.91

60

Nd

neodymium

144.24

61

Pm

promethium

[144.91, 144.91]

62

Sm

samarium

[150.36, 150.36]

63

Eu

europium

151.96

64

Gd

gadolinium

157.25

65

Tb

terbium

158.93

66

Dy

dysprosium

162.50

67

Ho

holmium

164.93

68

Er

erbium

167.26

69

Tm

thulium

168.93

70

Yb

ytterbium

173.05

71

Lu

lutetium

174.97

89

Ac

actinium

227.03

90

Th

thorium

232.04

91

Pa

protactinium

231.04

92

U

uranium

238.03

93

Np

neptunium

[237.04, 237.04]

94

Pu

plutonium

[244.06, 244.06]

95

Am

americium

[243.06, 243.06]

96

Cm

curium

[247.07, 247.07]

97

Bk

berkelium

[247.07, 247.07]

98

Cf

californium

[251.08, 251.08]

99

Es

einsteinium

[252.08, 252.08]

100

Fm

fermium

[257.10, 257.10]

101

Md

mendelevium

[258.10, 258.10]

102

No

nobelium

[259.10, 259.10]

103

Lr

lawrencium

[262.10, 262.10]



For notes and updates to this table, see www.iupac.org. This version is dated 28 November 2016.
Copyright © 2016 IUPAC, the International Union of Pure and Applied Chemistry.

(2) 识别化学元素的正则表达式

```
t_SYMBOL = (
    r"C[laroudsemf]?[Os]?[N[eaidbdpos]?[S[icernbmg]?[P[drmtboau]?]"
    r"H[eofgas]?[A[lrsgutcm]]B[eraik]?[Dy][E[urs]]F[erm]?[G[aed]]]"
    r"I[nr]?[Kr]?[L[iaur]]M[gndot][R[buhenaf]]T[icebmah]"
    r"U[V|W|Xe|Yb]?[Z[nr]]")
```

(2) 存储分子式的数据结构

```
class Atom(object):
    def __init__(self, symbol, count):
        self.symbol = symbol
        self.count = count
    def __repr__(self):
        return "Atom(%r, %r)" % (self.symbol, self.count)
```