

Molecular mass

The atomic composition of a molecule is represented by a molecular formula. Such a molecular formula represents the number of atoms of each element that occurs in the molecule. The elements have symbols that consist of a single capital letter, optionally followed by one or two lower case letters. The proper use of upper and lower case is important: Pb gives the element lead while PB represents a compound of phosphorus and boron. It is common for the elements C, H, N, O and S to be listed first, followed by the other elements in alphabetical order. This is demonstrated in the molecular formula of mustard gas ($C_4H_8S_2Cl_2$). The following table contains some examples of molecular formulas.

name	molecular formula
water	H_2O
carbon dioxide	CO_2
ethanol	C_2H_5OH
butyric acid	$C_4H_8O_2$
guanosine triphosphate	$C_{10}H_{16}N_5O_{14}P_3$

The *molecular mass* of a substance is the mass of a single molecule of the substance, expressed in atomic mass units (u). This molecular weight can be easily calculated as the sum of the atomic masses of the individual atoms of which the molecule is built up. For example, the molecular weight of water (H_2O) is the sum of the masses of the atoms of which a water molecule consists. The periodic table of elements shows that the atomic mass of hydrogen and oxygen, respectively, is equal to $1,00794\text{ u}$ and $15,9994\text{ u}$. The molecular mass of water is $2 \times 1,00794\text{ u} + 15,9994\text{ u} = 18,01528\text{ u}$.

Assignment

Given is a text file [periodic_system.txt](#), which contains a list with information about the elements in the periodic table. Each line - except for the first one which is used as a header - includes the following information about an element: *i*) atomic number, *ii*) symbolic representation, *iii*) English name, *iv*) Dutch name, and *v*) atomic mass. The information fields are separated by a tab. Asked:

1. Write a function `atomicmass` that takes the location of the text file [periodic_system.txt](#) as its argument. The function must return a dictionary as a result, which uses the symbolic representation of the elements as keys, and the atomic mass as the corresponding value.
2. Write a function `molecularmass`, which returns the molecular mass for a given molecular formula. In addition to the molecular formula itself, a second argument must be passed to the function: a dictionary that uses the symbolic representation of the elements of the periodic table as a key and links to their atomic mass. For simplicity, the various elements in a molecular formula are separated by a hyphen (-). For example, the formula for sulfuric acid is (H_2SO_4) displayed as the string `H2-S-O4`. The molecular mass of sulfuric acid can then be calculated as
$$2 \times \text{atomicmass}(\text{H}) + \text{atomicmass}(\text{S}) + 4 \times \text{atomicmass}(\text{O}) \\ = 2 \times 1,00794\text{ u} + 32,066\text{ u} + 4 \times 15,9994\text{ u}$$

u

The function should return the real value -1 if the given molecular formula contains the symbolic representation of an element which is not found in the given dictionary.

Example

In the following interactive session we assume that the text file [periodic_system.txt](#) is located in the current directory.

```
>>> index = atomicmass('periodic_system.txt')
>>> index['Si']
28.0855
>>> molecularmass('H2-S-O4', index)
98.07948
```

De atomaire samenstelling van een molecuul wordt weergegeven met een molecuulformule. Een dergelijke molecuulformule geeft het aantal atomen aan van elk element dat in het molecuul voorkomt. De elementen hebben symbolen die bestaan uit één hoofdletter, eventueel gevolgd door één of twee kleine letters. Het juiste gebruik van hoofdletters en kleine letters is van groot belang: Pb geeft het element lood aan terwijl PB staat voor een verbinding van fosfor en boor. Het is gebruikelijk dat de elementen C, H, N, O en S als eerste genoemd worden, gevolgd door de overige elementen in alfabetische volgorde. Dit is bijvoorbeeld te zien bij de molecuulformule van mosterdgas ($\text{C}_4\text{H}_8\text{SCl}_2$). Onderstaande tabel bevat enkele voorbeelden van molecuulformules.

naam	molecuulformule
water	H_2O
koolstofdioxide	CO_2
ethanol	$\text{C}_2\text{H}_5\text{OH}$
boterzuur	$\text{C}_4\text{H}_8\text{O}_2$
guanosinetrifosfaat	$\text{C}_{10}\text{H}_{16}\text{N}_5\text{O}_{14}\text{P}_3$

De *molecuulmassa* (of *moleculaire massa*) van een stof is de massa van één molecuul van die stof, uitgedrukt in atomaire massa-eenheden (u). Deze molecuulmassa kan makkelijk berekend worden als de som van de atoommassa's van de afzonderlijke atomen waaruit het molecuul is opgebouwd. Om bijvoorbeeld de molecuulmassa van water (H_2O) te berekenen, wordt de som genomen van de massa's van de atomen waaruit een molecuul water bestaat. Het periodiek systeem der elementen leert ons dat de atoommassa van waterstof en zuurstof respectievelijk gelijk is aan $1,00794 \text{ u}$ en $15,9994 \text{ u}$. De moleculaire massa van water is dus $2 \times 1,00794 \text{ u} + 15,9994 \text{ u} = 18,01528 \text{ u}$.

Opgave

Gegeven is een tekstbestand [periodiek_systeem.txt](#), dat een lijst met informatie over de elementen uit het periodiek systeem bevat. Elke regel — behalve de eerste die als hoofding geldt — bevat de volgende informatie over een element: *i*) atoomnummer, *ii*) symbolische voorstelling, *iii*) Engelstalige naam, *iv*) Nederlandstalige naam, en *v*) atoommassa. De informatievelden worden van elkaar gescheiden door een tab. Gevraagd wordt:

1. Schrijf een functie `atoommassa` waaraan de locatie van het tekstbestand [periodiek_systeem.txt](#) moet doorgegeven worden. De functie moet een dictionary als resultaat teruggeven, die de symbolische voorstelling van de elementen als sleutels gebruikt, en de atoommassa als corresponderende waarde.
2. Schrijf een functie `molecuulmassa` die de molecuulmassa teruggeeft voor een gegeven molecuulformule. Naast de molecuulformule zelf, moet nog een tweede argument aan de functie doorgegeven worden: een dictionary die de symbolische voorstelling van de elementen uit het periodiek systeem als sleutel gebruikt en koppelt aan hun atoommassa. Voor de eenvoud worden de verschillende elementen in een molecuulformule van elkaar gescheiden door een koppelteken (-). Zo wordt de formule voor zwavelzuur (H_2SO_4) weergegeven als de string `H2-S-O4`. De molecuulmassa van zwavelzuur kan dan berekend worden als

$$2 \times \text{atoommassa}(H) + \text{atoommassa}(S) + 4 \times \text{atoommassa}(O)$$

$$= 2 \times 1,00794 + 32,066 + 4 \times 15,9994 = 98,07948$$
 De functie moet de reële waarde `-1` teruggeven indien de gegeven molecuulformule de symbolische voorstelling van een element bevat dat niet wordt teruggevonden in de gegeven dictionary.

Voorbeeld

In onderstaande interactieve sessie gaan we ervan uit dat het tekstbestand [periodiek_systeem.txt](#) zich in de huidige directory bevindt.

```
>>> index = atoommassa('periodiek_systeem.txt')
>>> index['Si']
28.0855
>>> molecuulmassa('H2-S-O4', index)
98.07948
```