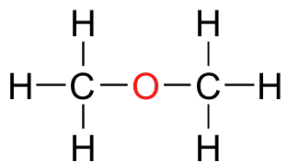
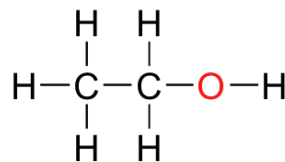


# Isomers

In chemistry, isomers are compounds that are the same in the sense that they contain the same number and the same kind of atoms, but differ in the way these atoms are interconnected and arranged. In other words, isomers are compounds with the same molecular formula but different structural formulas (a compound is also an isomer of itself). Ethanol (an alcohol) and dimethyl ether, for example, each consist of 6 hydrogen atoms, two carbon atoms, and an oxygen atom. Only the order of the chemical bonds is different. Both compounds share the same molecular formula ( $\text{C}_2\text{H}_6\text{O}$ ) but have different structural formulas.



dimethyl ether



ethanol

There are various standards that represent the structural formula of a compound as a string. You may assume that the atoms in such a string representation are symbolically named by an uppercase letter, possibly followed by a number of lowercase letters. The difference between uppercase and lowercase letters is important: the string PB represents one phosphorus atom (P) and one boron atom (B), while the string Pb represents one atom of lead. In addition, in between symbolic names of different atoms there might be other characters, such as in Na-OH or (CHHH)CO(CHHH). No other characters may occur in between the uppercase letter and the lowercase letters of the symbolic name of an atom.

## Assignment

1. Write a function `molecular` that takes the string representation of a structural formula of a particular compound as its argument. The function must return a dictionary, whose keys are the symbolic names of all atoms occurring in the given structural formula. The value that corresponds with a key indicates how many times the atom occurs in the formula. For example, if the string (CHHH)CO(CHHH) is passed, the function should return the dictionary {'H': 6, 'C': 3, 'O': 1}. Note that this dictionary actually is a representation of the molecular formula that corresponds to the structural formula given, hence the name of the function.
2. Use the function `molecular` to write a function `isomers`. This function takes two string representations of structural formulas as its arguments. The function must return a Boolean value that indicates whether or not the compounds that correspond with the given structural formulas are isomers.

## Example

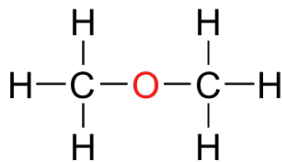
```
>>> structural1 = 'OCaOSeOO'
>>> structural2 = 'HHCHHCHHCHHCHH'
>>> structural3 = 'HHCHHHCCHHHCHH'

>>> molecular(structural1)
{'Ca': 1, 'Se': 1, 'O': 4}
>>> molecular(structural2)
{'H': 10, 'C': 4}
```

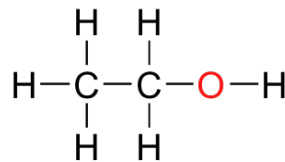
```
>>> molecular(structural3)
{'C': 4, 'H': 10}
```

```
>>> isomers(structural1, structural2)
False
>>> isomers(structural1, structural3)
False
>>> isomers(structural2, structural3)
True
```

Isomeren zijn stoffen die in die zin gelijk zijn dat ze hetzelfde aantal en dezelfde soort atomen bevatten, maar die van elkaar verschillen door de wijze waarop die atomen onderling verbonden of geschikt zijn. Anders gezegd, isomeren zijn stoffen met dezelfde molecuulformule maar die een andere structuurformule kunnen hebben (een stof is een isomeer met zichzelf). Ethanol (een alcohol) en dimethylether bestaan bijvoorbeeld elk uit 6 waterstofatomen, twee koolstofatomen, en een zuurstofatoom. Enkel de volgorde van de bindingen is verschillend. Ze hebben bijgevolg dezelfde molecuulformule ( $\text{C}_2\text{H}_6\text{O}$ ) maar een verschillende structuurformule.



dimethyl ether



ethanol

Er bestaan verschillende standaarden die de structuurformule van een stof als een string voorstellen. Je mag ervan uitgaan dat de atomen in die stringvoorstelling symbolisch benoemd worden door een hoofdletter, eventueel gevolgd door een aantal kleine letters. Het verschil tussen hoofd- en kleine letters is hierbij belangrijk: de string PB staat bijvoorbeeld voor één atoom fosfor (P) en één atoom boor (B), terwijl de string Pb staat voor één atoom lood. Voorts kunnen tussen de symbolische namen van verschillende atomen ook nog andere karakters voorkomen, bijvoorbeeld Na-OH of (CHHH)CO(CHHH). Tussen de hoofdletter en de kleine letters van de symbolische naam van eenzelfde atoom komen echter nooit andere karakters voor.

## Opgave

1. Schrijf een functie molecuulformule waaraan de stringvoorstelling van een structuurformule van een bepaalde stof als argument moet doorgegeven worden. Deze functie moet als resultaat een dictionary teruggeven, waarvan de sleutels bestaan uit de symbolische namen van alle atomen uit de gegeven structuurformule. De waarde die correspondeert met een sleutel geeft telkens aan hoeveel keer het atoom in de formule voorkomt. Voor de gegeven string (CHHH)CO(CHHH) moet de functie dan bijvoorbeeld de dictionary {'H': 6, 'C': 3, 'O': 1} teruggeven. Merk op dat deze dictionary dus eigenlijk een voorstelling is van de molecuulformule die correspondeert met de gegeven structuurformule, vandaar de naam van de functie.
2. Gebruik de functie molecuulformule om een functie isomeren te schrijven. Aan deze functie moeten twee stringvoorstellingen van structuurformules als argument doorgegeven worden. De functie moet een Booleaanse waarde teruggeven die aangeeft of de stoffen die overeenkomen met de gegeven structuurformules isomeren zijn of niet.

## Voorbeeld

```
>>> structuurformule1 = 'OCaOSeOO'
>>> structuurformule2 = 'HHCHHCHHCHHCHH'
>>> structuurformule3 = 'HHCHHHCCHHHCHH'

>>> molecuulformule(structuurformule1)
{'Ca': 1, 'Se': 1, 'O': 4}
>>> molecuulformule(structuurformule2)
{'H': 10, 'C': 4}
>>> molecuulformule(structuurformule3)
{'C': 4, 'H': 10}

>>> isomeren(structuurformule1, structuurformule2)
False
>>> isomeren(structuurformule1, structuurformule3)
False
>>> isomeren(structuurformule2, structuurformule3)
True
```