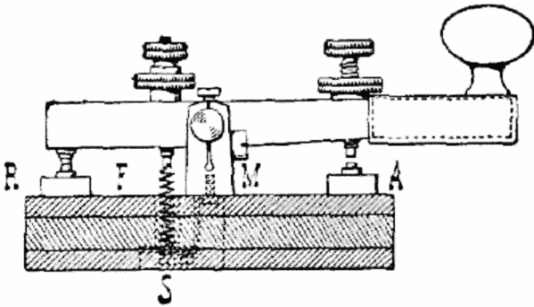


Isomorse

Morse code is a method for encrypted communication of textual information that gets transmitted as a series of on-off tones, lights or clicks at certain intervals. Fixed combinations of these signals represent the different letters, punctuation marks and digits. This Morse code was invented and developed in 1835 by Samuel Morse for use in telegraph systems. Using an electrical telegraph (see figure) it was possible to choose between two states: key down (= electric current) or key up (= no electric current) and duration (short or long). Telegraphy is generally considered to be the forerunner of later digital communication. It should be noted that during speed communication competitions between skilled Morse code operators and experts in sending text messages, [Morse code generally wins](#).



In the International Morse Code, short signals are denoted as dots (.) and long signals as dashes (-). Each character is then represented by a unique sequence of dots and dashes. As such, the letter S is represented in Morse code as ... and the letter O as ---. For emergency signals, Morse code can be sent by way of improvised sources that can be easily "keyed" on and off, making it one of the simplest and most versatile methods of telecommunication. The most common distress signal is SOS or three dots, three dashes and three dots, internationally recognized by treaty.

Assignment

The text file [morse.txt](#) contains a list of characters (first column) and their corresponding representation in Morse code (second columns). Both columns are separated by a tab. When you inspect this file, you will see for example that the letter C is represented in Morse code as -.-. and a comma is represented as ...-.

When transmitting messages in Morse code, a short break (letter space) is usually inserted in between two successive characters. Such a short break is denoted by a space character. In between words, a longer break (word space) is inserted, which is denoted by a forward slash (/). As such, the word EARN is represented as

...-.-.

and the word URN as

...-.-.

However, if we ignore the breaks (spaces), these two words are represented by the same pattern in Morse code. Words represented by the same Morse code pattern are called **isomorse**. Also note that the two words given as an example are **homophones**: they are pronounced the same but differ in meaning and spelling. Your task is to:

- Write a function `morsecodes` that takes the location of a text file. Each line of this file must contain a

character, followed by a tab and the representation of the character in Morse code. The function must return a dictionary that maps each character in the given file onto its representation in Morse code. In case the character is a letter, the mapping must use its uppercase variant as a key.

- Write a function `pattern` that takes two arguments: a word and a dictionary that maps characters onto their corresponding representation in Morse code. You may assume that each character in the given word occurs as a key in the given dictionary, with mapping of letters making use of their uppercase variant. The function has two additional parameters `complement` and `mirror`, that both take a Boolean value (default value: `False`). The function must compose the pattern that represents the given word in Morse code (without intermediate spaces in between representations of successive characters). Next, the function must determine the complement of the composed pattern by replacing each dot by a dash and *vice versa*, but only if the value `True` is passed to the parameter `complement`. Next, the function must mirror the pattern by reversing the order of its characters, but only if the value `True` is passed to the parameter `mirror`. The pattern obtained in this way must be returned as the result of the function.
- Write a function `isomorse` that takes three arguments: two words and a dictionary that maps characters onto their corresponding representation in Morse code. You may assume that all characters in the given words are used as keys in the given dictionary, with mapping of letters making use of their uppercase variant. In addition, the function has two optional parameters `complement` and `mirror` that take Boolean values (default value: `False`). The function must return a Boolean value that indicates whether the two given words are isomorse. This is the case if both words are represented by the same pattern in Morse code. In order to determine the pattern in Morse code for a given word, the function must make use of the given dictionary and the function `pattern`. In comparing the two given words, the pattern of the second word may have to be complemented and/or mirrored in case the value `True` is passed respectively to the parameters `complement` and `mirror`.
- Write a function `groups` that takes the location of two text files. The first file must contain a list of words, each on a separate line. Each line in the second file must contain a character, followed by a tab and the representation of that character in Morse code. You may assume that all characters in the given words also occur in the second file. The function must return a dictionary that groups the given words that are represented by the same pattern in Morse code, by mapping each pattern in Morse code that corresponds to a given word onto the set of all words that are represented by that pattern. To make it clear: this function never needs to compute complemented or mirrored patterns.

Example

In the following interactive session, we assume that the text files [morse.txt](#) and [words.txt](#) are located in the current directory.

```
>>> codes = morsecodes('morse.txt')
>>> codes['C']
'..-'
>>> codes['U']
'..-'
>>> codes['N']
'..'

>>> pattern('TAIPAN', codes)
'..-...-...'
>>> pattern('TAIPAN', codes, complement=True)
'..-...-...'
>>> pattern('TAIPAN', codes, mirror=True)
'..-...-'
```

```

>>> pattern('TAIPAN', codes, complement=True, mirror=True)
'-.--.-.-.-.'

>>> isomorse('TAIPAN', 'CUNETTE', codes)
True
>>> isomorse('TAIPAN', 'PYTHON', codes)
False
>>> isomorse('TAIPAN', 'ROUX', codes, complement=True)
True
>>> isomorse('TAIPAN', 'PATIENT', codes, mirror=True)
True
>>> isomorse('TAIPAN', 'TUDOR', codes, complement=True, mirror=True)
True
>>> isomorse('EDOMITE', 'EDOMITE', codes, mirror=True)
True

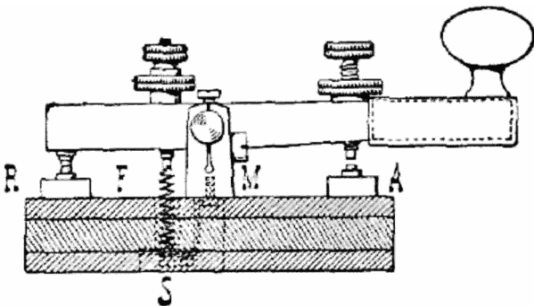
>>> groups('words.txt', 'morse.txt')
{'.--.-.-.-.-.': {'WILL', 'PITIED', 'PERINE', 'ADENINE'}, '-.-.-.-.-.-.-.-.': {'PYTHON'}, '-.-.-.-.-.-.': {'TREPAN', 'CUNETTE', 'TAIPAN'}}

```

Epilogue

The American inventor and businessman Thomas Alva Edison (1847-1931) took it a bit far: he proposed his wife in Morse code and nicknamed his children Marion Estelle en Thomas Alva Jr. as Dot and Dash.

Morse is een gecodeerde communicatievorm bestaande uit signalen die met tussenpozen worden uitgezonden. Vastgelegde combinaties van deze signalen stellen dan de verschillende letters, leestekens en cijfers voor. Deze **morsecode** werd in 1835 uitgevonden en ontwikkeld door Samuel Morse, met als doel ze te gebruiken voor het verzenden van telegrafische berichten. Met een telegraaf (zie figuur) kon men namelijk enkel kiezen tussen twee toestanden: sleutel naar beneden (= stroom) of sleutel naar boven (= geen stroom) en tijdsduur (kort of lang). Telegrafie wordt algemeen beschouwd als een voorloper van de latere digitale communicatie. Merk op dat in snelheidswedstrijden tussen ervaren morsecodeoperatoren en experts in het verzenden van SMS berichten, [morsecode steevast als overwinnaar](#) uit de bus komt.



In de internationale morsecode worden korte signalen genoteerd met een punt (.) en lange signalen met een koppelteken (-). Elk karakter wordt dan voorgesteld als een unieke opeenvolging van punten en koppeltekens. Zo wordt de letter s in morsecode voorgesteld als ... en de letter o als ---. Het is dan ook algemeen bekend dat ... --- ... de voorstelling in morsecode is van het SOS signaal.

Opgave

Het tekstbestand [morse.txt](#) bevat een lijst van karakters (eerste kolom) en hun corresponderende voorstelling in morsecode (tweede kolom). De kolommen worden gescheiden door een tab. In dit bestand zie je bijvoorbeeld dat de letter c in morsecode wordt voorgesteld als -.-. en dat een komma wordt voorgesteld als ---,.

Bij het versturen van berichten in morsecode wordt tussen elke twee opeenvolgende karakters normaalgezien een korte pauze ingelast (letterspatie) die wordt genoteerd met een spatie, en wordt tussen verschillende woorden telkens een langere pauze ingelast (woordspatie) die wordt genoteerd met een slash (/). Zo wordt het woord EARN bijvoorbeeld voorgesteld als

..-.-.-

en het woord URN als

..-.-.-

Als we de tussenpauzes (spaties) echter negeren, dan worden deze twee woorden in morsecode voorgesteld door hetzelfde patroon. We zeggen dan dat de twee woorden **isomorse** zijn. Merk op dat deze twee woorden ook **homofoon** zijn: ze zijn opgebouwd uit dezelfde klanken, maar hebben een verschillende betekenis. Gevraagd wordt:

- Schrijf een functie `morsecodes` waaraan de locatie van een tekstbestand moet doorgegeven worden. Elke regel van dit bestand moet een karakter bevatten, gevolgd door een tab en de voorstelling van het karakter in morsecode. De functie moet een dictionary teruggeven, die elk karakter uit het gegeven bestand afbeeldt op zijn corresponderende voorstelling in morsecode. Indien het karakter een letter is, dan moet bij de afbeelding gebruik gemaakt worden van zijn hoofdlettervariant.
- Schrijf een functie `patroon` waaraan twee argumenten moeten doorgegeven worden: een woord en een dictionary die karakters afbeeldt op hun corresponderende voorstelling in morsecode. Je mag ervan uitgaan dat alle karakters in het gegeven woord als sleutel gebruikt worden in de gegeven dictionary, waarbij voor de afbeelding van letters gebruik gemaakt wordt van hun hoofdlettervariant. De functie heeft ook nog twee optionele parameters `complement` en `spiegel`, waaraan een Booleaanse waarde kan doorgegeven worden (standaardwaarde: `False`). De functie moet het patroon samenstellen waarmee het woord wordt voorgesteld in morsecode (dus zonder spaties tussen de voorstelling van opeenvolgende karakters). Daarna moet de functie het `complement` bepalen van het gevormde patroon door elk punt te vervangen door een koppelteken `e` en *vice versa*, maar enkel indien de waarde `True` wordt doorgegeven aan de parameter `complement`. Daarna moet de functie het patroon spiegelen door de volgorde van de karakters om te keren, maar enkel indien de waarde `True` wordt doorgegeven aan de parameter `spiegel`. Het patroon dat op die manier verkregen wordt, moet door de functie als resultaat teruggegeven worden.
- Schrijf een functie `isomorse` waaraan drie argumenten moeten doorgegeven worden: twee woorden en een dictionary die karakters afbeeldt op hun corresponderende voorstelling in morsecode. Je mag ervan uitgaan dat alle karakters in de gegeven woorden als sleutel gebruikt worden in de gegeven dictionary, waarbij voor de afbeelding van letters gebruik gemaakt wordt van hun hoofdlettervariant. De functie heeft ook nog twee optionele parameters `complement` en `spiegel`, waaraan een Booleaanse waarde kan doorgegeven worden (standaardwaarde: `False`). De functie moet een Booleaanse waarde teruggeven die aangeeft of de twee gegeven woorden isomorse zijn. Dit is het geval als de twee woorden in morsecode voorgesteld worden door hetzelfde patroon. Om het patroon in morsecode van een gegeven woord te bepalen, moet uiteraard gebruik gemaakt worden van de gegeven dictionary en de functie `patroon`. Voor de vergelijking van de twee woorden moet het patroon van het tweede woord eventueel nog omgekeerd en/of gespiegeld worden indien respectievelijk de waarde `True` wordt doorgegeven aan de parameters `complement` en `spiegel`.
- Schrijf een functie `groepen` waaraan de locatie van twee tekstbestanden moet doorgegeven worden. Het eerste tekstbestand moeten een lijst van woorden bevatten, elk op een afzonderlijke regel. Elke regel van het tweede bestand moet een karakter bevatten, gevolgd

door een tab en de voorstelling van het karakter in morsecode. Je mag ervan uitgaan dat alle karakters die voorkomen in de gegeven woorden, ook voorkomen in het tweede bestand. De functie moet een dictionary teruggeven die de gegeven woorden met eenzelfde patroon in morsecode groepeert, door elk patroon in morsecode dat correspondeert met een gegeven woord af te beelden op de verzameling van alle gegeven woorden met datzelfde patroon. In dit geval moeten dus nooit complementaire of gespiegelde patronen bepaald worden.

Voorbeeld

Bij onderstaande voorbeeldsessie gaan we ervan uit dat de tekstbestanden [morse.txt](#) en [woorden.txt](#) zich in de huidige directory bevinden.

```
>>> codes = morsecodes('morse.txt')
>>> codes['C']
'-. .'
>>> codes['U']
'.. .'
>>> codes['N']
'-. '

>>> patroon('TAIPAN', codes)
'-.--.-.-.'
>>> patroon('TAIPAN', codes, complement=True)
'-.--.-.-.'
>>> patroon('TAIPAN', codes, spiegel=True)
'-.--.-.-.'
>>> patroon('TAIPAN', codes, complement=True, spiegel=True)
'-.--.-.-.'

>>> isomorse('TAIPAN', 'CUNETTE', codes)
True
>>> isomorse('TAIPAN', 'PYTHON', codes)
False
>>> isomorse('TAIPAN', 'ROUX', codes, complement=True)
True
>>> isomorse('TAIPAN', 'PATIENT', codes, spiegel=True)
True
>>> isomorse('TAIPAN', 'TUDOR', codes, complement=True, spiegel=True)
True
>>> isomorse('EDOMITE', 'EDOMITE', codes, spiegel=True)
True

>>> groepen('woorden.txt', 'morse.txt')
{'.--.-.-.-.': {'WILL', 'PITIED', 'PERINE', 'ADENINE'}, '-.-.-.-.-.': {'PYTHON'}, '-.-.-.-.-.': {'TREPAN', 'CUNETTE', 'TAIPAN'}}
```

Epiloog

De gepatenteerde uitvinder Thomas Alva Edison (1847-1931) had een kleine beroepsmisvorming: hij vroeg zijn vrouw ten huwelijk in morsecode en gaf twee van zijn kinderen Marion Estelle en Thomas Alva Jr. de troetelnaampjes Dot en Dash.