

Hapax Legomenon

When linguists study an old and dead language, they usually try guessing the meaning of a word based on the context in which the word occurs. By comparing the different contexts a word is used in, they can usually easily deduce the meaning of a word. Sometimes however, they come across words that are only used once in all texts known in that language. Such words usually can't be translated, because the context in which these words occur are barely known.

A word that occurs only once in the texts of a certain language, is called a *hapax legomenon*. This term is a transliteration of the Greek ἅπαξ λεγόμενον. The literal translation is '*what is said once*'. The term is not only used for words that occur once in a certain language. For example, it is also used for words that only occur once in an entire oeuvre of a certain author, or in a certain text.

Assignment

Write a function `hapax` that takes a string as an obligatory argument and prints a list as a result. This list must contain the *hapax legomena* from the given text, sorted in alphabetical order. To determine whether a word is a *hapax legomenon*, we consider the variants of words that consist solely of lowercase letters. We consider a word in a text as the longest possible succession of letters from the alphabet.

Example

```
>>> kapper = "Kapper Knap, de knappe kapper,  
... knipt en kapt heel knap, maar de knecht van  
... kapper Knap, de knappe kapper, knipt en kapt  
... nog knapper dan kapper Knap, de knappe kapper,  
... knipt en kapt"  
>>> hapax(kapper)  
['dan', 'heel', 'knapper', 'knecht', 'maar', 'nog', 'van']  
>>> betty = "Betty Botter bought a bit of butter.  
... The butter Betty Botter bought was a bit bitter  
... And made her batter bitter.  
... But a bit of better butter makes better batter.  
... So Betty Botter bought a bit of better butter  
... Making Betty Botter's bitter batter better"  
>>> hapax(betty)  
['and', 'but', 'her', 'made', 'makes', 'making', 's', 'so', 'the', 'was']
```

Wanneer taalkundigen een oude en dode taal bestuderen, proberen ze meestal aan de hand van de context waarin een woord verschijnt te raden wat het woord betekent. Door de verschillende contexten waar een bepaald woord gebruikt wordt te vergelijken, kunnen ze doorgaans zeer goed de betekenis van het woord afleiden. Soms komen ze echter woorden tegen die slechts eenmaal gebruikt wordt in alle gekende teksten van die taal. Zo'n woorden zijn meestal niet te vertalen, omdat de context waarbinnen de woorden gebruikt wordt amper gekend is.

Een woord dat slechts eenmaal voorkomt in de teksten van een bepaalde taal, wordt een *hapax legomenon* genoemd. Deze term is een transliteratie van het Griekse ἅπαξ λεγόμενον. De letterlijke vertaling is '*wat eenmaal gezegd wordt*'. De term wordt echter niet alleen gebruikt voor woorden die slechts eenmaal voorkomen in een bepaalde taal. Ze wordt bijvoorbeeld ook

gebruikt voor woorden die slechts eenmaal voorkomen in het volledige oeuvre van een bepaalde auteur, of in een bepaalde tekst.

Opgave

Schrijf een functie `hapax` die een string als verplicht argument neemt en een lijst als resultaat teruggeeft. Deze lijst moet de *hapax legomena* uit de gegeven tekst bevatten, gesorteerd in alfabetische volgorde. Om na te gaan of een woord een *hapax legomenon* is, beschouwen we de varianten van woorden die enkel bestaan uit kleine letters. De woorden van de tekst worden gevormd door de langst mogelijke opeenvolgingen van letters uit het alfabet.

Voorbeeld

```
>>> kapper = "Kapper Knap, de knappe kapper,  
... knipt en kapt heel knap, maar de knecht van  
... kapper Knap, de knappe kapper, knipt en kapt  
... nog knapper dan kapper Knap, de knappe kapper,  
... knipt en kapt"  
>>> hapax(kapper)  
['dan', 'heel', 'knapper', 'knecht', 'maar', 'nog', 'van']  
>>> betty = "Betty Botter bought a bit of butter.  
... The butter Betty Botter bought was a bit bitter  
... And made her batter bitter.  
... But a bit of better butter makes better batter.  
... So Betty Botter bought a bit of better butter  
... Making Betty Botter's bitter batter better"  
>>> hapax(betty)  
['and', 'but', 'her', 'made', 'makes', 'making', 's', 'so', 'the', 'was']
```