# 第二章

**6.在比较词表的讨论中，创建一个对象叫做translate，通过它你可以使用德语和意大利语词汇查找对应的英语词汇。这种方法可能会出现什么问题，你能提出一个办法来避免这个问题吗？**

**from** nltk.corpus **import** swadesh  
*# print swadesh.fileids()*  
it2en = swadesh.entries([**'it'**, **'en'**])  
**print** it2en[:5]

[(u'io', u'I'), (u'tu, Lei', u'you (singular), thou'), (u'lui, egli', u'he'), (u'noi', u'we'), (u'voi', u'you (plural)')]

de2en = swadesh.entries([**'de'**, **'en'**])  
**print** de2en[:5]

[(u'ich', u'I'), (u'du, Sie', u'you (singular), thou'), (u'er', u'he'), (u'wir', u'we'), (u'ihr, Sie', u'you (plural)')]

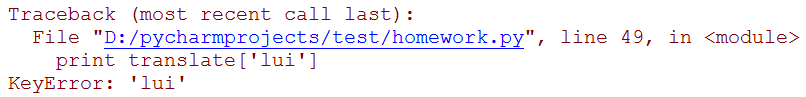
translate = dict(it2en)  
translate.update(dict(de2en))  
**print** translate[**'Hund'**]

I

**print** translate[**'cane'**]

I

**print** translate[**'lui'**]



原语言链表中有多对多关系的词，当输入translate['lui']时并不会正确显示he，而是会报错。

**解决方法：**

遍历语言链表，当检测到有多对多关系时，将该元素进行处理后再加入原语言链表。

**from** nltk.corpus **import** swadesh  
swadesh.fileids()  
it2en = swadesh.entries([**'it'**, **'en'**])  
de2en = swadesh.entries([**'de'**, **'en'**])  
  
**for** it, en **in** it2en:  
 **if ',' in** it:  
 words = it.split(**', '**)  
 **for** eachWord **in** words:  
 newWord = (eachWord, en)  
 it2en.append(newWord)  
  
**for** de, en **in** it2en:  
 **if ',' in** de:  
 words = de.split(**', '**)  
 **for** eachWord **in** words:  
 newWord = (eachWord, en)  
 de2en.append(newWord)  
  
translate = dict(it2en)  
translate.update(dict(de2en))  
**print** translate[**'lui'**]

he