# 词形还原

顾志浩 MF1933024

## 任务描述：

1. 输入一个单词
2. 如果词典中有该词，输出该词及其属性，转4，否则转3
3. 如果有该词的还原规则，并且词典中有还原后的词，则输出还原后的词及其属性，转4，否则，调用<未登录词模块>
4. 如果输入中还有单词，转1，否则，结束

## 算法实现：

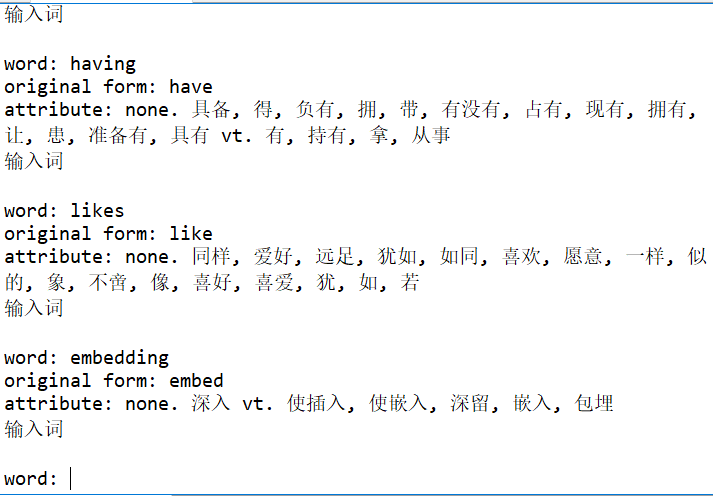
1. 首先加载字典dic\_ec.txt，字典中的内容为单词及其词形和意思。
2. 设计还原规则。对于一些单词，单词的变换就是其本身，有些词的变换是添加固定的结尾，或则替换结尾。所以在词的还原规则中，根据名词和动词设计了以下变换规则：
   1. \*ves -> \*f
   2. \*ves -> \*fe
   3. \*ies -> \*y
   4. \*es -> \*
   5. \*s -> \*
   6. \*ing -> \*
   7. \*??ing -> \*?
   8. \*ing -> \*e
   9. \*ying -> \*
   10. \*ed -> \*
   11. \*??ed -> \*?
   12. \*ed -> \*e
3. 对于词典中未收录的词，调用未登录模块

## 词典

1. Dic\_ec.txt，包含词、词形、词义的词典

## 测试

1. 截图



## 存在的问题

在本次实验中，我实现了一个词形还原工具，根据英语中的词变换规则，然而，英语中还有很多不规则的变换，无法根据一般的变换规则进行变换，所以暂时无法解决该问题。同时，给定的词典的编码很乱，经过变换，改成GBK编码，以“\t“进行分割，以便加载词典。