

## 任务三 语音合成（语音助手）

### 【任务描述】

文字与音频的互相转换是自然语言处理中很关键的技术点。“把文字转换成声音，让你的应用开口说话”就是利用语音合成技术将文本转换成音频。现在我国各大AI企业都有对应的免费开放API为开发者提供应用服务。例如，百度2108年6月份就发布了百度语音识别无限量使用。本任务完成从键盘输入中文，调用百度API将输入的中文转换成mp3音频文件同时播报这段音频。

### 【任务实施】

本任务使用任务2中已经申请的百度账号，请确保已经领取语音合成的免费额度。

**步骤1** 与任务2中一样，import利用百度账号创建一个client对象。

#### 新建AipSpeech

AipSpeech是语音合成的Python SDK客户端，为使用语音合成的开发人员提供了一系列的交互方法。

参考如下代码新建一个AipSpeech：

```
from aip import AipSpeech
```

**步骤2** 定义变量text，存储从键盘输入的内容。

```
text = input('请输入：')
```

**步骤3** 参考官网技术文档中“接口说明”，调用synthesis方法将text转换成音频并以mp3文件保存到本地。“per”参数可以用来设置合成语音的不同声音模式，“spd”用来设置语速，“vol”用来设置语调，这些参数虽然是可选设置，但是可以根据不同内容设置多样化的语音表达，赶紧试试吧。

- 合成文本长度必须小于1024字节，如果本文长度较长，可以采用多次请求的方式。文本长度不可超过限制

举例，要把一段文字合成为语音文件：

```
result = client.synthesis('你好百度', 'zh', 1, {  
    'vol': 5,
```

| 参数   | 类型     | 描述  | 是否必须 |
|------|--------|---|------|
| tex  | String | 合成的文本，使用UTF-8编码，<br>请注意文本长度必须小于1024字节         | 是    |
| cuid | String | 用户唯一标识，用来区分用户，<br>填写机器 MAC 地址或 IMEI 码，长度为60以内 | 否    |
| spd  | String | 语速，取值0-9，默认为5中语速                              | 否    |

**步骤4** 利用playsound播报MP3音频文件。

- 安装playsound

```
pip install playsound
```

- playsound的使用方法

```
from playsound import playsound  
playsound('audio.mp3')
```

本任务中result存储的是文字合成的二进制代码，可以用在result = client.synthesis () 之后增加一句print(result)语句，运行出现类似下图中的二进制码则表明文本合成语音成功了。

请输入: good morning

```
b'\xff\xf3(\xc4\x00\x0c@FT\x15[\x00\x00\x80\xc9\t\x03\xee\xef
\xc6\x90\xc0\xc1\x84\x8c\x89\x10\xe0*\x8e0\xf0\xc7\xc2@@\xe5\xc7Z\xf4\xcd\x88C\x87%
\xdf\xff\xfd\x9f\x90\xff\xff\xec\xbf\xfd\x7f\xdb\xfe\xff\xee\x90\xe3\x96*\x04\xea
\x0e\x92\xd3\xca\xff\xff\xf3(\xc4\n\x0e\xab\n\xbc\x01\x8b8\x00\x92>\xf8\xe12\x1f\xec
\xf6\xfe\xd9\xca:\x83\xff\xb6\xd9Q\x10\x17\x8dF\xc4\x7f\xb7\xe2I\xdb\xb1\xbf\xff\xfb
\x1cs\xe6\x9cw\xff\xff\xfd\x7f\xff\xff\xff\xff\xff\x5J\x0f\x99R\xfc\xabk/\xbe\xec\x99$
\x9b9\xff\xf3(\xc4\n\x0f(\xfe\xac\x01\xd9@\x00\x8f-\x1f\xa2\xa6\x8e\xca
\xa3\xa9\x86\x154\x04\xbb\x9fI\x9f;\x96\x00\xe0\x01\x01a\xb7\xff\xbb\xbf\xff\xc5RM
\xce\xe8Z\x16\xc5\x8e\x97\xb0\xa1g\xa7\xff\xff\xff\xff\xff\xff\xfc\xe4\x85\x15\x96T
\xb6\xdbm\xb6\x81\xff\xf3(\xc4\x08\x0e\xf0\xcf\x02^Sjr\xed\x8a\xe2\x04+\x01\x1d
\x14\xba\xe5\xc0\xb9\x9fn\xa5\x0e\xe1\xce\x05\x10\x94oR\x88\x01\xcf\xf3\t\t
```

本次任务完整参考代码如下:

```
from aip import AipSpeech
from playsound import playsound

# 填写自己申请的 ID 与 Key
APP_ID = '*****'
API_KEY = '*****'
SECRET_KEY = '*****'

client = AipSpeech(APP_ID,API_KEY,SECRET_KEY)
text = input('请输入: ')
result = client.synthesis(text,'zh',1,{ 'vol':5, 'per':0})

print(result)

# 识别正确放回语音二进制 错误放回dict
if not isinstance(result,dict):
    with open('audio.mp3','wb') as f:
        f.write(result)

playsound('audio.mp3')
```

## 【任务拓展】

随着AI技术的发展,语音合成技术已经被越来越多的应用,很多网站都在文章内容页面嵌入了语音朗读功能,AI合成主播也开始在新闻媒体中应用。我们也可以改进任务3,让机器给我们朗读一篇文章。下面通过语音合成技术将本地的Word文档变成“有声”读物,希望聪明的你举一反三,例如天气播报助手等。

**步骤1** 安装python-docx库。python-docx是一个用于创建和更新微软Word (.docx) 文件的Python库,具备设置段落、分页符、表格、图片、标题、样式等几乎所有的word文档中能常用的功能,但是主要用来创建文档。

```
In [18]: pip install python-docx
Collecting python-docxNote: you may need to restart the kernel to use updated packages.
  Downloading python-docx-0.8.10.tar.gz (5.5 MB)
Requirement already satisfied: lxml>=2.3.2 in c:\programdata\anaconda3\lib\site-packages (from python-docx) (4.5.2)
Building wheels for collected packages: python-docx
  Building wheel for python-docx (setup.py): started
  Building wheel for python-docx (setup.py): finished with status 'done'
  Created wheel for python-docx: filename=python_docx-0.8.10-py3-none-any.whl size=184495 sha256=76bc6a7f3dbb3671198f508263408d1ced6f80c5ab77abf45ccf2be104c36943
  Stored in directory: c:\users\h101\appdata\local\pip\cache\wheels\97\4c\2e\68066cbf12b9b2e66403da8982aaf4f656d9f5cb5dc3179e82
Successfully built python-docx
Installing collected packages: python-docx
Successfully installed python-docx-0.8.10
```

**步骤2** import Document , 利用Document打开本地的docx文档。

```
from docx import Document
path = 'new.docx'
document = Document(path)
```

**步骤3** 读取docx文档中的段落文字并保存到text中。

```
text = ''
for paragraph in document.paragraphs:
    print(paragraph.text)
    text += paragraph.text
```

接下来就是将text转换成语音, 操作与任务3一样, 完整的参考代码如下:

```
from aip import AipSpeech
from playsound import playsound

from docx import Document
path = 'new.docx'
document = Document(path)

# 填写自己申请的 ID 与 Key
APP_ID = '*****'
API_KEY = '*****'
SECRET_KEY = '*****'
client = AipSpeech(APP_ID,API_KEY,SECRET_KEY)

text = ''
for paragraph in document.paragraphs:
    print(paragraph.text)
    text += paragraph.text

result = client.synthesis(text, 'zh', 1, {'vol': 5, 'per': 0})
print(result)

# 识别正确放回语音二进制 错误放回dict
if not isinstance(result, dict):
    with open('audio.mp3', 'wb') as f:
        f.write(result)
playsound('audio.mp3')
```

## 农家书屋“听书”

“读者数字农家书屋”是基于百度语音服务的惠农扶贫项目，文化与科技相融合，最大限度的满足了基层用户的阅读需求。针对农村留守老人、儿童、残疾人士等不便阅读的人群，特推出“听书”功能，集成百度语音后端资源进行云语音识别（<http://ai.baidu.com/tech/speech/asr>），通过情感语音合成（<http://ai.baidu.com/tech/speech/tts>），有效保证优质的听书体验。



本任务实战代码如下,位于/xml/rw3.ipynb 同学们来运行一下吧