



## Mathematica 歌姬计划 (1) —— 与 Cortana 的联合



yxlllc

物理研究生，业余程序员

已关注

115 人赞同了该文章

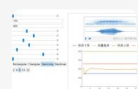
欢迎来到 Mathematica 歌姬计划，这一计划的目标是用 MMA 像 Vocaloid 那样调教自己想要的人声电子音乐。

### 前情提要

2016 年夏天，Wolfram 公司似乎意识到“玩家们”对音频信号处理方面的需求日益增长，在新推出的 11.0 版 Mathematica 中引入了 Audio 对象的概念，并加入了一系列以“Audio”为开头命名的音频处理函数。当时，我看到了 MMA 在音频方面有巨大的“可玩性”，于是开始研究如何利用这些功能制作电子音乐。一开始，我主要用表示 MIDI 音符的 SoundNote 函数结合变调函数 AudioPitchShift 和变速函数 AudioTimeStretch 做了一些类似鬼畜的东西。后来我觉得这些都没啥意思，要玩就玩个大的、有技术含量的（歌声合成引擎），在 2017 年 7 月我写下了这篇回答：

我想写一个歌声合成引擎，请教需要学习哪些音频、声学方面和编程方...

[www.zhihu.com](https://www.zhihu.com)



后来这一项目有了很多细节上的改进，工程也变得比较庞大。在 @LePtC 的建议下，我准备重新写一系列文章，从零开始系统性地介绍这一项目。

本文为本系列的第一篇文章，介绍如何在 Mathematica 中调用 Cortana 音源（中文和英文）进行语音合成。1.1 节介绍 MMA 语音合成系统的基本工作原理，1.2 节指导你检查系统是否能够开启 Cortana 音源，1.3 节提供了一键激活 Cortana 的工具以及“Speak 函数不接受中文输入”的解决方案

### 1.1 麦酱的声音从

赞同 115



15 条评论

分享

★ 收藏





如果想要做歌声合成，音源是很重要的。在 MMA 中，用 `Speak` 函数可以实现 TTS（即 text to string，文本转语音）。不过，Wolfram 公司并没有自己的语音合成引擎，MMA 的语言合成功能是通过接口实现的。

`Speak[expr]`  
播放 `expr` 的语音表示。

### 更多信息和选项

- `Speak["string"]` 播放 "string" 中的文本。
- `Speak[expr]` 对数学表达式、程序、图形和其它结构起作用。
- `Speak[HoldForm[expr]]` 播放 `expr` 的持有形式，并不计算表达式。
- `Speak` 尽可能产生自然的单词语速，而不是表达式结构的逐字表示。
- `Speak` 用您计算机的操作系统产生声音。

在 11.2 版本中，又新出了个实验性质的语音函数 `SpeechSynthesize`。

EXPERIMENTAL

SpeechSynthesize

11.2的新功能

`SpeechSynthesize[expr]`  
合成 `expr` 的内容。

`SpeechSynthesize[expr, file]`  
把 `expr` 的内容合成到位于 `file` 的文件中。

不过它也是用的接口，我们可以找到相关的 Link 文件：

名称	修改日期	类型	大小
Kernel	2017/11/27 17:05	文件夹	
LibraryResources	2017/11/27 17:06	文件夹	
PacletInfo	2017/10/3 16:24	Wolfram Mathemat...	1 KB
SpeechSynthesisTools	2017/10/3 16:24	Wolfram Mathemat...	4 KB

打开这个文件看一下，我们将发现如下线索：

```
dlls["Linux" | "Linux-x86-64"] = {
  "libflite.so",
  "libflite_cmu_grapheme_lang.so", "libflite_cmu_grapheme_lex.so",
  "libflite_cmulex.so",
  "libflite_cmu_indic_lang.so", "libflite_cmu_indic_lex.so",
  "libflite_cmu_time_awb.so", "libflite_cmu_us_awb.so",
  "libflite_cmu_us_kal.so", "libflite_cmu_us_kal16.so",
  "libflite_cmu_us_rms.so", "libflite_cmu_us_slt.so",
  "libflite_usenglish.so"
};

dlls["MacOSX-x86-64" | "MacOSX-x86"] = {};
dlls["Windows" | "Windows-x86-64"] = {"Flite"};
```

经过检索，我们可以 100% 确认 MMA 里的 `SpeechSynthesize` 函数用的是美国卡内基梅隆大学（简称 CMU）研发的 Flite 语音合成引擎，以下是链接和截图：

Speech Synthesizer

赞同 115

15 条评论

分享

收藏

...

CMU Flite: a small, fast run time synthesis engine

CMU Flite (festival-lite) is a small, fast run-time open source text to speech synthesis engine developed at CMU and primarily designed for small embedded machines and/or large servers. Flite is designed as an alternative text to speech synthesis engine to [Festival](#) for voices built using the [FestVox](#) suite of voice building tools.



所以无论如何，麦酱只能借别人的嗓子说话，好尴尬（囧.jpg）。

不过这并不影响我们计划的进行，虽然自己不会说话，找人合作也是不错的嘛。先来看看我们可以和谁合作吧。

如果用 `SpeechSynthesize` 函数即 CMU-Flite 引擎，用 `Voice` 选项可以设置音源，目前有5种音源，其中只有 "SLT" 是女声，其余都是男声，均只支持英文输入。

指定语音合成声音：

```
In[1]:= SpeechSynthesize["Hello Wolfram", Voice -> "KevinLenzo"] // Normal
```

[语音合成] [语音] [转换为普通表达式]



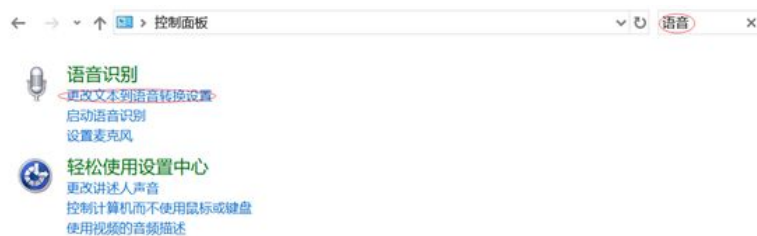
使用 `$SpeechSynthesisVoices` 获取可用声音列表：

```
In[2]:= $SpeechSynthesisVoices
```

[声音合成语音列表]

```
Out[2]:= {<|Generic -> {RMS, AlanWBlack, KevinLenzo, KevinLenzo16, SLT}|>}
```

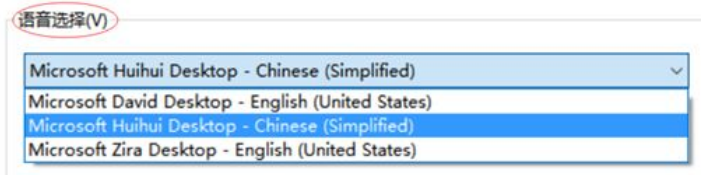
如果用 `Speak` 函数即操作系统的引擎，我们无法直接在 MMA 里设置音源选项。不过我们可以修改操作系统的默认配置。举 Windows 10 系统为例，我们首先打开控制面板，在搜索栏输入“语音”，然后找到“更改文本到语音转换设置选项”，如图所示：



点击进入后，选择“语音到文本转换”选项卡，如图所示：



然后展开“语音选择”这一栏，看看操作系统默认安装了哪些音源吧：



一般情况下，中文 Windows 10 系统默认的 TTS 音源是 Huihui，既支持中文也支持英文输入（虽然只标了 Chinese）。在调用 MMA 的 Speak 函数时，你听到的就是她的声音。其它标注English的音源只支持英文输入。

1.2 可不可以找 Cortana 姐姐？

“能不能给力一点啊”，麦酱说道，“这些人啊，听都没听说过，声音也不好听！” ( \_ \_ )

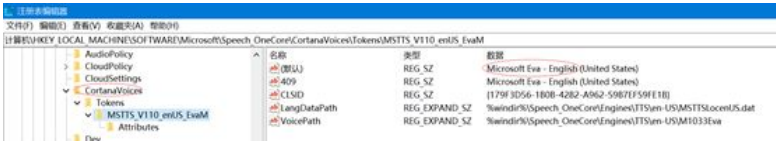
也是，如此傲娇的麦酱怎能甘于和这些既没名气也没实力的人合作呢，诶等一下，Windows 10 不是有小娜吗。



作为微软小冰的姐姐，《光晕》里士官长的助手（女朋友），Cortana 的名气还是响当当的。在声音性能方面，既然能像 Siri 一样流行于市面，自然是不用质疑的。以她为音源做歌声合成貌似是很不错的选择。

不过你也看到了，Windows 10 默认可选的 TTS 没有 Cortana。这难道意味着无法调用吗？当然不是！在查阅了一些资料之后，我发现有很多网友都曾经对此问题进行过深入的研究。最后大家一致得出的结论是，微软是有为 Cortana 预留外部接口的，只不过将它们隐藏了起来！

打开注册表编辑器，我们将很容易找到微软隐藏的线索：



顺藤摸瓜，根据注册表项里的 VoicePath，我们可以顺利找到英文 Cortana 和中文 Cortana（小娜）的声库文件。

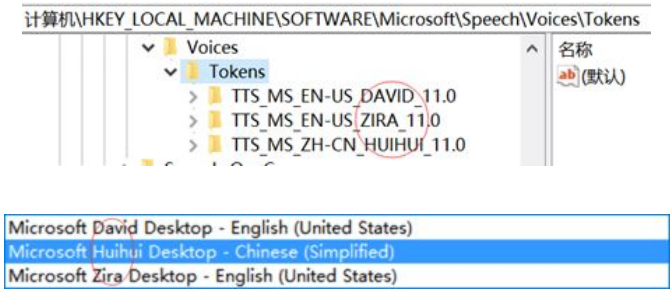
首先是英文 Cortana，微软用 Eva 作为编号：



其次是中文 Cortana（小娜），微软用 Hongyu 作为编号：



然后进一步研究发现，在控制面板里可选的语音引擎和以下注册表项密切相关：



分明就是一一对应关系！所以如果你的电脑中有上述这些声库文件（Hongyu 声库对应中文小娜，Eva 声库对应英文 Cortana），只需分别创建相应的注册表项就能开启他们了。

### 1.3 激活 Cortana 并让麦酱说中文

考虑到一步一步的操作比较繁琐，我写了一个 PowerShell 脚本直接批量生成注册表，以下是链接：

yxlllc/ActivateCortana

github.com

下载 ActivateCortana.ps1 这个文件后，以管理员身份打开 PowerShell，更改执行策略为 Unrestricted 后运行这个脚本。

对 PowerShell 不是很熟悉的同学，可以看下面的文章快速入门，只看“运行 Powershell 脚本”这一块就行了：

PowerShell 快速入门和运行脚本

赞同 115

15 条评论

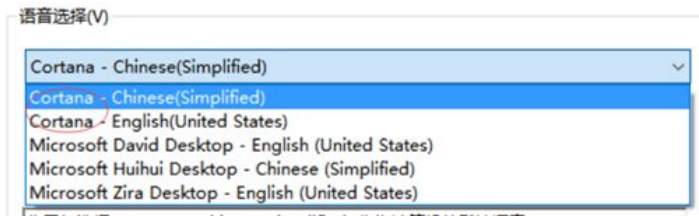
分享

收藏

...



运行成功后，重启计算机，这时再次展开控制面板>>语音识别>>更改文本到语音转换>>语音到文本转换>>语音选择，中文和英文的 Cortana 就可以选了：



我们选择 Cortana - Chinese(Simplified) 然后保存。打开 Mathematica，运行Speak["hello"] 测试一下，确认音源变成了 Cortana。

当然进行到这里还没有完。MMA 有个问题，就是 Speak 函数只能支持英文输入，就算用的是中文的语音引擎，比如你运行 Speak["你好"] 这样的东西必然会失败。为了让麦酱说中文，一个解决方案是利用 NETLink 直接调用 .NET Framework 合成 Cortana 的声音至音频流，MMA 代码如下：

```
Needs["NETLink`"];
LoadNETAssembly["System.Speech"];
Options[MySpeak] = {"Rate" -> 0, "Volume" -> 100};
MySpeak[string_, OptionsPattern[]] :=
Module[{synth},
  synth = NETNew["System.Speech.Synthesis.SpeechSynthesizer"];
  synth@Rate = OptionValue["Rate"];
  synth@Volume = OptionValue["Volume"];
  synth@Speak[string]]
```

相关功能可以见微软官方的说明文档，链接如下：

SpeechSynthesizer 类

[msdn.microsoft.com](https://msdn.microsoft.com)

然后用这个新定义 MySpeak 函数就可以说中文了，比如运行 MySpeak["你好"]。并且相比原来的 Speak 多出的功能是，我们可以设置发音速度和音量，比如：

```
MySpeak["你好", "Rate"->0,"Volume"->100]
```

其中 "Rate" 选项控制速度，取值范围是 -10 到 10 的整数，缺省值 0，而 "Volume" 选项控制音量，取值范围是 0 到 100 的整数，缺省值 100。

如果以上测试都没有问题，这意味着麦酱可以和 Cortana 尽情地合作和玩耍了，Enjoy!

发布于 2018-02-01

「真诚赞赏，手留余香」



1 人已赞赏




- Wolfram Mathematica
- 语音合成
- Cortana





文章被以下专栏收录

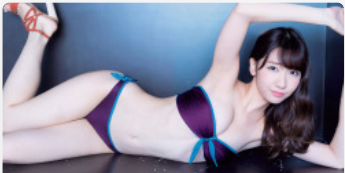


**Mathematica 还能这样玩**

Mathematica（mma、麦酱），宇宙第一计算姬（钦定的），投喂 CPU 时间就能把答...

已关注

推荐阅读



禅与奶罩识别艺术

王司图 发表于集智



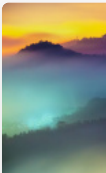
Mathematica 歌姬计划（2）—— 听力训练

yxlll... 发表于Mathe...

《舌尖上的胖胖》

如果都请你吃过，讲真那就嫁了吧！

馒头音乐



1分34秒

江淮乘用

15 条评论

切换为时间排序

- 写下你的评论...
- 

**王瑞扬**

我觉得距离歌姬计划还差很远。现在只是成功调用了 API，但是 API 有没有 tone 调节之类的东西很难说吧

3

1 年前
- 

yxlllc (作者) 回复 王瑞扬

所以是（1）啊

4

1 年前
- 

**挖掘机小王子**

加油！

赞

1 年前
- 

**Benjamin**

Mathematica 这么好玩吗？

赞

1 年前
- 

**wjxway**

6666

赞

1 年前
- 

**晨光飘散**

那个，这个要怎么停止语音，我导入了18万字的小说，现在停不下来了，退出了内核都没有用.....

5

1 年前
- 

赵雪冰 回复 晨光飘散

哈哈哈哈哈

赞

1 年前

赞同 115

15 条评论

分享

收藏

...



yxlllc (作者) 回复 晨光飘散

1 年前

额，这个确实是MMA的一个问题，貌似是因为它新建了一个系统进程，所以你退出内核也没有，只能任务管理器解决了



赞



杨树 回复 晨光飘散

1 年前

哈哈哈哈哈哈哈哈哈哈



赞



晨光飘散

1 年前

任务管理器解决了.....



赞



无虞

1 年前

那个.....我的mma用Speak["你好"] 的确失败了，但是Speak[你好]却成功了←\_←



2



yxlllc (作者) 回复 无虞

1 年前

居然还有这种操作？！



1



无虞 回复 yxlllc (作者)

1 年前

是啊很神奇的.....(:3」 ∠)\_估计是mma对中文的支持有某些有趣的特性或者干脆这就是gub了（笑）



赞



无虞 回复 yxlllc (作者)

1 年前

“Speak[expr] 对数学表达式、程序、图形和其它结构起作用”←大概就是因为这个“其他结构”吧



赞



蛇纹身的程序员

1 年前

这个引擎没有音高和时值控制，只能靠后期做pitch shift和TSM，效果估计不会怎么好。。



赞

▲ 赞同 115



● 15 条评论

🔗 分享

★ 收藏

