

项目研究报告

项目名称：物联网应用中的策略自动生成器

项 目 编 号 XYB2018275

项 目 组 长 B16070504 张美楠

项 目 组 成 员 B16070507 郑潍雯

B16070501 王骄阳

指 导 教 师 闵丽娟

指 导 单 位 物联网学院

**南 京 邮 电 大 学**

**2019 年 05 月 02 日**

**摘要**

项目围绕实现物联网应用领域中的只能家居方面的策略的自动生成展开，旨在将研究小组的项目开发成果，策略自动生成器的软件开发过程进行描述。 开发过程主要包括：基于Python语言实现的各大中文语义开发平台的中文自然语言分析，其中BosonNLP开发平台占据主要部分；基于SQL数据库实现的智能家居领域语言的关键词，近义词等的统计工作；科大讯飞推出的以语音为核心的人机交互解决方案-AIUI开放平台，实现智能手机上的语音识别功能；以及基于当前app主流开发工具Android Studio开发的应用程序-策略自动生成器。 项目主要解决的问题，捕获关于智能家居领域家电控制方向的自然语言。理解语义并作出相应的回馈指令，从而实现语音控制智能家居的目的。

关键词：物联网，策略自动生成器，BonsonNLP，SQL数据库，AIUI

**一、引言**

**1 项目背景**

随着现阶段人工智能技术的不断发展，让机器读懂人类语言，并在日常生活中为人类服务，从而解放双手成为现阶段计算机应用技术的突破口。自然语言处理（NLP）技术的重要组成部分，语义识别成为当前的一大热门技术，因此通过语音识别的策略自动生成技术必将有更大的市场。目前市面上的策略大都是已被编辑的，缺少了一定的灵活性。并且，对于以后更加高度智能化、人性化的生活中，由用户直接对话机器，由机器提取关键字自动生成策略，再发送指令、执行指令的方式定是更加智能化、人性化的。

我们对目前存在的涉及到策略自动生成的一些产品进行一定程度的调研分析，结合实际生活经验，发现当前的策略生成的相关技术方面存在着死板，缺少灵活性，繁复，缺少了对相似关键字的扩展存储以及模式识别技术的应用等不足。在这样的情况下，对整个项目软件进行需求架构分析后，开发出一款具有语音识别，文字转化，中文分词，关键字提取，策略生成等功能的物联网应用生成器将对日常生活以及技术研究有重大的意义。

**2 项目形式**

本项目是一项安卓软件开发项目。

**二、项目需求说明**

**1.1 面向的用户**

本项目软件面向所有居家用户，在家中通过物联网应用中的策略自动生成器来实现通过语音控制联网家具。

**1.2 实现目标**

实现策略的自动生成，并把生成的策略信息存储到数据库中，供物联网应用中的策略引擎来调用。研究内容包括：语音输入和预处理、关键词的抽取和匹配、语料库收集数据、BosonNLP平台数据分析、AIUI语音接入等。

主要实现目标有以下几点：

1、能够通过语音识别，识别出用户的需求，将语音转化为文字。

2、进行中文分词，在处理过的文字中进行关键字的提取。

3、由识别出的关键字进行相应的处理，最终达到策略自动生成的目的。

**1.3 项目开发要求**

1. 项目开发规范统一：模块划分、代码编写均要规范文档。
2. 程序优化、安全并要有良好的可扩展性。
3. 用户界面简洁美观，便于操作，简单实用。
4. 建立友好的用户界面，与用户进行良好的交互。
5. 软件性能良好，运行流畅。

**1.4 开发工具**

1、Microsoft Visual Studio 6.0

2、Android Studio 2.0

3、Python3.0

**三、项目核心功能的实现**

**1 NLP以及简单的数据库**

自然语言处理（NLP）是人工智能和语言学领域的一个分支学科，同时也是机器学习中最热的一个方向之一。试想想，机器可完全明白人类的语言，并且可以娴熟地使用自然语言与人类沟通，这不就是人工智能的目标？幸好，现在各种文本预处理、文本挖掘技术已日渐成熟，可以帮助我们从文本中提取出更多有意义的信息。 此项目主要使用了一款NLP开放平台-玻森，此团队专注于中文语义分析技术，其通过自主研发的中文分词、句法分析、语义联想和实体识别技术，为开发者提供了简单、强大、可靠的中文语义分析云端API。

通过此平台，项目组的成员主要进行了以下一些测试，关键词提取“现在外面的天气真热啊”，利用Python执行代码之后得到的关键词有“天气，热”。语义联想“打开空调”，执行后得到的相近语义有“翻开，拆开等”。针对各种类型的家电产品，首先通过各种中文分析工具，提取出每一条控制语句中的关键字，并加之适当的语义分析，将统计结果录入数据库，建立一套关于针对智能家居控制领域的语法，便于后续程序的开发。数据库可以直观的理解为存放数据的仓库，那么我们将家电的控制指令中的关键字提取出来后，将其存储于数据库。想要达到完全智能化的控制，需要通过大量的机器训练或是提供大量的数据。

**2 语音识别平台**

众所周知，科大讯飞、云知声、思必驰是几家专注于语音识别技术的公司，依靠垂直语音技术领域，长期盘踞主流语音市场，拥有先发优势。那么此次软件开发过程中，小组成员经过调研后，毫不犹豫的选择接入了科大讯飞的语音识别平台技术。 AIUI 是科大讯飞2015年推出的一套以语音为核心的人机交互解决方案，意在使应用和设备能够快速具备能听会说，能理解会思考的能力。此策略自动生成器的最大的功能之一，语音控制智能家居的能力通过接入AIUI得到了简单方便的实现。

**3 程序开发**

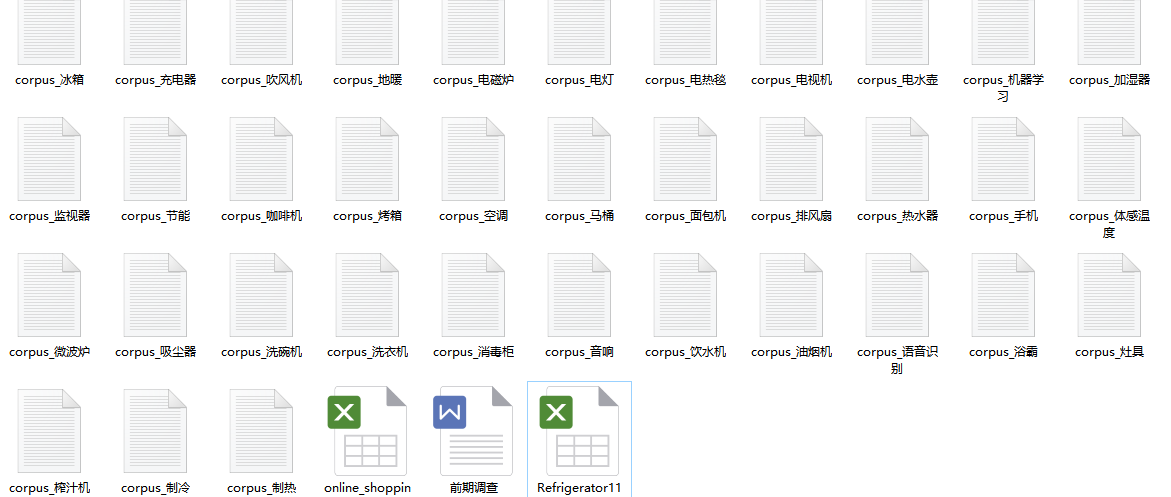
搭建程序的界面，实现其内部的逻辑功能，建立前后台的连接，引入各个平台的API，相信是众多app开发者最熟悉的操作内容。首先在开发工具的选择上，小组选择了当前主流的集成开发工具AS，其拥有完善的插件管理，支持多种代码管理工具，智能且有GOOGLE官方支持等多项优点。 在项目的实际开发过程中，界面的设计上基本没有遇到什么问题，实际接入AIUI平台的SDK时，由于项目开发的不熟练，以及官方使用文档的理解能力不够，开发初期饶了很多的弯路，但是好在经过大家的共同努力，基本达成了预期的目标成果。用户输入一段及其控制语音指令，程序反馈出对应的机器操作指令。实际软件中已经基本实现了冰箱，空调，电灯，电视机的开关等控制，欢迎感兴趣的小伙伴体验这款（物联网应用中的）策略自动生成器软件并提供您宝贵的建议。

**四、项目数据及软件测试**

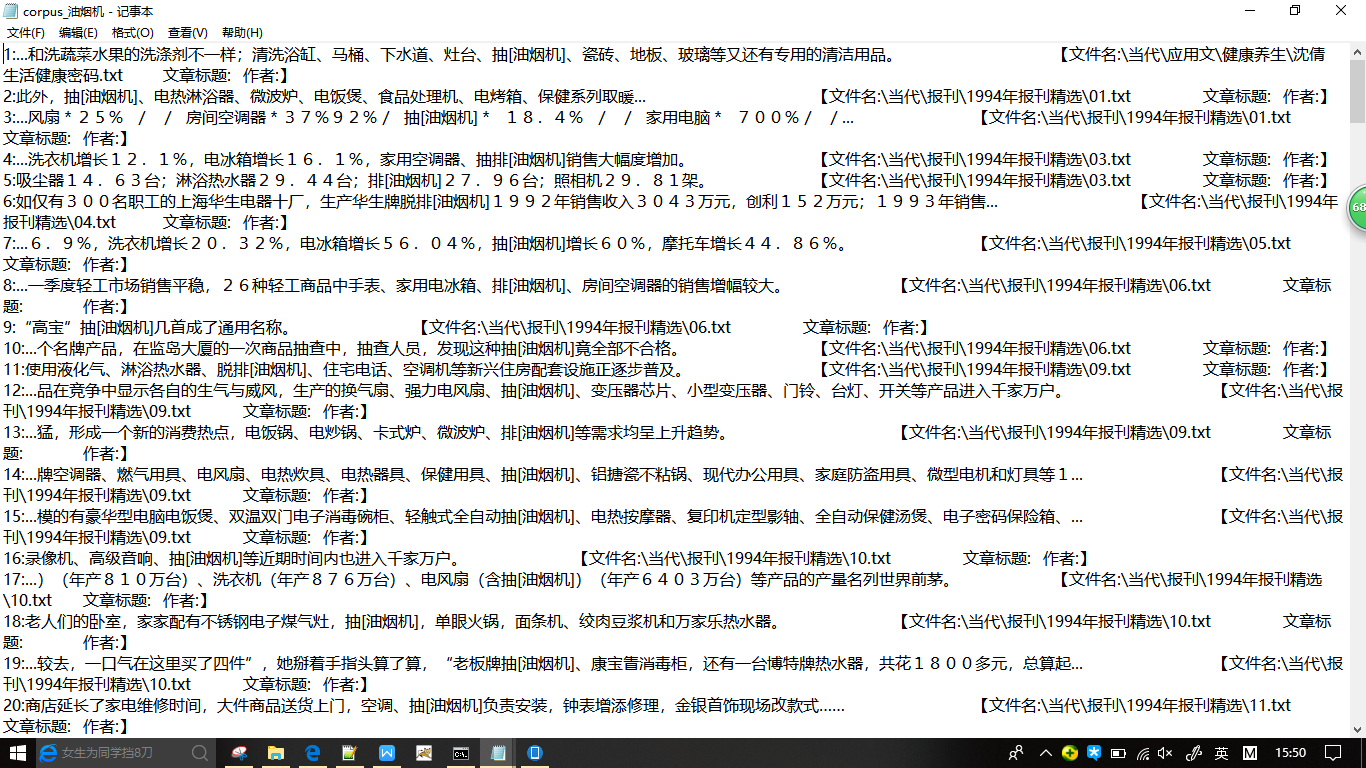
**1 语料库数据收集**

本项目所收集的语料内容丰富，最丰富的词条数目接近万条。

（1）各类电器收集的数据语料如下：

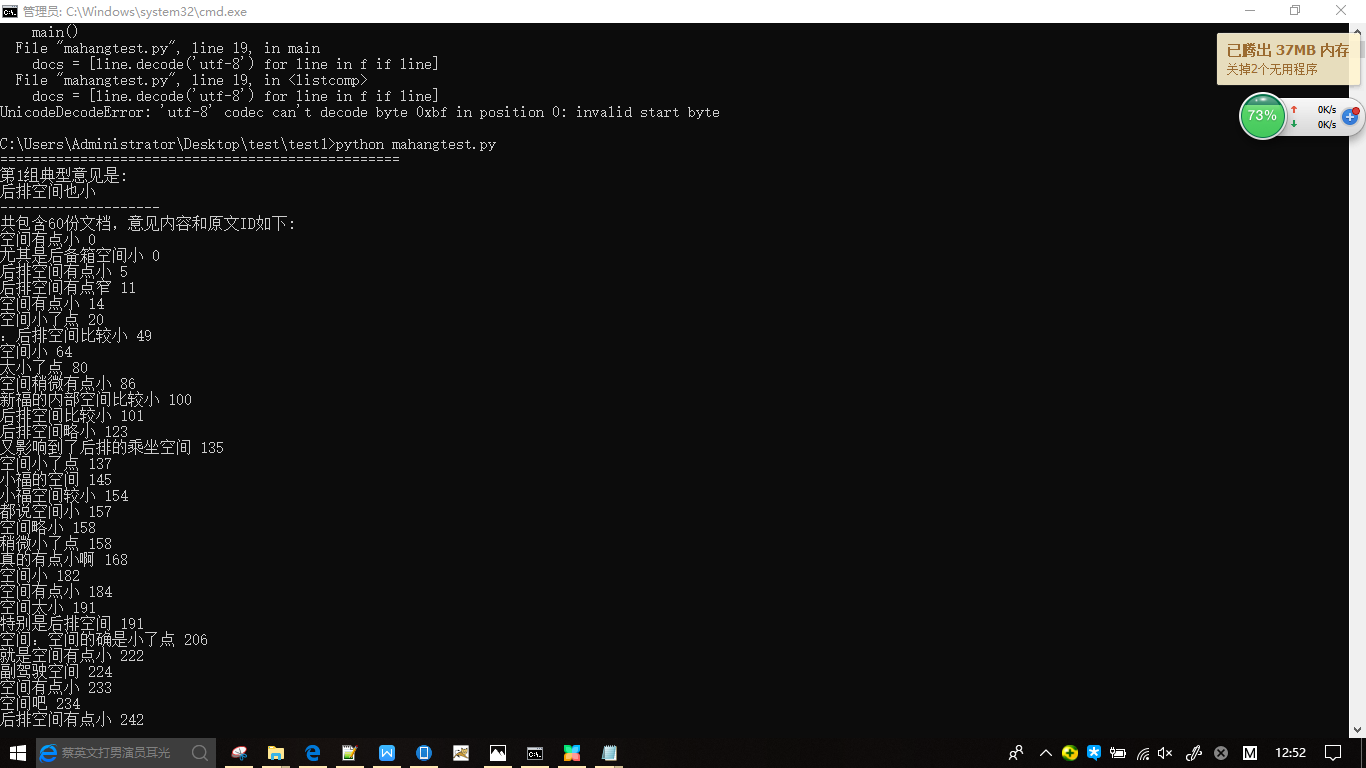


（2）油烟机数据语料内部示例：

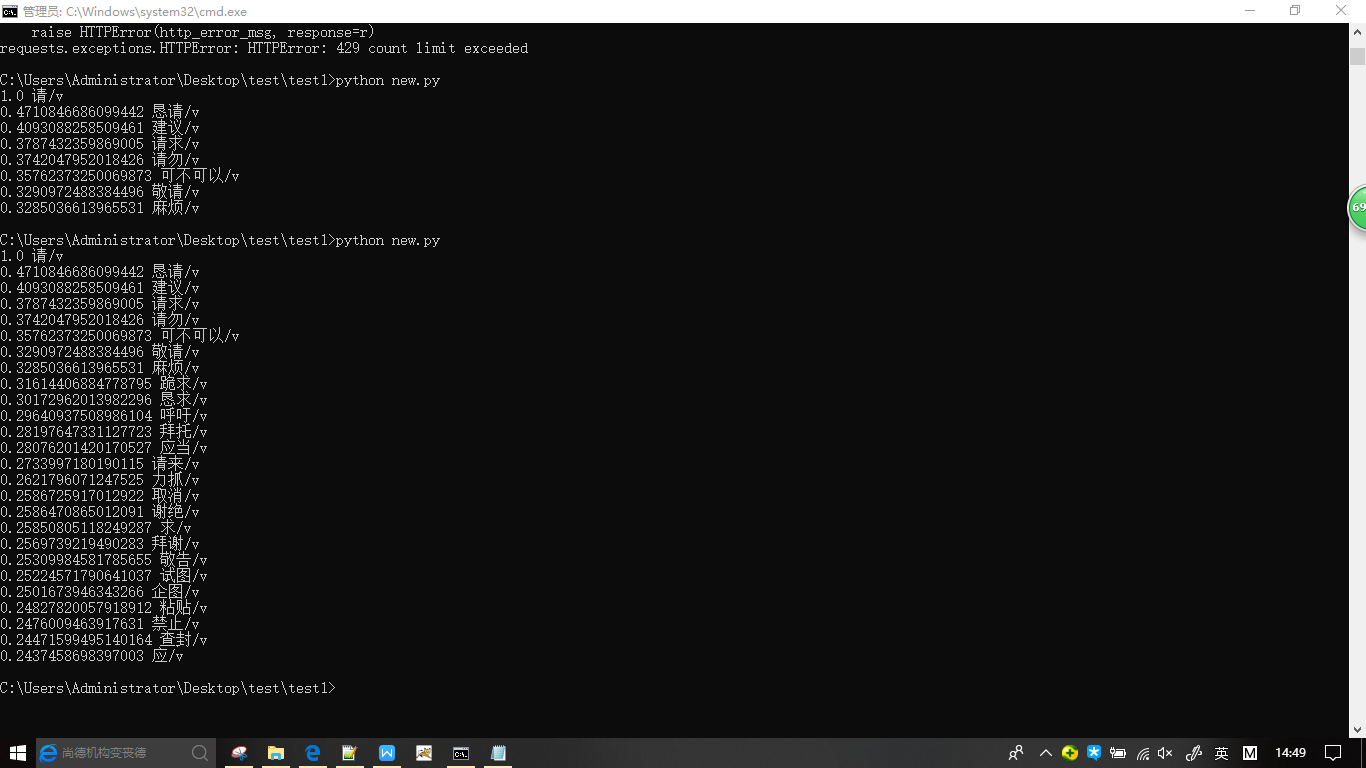


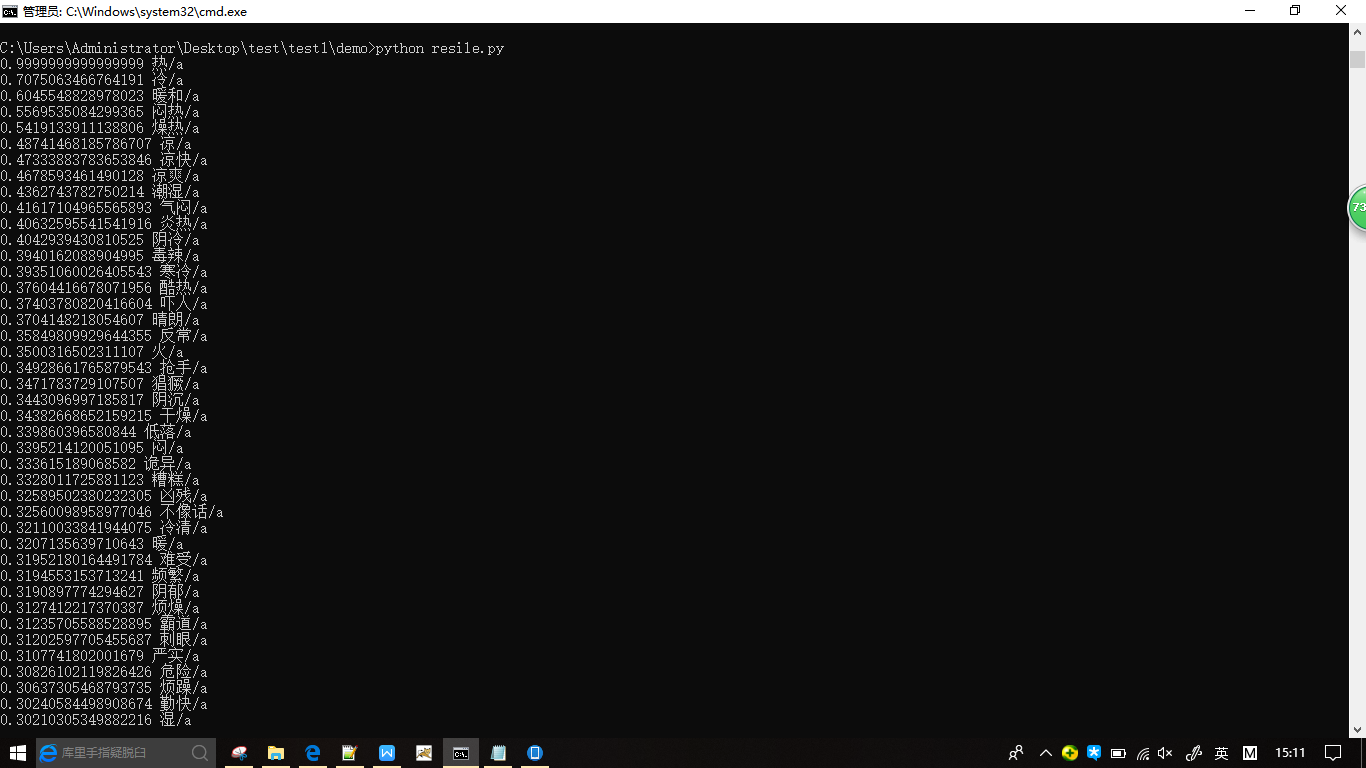
**2 数据分析测试**

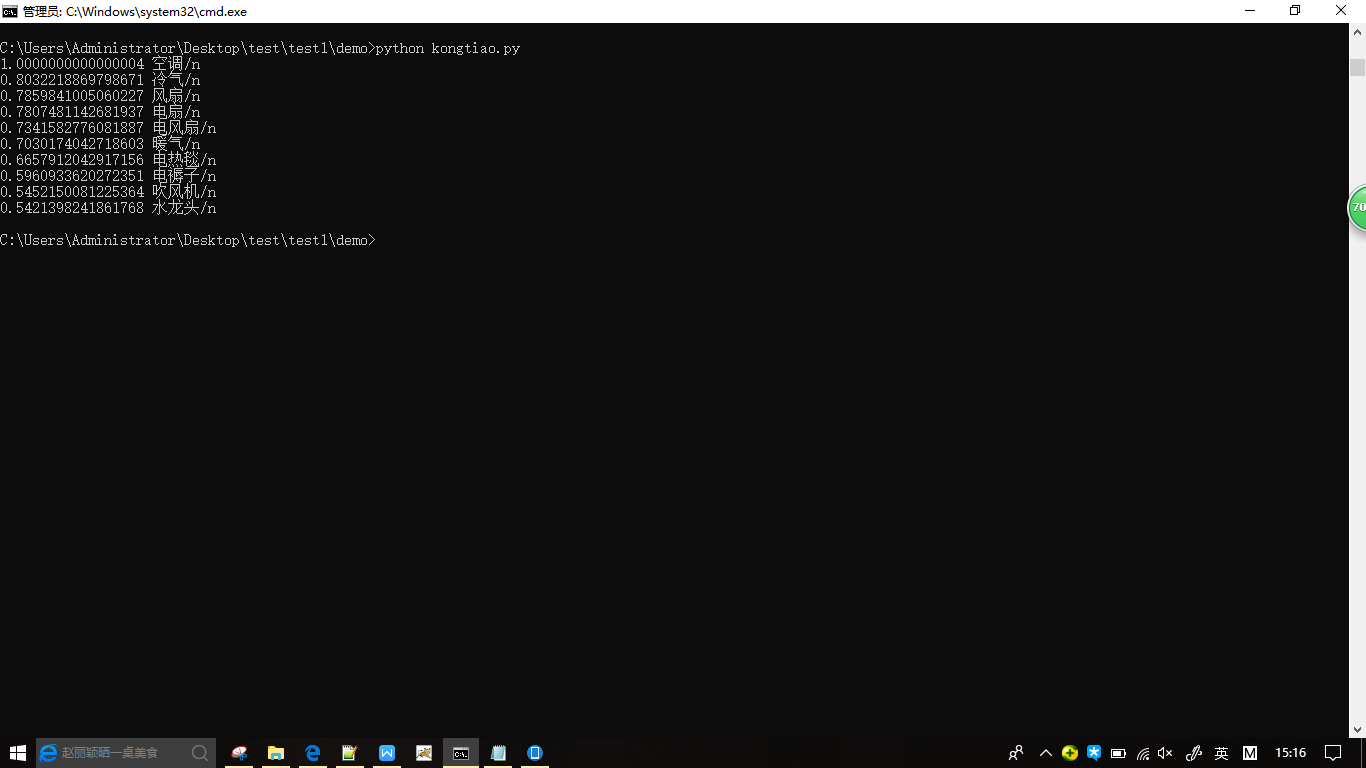
（1）BosonNLP的测试数据：在MH370的微博中发现不同话题



（2）各类词语的语义联想如：“请”，“热死了”，“空调”等

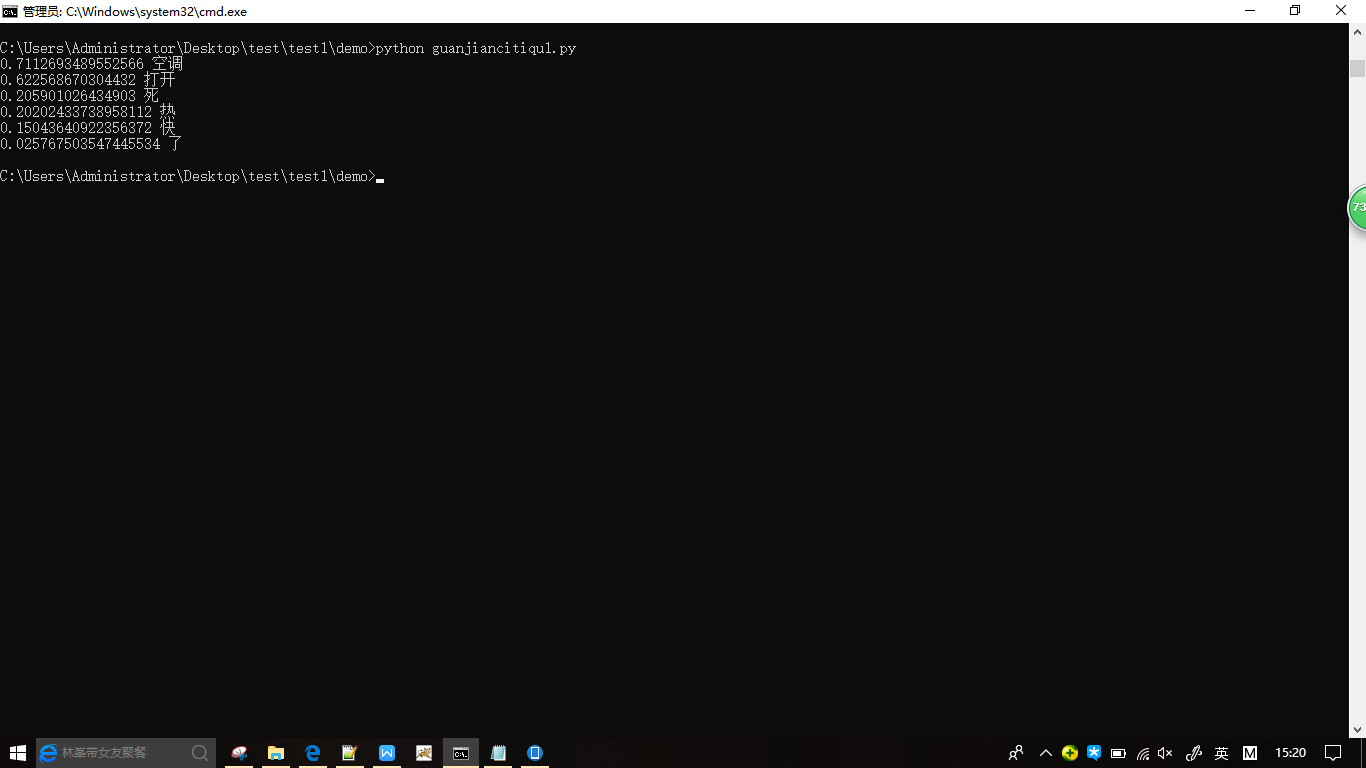




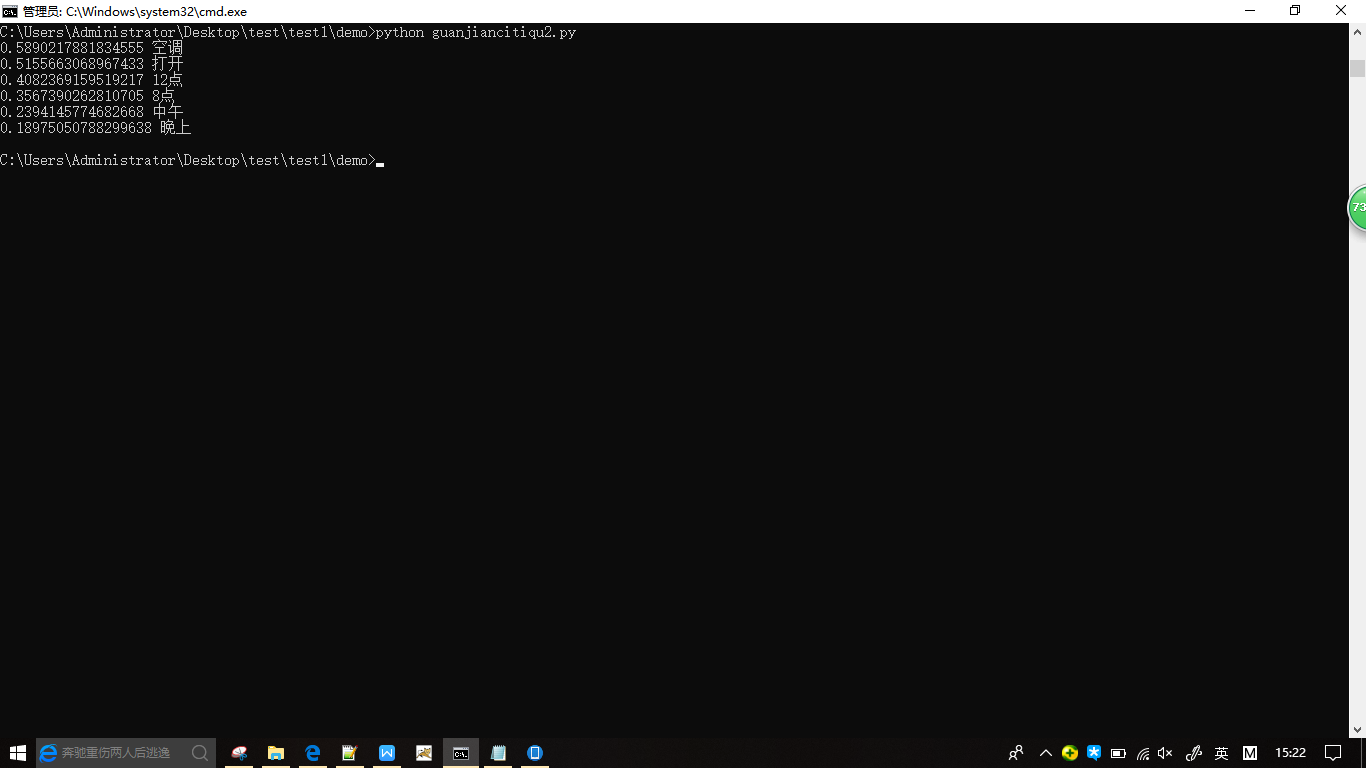


（3）关键字提取

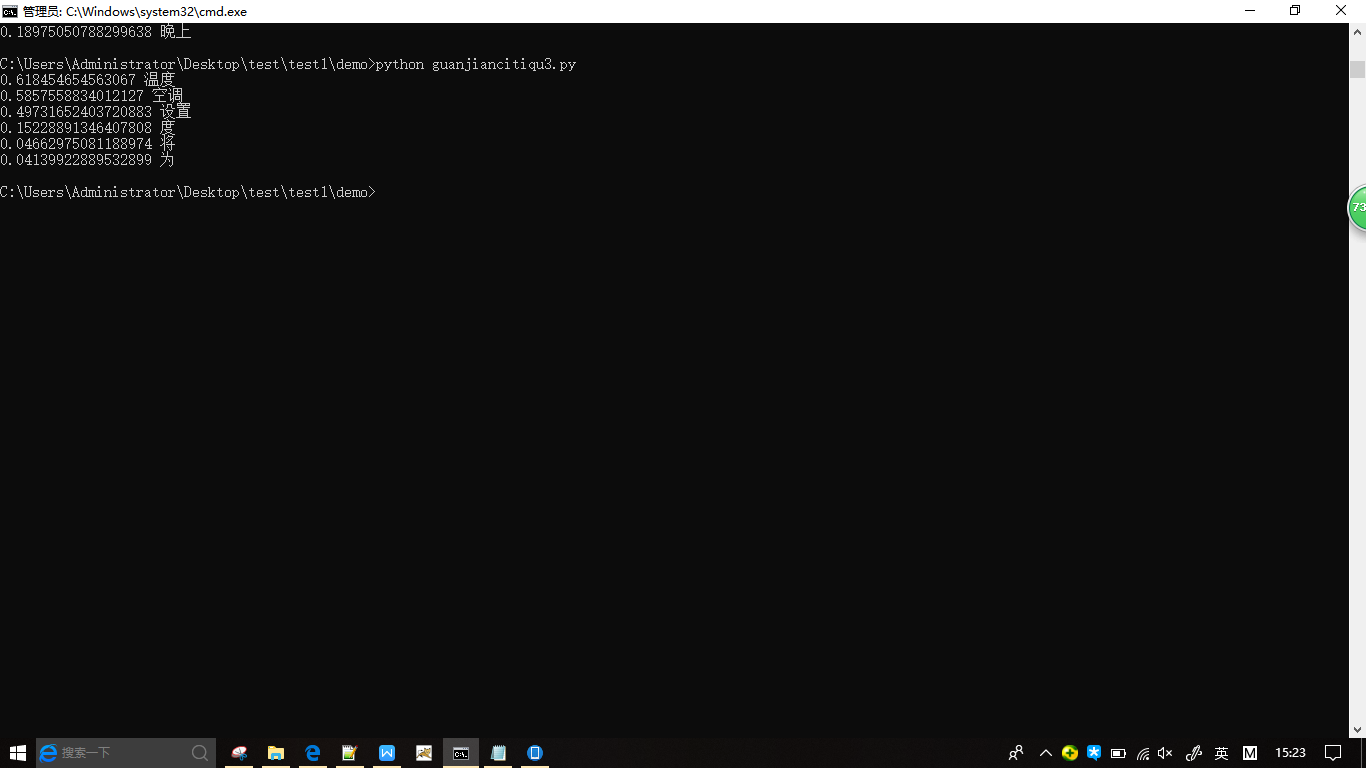
“热死了，快打开空调”的关键词提取：



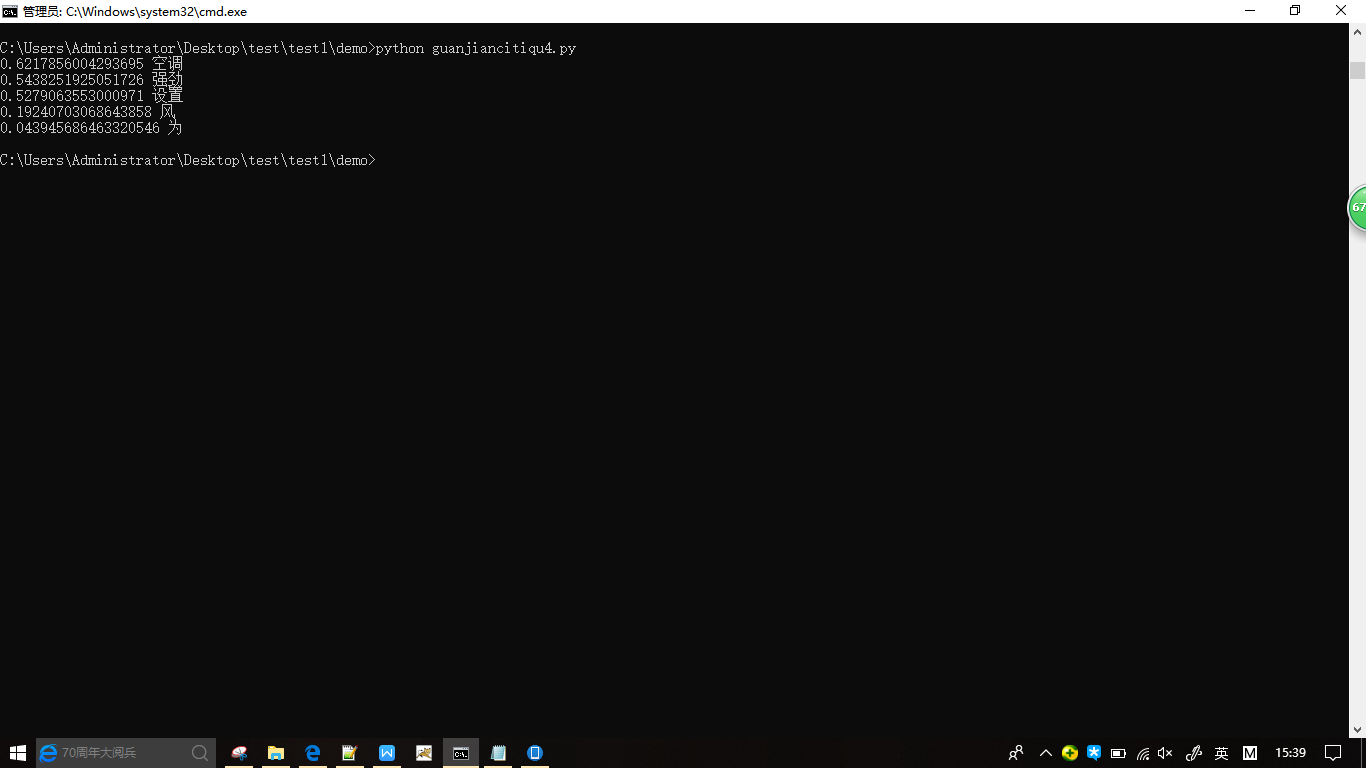
“晚上8点半打开空调,中午12点打开空调”的关键词提取：



“将空调温度设置为18度”的关键词提取：

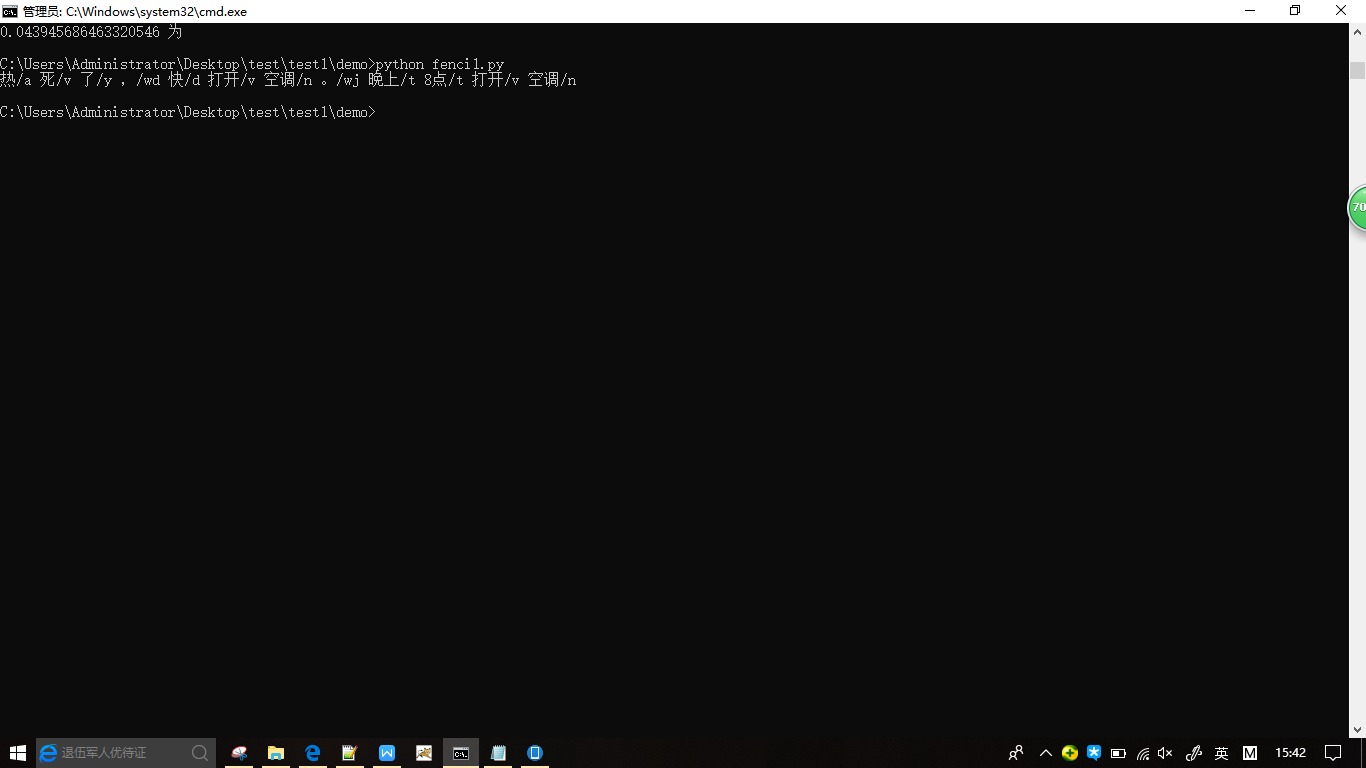


“空调设置为强劲风”的关键词提取

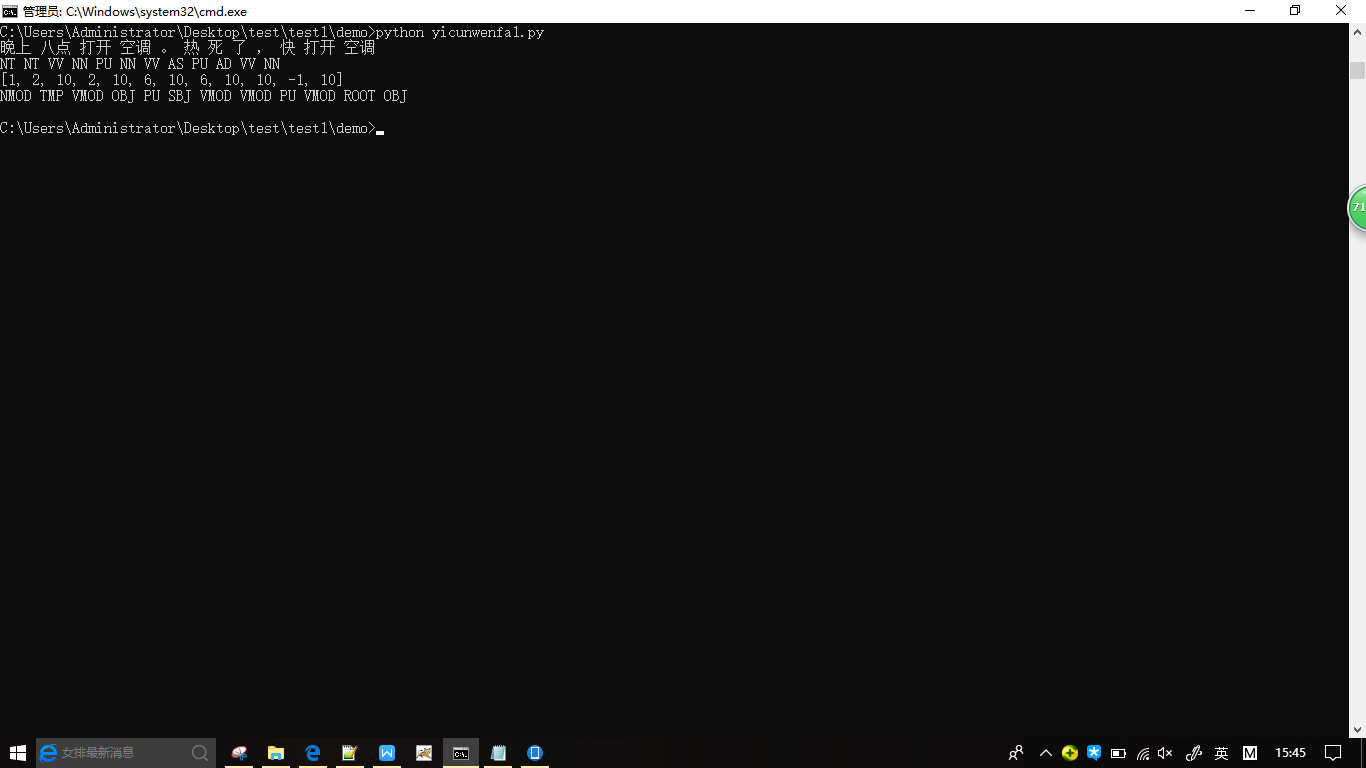


（3）分词以及语法分析

“热死了，快打开空调。晚上8点打开空调”的分词：



“晚上八点打开空调。热死了，快打开空调”的依存文法分析：



**3 软件运行测试**

（1）软件主界面，用户通过点击‘开始录音’按钮输入控制指令：

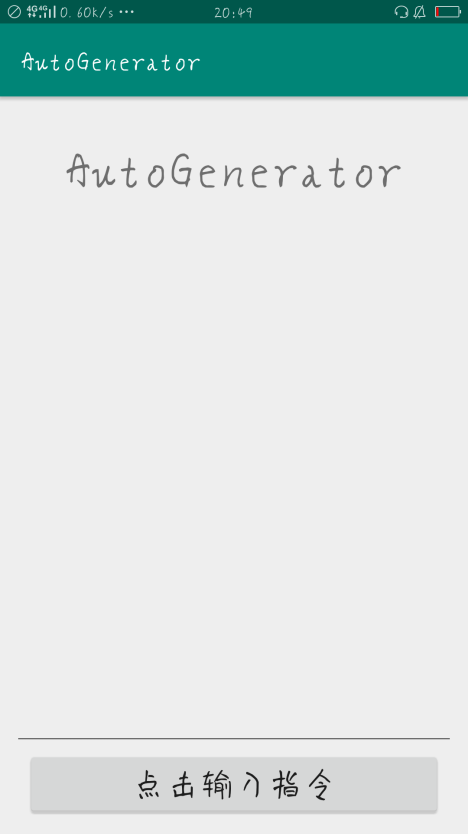


图1 策略自动生成器Auto Generator主界面图

（2）Auto Generator部分反馈指令图

语音输入“把冰箱温度设置成5度”之后的指令反馈图如下图2：

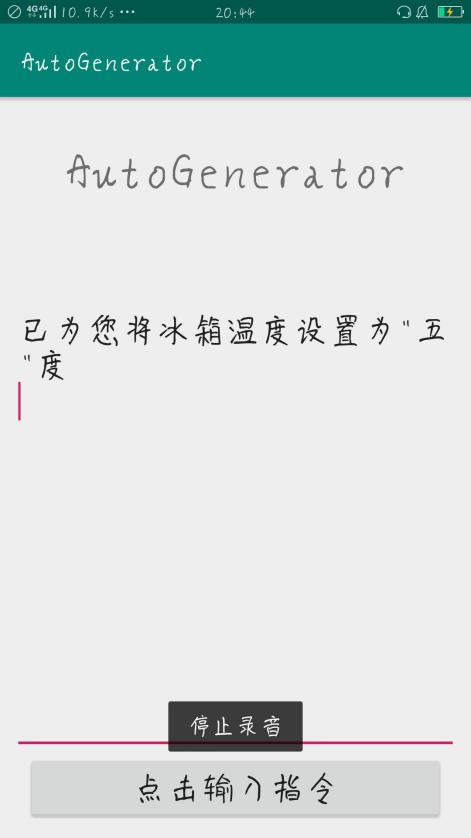


图2 Auto Generator指令反馈图

语音输入“请打开空调”/“打开空调”/“把空调打开”/“热死了，快打开空调”之后的指令反馈图如下图3：



图3 Auto Generator指令反馈图

语音输入定时开关机操作如“晚上8点半打开空调”/“3小时后将空调关机”之后的指令反馈图如下图4：

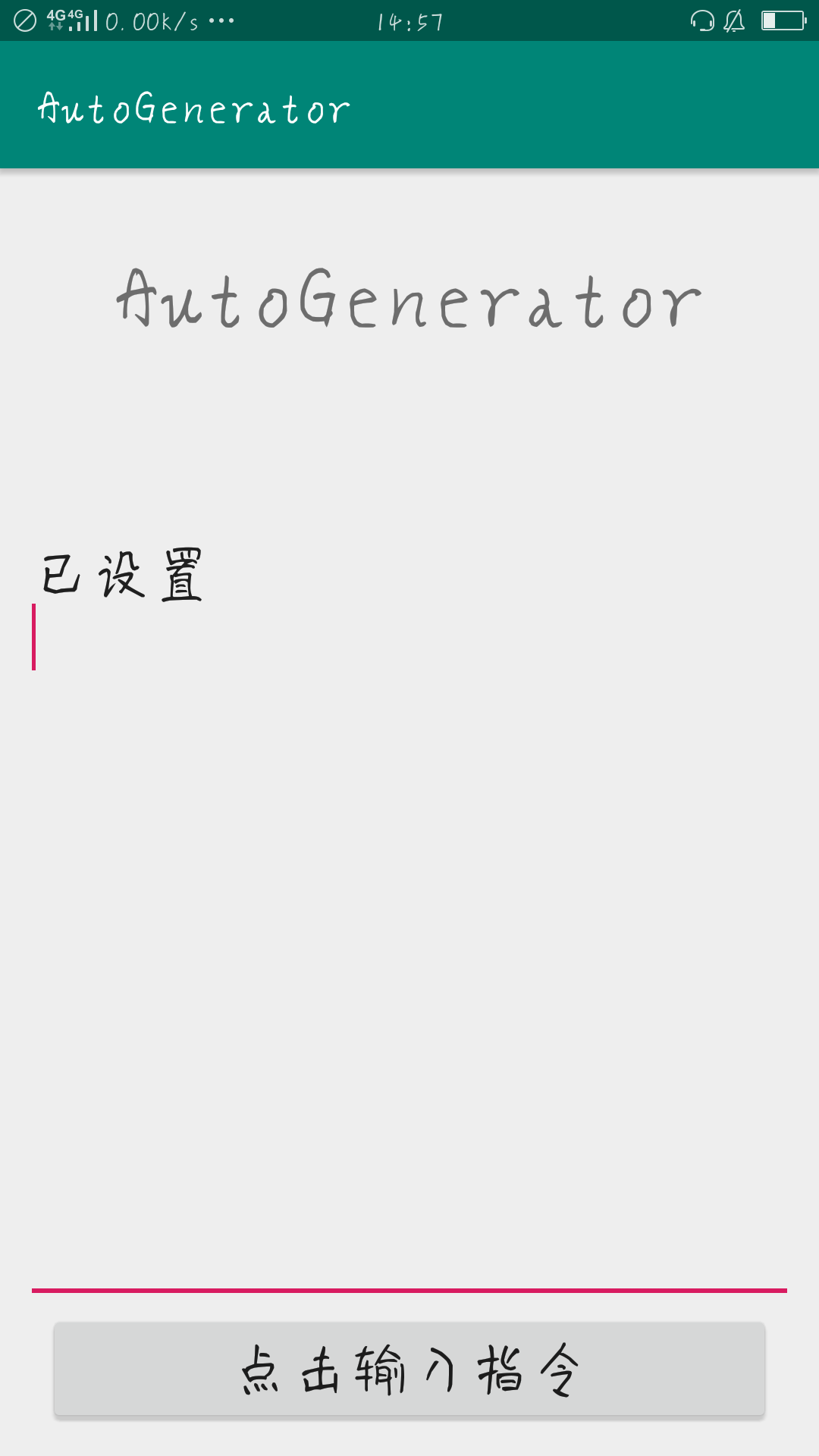


图4 Auto Generator指令反馈图

语音输入模式设定操作如“打开空调制冷模式”之后的指令反馈图如下图5：

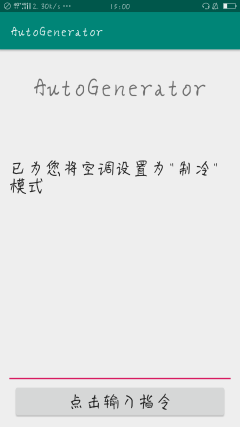


图5 Auto Generator指令反馈图

语音输入温度设定如“将空调温度设置为18度”之后的指令反馈图如下图6：

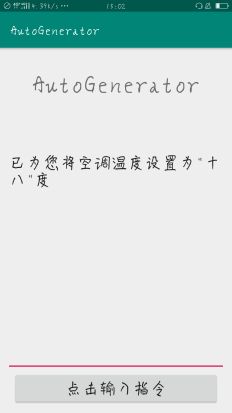


图6 Auto Generator指令反馈图

语音输入风速设置如“空调设置为强劲风”之后的指令反馈图如下图7：

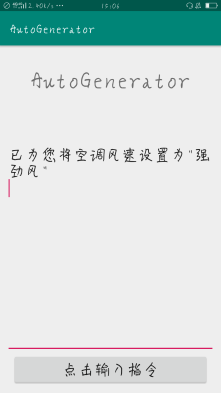


图7 Auto Generator指令反馈图