**G3.29**

**HỌ TÊN: Nguyễn Diệu Hương MSSV: 21110489 KÝ TÊN: HUONG**

**TRƯỜNG: HCMUTE**

**MÔN: Lập trình Python**

**NGÀY: .01/03/2023 (BUỔI HỌC SỐ 03)**

============

**Buổi 3 = Bài tập 2 = Lập trình Python xử lý giọng nói (Trợ lý ảo): Python Audio Pro**

**1. Giới thiệu chủ đề**

# 1.1. Tài liệu và nguồn thực nghiệm

* **SÁCH VÀ TÀI LIỆU (BẢN IN)**

[1]. Interactive Audio Special Interest Group, “*Interactive 3D Audio Rendering Guidelines - Level 2.0,*” Standard, MIDI Manufacturers Association, Los Angeles, CA, 1999.

* **TẬP MULTIMEDIA MẪU**

[https://www.freesoundeffects.com](https://www.freesoundeffects.com/)  OR https://freesound.org : Âm thanh

# 1.2. Giới thiệu Các thư viện Python liên quan

NAP CÁC THƯ VIỆN AUDIO speech\_recognition gTTS playsound

Ngoài ra, thêm:

Os : thư viện hỗ trợ lưu file, mở file, đường dẫn,..

time: thư viện lấy thời gian = đồng hồ

google\_trans\_new: thư viện hỗ trợ mô phỏng Google Translate.

# 1.3. Giới thiệu bài toán

**===BÀI MẪU===**

**Lập trình App xử lý "lời nới" (Speech):**

+ Nghe tiếng Việt => Text

+ Trả lời bằng tiếng Việt : Text => Nói tiếng Việt

## [1] Nghe tiếng Việt => Text

B1: Cài đặt thư viện

conda install …. Các thư viện sau:

------------------gốc hướng dẫn / dùng python trong visual (pip3) pip3 install SpeechRecognition

pip3 install pyaudio

B2 : Nạp thư viện vào App

import speech\_recognition as sr

B3 : Chọn nguồn âm thanh nhập :

+ Từ file âm thanh

+ Từ Microphone

# initialize the recognizer r = sr.Recognizer()

B4 : Xử lý nhận diện: từ MIC OR Files

**CODE (tham khảo)**

# -\*- coding: utf-8 -\*-

"""

Created on Tue May 18 01:33:29 2021

@author: VOXUAN

""" import speech\_recognition as sr

# NGHE NÓI => VIẾT RA CHỮ VIỆT r = sr.Recognizer() with sr.Microphone() as source:

print("Adjusting noise ")

r.adjust\_for\_ambient\_noise(source, duration=1)

print("Nói bằng tiếng Việt đi bạn 5s sau sẽ in ra Text...")

# read the audio data from the default microphone audio\_data = r.record(source, duration=5) print("Kết quả nhận diện...") # convert speech to text try:

text = r.recognize\_google(audio\_data,language="vi") except:

text = "bạn nói gì mình không hiểu!" print("Bạn đã nói là: {}".format(text))

|  |  |  |
| --- | --- | --- |
|  | **[2] Text => đọc tiếng Việt** |  |

B1: Cài đặt thư viện

|  |
| --- |
| conda install …. Các thư viện sau:  -----------------------gốc hướng dẫn / dùng python trong visual (pip) |
| pip install gTTS  pip install SpeechRecognition # đã cài đặt ở trên pip install playsound |

B2 : Nạp thư viện vào App

**import** speech\_recognition **as** sr **from** gtts **import** gTTS

**import** os

**import** time

**import** playsound

B3 : Xử lý

**FULL CODE (Tham khảo)**

# -\*- coding: utf-8 -\*-

"""

Created on Sat Feb 26 10:18:52 2022

@author: VOXUAN

"""

# B1: NẠP THƯ VIỆN

import speech\_recognition as sr

from gtts import gTTS import playsound

# B2: CHỌN PHƯƠNG ÁN NHẬP ÂM THANH TỪ MICROPHONE

r = sr.Recognizer()

# B3: XỬ LÝ NHẬN DIỆN MIC

#PHẦN 1: IN RA TEXT THEO LỜI NÓI = Nghe tiếng Việt => Text

with sr.Microphone() as Source: #hiệu chỉnh mic để chuẩn bị nói print("Hieu chinh nhieu trươc khi noi!")

r.adjust\_for\_ambient\_noise(Source, duration=1)

#nhận lời nói của người dùng từ MIc mặc định lưu dữ liệu âm thanh vào audio\_data print("Nói tiếng Việt đi, sau 5s sẽ in ra văn bản!") audio\_data = r.record(Source, duration = 5)

#In ra văn bản text

print("KẾT QUẢ NHẬN DIỆN ..................")

#chuyển lời nói thành văn bản try:

text = r.recognize\_google(audio\_data,language="vi") except:

text = "Quý vị nói gì nghe không rõ...!"

#in kết quả ra

print("Quý vị đã nói là : ",format(text)) #PHẦN 2: XUẤT RA LỜI NÓI THEO VĂN BẢN ĐÃ NHẬP = Trả lời bằng tiếng Việt : Text => Nói tiếng Việt

# ĐỌC CHỮ VIỆT => NÓI RA TIẾNG VIỆT

def HoTen(text):

Ten = gTTS(text=text, lang = 'vi')

filename = 'sttHoTen.mp3' Ten.save(filename)

playsound.playsound(filename)

HoTen("số thứ tự 29, họ tên: Leora..là sinh viên của TRƯỜNG ĐẠI HỌC Sư Phạm Kỹ Thuật Thành Phố Hồ Chí Minh.")

# Các Bài tập tương tự

Bài mẫu yc1: Nhập Y (Yes) để nói và in ra câu bạn vừa nói, nhập N (No) để không thực hiện cho đến khi nào bấm Y

#Nhập Y (Yes) để nói và in ra câu bạn vừa nói

#Nhập N (No) để không thực hiện cho đến khi nào bấm Y

# B1: NẠP THƯ VIỆN

import speech\_recognition as sr

from gtts import gTTS

import playsound

# B2: CHỌN PHƯƠNG ÁN NHẬP ÂM THANH TỪ MICROPHONE

r = sr.Recognizer()

# B3: XỬ LÝ NHẬN DIỆN MIC

with sr.Microphone() as Source:     #hiệu chỉnh mic để chuẩn bị nói

    n=input (("ban muon in hay chua: (Y/N)"))

    while (n=='N'):

        n=input (("ban muon in hay chua:(Y/N)"))

        if (n=='Y'): continue

    if n=='Y':

        print("Hieu chinh nhieu truoc khi noi!")

        r.adjust\_for\_ambient\_noise(Source, *duration*=1)

        #nhận lời nói của người dùng từ MIc mặc định lưu dữ liệu âm thanh vào audio\_data

        print("Nói tiếng Việt đi, sau 5s sẽ in ra văn bản!")

        audio\_data = r.record(Source, *duration* = 5)

        #In ra văn bản text

        print("KẾT QUẢ NHẬN DIỆN ..................")

        #chuyển lời nói thành văn bản

        try:

            text = r.recognize\_google(audio\_data,*language*="vi-VI")

        except:

            text = "Quý vị nói gì nghe không rõ...!"

        #in kết quả ra

        print("Quý vị đã nói là :  ",format(text))

Bài mẫu yc2: Cho chọn nhiều ngôn ngữ

# IN RA TEXT THEO LỜI NÓI PHÂN LOẠI NGÔN NGỮ

# B1: NẠP THƯ VIỆN

import speech\_recognition as sr

from gtts import gTTS

import playsound

# B2: CHỌN PHƯƠNG ÁN NHẬP ÂM THANH TỪ MICROPHONE

r = sr.Recognizer()

# B3: XỬ LÝ NHẬN DIỆN MIC

#lựa chọn ngôn ngữ nói:

print("Chose language: 1. Viet,    2. 日本語,   3. English")

x=input("Numbes: ")

def Lag(*x*):

    switcher={

            "1":"vi-VI",

            "2":"ja-JP",

            "3":"en-US",

            }

    return switcher.get(x, "0")

if Lag(x)=="0" : exit("ERROR")

with sr.Microphone() as Source:     #hiệu chỉnh mic để chuẩn bị nói

    print("Start: ")

    r.adjust\_for\_ambient\_noise(Source, *duration*=1)

    #nhận lời nói của người dùng từ MIc mặc định lưu dữ liệu âm thanh vào audio\_data

    if x=="1"  : print("Say with me in Vietnamese")

    elif x=="2" : print("Say with me in Japanese")

    else: print("Say with me in English")

    audio\_data = r.record(Source, *duration* = 5)

    #In ra văn bản text

    print("RESULT ..................")

    #chuyển lời nói thành văn bản

    try:

        text = r.recognize\_google(audio\_data,*language*= Lag(x))

    except:

        text = "I DON'T UNDERSTAND WHAT YOU SAY...!"

    #in

    print("You said with me:  ",format(text))

Bài mẫu yc3: lưu file theo tên nhập vào

import speech\_recognition as sr

import speech\_recognition as sr

from gtts import gTTS

import os

import time

import playsound

def HoTen(*text*):

    Ten = gTTS(*text*=text, *lang* = 'vi')

    name=input("nhap ten file (vd: test.mp3): ")

    filename =os.path.abspath('')+'\\'+name

    Ten.save(filename)

    playsound.playsound(filename)

HoTen("số thứ tự 29, họ tên: Leora .Là sinh viên của TRƯỜNG ĐẠI HỌC Sư Phạm Kỹ Thuật Thành phố Hồ Chí Minh.")

**Dựa vào Ứng dụng mẫu: “Trợ lý ảo” || Voice Assistant sau, thực hiện các bài tập sau**

**Bài 1:** Viết chương trình “thư ký” soạn thảo văn bản theo lời đọc của người dùng: dạng tương tự học sinh viết chính tả.

**Bài 2:** Viết lại chương trình mẫu với nhiều dạng khác:

+ Cho phép chọn Ngôn ngữ các quốc gia khác nhau: thay vì chỉ tiếng Việt, ví dụ: tiếng Ả rập, tiếng Anh,

Tiếng Hàn, tiếng Hoa, tiếng Nhật.*[mượn translate.google.com : input speech]*

+ Tiếng Việt: chọn giọng nam or nữ || giọng miền Nam or Miền Bắc or…

*[nạp thư viện khác của anh/chị Việt Nam: khác thư viện gtts]*

# IN RA TEXT THEO LỜI NÓI PHÂN LOẠI NGÔN NGỮ

import speech\_recognition as sr

from gtts import gTTS

import playsound

r = sr.Recognizer()

#lựa chọn ngôn ngữ nói:

print("Chose language: 1. Viet,    2. 日本語,   3. English")

x=input("Numbes: ")

def Lag(*x*):

    switcher={

            "1":"vi-VI",

            "2":"ja-JP",

            "3":"en-US",

            }

    return switcher.get(x, "0")

if Lag(x)=="0" : exit("ERROR")

with sr.Microphone() as Source:     #

    print("Start: ")

    r.adjust\_for\_ambient\_noise(Source, *duration*=1)

    if x=="1"  : print("Say with me in Vietnamese")

    elif x=="2" : print("Say with me in Japanese")

    else: print("Say with me in English")

    audio\_data = r.record(Source, *duration* = 5)

    #In ra văn bản text

    print("RESULT ..................")

    #chuyển lời nói thành văn bản

    try:

        text = r.recognize\_google(audio\_data,*language*= Lag(x))

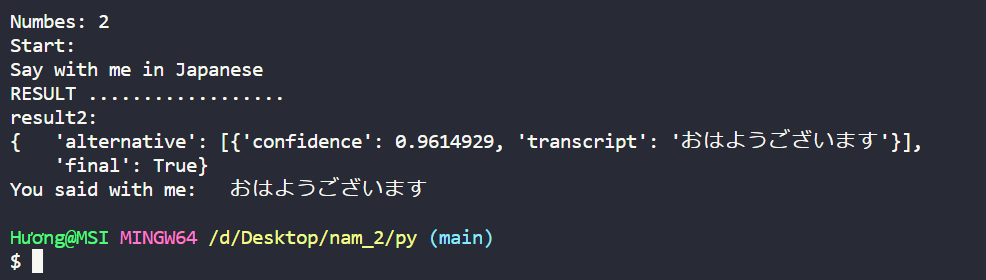
    except:

        text = "I DON'T UNDERSTAND WHAT YOU SAY...!"

    #in

    print("You said with me:  ",format(text))

kết quả:



Hương@MSI MINGW64 /d/Desktop/nam\_2/py (main)

**Bài 3:** Viết lại chương trình: nhận lệnh và xuất ra mã lệnh một đoạn chương trình mẫu bằng các ngôn ngữ lập trình khác nhau (cho người dùng chọn); VD: C#.NET, Python, Java, ..

**Bài 4:** Viết lại chương trình: mô phỏng google dịch

#GOOGLE DỊCH

import speech\_recognition as sr

from gtts import gTTS

import playsound

import os

from googletrans import Translator

r = sr.Recognizer()

#LỰA CHỌN NGÔN NGỮ NÓI ĐẦU VÀO

print("Chose language to translate: 1. Viet,    2. 日本語,   3. English")

x=input("Language in: ")

print("Chose language to out: 1. Viet,    2. 日本語,   3. English")

y=input("Language out: ")

def Lag(*x*):

    switcher={

            "1":"vi",

            "2":"ja",

            "3":"en",

            }

    return switcher.get(x, "0")

if Lag(x)=="0" : exit("ERROR")

if Lag(y)=="0" : exit("ERROR")

#XỬ LÝ ĐẦU VÀO

with sr.Microphone() as Source:

    print("Start: ")

    r.adjust\_for\_ambient\_noise(Source, *duration*=1)

    #nhận lời nói của người dùng từ MIc mặc định lưu dữ liệu âm thanh vào audio\_data

    if x=="1"  : print("Say with me in Vietnamese")

    elif x=="2" : print("Say with me in Japanese")

    else: print("Say with me in English")

    audio\_data = r.record(Source, *duration* = 5)

    #In ra văn bản text

    print("RESULT ..................")

    #chuyển lời nói thành văn bản đưa vào để dịch

    try:

        text = r.recognize\_google(audio\_data,*language*= Lag(x))

    except:

        exit  ("I DON'T UNDERSTAND WHAT YOU SAY...!" )

    #in lời nói đầu vào dưới dạng văn bản

    print("You said with me:  ",format(text))

#DỊCH NGÔN NGỮ THÀNH NGÔN NGỮ ĐÃ CHỌN

translator = Translator()

result = translator.translate(text, *src*= Lag(x), *dest*=Lag(y))

#in ra sau khi dịch

print(result.text)

#nghe cách đọc

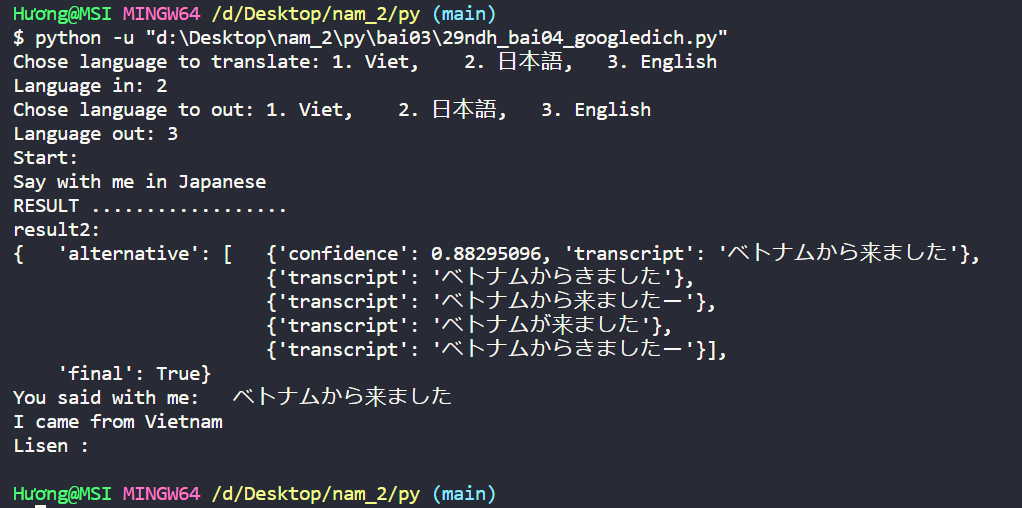
listen = gTTS(result.text, *lang* = Lag(y))

filename =os.path.abspath('')+'\sttListen.mp3'

listen.save(filename)

print("Lisen : ")

playsound.playsound(filename)

Kết quả: