

1. 動機

普段、Siri などの AI と会話していてなぜ会話が成り立っていないことがあるのか疑問に思ったので原因を調べようと思いました。

2. 使用機器

ソフト: IDLE (Python 3.6 64-bit)

OS: Windows 10

プログラムのダウンロードページ

<http://www.shuwasystem.co.jp/books/pythonpermas>

3. Python とは

Python とは、簡潔で読みやすい文法が特徴的な汎用の高水準プログラミング言語の一つです。Python では、豊富なライブラリが公開されています。

(ライブラリとは、多くのアプリケーションで利用するためにまとめられたプログラム集です。) ライブラリを用いることで、欲しいプログラムが簡単に作成できます。

チャットボットとは

「チャットボット (chat bot)」とは、「チャット」と「ボット」を組み合わせた言葉で、人工知能を活用した「自動会話プログラム」のこと。

「チャット」は、インターネットを利用したリアルタイムコミュニケーションのことで、主にテキストを双方向でやり取りする仕組みのことです。

「ボット」は「ロボット」の略で、人間に代わって一定のタスクや処理を自動化するためのプログラムのことです。

今回使った外部ライブラリについて

Janome

Janome は Python の形態素解析エンジン。日本語のテキストを形態素ごとに分割して品詞を判定したり分かち書き (単語に分割) したりすることができる。MeCab などの外部エンジンは必要なく pip コマンドでインストール可能。

Beautiful Soup

Beautiful Soup は HTML データの構文を解析するためのパッケージです。HTML を解析するために、HTML タグや CSS のセレクタで抽出する部分を指定することができます。

それぞれのクラスが返事をする確率

パターンクラス 50%

テンプレートクラス 20%

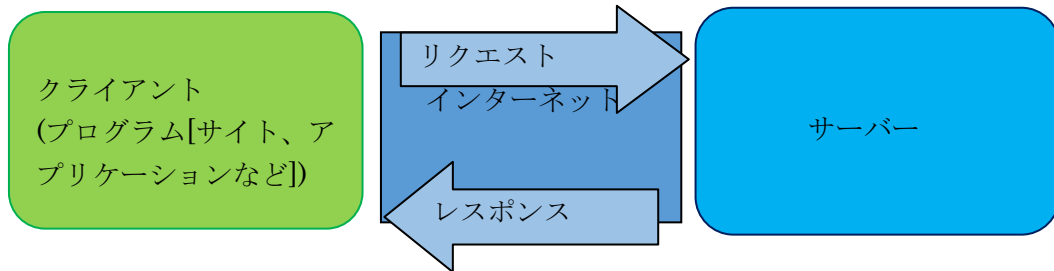
ランダムクラス 25%

リピートクラス 5%

Web API について

Web API とは、コンピュータプログラムの提供する機能を外部の別のプログラムから呼び出して利用するための手順・規約（API：Application Programming Interface）の類型の一つで、HTTP など Web の技術を用いて構築されたもののこと。Web サイトに外部のサイトの提供する機能や情報を組み込み、アプリケーションソフトから Web 上で公開されている機能や情報を利用する際などに用いられる。

Web API の図



実行した画面▽



4. 結果

文章の解析やパターンの解析で会話を成り立たせようとする、どうしても矛盾がおきる。あるいは膨大な会話のデータが必要になる。(LINEの「りんな」は自身と他のユーザーと会話しているデータを集め、長く続いた会話を取り込んでいる)

5. 感想

普段使っているスマートフォンのSiriやLINEのりんななどのチャットボットが本当の人間のように対話出来ない理由を今回の実習を通して理解できて良かったです。Pythonのことについて調べていて、DNA配列の分析、音声認識や文字認識などのパターン認識、など、Pythonによるデータの解析は色んなところで使われているんだなと思いました。今回のチャットボットの文章の作成もBeautiful Soupを使った文章のデータ解析を行っているので意味のある文章が出来ました。