

チャットボットへの試み

芝浦工業大学 数理科学研究会
堀井雷生

平成 29 年 11 月 3 日

研究背景

最近、インターネット内で行われる通信方法について触れる機会があり、試しに自分で実装したいという思いからプログラム間での通信を行ってみた。ついでにその通信を使って何かしてみたいと感じ、その何かに対する答えとして一番最初に思いついたのがチャットボットだったので、プログラムを使って異なるホスト間でのチャットボットを行えるようにすることを本研究の目的として、その制作に取り組んだ。

1 TCP/IP 階層モデル

現代において、異なる機種間のコンピュータ同士が通信できるのは機種に関係なく、ネットワークの構造やプロトコルと呼ばれるコンピュータ同士が通信をする際の手順や規則が統一されているためであり、そのモデルとなるものが ISO(International Organization for Standardization: 国際標準機構) によって定められた OSI 参照モデルである。OSI 参照モデルは 7 階層で構成されており、これを第 4 層: アプリケーション層, 第 3 層: トランスポート層, 第 2 層: インターネット層, 第 1 層: ネットワークインターフェース層の 4 階層にまとめたものが TCP/IP である。データを送信の際は元となるデータに第 4 層のデータを付加し、第 1 層までそれぞれの層のデータを付加してから相手に送信する。受信側は受け取ったデータを送信側とは逆の順序で読み、データを受け取る。このうち、トランスポート層に属するのが TCP(Transmission Control Protocol), UDP(User Datagram Protocol) であり、送信先の相手の存在の有無を確認せずにデータの送信を行う UDP であるのに対して、TCP では送信先と確実に接続が出来た場合にのみデータを送信する。信頼性は高いがその分時間がかかる。

2 マルコフ連鎖

文については、もともとある文を形態素解析ツール Mecab を用いて品詞に分解し、それらをマルコフ連鎖を用いて並べることによって生成することにした。マルコフ連鎖はある初期状態分布 S_0 に遷移確率行列 P をかけると次の状態分布 S_1 は

$$S_1 = S_0 P$$

となり、次の状態が前の状態に依存する性質を持つ。これにより、文の先頭に来る品詞さえ決めておけばそこから後ろの品詞は勝手に決まるのではほぼ自動的に文を作ることが出来る

3 通信手順

TCP による通信では特定のポートで接続要求を待つサーバとサーバが持つポートに接続要求を行うクライアントの 2 つの役割が必要であり、以下のような手順で通信を行っている。

サーバプロセス	クライアントプロセス
1 ソケット作成	1 ソケット作成
2 接続待ちにする IP アドレスとポート番号の設定	2 接続相手の IP アドレスとポート番号の設定
3 ソケット命名	3 接続要求
4 接続待機	4 通信 (文を送信)
5 接続受付	5 終了
6 通信 (文を受信)	
7 終了	

図 1: サーバ、クライアント間の通信

ここでのソケットは、サーバとクライアントを仮想的に結びつけるの役割を持つもののことを指している。また、今回は上記の手順を元にクライアントにはサーバの機能を、サーバにはクライアントの機能を持たせ、相互に文のやりとりが出来るよう改良し通信を行った。

今後の課題

インターネットにおいてどのように通信が行われているのかやマルコフ連鎖などについて、その片鱗に触れることが出来て良い経験になった。今後異なるホストでの通信やより論理性のある自動文の生成を行えるようにしてみたい。

参考文献

- [1] http://www.sbcr.jp/books/img/Linuxnet_02.pdf, 最終アクセス: 10 月 27 日.
- [2] <http://ascii.jp/elem/000/000/434/434183/index-2.html>, 最終アクセス: 10 月 29 日.
- [3] Raymond A. Barnett, Michael R. Ziegler, Karl E. Byleen, 初歩からの数学 論理確率とマルコフ連鎖, 丸善出版, 2015.