ML のアーカイブを用いた FAQ 自動生成システム

02H080 西川一美

1　はじめに

現在 IT 産業の発展により、インターネットを利用したサービスも数多く生まれた。その中でよく利用されているのは電子メールである。というのも、最近では携帯電話を使って、いつでもどこでも利用できたり、フリーメールアドレスを簡単に取得できるようになったことで身近なものになってきたからである。今日では、電子メールは様々な世代、場面 (趣味、仕事) で活用されている。この電子メールを使ったサービスの一つとして、メーリングリストがあげられる。

本研究では、メーリングリストのアーカイブを用いた機能の構築を行う。対象とするグループは、大学という限られた空間で利用されるもので、対象者は学生である。目的は、授業で役立つシステムの構築である。

本研究で構築する機能とは、過去のデータを使い、質問メールとそれ以外を分別するシステムである。

最終的には、取り出したデータをリスト化して、疑問を解決できる機能を実装化させることが目標である。

2　概要

授業で使用するメーリングリストでは、質問メールが何通も届く。すべてのメールをチェックするのは大変である。また、同じ内容の質問に何度も答えることは手間のかかる行為である。

そこで、その為にメーリングリストのデータベースから質問メールとそれ以外を分け、質問とそれに連なるスレッドのリストを作成して FAQ を生成する。学生の質問メールを受けて、もし過去に同じ内容のものがあれば、過去に同じ質問があったことを教える文と該当する情報 (スレッドリスト) を過去のデータから取り出してメール本文を生成し、発信する。しかし、過去のデータから該当するものがない場合は解答文を作成して発信する。

3　まとめ

この研究テーマに決定した動機は、「授業で使っているメーリングリストに便利な機能を実装したい」であった。そして、メールで質問すれば自動で返答が送られ、過去に同じ内容の質問があればリストが届くようなシステムを目指し、研究テーマになった。本研究では、目指したシステムとは少し異なり、システム一部、重要な分別プログラムの作成だけを行った。

今回のシステムは、授業とはまた違う方面で使用しているメーリングリストのデータをサンプルデータとして使用した。そのため、実際に授業で使う場合とは異なるパターンの文章が多々あると見受けられる。そのため、授業用のメーリングリストとして実装するためには、さらなるパターンの検証とフレーズの追加・削除が必要だと考える。