

帰納学習に基づく電子メールエージェント*

6Y-4

～ システムの概要とデータの収集 ～

溝口 文雄† 大和田 勇人† 徐 翠君† 柳田 正博†

東京理科大学 理工学部†

1 はじめに

最近のインターネットの爆発的な普及を背景に情報化社会が確立しつつある。このような情報化社会において人間の情報活動を支援するエージェントの役割が関心を集めている [1]。一般的にエージェントが備えるべきとされている性質には、自律性、適応性、局所性、可動性、協調性などがあげられるが、本稿においては「ユーザーの意志を代行する代理人」という意味のソフトウェアを“エージェント”と定義している。

本稿では、我々が開発した電子メールを対象にその操作を支援したり、代行したりするエージェントについて報告する。このエージェントの最大の特徴は帰納学習機能を備えることによる個人依存型の電子メール操作環境を提供することにある。

2 電子メールエージェントの構成

本稿で報告するエージェントは、・フィルタリング情報や論理規則の管理、・メッセージの表示要素の生成、・データの抽出・加工、・学習規則の生成と適用を行ない、ファイル処理モジュール、インタラクションモジュール、通信モジュールから構成されている (図 1)。そして学習システム及びデータベースサーバとインタラクションをとりつつユーザー依存型の有益な電子メール操作環境を提供することを目的としている。

以下の節では、エージェントの行なう処理について、順に述べていく。

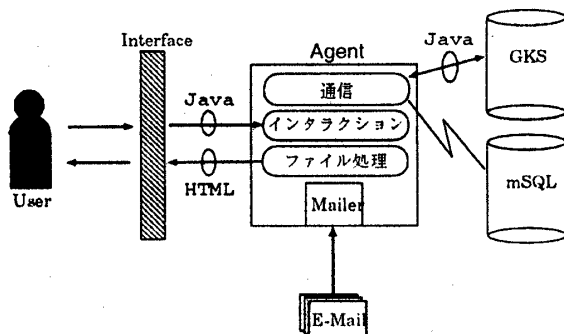


図 1: 電子メールエージェントの構成

* E-mail Agent using Inductive Learning

† Fumio MIZOGUCHI, Hayato OHWADA, Tsui-chun HSU, Masahiro YANAGIDA

‡ Faculty of Sci. and Tech. Science University of Tokyo

3 データ処理

電子メールを有効利用するために分類整理は必須である。そのため電子メールのフィルタリング情報と属性を考慮して、ユーザーの好みにあうようにカスタマイズが必要がある。ここで、フィルタリング情報とは、電子メールから抽出される、送信者の名前やメッセージピックなどの情報で、ユーザー側では変更できない。また属性とは、分類先、興味の有無などユーザーに依存した情報で柔軟に設定・変更ができる。そしてこれらの抽出されたフィルタリング情報や属性はメールに一意に与えられたファイル名をキーとしてデータベース化されて蓄積されるため、メールの再利用時には SQL を用いての検索操作も可能となっている。

本研究においてはフィルタリング情報と属性は次のようなものを考える。

• フィルタリング情報

User	送信者
Domain	所属
Subject	メッセージピック
Returned	返信
has_word	本文からの抽出語句

• 属性

Category	分類カテゴリ
Priority	優先度
Preference	嗜好
Worth reading	興味

エージェントは、すでに処理済みの電子メールからのフィルタリング情報やユーザーがあらかじめ与えておいた属性から学習によって論理規則を導出し、これを適用することによって、新着のメールの属性を推定してアイコンや文字列でユーザーに示す。現在、Category と Priority の表示を実装している。

4 学習機能

本エージェントが呼び出す学習システムは帰納論理プログラミングの枠組みによるもので、我々の研究室で開発された GKS/Java (以下 GKS)¹である。本システムは、Java²言語による小型ソフトウェア・モジュール (アプレット) として実装されており、ユーザーから学習規則生成の要求をされるとエージェントは GKS ア

¹ GaKuShu on Java Language

² ネットワーク・コンピューティングを意識して設計された言語環境、Sun Micro Systems 社が開発した。

プレットを呼びだし、抽出処理によって得られたデータから節形式で表現される論理規則を導出する。この規則を新着メールに適用することによってメール操作を軽減することが可能となっている(例えば自動分類)。

5 ユーザーインターフェース

ユーザーインターフェースには、Java に対応した WWW ブラウザ、Netscape Navigator³を用いる。これは GUI の利用により非常に操作性に優れ、インターネット利用上のツールとして最も広くユーザーに受け入れられているため、インターフェースに採用することでユーザーに使いやすい環境を提供できるからである。また、Java に対応したブラウザであるため Java 言語を用いることにより、HTML として記述されたメールのなかにさまざまなインタラクティブ性を提供することができる。本エージェントではこれを利用して、ユーザーに依存した電子メールの属性を抽出する。

開発中のエージェントはメールの未既読によって新規メールモードと検索モードの2つの顔を持つ。新規メールモードでは、対象を新着メールとし、メッセージの表示と属性情報の収集を行なう。図2はメッセージ表示のインターフェースで、メッセージの上にメールの操作に必要なコマンド群を備えている。

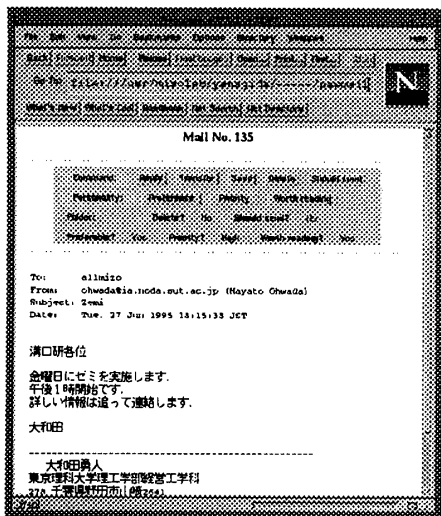


図2: 新規メールモードのインターフェース

新着メールに対してエージェントは学習によって得られた規則を利用して属性の推定を行なうが、ユーザーが属性の訂正や修正、追加あるいは新たに属性を収集したい場合、Java 言語によってメッセージの上部のコマンド群内に実装された属性収集インターフェースを用いて行なうことができる。Java 言語で実装することで、ユーザーとのインタラクションを容易に実現し、マン・マシンによる知的作業の拡張やコラボレーションを現実のものとする可能性も秘めている。

また、検索モードでは図3のインターフェースで収集された属性に基づいて電子メールを分類し、再利用時

には容易に検索できるようになっている。図3に既存のメールを属性“分類カテゴリ”で整理して表示したものである。ここでは、整理された抽出属性以外の属性はアイコンで表示され、フィルタリング情報にハイパーリンクを形成する。メール自体はHTML形式のファイルに変換されて保管されているので、メールへのアクセスはリンクを張ることにより容易に実現できる。図3では、フィルタリング情報である“User@Domain”にリンクを生成している。

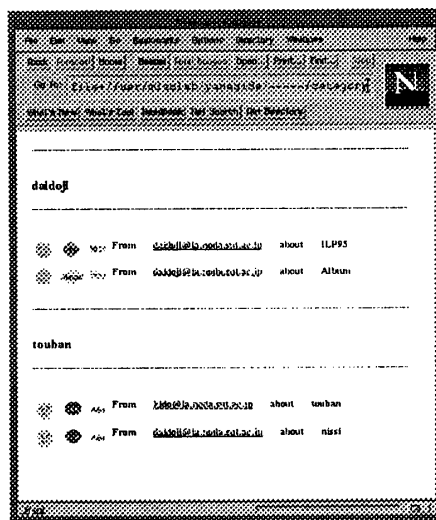


図3: 検索モードのインターフェース

メールに対する返信やメールの送信を行なう場合の入力インターフェースとしては現在 mule を採用しているが、返信の場合にはフィルタリング情報を引き渡すことによってエージェントが宛先などを入力するのでユーザーが入力する必要はない。

6 おわりに

本稿では、帰納学習に基づくエージェントにおけるシステム概要とデータの収集について報告した。本研究で設計した電子メールエージェントは電子メールに属性を付加して HTML 形式で保管、管理しているため、メールへのアクセスが容易である。しかし、HTTP をプロトコルにすることによるユーザー認証の問題などがまだ未解決であり、また多くの解決すべき課題が残されている。このような問題点を追求し、解消していくことが必要である。

参考文献

- [1] Tom Mitchell, Rich Caruana, Dayne Freitag, John McDermott and David Zabowski: Experience with a Learning Personal Assistant, *COMMUNICATIONS OF ACM*, Vol.37, No.7, pp.81-91,1994.
- [2] 溝口文雄, 柳田正博, 大和田勇人: 帰納学習による分類機能をもった電子メールクライアントの開発, 情報処理学会第51回全国大会論文集, 3, pp.213-214,1995.

³ ©Netscape Communications Corporation