

## Probeにおける語彙モデルとデザインパターンの利用と考察\*

5 C-2

青木 寛 山腰 哲 岩田 真明 木村 耕

電気通信大学 情報工学科

## 1 はじめに

Probe(同時発表)とは、エージェントがネットワーク上に複数ある部品検索サーバ間を渡り歩き、開発者の要求に適した部品を検索するシステムである。部品検索サーバには開発済みのPMS/Jを利用して[1]。

Probe内で活動するエージェントは心的状態の変化により処理の過程を蓄積することができるが、この機能だけでシステムの効率を改善していくのは難しい。そこで以下に説明する語彙モデル、デザインパターンを利用した二つの手法により、システムの効率化を支援する。語彙モデルでは検索に関する効率化を、デザインパターンではより大局的な効率化を図った。

## 2 語彙モデル

エージェントは開発者が入力したキーワードをもとに部品の検索をおこなうのだが、PMS/Jでは文字列のマッチングによって検索を行っている。これでは意味的に近い単語であっても字面が異なると全く検索にかからない。

これを補うために語彙モデルを用いた。語彙モデルとは単語間の関連に重み付けしたネットワークのことである。関連に幾つか種類を持たせ、その種類と重みによって単語間の意味的な関係やその度合を表している。

## 2.1 重み変更アルゴリズム

学習に際して重みを変更する必要があるが、変更アルゴリズムはパーセプトロンの学習モデル(デジタルモデル)をヒントにした。パーセプトロンとは神経細胞の結合をモデル化したパターン識別装置であり、学習機能を備えている[2]。

\*A study of Vocabulary Model and Design Pattern for Probe by Hiroshi Aoki, Akira Yamakoshi, Masaaki Iwata and Koh Kimura, Department of Computer Science and Information Mathematics, The University of Electro-Communications(UEC).

重み変更のアルゴリズムは次の漸化式で表される。

$$w(t+1) = \begin{cases} w(t) & \text{if } x(t) \in X_0 \wedge w(t)^t x(t) > r \\ w(t) & \text{if } x(t) \in X_1 \wedge w(t)^t x(t) < r \\ w(t) - \epsilon x(t) & \text{if } x(t) \in X_0 \wedge w(t)^t x(t) \leq r \\ w(t) + \epsilon x(t) & \text{if } x(t) \in X_1 \wedge w(t)^t x(t) \geq r \end{cases}$$

ここで、 $w(t)$ は重みを表すベクトル、 $x(t)$ はパスを表すベクトル、 $X_0$ は望ましくない単語へのパスの集合、 $X_1$ は望ましい単語へのパスの集合である。 $r$ は検索範囲を表す定数である。 $X_0$ 、 $X_1$ 、 $r$ はユーザによって与えられる<sup>1</sup>。

## 2.2 利用方法

Probeにおいて、検索キーの意味的な拡張に語彙モデルを利用している。部品の検索は

1. 入力されたキーで検索する
2. 検索キーから連想される単語で検索する

という手順で行う。手順2により、意味的に近い単語をキーワードとして持つ部品も検索することができる。

## 2.3 動作例

具体例をとって語彙モデルの有効性を検証する。語彙のサンプルとしては、エディタ作成時に用いられる単語を使用した(図1)。このモデルで学習前後において、検索により獲得できた有効な部品数の比較を表1に示す。

## 3 デザインパターン

デザインパターンとは、一般的な設計構造のキーとなる側面に名前を付け、抽象化、識別化し、再利用可能なオブジェクト指向設計を生み出すのに有用となるようにしたものである。また、デザインパターンはそれにかかわっているクラスやインスタンス、それらの役割や協調関係、責任の分担を規定す

<sup>1</sup>  $X_0$ 、 $X_1$ に関しては、ユーザが実際に集合が与えるわけではなく、教師信号を与えてもらう。

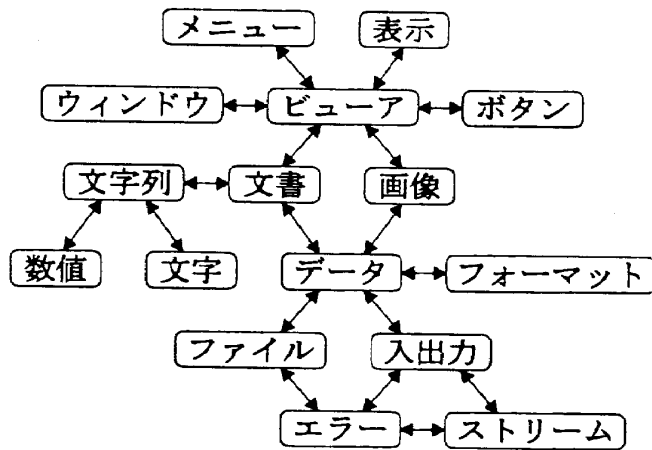


図 1: 語彙モデルの例

表 1: 学習回数と有効な部品数

連想もと	学習回数		
	0	100	200
文字列	2	3.73	6.87
フォーマット	1	3.72	5.79
ファイル	1	3.55	5.86

数値は 1000 回実験した平均値

る。それぞれのデザインパターンは、オブジェクト指向設計における特定の問題や課題に焦点を絞っている [3]。以下に Probe で採用した適用方法について述べる。

### 3.1 検索方法

デザインパターンの検索は過去に開発された事例をもとに行われる。デザインパターンの検索を行うとき、まずキーワードから過去事例を検索する。過去事例には設計情報として適用したデザインパターンをはじめ様々な情報を持っている。この適用したデザインパターンを提示することにより、開発者が必要としている設計情報を提供することができる。

### 3.2 利用方法

上記の検索手順 (2.2 節) により適当な部品が見つからなかったとき、デザインパターンと使用例を提示する。開発者はデザインパターンを参照することにより、可読性、再利用性の高いアプリケーションを開発することができる。

### 3.3 デザインパターンの効果

部品が見つからなかった場合、開発者はデザインパターンをもとにアプリケーションの開発を行う。デザインパターンを適用したアプリケーションには、再利用性の高い構成部品が含まれていると考えられる。この構成部品を PMS/J に登録することにより、より適切な部品を提供できるようになる。

## 4 考察

### 4.1 語彙モデル

現在、重み変更アルゴリズムはデジタルモデルのパーセプトロンの学習を参考にしている。これは学習の効率という点で問題がある。このため、アナログモデルの学習や、その他の学習モデルを利用してその効率の違いを比較実験してみる必要がある。

### 4.2 デザインパターン

現システムでは、デザインパターンの利用は過去事例からのキーワードによる検索のみにとどまっている。現在の検索方法では、過去事例が存在しない場合にはデザインパターンを提示することができない。このため、他の経路でデザインパターンを提供する方法が必要である。ひとつの方法として、メタパターンを利用した検索システムの実現について現在考察中である。

## 5 おわりに

本論文では Probe において語彙モデル、デザインパターンを利用することが、検索システムの効率化を促進させることを述べた。現在、これらの利用に関してさらに効果的な方法を模索中である。

## 参考文献

- [1] 植木紀昌、青木寛、野村竜太郎、木村耕、“Java 部品管理システムを利用したソフトウェア開発環境の実例”、情報処理学会、第 54 回全国大会、講演論文集 (分冊 1) pp227-228(1997-3)
- [2] 上坂 吉則、尾関 和彦 著、“パターン認識と学習のアルゴリズム”、文一総合出版 (1990)
- [3] E.Gamma、R.Helm、R.Johnson、J.Vlissides 著、本位田 真一、吉田 和樹 監訳、“オブジェクト指向における再利用のためのデザインパターン”、ソフトバンク (1995)