

ソフトウェアパターンの自動的な体系化の試み

久保 淳人[†] 鷲崎 弘宜[†] 深澤 良彰[†]

本稿では、ソフトウェアパターン間の関係を分析しソフトウェアパターン体系を自動的に構築するシステムを提案する。

An Approach to Automatic Systematization of Software Patterns

ATSUTO KUBO,[†] HIRONORI WASHIZAKI[†] and YOSHIAKI FUKAZAWA[†]

In this paper, we propose a system that systematizes software patterns automatically.

1. はじめに

WWW(World Wide Web)には、既に多数のソフトウェアパターン(以下パターン)文書が存在する。これらのパターンを体系化すれば、目的とするパターンの検索と、関連する複数のパターンの連続的な適用を効率よく行うことが出来る。しかしながら、デザインパターンの分類および体系化については幾つかの試みが存在する²⁾⁶⁾ものの、既に1000以上のパターンが発表され、なお増え続けている現在では、人手によるパターンの網羅的な体系化は事実上不可能である。

本稿では、ソフトウェアパターン間の関係の分析とソフトウェアパターン体系の構築を自動的に行うシステムを提案する。本システムを用いることによって、パターン間の関係の分析を一部自動化する。また、それまで人間が気付かなかったパターン間の関係についての示唆が、本システムから提供されることが期待できる。

2. 本システムの概要

本システムは、入力されたパターン文書集合について、パターン間の関係を抽出する。HTMLを用いて主に自然言語で記述されているパターン文書を扱う。

第1段階では、HTMLによるマークアップなど、入力されたパターン文書の物理構造を解析する。パターン文書の多くが見出しと本文の対が複数集まる構造であることを考慮して、出力は見出しと本文の対の集合

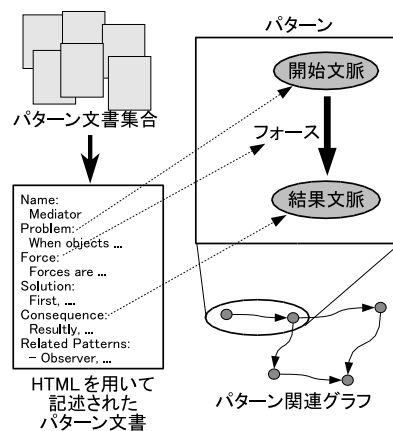


図1 分析の流れ

とした。

第2段階では、見出し-本文対の集合が既知のパターン記述形式に適合している場合に、その記述形式に沿って見出しと本文の対をシステム内部のモデル(3で詳説する)に再編成する。パターンフォームリポジトリには、既存のパターンフォームとの適合性を判断するための情報があり、必要に応じて利用する。

第3段階では、得られたパターン間の関連を分析する。詳細は次節で解説する。

3. パターンの形式化

複数のパターンを連続して利用する場合に、パターンを適用した結果として得られる文脈として、他のパターンが想定する問題を得ることがある⁵⁾。パターン適用前の文脈を開始文脈と呼ぶことにする。また、

[†] 早稲田大学理工学部
School of Science and Engineering, Waseda University

表 1 開始文脈が類似するパターンの対 (一部)

パターン 1	パターン 2	類似度
AttributeObject	CollectionObject	0.86
AttributeObject	RelationshipObject	0.79

表 2 結果文脈が類似するパターンの対 (一部)

パターン 1	パターン 2	類似度
AttributeObject	Command	0.24
Command	Prototype	0.12

フォースはパターンの適用の際考慮する事柄である。

これを踏まえて、本システムにおいては、パターンを開始文脈、フォース、結果文脈の組として扱う。C を文脈の集合、 Λ はフォースの集合として、パターンの集合 P を以下のように定義する。

$$P = \{(c_i, c_j, \lambda_k) | c_i, c_j \in C, i \neq j, \lambda_k \in \Lambda\}$$

$$C = \{c_1, c_2, \dots, c_n\}$$

$$\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_m\}$$

このとき、パターンの集合は、文脈を頂点、パターンを有向辺とする多重有向グラフを形成する。これをパターン関連グラフ (Patterns Relationship Graph, PRG) と呼ぶことにする。

このように定義したパターン間の関係は幾種類か考えられるが、簡単のため、本システムではまず類似関係のみを扱うことにする。類似度の算出には tf*idf 法⁴⁾を用いる。パターン全体の類似関係の他に、パターンの各項目間で、例えば以下のような類似関係が考えられる。

開始文脈同士の類似 類似した問題を扱うパターンの対

結果文脈同士の類似 適用の結果が類似したパターンの対

一方の結果文脈と他方の開始文脈との類似 前者の後に後者を連続して適用可能なパターンの対

フォース同士の類似 フォースが互いに類似するパターンの対

4. 実 験

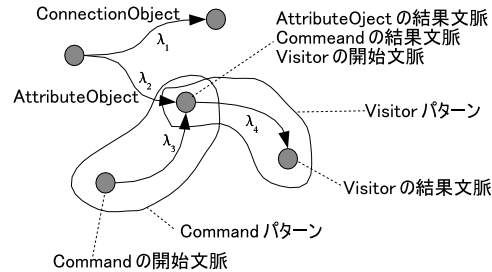
システムの有効性を検証するため、実装と簡易実験を行った。システムに文献 1) のパターン文書 (42 個) を入力し分析した。開始文脈同士、結果文脈同士、一方の結果文脈と他方の開始文脈が類似していると判定されたパターンの対をそれぞれ表 1 表 2 表 3 に示した。また、分析結果から作成した PRG の一部を図 2 に示した。

図 2 からは、Command パターンに続いて Visitor パターンを適用できる可能性と、AttributeObject パターンに続いて Visitor パターンを適用出来る可能性

表 3 結果文脈と開始文脈が類似するパターンの対 (一部)

パターン 1(結果文脈)	パターン 2(開始文脈)	類似度
Command	Visitor	0.39
Command	Iteraor	0.18

図 2 パターン関連グラフの例



があることが分かる。しかしながら、この PRG は恣意的な抽出の結果作成されたものであるので、客観的な評価方法を早急に確立する必要がある。

5. 今後の課題

まず、客観的な評価方法の確立が求められる。また、今回の実験では単一起源のパターン文書集合を入力としたが、複数起源のパターン文書集合を入力した場合についても評価したい。また、今回はフォースについての分析を行わなかったが、フォース同士の関係については、Ong らがフォース間の対立や協調・導出といった関係を示している³⁾。

参 考 文 献

- 1) An Object-Oriented Pattern Digest - Software Design Patterns Catalog, All Patterns Books and Resoures, <http://patterndigest.com/>
- 2) Erich Gamma ほか (著), 本位田 真一, 吉田 和樹 (監訳), デザインパターン - オブジェクト指向における再利用のための, ソフトバンク (1995).
- 3) Han-Yuen Ong, Michael Weiss and Ivan Araujo, Rewriting a Pattern Language to Make it More Expressive, In Proc. of ChiliPLoP2003 (2003).
- 4) 徳永 健伸, 情報検索と言語処理, pp11-33, 東京大学出版会 (1999)
- 5) 鷺崎 弘宜, 深澤 良彰, ソフトウェアパターン研究の現在と未来, 情報処理学会第 141 回ソフトウェア工学研究会 (2003).
- 6) Walter Zimmer, Relationships between Design Patterns in James O. Coplin, Douglas C.Schmidt (Eds.) Pattern Languages of Program Design (Proc. PLoP'94), Addison-Wesley, Reading.