

平成 21 年 6 月 16 日現在

研究種目： 若手スタートアップ
 研究期間：2007～2008
 課題番号： 19800044
 研究課題名（和文） ソフトウェアパターンの利用支援に関する研究
 研究課題名（英文） A Study on Utilization of Software Patterns

研究代表者

久保 淳人 (KUBO, Atsuto)
 早稲田大学・理工学術院・助手
 研究者番号：10454026

研究成果の概要：熟練ソフトウェア開発技術者の持つノウハウを文書化する記述形式としてソフトウェアパターンがある。現在、ソフトウェア開発に関する知識がパターンとして蓄積されているが、その効率的な利用方法は提案されていない。本研究では、自然言語処理や情報検索技術を用いたソフトウェアパターン利用支援手法を提案した。本研究の成果はソフトウェア開発者が直面した問題に適したパターンを効率的に発見・適用することを支援し、その結果として効率的なソフトウェア開発の実現が期待される。

交付額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	840,000	0	840,000
2008年度	490,000	147,000	637,000
年度			
年度			
年度			
総計	1,330,000	147,000	1,477,000

研究分野： 総合領域（ソフトウェア開発工学）

科研費の分科・細目：情報学・ソフトウェア

キーワード： ソフトウェア開発工学、ソフトウェアパターン、ソフトウェア再利用、自然言語処理、情報検索

1. 研究開始当初の背景

(背景) ソフトウェア工学における目標は、大規模で高品質なソフトウェアを効率よく開発するための知識や枠組みの構築である。ソフトウェアパターンはソフトウェア工学の一分野であり、ソフトウェア開発における問題と、その解決に至った理由、具体的な解法を合わせて記述したものである。例えば「(問題)ソート機能を実現するとき(理由)ソート対象の順序がランダムに近く(理由)時間

効率を重視するならば(解法)クイックソート・アルゴリズムを用いるとよい」という経験則はパターンとして記述できる。図1のように、熟練技術者が持つ経験則やノウハウをソフトウェアパターンとして記述することで、他の技術者に対して知識の伝達を行うことができる。他の技術者はソフトウェアパターンを読み、直面した問題に対して適用可能か判断を行い、実際に適用して問題を解決する。ソフトウェアパターンは、記述項目がほぼ統一されており、カタログ化

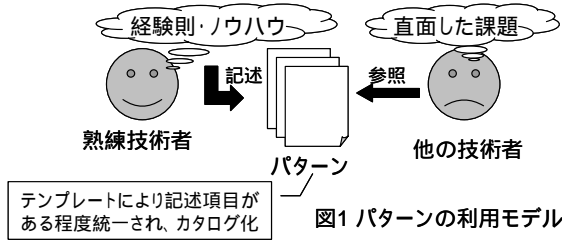


図1 パターンの利用モデル

することで相互比較が容易である。以上の特徴から、ソフトウェアパターンを用いることで技術者の学習・教育やソフトウェア開発組織におけるノウハウの共有を効率的に行うことができる。

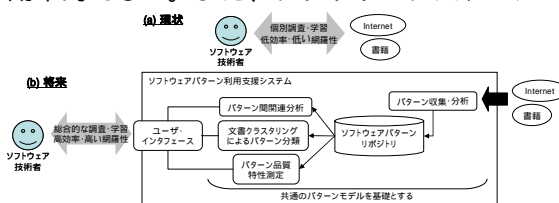
(研究動向) Gammaらによるオブジェクト指向デザインパターンの発表以来、多数のソフトウェアパターンが発表された。ソフトウェアパターンが扱う分野は、要求獲得、分析、設計、実装、テスト、保守、ソフトウェア開発組織など、ソフトウェア開発におけるほぼ全域にわたり、2000年の時点で少なくとも700個のソフトウェアパターンが存在し[Rising]現在も増加している。近年は、各問題領域に特化したソフトウェアパターンが多く発表されている。しかしながら、発表されたソフトウェアパターンを収集・分析・提示する仕組みが存在せず、パターン利用支援は立ち後れている。

[Rising] L. Rising, Pattern Almanac 2000, Addison-Wesley, 2000.

[Zimmer] W. Zimmer, Relationships Between Design Patterns, Pattern Languages of Program Design, Addison-Wesley, pp.345-364, 1996.

2. 研究の目的

ソフトウェアパターン利用支援における問題点として、ソフトウェアパターン情報基盤の不在がある。図2(a)に示すように、現状では発表されたソフトウェアパターンはWorld Wide Web上や書籍に分散して存在するためソフトウェアパターンの効率的な利用が困難である。ソフトウェアパターン利用者は、汎用の検索エンジンや書籍を利用してソフトウェアパターンを調査・学習するため効率的でない。また、ソフトウェアパターン



が集約されていないため、あるソフトウェアパターンと別のソフトウェアパターンの協調・対立といったパターン間の関係性の考察や、ソフトウェアパターンの分類が困難で、ソフトウェアパターンが有効に利用されない。

そこで、図2(b)のような、ソフトウェアパターンに関するソフトウェアパターン情報基盤を作成し、技術者は情報基盤を通してパターンを利用するかたちとすることで、ソフトウェアパターン利用支援を行う。

2. 研究の方法

以下の技術や手順を用いてソフトウェアパターンの収集・分析を行う。

ソフトウェアパターンの共通モデルの作成。ソフトウェアパターンは問題、制約条件、解決といった共通の要素を含む。大多数のソフトウェアパターンに共通して含まれる要素を調査し、ソフトウェアパターンの共通形式を提案する。この成果は、下記リポジトリ構築の基礎となる。

ソフトウェアパターン・リポジトリの構築。ソフトウェアパターンに関する大規模な情報基盤が存在しないため、多数のソフトウェアパターンを共通の書式で蓄積したリポジトリを構築する。

文書クラスタリング手法を用いたソフトウェアパターンの自動分類。文書クラスタリング手法をパターン文書集合に対して適用する。上記共通モデルにおける問題、解決、制約条件等、ソフトウェアパターンの特定の側面に着目した整理・分類が可能になる。

自然言語処理と情報検索技術を用いたパターン間関連の自動分析手法。パターンの増加にしたがってパターン間関連の個数が加速度的に増加するので、パターン間関連に関する考察のための初期の絞り込みとして、文書類似度を応用したパターン間関連の分析手法を提案する。

ソフトウェアパターンの品質測定法。各パターンが扱う問題や解法の抽象度等、種々の品質特性を測定する手法を提案する。

4. 研究成果

研究成果を以下に示す。

ソフトウェアパターンに関する共通モデルを作成した。まず、世界的に著名なソフトウェアパターンの議論の場である国際会議 The conferences on

Pattern Languages of Programming (PLoP) で過去に提案されたソフトウェアパターンの記述形式を調査し、多くのパターン記述の最大公約数的なパターン文書モデルを作成した。パターン文書モデルは以下の文書片の集合である：名前、文脈、問題、制約(フォース)、解法、結果文脈、関連パターン。

パターン文書モデルにしたがってパターンリポジトリを構築した。1つのパターン記述は上記のパターン文書モデルで定義される項目の集合として格納され、項目ごとの検索が可能である。構築されたパターンリポジトリは他の手法の分析対象として利用される。

文書クラスタリング手法をパターン文書に対して適用した。学会発表4および1では、Gammaらのデザインパターン文書についてクラスタリング分析を行い、前者で階層型クラスタリング、後者でk-means法によるクラスタリングを適用した。分類要素であるパターン間の距離関数としては、TF-IDF法による文書ベクトルの余弦値類似度を用いた。階層型クラスタリングにおいては、chaining効果のため部分的に良好な分類結果(たとえば「生成に関するパターン群」)の他は統一性の少ないグループが得られた。k-means法による分類では、生成に関するパターンのグループ、構造に関するパターンのグループといった共通の特徴を持つパターン集合に、多くのパターンが分類されるという結果を得た。

自然言語処理と情報検索技術を用いたパターン間関連の自動分析手法を用いて実際のソフトウェアパターン文書の分析実験を行った。学会発表5、9では、複数の著者によって記述されたセキュリティパターン群に対して自動分析手法を適用した。通常、異なる著者によるソフトウェアパターン間の関連性については分析がなされないが、提案手法を用いることで、異なる著者によるソフトウェアパターン間について、存在する可能性が高い関連候補を自動的に抽出することができた。

ソフトウェアパターンに関する測定法を提案した。投稿論文1および学会発表6では、では、パターン文書中に現れる他パターンへの参照をパターン間のリンクとみなし、Google PageRankを拡張した技法によって収束計算を行うことで各パターンの重要度を算出する手法を提案した。また、学会発表8では、パターン間の半順序関連に注目し、あるパターンが推移的に参照するパターン

数、あるパターンを推移的に参照するパターン数から、パターンが採用する解法の抽象度を測定する手法を提案した。提案手法をGammaらのデザインパターンおよびBuschmannらのアーキテクチャパターンに適用した結果、アーキテクチャパターンの抽象度は高く、またGammaらのデザインパターン群でも他のパターンを使って複合的な解法を提供するパターンの抽象度が高く算出される傾向を確認した。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 1 件)

中山弘之, 久保淳人, 鷲崎弘宜, 深澤良彰: 重要度算出に基づくソフトウェアパターン検索システム, コンピュータソフトウェア, Vol.25, No.2, pp. 114-134 (2008). 査読あり

[学会発表](計 10 件)

久保淳人, 鷲崎弘宜, 深澤良彰: クラスタリング技法を用いたソフトウェアパターン分類, ウィンターワークショップ 2009・イン・宮崎 (2009).

久保淳人, 鷲崎弘宜, 吉岡信和, 井庭 崇, 大久保隆夫: 第15回プログラムのパターンランゲージ会議 (PLoP2008) 参加報告, 情報処理学会ソフトウェア工学研究会 第162回研究会 SE-164 (2009).

久保淳人, 鷲崎弘宜, 深澤良彰: ソフトウェアパターン選択支援の現状と展望, ソフトウェア開発のパターンとアーキテクチャ (ソフトウェアエンジニアリングシンポジウム 2008 併設ワークショップ) (2008).

久保淳人, 鷲崎弘宜, 深澤良彰: クラスタリングを用いたソフトウェアパターンの分類支援手法, ウィンターワークショップ 2008・イン・道後 論文集, SIGSE, pp. 85-86 (2008).

Fernandez, E., Washizaki, H., Yoshioka, N., Kubo, A. and Fukazawa, Y.: Classifying security patterns, The 10th Asia Pacific Web Conference (APWeb2008), Shenyang, China (2008).

Kubo, A., Nakayama, H., Washizaki, H. and Fukazawa, Y.: PatternRank: A

Software Pattern Search System Based on Mutual Reference Importance, 15th Pattern Languages of Programming (PLoP2008), Nashville (2008).

Fukaya, K., Kubo, A., Washizaki, H. and Fukazawa, Y.: Design Pattern Detection Using Source Code of Before Applying Design Patterns, the 1st International Workshop on Software Patterns and Quality (SPAQu'07) (2007).

Kubo, A., Washizaki, H. and Fukazawa, Y.: A Metric for Measuring Abstraction Level of Design Pattern, the 14th Conference on Pattern Languages of Programs (PLoP2007), Monticello, Illinois (2007).

Kubo, A., Washizaki, H. and Fukazawa, Y.: Extracting Relationships among Security Patterns, the 1st International Workshop on Software Patterns and Quality (SPAQu'07) (2007).

深谷和宏, 久保淳人, 鷺崎弘宜, 深澤良彰: パターン適用前の状況を活用したデザインパターン検出, ソフトウェア工学の基礎ワークショップ(FOSE2007), pp. 95--104 (2007).

〔図書〕(計 1 件)

鷺崎弘宜, 丸山勝久, 山本里枝子, 久保淳人: ソフトウェアパターン, 近代科学社 (2007).

〔その他〕

6. 研究組織

(1) 研究代表者

久保 淳人 (KUBO ATSUTO)

早稲田大学・理工学術院・助手

研究者番号: 10454026

(2) 研究分担者

なし

(3) 連携研究者

なし