

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成25年 3月31日現在

機関番号：33917

研究種目：基盤研究（C）

研究期間：2010～2012

課題番号：22500037

研究課題名（和文） 確率的言語モデルを用いたソフトウェア解析の研究

研究課題名（英文） Studies on Software Analysis using Probabilistic Language Models

研究代表者

沢田 篤史（SAWADA ATSUSHI）

南山大学・情報理工学部・教授

研究者番号：40273841

研究成果の概要（和文）：本研究は、ソフトウェア開発者の知的作業に起因して生ずるプロダクトの曖昧性を、確率的要素を含むモデルで表現し、ソフトウェア解析・進化のプロセスを体系的に支援するための方法論とツール基盤を構築した。ソフトウェア進化研究の包括的な調査に基づいて本研究の位置づけを行うとともに、再利用可能なソースコード解析ツール基盤の開発やソフトウェアアーキテクチャ設計判断過程のモデル化を行った。

研究成果の概要（英文）：In this study we have described the intellectual ambiguity within software products as probabilistic software models and developed a foundation of the software analysis and evolution tools which are able to use these models. In order to locate our study in the context of software maintenance, we have carried out a comprehensive literature review in the area of software analysis and evolution. Based on the result of the review, we have developed a reusable foundation for source code analysis tools and also modeled the decision making processes for the software architecture design.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2010年度	1,400,000	420,000	1,820,000
2011年度	1,000,000	300,000	1,300,000
2012年度	900,000	270,000	1,170,000
年度			
年度			
総計	3,300,000	990,000	4,290,000

研究分野：総合領域

科研費の分科・細目：情報学・ソフトウェア

キーワード：ソフトウェア工学，ソフトウェア解析，ソフトウェア進化

1. 研究開始当初の背景

ソフトウェアの開発や保守にかかる行為は、人間の知的作業であり本質的に属人性を内在する。その結果として生成されるソフトウェアプロダクトは曖昧性を持つ。国内外を問わず盛んに研究が行われてきた開発方法論やソフトウェアパターンなどをはじめとする従来のソフトウェア開発支援技術では、ソ

フトウェア開発プロセスから属人性を排し、プロダクトの形式を厳密に規定することでソフトウェアの曖昧性に対処してきた。これらは、ソフトウェアの持つ曖昧性を抑制することで開発の機械的な支援を実現するアプローチであるといえるが、開発するソフトウェアの領域知識やプロダクトの形式、プロセスにおける定型作業の細かな差異に適

応するためには、それぞれの要求に応じた形式化から方法論の構築、ツールの実現をいわば場当たりに繰り返す必要がある。

我々は、ソフトウェア解析や進化の領域におけるこの問題に対処するために、自然言語処理において用いられる、自然言語の曖昧性や冗長性に対処するための確率的情報理論が有効に機能する可能性があるという着想に至った。

2. 研究の目的

本研究の目的は、ソフトウェアプロダクトに内在する曖昧性を、開発者の知的作業に起因して生じる確率的な現象としてとらえ、自然言語処理等で用いられる確率モデルをソフトウェア解析と開発支援に活用する体系的な手法を確立することである。

そのために本研究では、(1)オープンソースソフトウェアを代表とするソフトウェア資産の保守の文脈において、ソフトウェア解析やソフトウェア進化の支援方式について包括的に調査し、各成果がどのように知的作業に起因する曖昧性に対処しているのかを整理すること、(2)ソースコードやソフトウェアモデルの構築時に生じるプロダクトの曖昧性を、確率的要素を含むソフトウェアモデルすなわち設計判断モデルとして表現すること、(3)ソースコード解析のための確率的ソフトウェアモデルを導入可能なツール基盤を構築し、手法の有効性を検証すること、の三項目を重点的な目標とした。

3. 研究の方法

目的として掲げた次の三項目について、それぞれ並行・連携して研究を遂行した。

(1) ソフトウェア進化・解析に関する研究動向の包括的調査

国内外の論文誌や会議録で発表された研究論文を包括的に調査し、ソフトウェア進化、保守、解析の文脈における研究活動の分類を行う。さらに、特にオープンソースやソフトウェアモデルを対象としたソフトウェア解析の分野で行われている種々のアプローチにおける本研究の位置づけを明らかにする。

(2) ソフトウェア解析と設計判断を支援するモデルの構築

ソフトウェア開発において本研究の支援が重要であると考えられるソフトウェアアーキテクチャ設計のプロセスにおいて、その設計判断を支援する意思決定モデルを構築する。

(3) 確率モデルを導入可能なツール基盤の構築

ソフトウェア解析および検査のプロセスを対象として、確率的なモデルを導入可能なツール基盤を体系的に設計する。この基盤に基づいてソースコード検査を行うためのツールを実装し、有効性を評価する。

4. 研究成果

前掲した三つの重点目標について、本研究では、それぞれ次のような成果を得た。

(1) ソフトウェア解析・進化に関する包括的調査

ソフトウェア保守・進化・解析に関し、国内外の論文誌や会議録に掲載された研究論文約150編を対象として文献調査を行った。その結果、これらの研究を分類するために図1に示す三つの観点を新たに提案した(雑誌論文①, 学会発表⑤)。

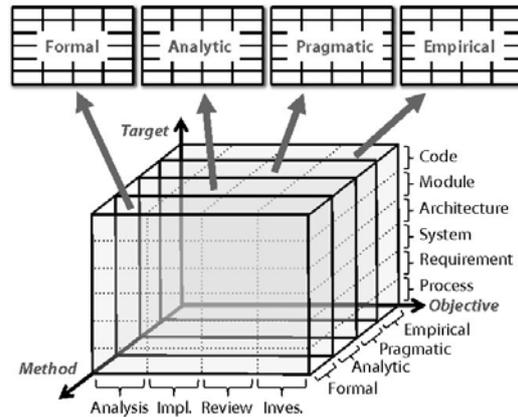


図1 保守・進化・解析研究の分類

すなわち、手法 (Method) の観点から、形式的手法 (Formal)、解析的手法 (Analytic)、実践的手法 (Pragmatic)、実証的手法 (Empirical) というカテゴリを、対象 (Target) の観点から、コード (Code)、モジュール (Module)、アーキテクチャ (Architecture)、システム (System)、要求 (Requirement)、プロセス (Process) というカテゴリを、さらに目的 (Objective) の観点から分析 (Analysis)、実現 (Implementation)、検査 (Review)、解明 (Investigation) というカテゴリをそれぞれ設け、これに基づいて研究の分類を行った。その結果、図2に示す通り、近年の動向として実証的手法を適用する研究が増加しており、またそれら研究の対象として、ソースコードおよびソフトウェアアーキテクチャに注目が集まっている傾向が明らかになった。

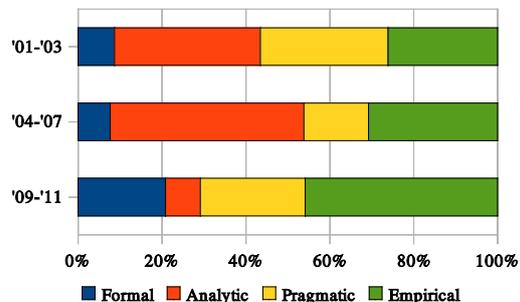


図2 適用手法の動向

この結果から、国内外の研究動向からみて、本研究でも多くの事例から設計判断や解析に関するモデルを構築し、ソースコード解析やソフトウェアアーキテクチャ設計・保守を中心的な対象とするアプローチを選択することが妥当であることが示された。

本調査研究の成果は、本研究の位置づけのみならず、ソフトウェア工学分野の研究動向を示す重要なものであることから、調査範囲をさらに広げるとともに系統的な文献調査手法を適用するなど、今後も継続する予定である。

(2) ソフトウェア解析と設計判断を支援するモデルの構築

ソフトウェア開発において特に知的な判断やトレードオフの評価が求められる、ソフトウェアアーキテクチャの設計および保守を、実践的かつ実証的な手法により支援するために、アーキテクチャ設計マップを提案した(雑誌論文⑨, 学会発表⑨⑩)。ソフトウェアアーキテクチャの設計では、構築されるシステムに求められる非機能要求を充足することが大きな関心事の一つとなる。その設計判断においては、非機能、すなわち品質要求同士の衝突や干渉が起こる場合があり、一定の要求については妥協することも含め、それらの間の調整は設計者の経験によって行われる。アーキテクチャ設計マップでは、主に組込みシステム分野において過去に行ったソフトウェア開発経験とそこから得られた知見(学会発表②③⑥, 図書①)に基づいて、図3に示すメタモデルに従い、品質マップおよび詳細設計マップを提示した。これにより、品質のトレードオフを考慮した設計判断プロセスの支援(図4)とその詳細な記録を可能とした。

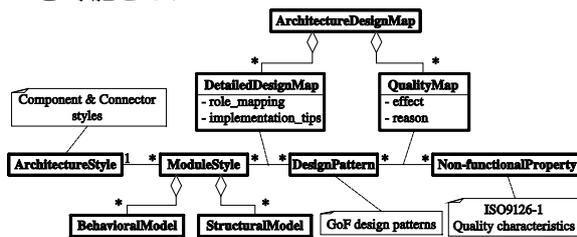


図3 アーキテクチャ設計マップ

このマップのトレードオフ分析に関わる部分へ確率的な判断基準を導入することで、過去の実例や実績を反映した知的な判断を支援することが可能となる。

アーキテクチャ設計マップにおける品質のトレードオフを考慮した設計判断支援の考え方は、アーキテクチャ設計に限ったものではなく、ソースコードの解析や保守(雑誌論文①④⑤⑩), 並行システムやソフトウェアアーキテクチャの検証(学会発表③④⑦⑧⑩), さらにはSOAなどのシステムを対象とした動的構成手法(雑誌論文⑫)の基盤とし

ても有効であると考えており、今後も多くの事例の収集に基づき、マップの対象の拡張と内容の補強を継続的に行う予定である。

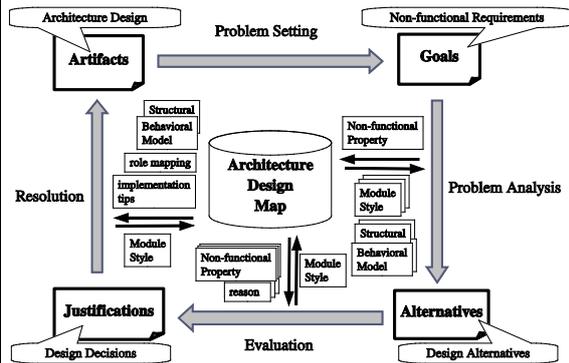


図4 アーキテクチャ設計プロセス

(3) 確率モデルを導入可能なツール基盤の構築

ソースコード解析や検査においては、確率モデルをはじめとする実証的なアプローチが広く行われている。個々の研究成果の有用性は示されているが、実用的見地からそれらを容易に統合可能なツール基盤が求められている。

本研究では、ソースコード解析のためのソフトウェアアーキテクチャを、解析論理の保守性および再利用性に考慮して設計した。このアーキテクチャを前提にソースコード解析のために共通に利用可能なコンポーネントを整備し、ツール基盤として提案するとともに、Java言語を対象としたソースコードインスペクションツールJCI(図5)を実装し、ツール基盤の有効性を評価した(雑誌論文②⑧, 学会発表①)。

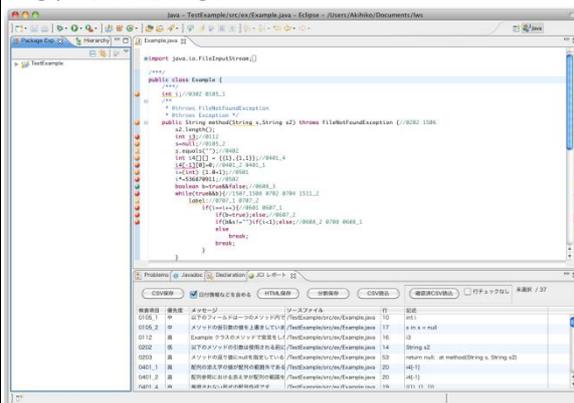


図5 JCI: Java Code Inspector

図6に、本研究で設計したツール基盤のアーキテクチャを示す。

ツール基盤のアーキテクチャは、ソースコード解析のための記号表、構文木、および解析のための論理を実装する解析コンポーネントからなる。アーキテクチャは、それ自身の理解性や保守性を高めるためにGoFデザイン

パターンを適用して設計した。すなわち、構文木には Composite パターンを、解析コンポーネントは Visitor パターンを用いた。

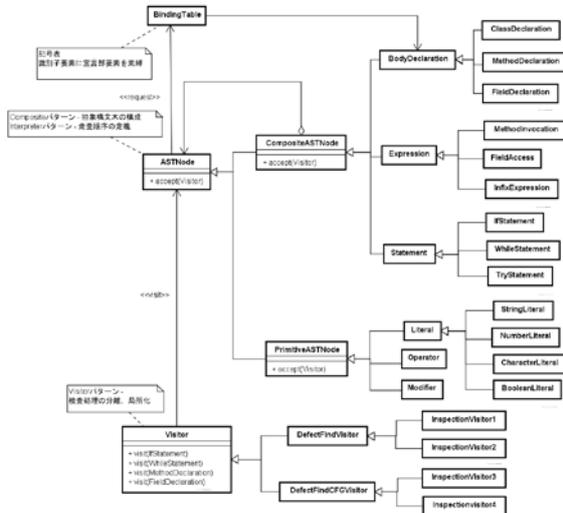


図 6 ツール基盤のアーキテクチャ

このアーキテクチャに基づいて具体的なツールを構築するには、解析コンポーネントとしての Visitor をホットスポットとして拡張し、それぞれの応用に向けた論理を実装すれば良い。

本研究で実装した Java 言語のソースコードインスペクションツール JCI では、検査論理を Visitor として実装した。実装した検査論理は、有限オートマトンによって記述され、それを解釈することでソースコードから欠陥を発見する。この検査論理を記述した有限オートマトンの遷移条件に確率を導入することで、過去の事例に基づく確率的な欠陥検査を実現することができる。

提案したツール基盤のアーキテクチャは、ソースコードのみを対象としたものではなく、形式的に記述されたプロダクトに対し、何らかの文法的処理に基づいてアクションを行う CASE ツールに共通する基盤となりうる。このことから、今後もこの基盤に基づいて各種 CASE ツールを構築するとともに、その改良を行う予定である。

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 12 件)

① C. Wijesiriwardana, G. Ghezzi, E. Giger, A. Sawada, H. Gall, Dependency Based Approach for Software Analysis Web Services Replacement, Proc. the 19th Asia-Pacific Software Engineering Conference, 査読有, 2012, pp. 294-299

DOI: 10.1109/APSEC.2012.130

② 大森隆行, 丸山勝久, 林晋平, 沢田篤史, ソフトウェア進化研究の分類と動向, コンピュータソフトウェア, 査読有, Vol. 29, No. 3, 2012, pp. 3_3-3_28

DOI: 10.11309/jssst.29.3_3

③ 吉田敦, 蜂巢吉成, 沢田篤史, 張漢明, 野呂昌満, 属性付き字句系列に基づくソースコード書き換え支援環境, 情報処理学会論文誌, 査読有, Vol. 53, No. 7, 2012, pp. 1832-1849

<http://id.nii.ac.jp/1001/00083176/>

④ A. Sawada, M. Noro, H. Chang, Y. Hachisu, A. Yoshida, A Design Map for Recording Precise Architecture Decisions, Proc. the 18th Asia-Pacific Software Engineering Conference, 査読有, 2011, pp. 298-305

DOI: 10.1109/APSEC.2011.36

⑤ 沢田篤史, 野呂昌満, 蜂巢吉成, 張漢明, 吉田敦, 長大介, 浦野彰彦, ソースコードインスペクションツールのためのソフトウェアアーキテクチャの設計と進化, コンピュータソフトウェア, 査読有, Vol. 28, No. 4, 2011, pp. 4_241-4_261

DOI: JST. JSTAGE/jssst/28.4_241

⑥ 蜂巢吉成, 吉田敦, 野呂昌満, 沢田篤史, 張漢明, HTML 要素の状態を考慮した CSS の拡張方法の提案, 電子情報通信学会論文誌 D, 査読有, Vol. J94-D, No. 11, 2011, pp. 1931-1934

⑦ 蜂巢吉成, 野呂昌満, 沢田篤史, 張漢明, 吉田敦, コンパイル方式による XQuery 問い合わせプログラム生成方法, ソフトウェア工学の基礎 XVIII — 日本ソフトウェア科学会 FOSE 2011, 査読有, 2011, pp. 21-30

⑧ 曾我展世, 吉田敦, 蜂巢吉成, 沢田篤史, 張漢明, 野呂昌満, 表現の違いを考慮したマクロ逆置換方法の提案, ソフトウェアエンジニアリングシンポジウム 2011 論文集 (情報処理学会シンポジウムシリーズ), 査読有, Vol. 2011, 2011, pp. 1-6 <http://id.nii.ac.jp/1001/00077559/>

⑨ R. Yokomori, H. Siy, N. Yoshida, M. Noro, K. Inoue, Measuring the Effects of Aspect-Oriented Refactoring on Component Relationships: Two Case Studies, Proceedings of the 10th Annual Aspect-Oriented Software Development Conference (AOSD2011), 査読有, 2011, pp. 215-226

DOI: 10.1145/1960275.1960301

⑩ 張漢明, 野呂昌満, 沢田篤史, 蜂巢吉成, 吉田敦, モデル検査を用いた振舞検証の実用化技術に関する考察—網羅性に着目して—, ソフトウェア工学の基礎

XVII—日本ソフトウェア科学会 FOSE2010, 査読有, 2010, pp.107-122

- ⑪ 浦野彰彦, 沢田篤史, 野呂昌満, 蜂巢吉成, 張漢明, 吉田敦, デザインパターンを用いたソースコードインスペクションツールのソフトウェアアーキテクチャ設計, ソフトウェア工学の基礎 XVII—日本ソフトウェア科学会 FOSE2010, 査読有, 2010, pp.15-24
- ⑫ 吉田敦, 蜂巢吉成, 沢田篤史, 張漢明, 野呂昌満, 属性付き字句系列に基づくプログラム書換え支援環境の試作, ソフトウェアエンジニアリング最前線—情報処理学会 SES2010, 査読有, 2010, pp.119-126

[学会発表] (計 10 件)

- ① 張漢明, 野呂昌満, 沢田篤史, 吉田敦, 蜂巢吉成, アーキテクチャ指向開発における形式手法適用に関する考察, 情報処理学会第 28 回組込みシステム研究発表会, 2013 年 3 月 13 日, 対馬市交流センター (長崎県対馬市)
- ② 江坂篤侍, 野呂昌満, 沢田篤史, SOA に基づくシステムのためのアプリケーションプラットフォームのプロダクトライン化に関する研究, 情報処理学会第 179 回ソフトウェア工学研究発表会, 2013 年 3 月 12 日, 化学会館 (東京都千代田区)
- ③ 張漢明, 沢田篤史, フォールトパターンを用いたデバッグ支援, 日本ソフトウェア科学会第 19 回ソフトウェア工学の基礎ワークショップ, 2012 年 12 月 13 日, ゆふいん山水館 (大分県由布市)
- ④ 張漢明, 野呂昌満, 沢田篤史, 横森励士, 吉田敦, 蜂巢吉成, 並行システム記述におけるフォールトパターンに関する考察, 電子情報通信学会ソフトウェアサイエンス研究会, 2012 年 11 月 1 日, 広島市立大学 (広島県広島市)
- ⑤ 吉田信明, 張漢明, 沢田篤史, 中村行宏, ネットワークの振る舞いに依存するシステムへの形式手法の適用の検討, 電子情報通信学会ソフトウェアサイエンス研究会, 2012 年 11 月 1 日, 広島市立大学 (広島県広島市)
- ⑥ 大森隆行, 丸山勝久, 林晋平, 沢田篤史, ソフトウェア進化研究に関する動向調査 - IWPSE シリーズを題材に -, 電子情報通信学会ソフトウェアサイエンス研究会, 2012 年 3 月 14 日, てんぶす那覇 (沖縄県那覇市)
- ⑦ 張漢明, 野呂昌満, 沢田篤史, 吉田敦, 蜂巢吉成, 横森励士, パターンに基づく CSP 記述の検査に関する考察, 電子情報通信学会ソフトウェアサイエンス研究会, 2011 年 10 月 28 日, 北陸先端科学技術大

学院大学 (石川県能美市)

- ⑧ 張漢明, 野呂昌満, 沢田篤史, 蜂巢吉成, 吉田敦, 階層分割に基づく組込みソフトウェアの振舞い検証の支援について, 情報処理学会第 18 回組込みシステム研究発表会, 2010 年 8 月 10 日, はこだて未来大学 (北海道函館市)
- ⑨ 沢田篤史, On the Job Learning ~産学連携による新しいソフトウェア工学教育手法~ (招待講演), 2010 年 6 月 26 日, 慶應義塾大学日吉キャンパス (横浜市港北区)
- ⑩ 浦野彰彦, 沢田篤史, 野呂昌満, 蜂巢吉成, デザインパターンを用いた CDI ツールのアーキテクチャとその実現, 情報処理学会第 168 回ソフトウェア工学研究会 / 第 17 回組込みシステム研究会 合同研究発表会, 2010 年 6 月 2 日, 化学会館 (東京都千代田区)

[図書] (計 1 件)

- ① 沢田篤史, 平山雅之 (編著), 三橋二彩子, 丸山勝久, 小野洋彦, 岸知二, 片山徹郎, 青木利晃, 片岡欣夫, 小笠原秀人 (著), CQ 出版社, 組込みソフトウェア開発技術, 2011, 352

[産業財産権]

○出願状況 (計 0 件)

○取得状況 (計 0 件)

[その他]

6. 研究組織

(1) 研究代表者

沢田 篤史 (SAWADA ATSUSHI)
南山大学・情報理工学部・教授
研究者番号: 40273841

(2) 研究分担者

野呂 昌満 (NORO MASAMI)
南山大学・情報理工学部・教授
研究者番号: 40189452

張 漢明 (CHANG HAN-MYUNG)
南山大学・情報理工学部・准教授
研究者番号: 90329756