

氏名（本籍）	にしもと まさし 西本 匡志（広島県）
学位の種類	博士（情報工学）
学位記番号	甲第 135 号
学位授与年月日	2019 年 9 月 25 日
学位授与の要件	広島市立大学大学院学則第 36 条第 2 項及び広島市立大学学位規程第 3 条第 2 項の規定による
学位論文題目	イベント駆動型アプリケーションの開発を支援するための API 利用パターン情報抽出とその応用に関する研究 Extracting and Utilizing API Usage Patterns to Support Developments of Event-Driven Applications
論文審査委員	主 査 教 授 弘中 哲夫 副 査 教 授 若林 真一 副 査 教 授 井上 智生 副 査 講 師 川端 英之

論文内容の要旨

今日のスマートフォンや Web サービスなどのアプリケーション開発は、ライブラリやフレームワークで提供される Application Programming Interface (API) を利用せずに行うことは難しい。例えば、スマートフォンの地図アプリケーションは、ネットワーク経由での地図情報の取得や GPS による現在位置の取得、画面への各種情報の表示などの機能から構築されている。これらの機能はいくつかの API を組み合わせることによって実装されている。開発者は API を利用することによって、幅広い分野における専門性の高い知識を持たずに、気軽にハードウェアの制御やネットワーク通信、高機能な UI 部品を用いた画面描画などを実装できる。アプリケーション開発では、開発者が適切に API を使いこなすことが求められている。

アプリケーションを構成する個々の機能は API の組み合わせによって実装されるので、効率的なアプリケーション開発のためには、開発者は API の使い方に長けていることが望ましい。しかしながら、API の使い方に関する知識は簡単に得られるとは限らない。例えば、API の全てのメソッドに関する適切な使い方が API のドキュメントに記載されていない場合がある。また、API の仕様が頻繁にアップデートされるので、API のバージョンが異なれば、同じ名前を持つ API のメソッドであっても使用法が異なることがある。これに対し、API の使用法等の情報を得るために、オープンソースリポジトリのマイニングによって、ソースコードで協調して呼び出されている API メソッドの組み合わせをパターンとして自動抽出する方法が注目されている。そして、アプリケーションの開発支援のために抽出された API 利用パターン情報を使うシステムやツールに関する研究が盛んに行われている。例えば、API の使用例を探すためのオープンソースリポジトリに対するコード検索シ

システムやソースコードの文脈を考慮して特定のオブジェクトに対して適切な API のメソッド呼び出しを推薦するツール、既存ソースコードの内容理解を支援するためのツールなどに関する研究開発事例がある。

しかしながら、API 利用パターン情報を用いた開発支援に関する既存の手法やツールは、開発への効率化の要求が高まりつつあるイベント駆動型アプリケーションの開発支援のために十分な機能を有しているとは必ずしもいえない。スマートフォン用のソフトウェアや Web サービスに代表されるイベント駆動型アプリケーションでは、一連の協調して動作する API のメソッド呼び出しが複数のユーザ定義メソッド内に配置され、それぞれが非同期的に実行される。従って、イベント駆動型アプリケーションの開発に有用な情報を得ることを目的とした API 利用パターン情報の抽出には、複数のユーザ定義メソッドを横断的に解析して関連の深い API メソッドの情報を収集することが不可欠である。しかし、既存の API 利用パターン情報の抽出手法では、複数のユーザ定義メソッド内に配置される関連の深い API メソッド同士を API 利用パターン情報として抽出する機能が十分とは言い難い。

本論文では、イベント駆動型アプリケーションの開発支援に適した API 利用パターン情報の抽出手法とその応用について議論する。我々は、ソースコード中の複数のユーザ定義メソッドを横断的に解析して、協調して動作する API をまとめた API 利用パターン情報の抽出を可能にするために、従来手法とは異なる二つのアプローチによるプログラム解析手法を提案する。そのうちの一つは API メソッドの名前や所属クラスの類似性に着目する方法であり、API 設計者の意図を汲んで API メソッド同士の関連付けを試みるものである。もう一つはデータ依存関係に基づいて互いに異なるメソッド中に存在する API 呼び出し同士を関連付ける方法である。いずれも、API メソッド呼び出し系列に基づく既存の抽出手法では扱いきれないイベント駆動型アプリケーションからの情報抽出、および、それを用いた開発支援システムの実現を可能にする。

API 利用パターン情報の表現における最も単純な方法として、API メソッド集合による表現が挙げられる。組み合わせで使用されることが多い API メソッドの集合を予め整理しておき検索可能な形でユーザに提供することはアプリケーション開発支援環境の一機能として有用である。本論文ではまず、イベント駆動型アプリケーションの開発支援に有用な API メソッド集合をソースプログラムから自動抽出する手法を提案し、その有効性について論じる。提案手法の特徴は次の 2 点である。ソースプログラムから API メソッド集合を抽出するにあたり、複数のユーザ定義メソッド中に分散して記述されているメソッド呼び出し同士を関連付けようとする点、また、所属クラスが互いに類似しているメソッドを同一の集合にまとめる方式をとっている点である。提案手法に基づいて収集された API メソッド集合のデータを利用するにあたっては、文書検索で広く利用されている TF-IDF による重み付けや自然言語処理の基本技術である形態素解析を応用した API メソッド集合のインデクシングにより、データベース化した API メソッド集合情報に対しキーワードを用いた問い合わせが可能な API メソッド集合検索システムを設計する。本論文では、Web インターフェースを備えた API メソッド集合検索システムのプロトタイプを実装する。そして、オープンソースリポジトリから収集した大規模データを用いた実験の結果に基づいて、提案手法がアプリケーション開発支援のために有用な情報を提示できることを示す。

API 利用パターン情報の表現方法としては、単なる API メソッド集合情報ではなく、抽出されたプログラムの構造等の情報を添える方法も考えられる。例えば、ある API メソッドが生成したオブジェクトを別の API メソッドが参照している場合のデータ依存関係や、個々の API メソッドがどのようなメソッド内で呼び出されていたかを表す構造に関する情報などは、API 利用パターン情報として有用である。本論文では、API 利用パターン情報のもう一つの表現方法として、単なる API メソッド集合ではなく API メソッドの関係を詳細に把握する表現方法とその抽出手法を提案する。この手法で用いる API 利用パターン情報は、コードスケルトンの構成要素として API メソッドをまとめて扱う方法である。すなわち、プログラム部品を合成するためにそれを利用できるという大きな特徴がある。コードスケルトンを構成する個々の API メソッドが出現していたメソッド名も合わせて API 利用パターン情報として管理しておけば、あるソースコードに対してそのコードスケルトンが表現する記述を半自動的に書き加えることが可能になる。本論文では、アプリケーション開発環境のエディタの拡張機能として、API 利用パターン情報の検索機能に加え、検索して得られた情報に対応するコードスケルトンを編集中のソースコードに自動的に反映させることができるシステムを提案し、その設計やプロトタイプシステムの実装について論じる。本プロトタイプシステムを使い、Google が提供する Android プログラムのサンプル集 (Google Samples) に対するコード復元実験、および、アプリケーション開発に提案システムを用いたユーザスタディにより、提案手法の有効性を示す。

イベント駆動型アプリケーションの開発支援に効果的な API 利用パターン情報は、ソースコードに対するコードリーディングにも応用できる。我々の提案する API 利用パターン情報抽出手法は、ソースコード中に散在する API メソッド呼び出しの中から協働し合う API 同士を関連付けることができる。この能力は、オープンソースリポジトリから頻出パターンを抽出するだけでなく、一つのソースコードに対してそれを構成する諸機能それぞれを分離するために用いることができる。すなわち、手元のソースコードに対して API 利用パターン情報抽出を適用することにより、実装されている機能を列挙し、各機能がどのようなコード記述で実現されているかを切り出して表示することができる。このようなイベント駆動型アプリケーションに対応した API 利用パターン情報抽出手法の実現によりはじめて可能になる。本論文では、手元のソースコードから抽出された機能を要約した視覚表現の一覧を提示して、機能毎にソースコードの記述箇所を強調表示するシステムを提案し、その設計について論じる。そして、Google Samples に含まれる Android アプリケーションに対する提案システムの適用結果を踏まえ、ソースコードからの機能の切り出し方や提案システムの使い勝手などに対して主観的な観点での考察を行い、ソースコード理解の支援に有用性があることを示す。

本論文では、API 利用パターン情報の抽出手法として、所属クラスに基づいて API メソッドを関連付ける方法、および、データ依存関係とプログラムの構造に基づきコードスケルトンの構成要素として API メソッドをまとめて扱う方法を提案し、それぞれプロトタイプシステムを実装し、オープンソースリポジトリのデータを用いて実証的に評価している。いずれの手法も、イベント駆動型アプリケーションからの情報抽出、および、イベント駆動型アプリケーションの開発の支援に有用な情報提示を実現するものである。本論文で提

案している API 利用パターン情報の抽出とその応用法は、従来の方法では実現が難しかったイベント駆動型アプリケーションの開発支援に進展をもたらすものである。

論文審査の結果の要旨

2019年8月9日(金)14時40分から16時10分まで博士学位論文発表会(公聴会)を開催した。申請者が論文内容について発表を行い、聴講者や審査委員が論文内容に関する質疑応答および議論を行った。発表会終了後の16時20分から16時35分まで審査委員会を開催し、論文の可否に関する審議を行った。

本論文では、イベント駆動型アプリケーションの開発支援に活用できる効果的なAPI利用パターン情報の抽出手法とその応用システムを提案し、その有効性を示している。イベント駆動型プログラムは深く関係のあるAPIメソッド同士が複数のユーザ定義メソッドに分かれて記述される。そのため、ユーザ定義メソッド内に閉じて解析するAPI利用パターン情報抽出手法では十分ではない。これに対して、提案手法はソースコード中に含まれる複数のユーザ定義メソッドを横断的に解析してAPI利用パターン情報を抽出している。提案手法ではプログラムの解析にAPIメソッドの所属クラス名の類似性やデータの依存関係などを用いて、既存手法では得られなかった密接な関連を持ったメソッド同士からなるAPI利用パターン情報が抽出できることを示している。また、抽出したAPI利用パターン情報の応用として、(1)キーワードによるAPI利用パターン情報の検索システム、(2)検索システムで得られたものに対応するプログラム記述を開発中のプログラムに自動挿入できる開発支援システム、(3)API利用パターン情報抽出技術を用いた自動要約機能付きソースコード理解支援システムの3つのイベント駆動型アプリケーション開発支援システムを提案し、有効性を示している。

これらの研究成果は、ジャーナル論文2編(すべて第一著者)、査読付き国際会議論文3編(2編は第一著者、残りは第二著者)として公表されている。