

平成 22 年度 情報工学コース卒業研究報告要旨

阿草研究室 研究室	氏 名	藤 谷 隆 宏
卒業研究題目	シーケンシャルパターンマイニングによる JavaScript の代替コード発見	
<p>JavaScript は動的なコンテンツを容易に作成する方法として多くの Web サイトで用いられている。しかし標準化が遅れたこともあり、複数の処理系が提案されている。これらの処理系はそれぞれ利用できる関数やオブジェクトが異なるという問題がある。Web サイトは様々な Web ブラウザを介して閲覧されるため、開発者は各ブラウザの各バージョン毎にその処理系を特定し、想定する全てのブラウザで正しく実行されるように JavaScript を記述する必要がある。この問題をクロスブラウザ問題と言う。</p> <p>クロスブラウザ問題を簡単に解決するために、prototype.js や jQuery などのライブラリが存在するが、ライブラリには多数の互換情報が含まれているため、開発者が必要な互換情報以外にも読み込まなければならず、それにより Web ページの動作速度が遅くなってしまいう可能性がある。またライブラリが必要な互換情報を提供していない場合もあるため、開発者は JavaScript の書籍や、Web 上のサンプルコード等から情報を得る必要がある。</p> <p>この問題を解決するために代替コードと呼ばれる、異なる実行環境でも同等な振舞いをするコードを自動的に発見する先行研究がある。しかし先行研究の手法で発見できるコードは関数の 1:1 関係のコードのみであった。本研究では、先行研究で発見できない代入文を含む代替コードや M:N 関係の代替コードの発見することを目的としている。</p> <p>本研究では先行研究を改良し、シーケンシャルパターンマイニングを用いることで、より多くの代替コードを発見する手法を提案する。提案手法ではまず、Web 上に散在する JavaScript コードを収集し、それぞれのコードを JavaScript の構造検索を可能にするための構文要素をマークアップした XML に変換し、それらを蓄積したソースコードリポジトリを構築する。先行研究では 1:1 関係のトランザクションを抽出しアイテムセットマイニングを適用していたが、M:N 関係のトランザクションの抽出を行うためシーケンシャルパターンマイニングを用いた手法を適用する。</p> <p>さらに本研究では提案手法の有用性を評価する実験を行った。Web 上をクローリングして得たファイルをそれぞれ XML に変換し、ソースコードリポジトリを構築した。それらを解析し評価実験を行った結果、先行研究では検出できなかった代入文を含む代替コードや、M:N 関係の代替コードを検出することに成功した。Web をクローリングして得た 60,572 ファイルから、56,575 個の代替コードを含む if 文を抽出し、それを元に作成した 39,215 トランザクションのデータに対してマイニング手法を適用した。その結果、先行研究では得られなかった 14 個のルールを含めた 20 個のルールを得た。得られたルールを評価することにより、本研究の有用性を確認した。</p>		