

コードクローン研究のあゆみ

肥後 芳樹 大阪大学大学院情報科学研究科



〔受賞論文〕

- ・ 産学連携に基づいたコードクローン可視化手法の改良と実装
- ・ 肥後芳樹, 吉田則裕, 楠本真二, 井上克郎(大阪大学大学院情報科学研究科)
- ・ 情報処理学会論文誌, Vol.48, No.2, pp.811-822 (2007)

このたび、標記の論文で本会論文賞をいただくこととなり非常に驚いております。このような賞を賜るとは夢にも思っておりませんでした。大変名誉なことですので、本会およびこの賞の名に恥じぬよう、今後もよりいっそう研究に精進していきたいと思っております。

コードクローンに対する研究は、この論文を執筆するよりもかなり以前、1999年ごろから取り組んでいました。研究を開始したきっかけは、とあるソフトウェア会社との雑談です。「何十年も保守し続けてきた巨大なソフトウェアの管理が手に負えなくなった、特に、コピーされたモジュールが散在しており、1つの変更がどこに影響を与えているか、どこがコピーされた部分かが把握できておらず困っている」、というものでした。これが元になり、当時博士後期課程に在学していた神谷年洋氏(現、産業技術総合研究所勤務)を中心にソースコードの重複部分(コードクローン)を検出するツールを開発し、また博士前期課程に在学していた植田泰士氏(現、宇宙航空研究開発機構勤務)を中心にコードクローン情報を可視化するツールを開発しました。これらのツールを用いることにより、ソフトウェア内にどのようにコードクローンが分布しているのか、またそれらはどのような特徴を持っているのかが、短時間で分析できるようになりました。

この技術は実用的なものに発展し得ると考え、2002年ごろより、これらのツールを国内外のソフトウェア会社・ソフトウェア開発者に無料で配布し、利用させていただいています。また、セミナーを開催し、ツールの開発者(大学)と利用者(産業界)との意見交換の場を設けています。このような産学連携の結果、検出・可視化手法をより実用的な技術として改良できたと手応えを感じていました。

そのとき偶然にも、筆者自身が産業ソフトウェアを分析する機会をいただくことができました。分析の対象は、情報処理推進機構(IPA)ソフトウェアエンジニアリング

センター(SEC)の先進ソフトウェア開発プロジェクトにおいてベンダ5社が開発していたプローブ情報システムでした(プローブ情報システムとは、センサなどの計測機器が収集した情報をネットワークを通じてセンタに転送し、その情報を分析・蓄積・加工などすることによって種々の有用な情報を提供するシステムです)。全社合計の開発規模は数十万ステップと比較的大きな規模でしたが、改良した検出・可視化手法を用いた分析は非常にスムーズに進み、当時博士後期課程に在学していた筆者と、博士前期課程に在学していた吉田則裕氏の2人による約10時間の分析で大まかにシステムの状態を把握することができました。改良した手法とこの分析結果をまとめたものが標記の論文です。筆者が思う以上の評価が得られたことを非常に嬉しく思います。

最後になりましたが、現在、ソフトウェアの分析手段としてコードクローン検出および分析技術は高い注目を集めています。ソフトウェア工学の著名な会議であるICSE(International Conference on Software Engineering)の2007年度大会では、ソフトウェア保守やテスト、アスペクト指向プログラミングなどと並んで、コードクローン専門のセッションが設けられたことから、その関心の高さを窺うことができます。これからもこの分野の発展に少しでも寄与できるように願いつつ、筆をおきたいと思っております。

(平成20年4月28日受付)

肥後 芳樹(正会員) higo@ist.osaka-u.ac.jp

平成18年大阪大学大学院博士後期課程修了。平成19年同大学院情報科学研究科助教。博士(情報科学)。ソースコード分析、特にコードクローン分析やリファクタリング支援に関する研究に従事。