

プログラミングと性格に関する一考察

高澤 亮平[†] 坂本 一憲[†] 鷺崎 弘宜[†] 深澤 良彰[†]

早稲田大学[†] 国立情報学研究所[‡]

1 はじめに

「ソースコードには性格が表れる」という事が開発者の間では経験則として知られているが、科学的な根拠は見つかっていない。

そこで本論文では、競技プログラミングのイベントの参加者を対象にアンケートを行い、性格分析を行った。また、コンテストに提出されたソースコードに対してメトリクス測定を行い、性格分析の結果との相関を調査した。

本論文におけるリサーチクエスチョンは以下の通りである。

- RQ1. エゴグラムによる性格分析を行うことで、プログラマ特有の傾向が見られるか？
 RQ2. 各個人の性格分析の結果と、成果物としてのソースコードのメトリクス測定の結果との間に相関が見られるか？

2 背景

2.1 エゴグラム

エゴグラムとは、交流分析と呼ばれる心理学理論を用いて、自我状態を定量化する分析手法である [1]。自我状態を「支配性 (Critical Parent, CP)」「寛容性 (Nurturing Parent, NP)」「論理性 (Adult, A)」「奔放性 (Free Child, FC)」「順応性 (Adapted Child, AC)」の 5 つの要素に分類して点数化し、グラフ化する。それぞれの数値の大きさや、各要素の高低で精神状態の判定を行う。

2.2 競技プログラミング

競技プログラミングとは、与えられた問題に対して、正しい出力をするプログラムを構築する速さや正確さなどを競う競技である。競争という概念をプログラミングに持ち込むことで、参加者の意欲を向上させる効果が示されている [2] だけでなく、近年ではリクルーティングにおいて、志望者の技術力を測る目的で利用される場合も多くなっている。

[†]「An empirical study of relationship between personality and programming」

[†] Ryohei Takasawa, Hironori Washizaki, Yoshiaki Fukazawa, Waseda Univ.

[‡] Kazunori Sakamoto, National Institute of Informatics

3 手法

3.1 アンケート調査

2014 年 11 月に開催された競技プログラミングのイベントである CODE FESTIVAL [3] において、参加者 97 人を対象にエゴグラムに基づく性格分析を行った。参加者には質問紙法と呼ばれる手法を用い、50 問からなるアンケートに答えてもらい、その回答を集計し、点数化することでエゴグラムを作成した。

3.2 メトリクス測定

CODE FESTIVAL で用意された問題に提出されたソースコードのうち、使用人数 74 人と最も多かった言語である C++ で記述されたソースコード 3339 ファイルについて、SourceMonitor [4] を用いてメトリクス測定を行った。また、各個人のメトリクスについて集計したものと、エゴグラムの各要素との相関を調査した。測定したメトリクスは、ファイル数、コード行数、コード中のコメント率、平均メソッド行数、平均ネスト深さ、平均複雑度の 6 つである。

4 結果

4.1 性格分析

図 1 に、参加者によるアンケートの結果を集計した箱ひげ図を示す。また、図 2 に日本人の平均的なエゴグラムとされるものの値と、CODE FESTIVAL 参加者の値の平均値との比較を示す。

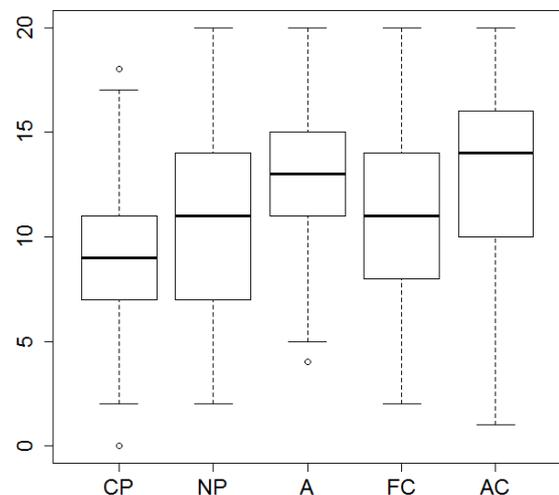


図 1 参加者による回答結果

	CP	NP	A	FC	AC
ファイル数	0.17399	0.058955	0.110106	0.141162	-0.16551
コード行数	-0.05792	-0.00948	-0.05493	0.020701	0.074632
コメント率	0.046214	<u>0.247569</u>	-0.04068	0.131956	0.036603
平均メソッド行数	<u>-0.20837</u>	-0.07712	-0.15308	-0.11588	0.022718
平均ネスト深さ	0.059564	0.126029	0.109228	-0.07766	-0.05983
平均複雑度	-0.00503	0.15164	-0.00705	0.077003	-0.06132

表1 メトリクスとエゴグラムの相関係数

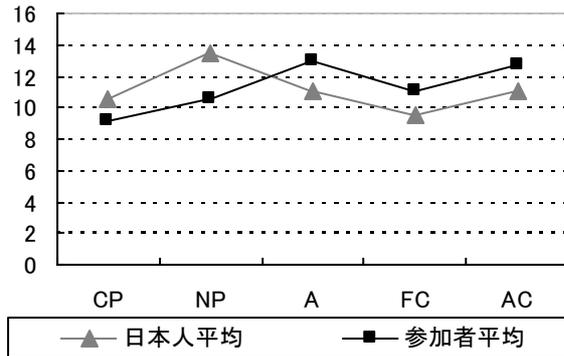


図2 エゴグラムの比較

NP が優位となる日本人平均に比べ、参加者平均ではAが優位となった。Aは理性や論理性を表す値であり、プログラマとしての特性が反映されているのではないかと考える。

他の要素においては、CP, NPの「親らしさ」にあたる要素が低い傾向が見られ、責任感や他者への配慮が少し乏しい傾向にあると思われる一方、FC, ACの「子供らしさ」に当たる要素が高く、協調性があり、好奇心が旺盛である傾向を見ることができた。

4.2 メトリクス測定

表1にメトリクス値とエゴグラムの各要素の相関係数を示す。ほとんどの要素では相関が見られなかったが、表中の下線部に示す2つの点で弱い相関を確認することができた。

1つ目は、CP(支配性)とメソッドの平均行数の弱い負の相関である。CPは理想や責任感を反映する値であり、このことが機能を分割する、メソッドの凝集度を低くする、などの一般に良いとされるコーディング行動を遵守することに繋がっているのではないかと考える。

2つ目は、NP(寛容性)とコードのコメント率の弱い正の相関である。ソース中にコメントを記述しておくことは自分のみならず、他の開発者に良い影響を与えることが示されており[5]、他者への思いやりを反映するNPの高さが反映されているのではないかと考える。

5 まとめ

本論文では、プログラマを対象としたアンケート調査を実施し、性格の分析を行った。また、ソースコードのメトリクス測定を行い、特徴との関係の分析を行った。リサーチクエスションに対する回答は以下の通りである。

- A1. 日本人の平均に対し、論理性が優位となるなどの相違が見られた。
- A2. メソッドの行数と支配性、コメント率と寛容性の間に弱い相関が見られた。

今後の展望としては、プログラミング教育への応用が考えられる。性格分析により個人のコーディングの癖が把握できれば、それに応じて指導を行うことでより高い教育効果が期待できるのではないかと考える。

謝辞

本研究は株式会社リクルートホールディングスの助成を受けたものです。本論文のデータ収集に協力していただいたCODE FESTIVALの参加者の皆様、コンテストシステムを提供していただいたAtCoder社の皆様、運営面で協力いただいた早稲田大学の佐藤靖治氏、小林純一氏、およびスタッフの皆様には感謝いたします。

参考文献

- [1] 杉田峰康, 講座サイコセラピー 交流分析, 日本文化化学社, 1985.
- [2] Shilov, N. V. and Yi, K.: Engaging students with theory ACM collegiate programming contest, Commun. ACM, Vol. 45(2002), pp. 98-101.
- [3] CODE FESTIVAL 2014 - http://recruit-jinji.jp/code_fes2014/
- [4] SourceMonitor - <http://www.campwoodsw.com/sourcemonitor.html>
- [5] Oliver Arafat, Dirk Riehle.: The commenting practice of open source, In Proc. Of the 24th OOPSLA(2009), pp. 857-864.