



「プログラミングコンテストの解答を利用した同一アルゴリズムの別実装の比較」

私は情報科学研究科に進学し、同一アルゴリズムを複数人がプログラムで実装した際の、それぞれのコードの可読性・実行速度といった優位性を評価し定量化する方法を研究したいと考えています。プログラミングコンテストは短時間にある問題を解き、それを解くまでに時間などを競うコンテストです。近年、日本でもプログラムの授業が義務教育において導入が検討され[1]、ITエンジニアの育成の重要性がますます高まっています。プログラミングコンテストを利用した学習は、サービスによってはランキング上位のコードを閲覧することもできるので、プログラミング初学者にとって非常に有益です。しかしながら、プログラミング初学者は優れたコードがどのような特徴を一律に持っているのか、包括的に知ることは出来ません。

先行研究ではプログラミングコンテストの初心者と上級者の間で比較を行い、修正量の調査を行っていました[2]。しかし、同一の問題を解くアルゴリズム実装間の調査は行えていません。上級者が書く優れたコードの中にも、可読性、実行時間など様々な特性を持つものが存在します。そのコードがどういった点で優れているのか定量化できれば、同一のアルゴリズムを解く自身のプログラムと比較が出来ます。

具体的なアプローチとしては、競技プログラミングの同一アルゴリズム(出題された同一の問題)に対する回答のソースコード、実行時間等のデータセットを元に、優位性のあるコードの特徴量はどこで抽出できるかの研究を行います。ある問題を解くために書かれたコード間において実行時間や使用メモリ、プログラムの構造等で比較して、可読性や実行における良性能を持つコードの特徴をそれぞれのプログラミングコンテストの問題毎に調査し、コードを一律に評価できる特徴量の特定を行います。



以上のアプローチを用いることにより、プログラミングコンテストにおけるコードの可読性と実行速度を持つコードの定量化を行えます。また、この定量化した数値を用いることでプログラミング初心者が過去の問題を解いた際に可読性を上げたり、性能を上げる提案をリアルタイムで行えるようにしたいと考えています。そのために、コードを特徴付ける手法の調査を行い、様々な観点から正確に定量化を行える様に研究を進めていきたいと考えています。

[1] 文部科学省：

http://www.mext.go.jp/a_menu/shotou/zyouhou/detail/1375607.htm

[2] 堤 祥吾, プログラミングコンテスト初級者・上級者間におけるソースコード特徴量の比較