

令和 元年 6 月 13 日現在

機関番号：35403

研究種目：基盤研究(C) (一般)

研究期間：2016～2018

課題番号：16K01147

研究課題名(和文) プログラム読解学習のための学習教材の自動生成方式と学習者の技能定量化手法の提案

研究課題名(英文) Proposing an automatic generation method of learning materials for source code reading-based learning and learner's skill quantification method

研究代表者

松本 慎平 (Matsumoto, Shimpei)

広島工業大学・情報学部・准教授

研究者番号：30455183

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：プログラミングが不得手な学習者を効果的に支援するためには、学習活動のセンシングから潜在的な知識構造を推定し、プログラミングに要求される多様な技能要素に照らし合わせながら、学習者に適した教材を提供すべきである。研究代表者らはこれまで、プログラミングが学習経験に強く関係する暗黙的思考技能である点に着目し、学習者の特徴や技能が眼球運動に反映されることを確認したが、眼球運動の履歴を定量的に評価するための手法の構築が望まれている。そこで本研究では、プログラム構造の観点から読解学習教材用のソースコードを自動生成する方式、確率モデルを用いてソースコード読解中の視線移動を定量的にモデル化する手法を提案した。

研究成果の学術的意義や社会的意義

プログラミング読解の様子を眼球運動から調査する研究は既に報告されているが、定性的な観点でのみの評価にとどまっている。特に、プログラミングの習得知識や技能水準を眼球運動の結果から定量的に定義することを試みた研究は十分に行われていない。とりわけ、プログラミングを不得手とする学習者層を対象として、プログラム読解過程に着目し、「不得手の契機となる原因」や「理解を妨げる要因」を追求するにあたって貴重な材料となり得る暗黙的技能を、プログラム構造と眼球運動との関係性から明らかにしようとした研究は見当たらない。この点の報告に、本研究の学術的貢献がある。

研究成果の概要(英文)：To efficiently support novice programming learners feeling programming difficult, clarifying the cause of preventing programming understanding, and developing a new teaching method appropriate for their understanding degree would be necessary. To analyze each student's understanding level, this study focuses on eye movement which is strongly affected by accumulated programming experience, and proposes program structure based analysis method during program reading. The proposed method is inspired by the concept of program slicing. This paper especially examines the relationship between the data dependence relationship in the program consisting only of assignment operations and arithmetic operations and the gaze transition pattern represented by a simple Markov process. From the experimental results, we found meaningful relationship between the two.

研究分野：プログラミング教育

キーワード：眼球運動 データ依存関係 マルコフモデル ソースコード読解 プログラミング 学習支援 プログラムスライシング

1. 研究開始当初の背景

昨今、プログラミングは特に重要な技能として社会的にも広く認識されている。一方で、プログラミングは、計算機システムの動作原理の理解、経験に裏付けられる創造力や表現力に加え、記憶力、数学力、国語力、論理的思考力、構造的把握力など、多岐に渡る基本素養を複雑に組み合わせた技能であると考えられる。しかしながら、プログラミングの技能標準やプログラミングの学習指針は十分に検討されていないため、プログラミング力の養成指針や到達目標、また、プログラミングで最も重視する基本素養は、多くの教育現場では未だ教授者の教育観に強く依存している。学習者のプログラミング力は、成績という観点では、教授者の教育観に基づく一次的尺度で決定される。この場合、数理的論理思考力を強く要求する場合が多く、これを得意とする学習者層は教授法や内容に依存せず優れた評価を獲得できる。反面、その他の学習者は、プログラミングに適正のない層として評価される。したがって、学習者の到達度は、二つの極端な層に分布する傾向にあり、この傾向は、年齢や性別、教育水準の差に関わらず、等しく発生し普遍的なものであると指摘されることも多い。

プログラミングを対象とした研究は、①学習者の分析と特徴の発見、②プログラミング教授法・教材構築、③プログラミング学習用ソフトウェアの開発、の3種類に大別される。①において、学習者が陥りやすい誤りパターンの観測や、学習者の理解を妨げる項目とその対策などが、アンケートや学習履歴などから明らかにされている。しかし、学習者個々の潜在意識や思考過程、暗黙的知識の具体化は、アンケートや反応パターンだけでは十分ではない。②については、オンラインジャッジと関連した練習問題を豊富に用意し、概念の説明に工夫が凝らされ、かつ、質量共に充実した教材が多く開発されている。しかし、専門家の観点から分かりやすさを追求するだけでは、プログラミングを不得手とする学習者にとっては十分ではない。③については、講義支援や自学学習環境構築が主な研究課題とされ、LMS 開発が主流である。③の中では、ビジュアルプログラミングの仕組みを導入し、娯楽的要素に着目した機能開発や直感的な概念提示が行われる場合が多い。しかし、③も②と同様の問題を抱えている。

以上背景のもと、研究代表者らはこれまで、プログラミング読解過程の眼球運動に基づく分析を進めている。また、学習者の到達水準に応じて多様な学習の方向性を与えるプログラミングのスキル標準を構築している。プログラミング読解の様子を眼球運動から調査する研究は既に報告されているが、プログラミングを不得意とする学習者層を対象として、成熟的に学習を進めるための学習支援法・教授法を構築する目的で眼球運動を動員した研究は、研究代表者らの調査の限りでは確認されていない。先行研究により、「代入と順次実行」の基礎的技能に限り研究代表者らの成果の有効性が確認されたが、眼球運動の履歴は定性的な観点でのみ解釈された分析結果であるため、学習者の更なる支援の実現に向けては、眼球運動の履歴を定量的に評価するための手法構築が望まれている。

2. 研究の目的

本研究では、プログラミングが知識と経験の両方の技能を要求する認知活動であることに加えて、学習経験に基づく論理的思考技能を基礎にしている点に着眼する。このことに基づき、蓄積された経験が強く表出される眼球運動を動員し、かつ、プログラム読解中のスライシングの観点から知識や技能の推定を試みるものである。本研究では、(1) データ依存関係、制御依存関係といったプログラム構造の観点から、プログラムの知識・技能とプログラム構造とを対応付けるフレームワークの提案、(2) プログラムを特徴付ける各パラメータに基づきプログラム読解学習教材としてのソースコードを自動生成する機能の実装、(3) プログラミング読解過程中の眼球運動を取得し、視線の推移を隠れマルコフモデルを中心とした確率モデルで記述するための手法の提案、(4) 確率モデルとプログラム構造との構成的差違を定量的に評価し、知識・技能の習得度を定義するための手法の提案、を目的とする。研究の概要を図1に示す。

3. 研究の方法

本論文では、プログラム外の要素によらないプログラムの持つ内的構造が、大学などの教育機関において初学者の最初に躓く要因であると捉え、プログラムに対する理解度の差がソースコードの読解技能の差として視線に現れる点に着眼する。このことに基づき、蓄積された経験が強く表出される視線運動を動員し、かつ、プログラム読解中のスライシングの観点から知識や技能の推定の基礎となる知見、すなわち、プログラムを適切に読解できていれば視線運動にプログラムソースの構造が反映されるということを明らかにするものである。

プログラム読解の従来の視線運動の分析は、時系列データ、あるいは視線運動の総量に基づ

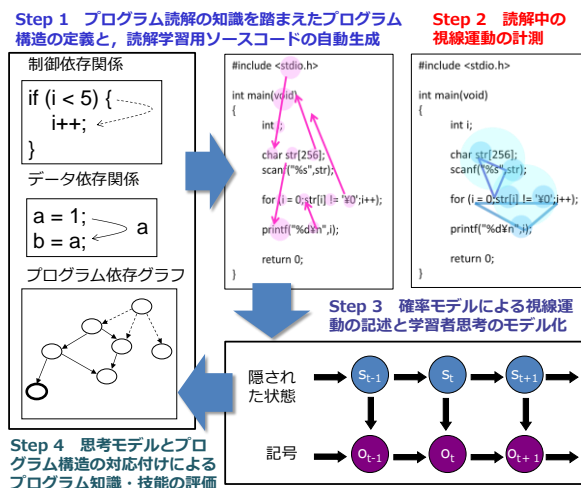


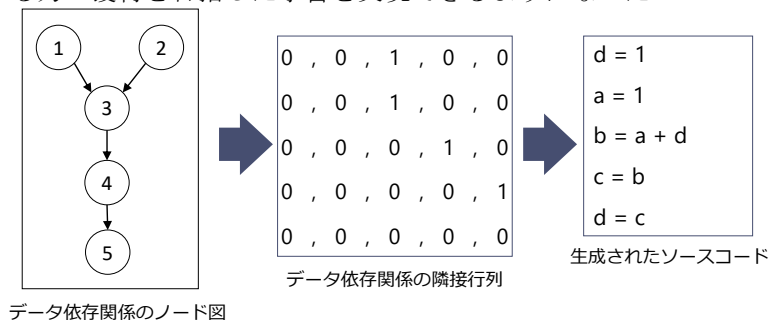
図1 本提案の概要

く分析であった。ただし、被験者によって読解時間が異なっていることや、読解方法に特徴があることから、時系列パターンをそのままの形で利用し、各被験者の理解過程を推定することは容易ではない。そこで、「プログラムを読解するためには、データの依存関係の有無を踏まえてプログラムを読む必要があり、このデータ依存関係を踏まえた読む活動は視線運動に現れる」との仮説を立て、これを本研究では検証する。これは、対象の構造に基づく視線運動の時系列分析の基礎となると考えている。1点を注視する視線の停留運動 fixation に着目し、視線運動の時系列変動を量子化する。そして、量子化されたデータを定量的に解析するための手法として、マルコフモデルを動員する。

被験者がすべての変数の関係を把握するまでの過程を測るために、被験者は提示されたソースコードを読解し、“提示されたソースコードの実行が終了した際、ソースコードに含まれる全ての変数の値はどのようになっているのか”を明確に把握できるまで、プログラム読解中の視線運動を計測する。読解時間に制約を設けず、被験者が納得するまで計測を継続させる。また、プログラム構造の中でもデータ依存関係にのみ焦点を当て、代入文だけで構成される単純な命令文のみを実験課題とする。この代入文は、任意の規則で自動的に生成されたものを利用し、問題作成者の恣意性に一切関与されない形で提供される。以上により、被験者の技量や理解の仕方に強く影響を受けない形で、視線運動の履歴を評価できると考えている。加えて、被験者は理解できるまで何度も読解するため、マルコフモデルを適用可能であると考えている。

視線分析に用いるためのソースコードは、研究代表者らがこれまで開発を進めている読解学習システムにより得ることができる。読解学習システムでは、変数や命令の数など問題生成に関する制約を設定した後、任意の規則に基づいてソースコードを生成することができる。本研究では、従来の読解学習システムを拡張し、データ依存グラフを反復学習できるような機能を追加した。これにより、視線分析をより円滑に遂行できるばかりでなく、データ依存グラフを把握しながら効率良く読解する力の獲得を目指した学習を実現できるようになった。

与えたデータ依存グラフに沿って学習問題が自動作成する機能の概要を図2に示す。生成方法として、まず隣接行列を使用してデータ依存グラフを表す。なお、生成の条件として、変数の再利用率の指定、演算子の制限を行うことができる。



データ依存関係のノード図

図2 データ依存グラフに沿って学習問題を自動作成する機能

4. 研究成果

視線分析の結果、ソースコードの内容を正確に答えることができている学習者においては、データ依存関係の影響が視線に表れその程度を量的に評価可能であることを明らかにすることができた。また、先行研究で報告されてきたソースコード理解過程での視線の動きをソースコードの構造の観点から説明できることを示した。その他、fixationの回数と regressionの回数を調査した結果、今回のような単純なプログラムにおいても同様な傾向を確認することができた。データ依存関係がプログラム理解の効率や度合いに影響を与える点については、従来、被験者に対する主観アンケートやテストなどから得た反応パターンを分析した結果に基づいて述べられてきた。本研究の取り組みは従来のアプローチで明らかにされてきた知見を視線運動から追跡可能であることを示したものであり、視線運動をプログラミング理解の推定に活用する上で基礎となる成果である。本研究で対象とした読解課題は単純なものであり、理解できる学習者の傾向に限定されるが、今回の実験で課題に対するデータ依存関係に基づいた視線傾向を示すことができた。このことから、今後プログラムを適切に理解できない学習者の視線傾向も明らかになれば、その傾向から外れる学習者が依存関係を把握できていないなどの判別や、その学習者に対してどのように読めばよいかなどの支援が可能となる。また、規模を大きくし複雑化させることで、従来の方法論だけでは明らかにすることが困難であった暗黙的な思考過程の解明に貢献可能であると考えられる。スライシングが読解に影響を与えているということは、従来から指摘されてきた。西松ら(西松 他, フォールト位置特定におけるプログラムスライスの実験的評価, 電子情報通信学会論文誌 D-I, J82-D-I(11), pp.1336-1344, 1999)は、着目する変数の値に影響を与え得る文(スライス)をハイライトで表示する機能を有するデバッグを開発し、デバッグ時間の短縮に寄与したことを報告した。この点について、ハイライトは効率的な読解を促し、その結果としてデバッグ時間の短縮に寄与したと考えられることから、本研究で示した結果は西松らの報告を支持するものであり、適切なものと評価できる。

本研究で対象とした読解課題は非常に単純なものであるが、今後学習者の技量や読解教材の質を評価する上で基本となり得る知見であると考えられる。今後、理解できない学習者の読み方が、データ依存関係から逸脱したものになるとわかれば、この知見からプログラムの依存関係を適切に把握して標準的に読めているかといった判断ができるようになると思われる。たとえば、データ依存関係を意識させるような教授法構築や、得られた知見を応用し学習支援システムの開発が期待される。データ依存関係が読解パターンに表出する点から、学習者はデー

タ依存関係を手掛かりとしてソースコードを読解していると考えられる。よって、西松らが示したスライス箇所のハイライト機能を提示することで、プログラム読解を支援する教育・実現化ツールの開発が期待される。また、データ依存関係を読解から追跡することの必要性を確認できたため、プログラミングを不得手とする学習者層を対象として、“不得手の契機となっている原因”や“理解を妨げている要因”を視線追跡から推察可能であると考えられる。たとえば、学習者の理解が不十分な場合、視線運動を取得し、その原因をデータ構造の観点から説明できれば、外的な構造に頼らない、内的な構造にのみ着眼した新たなプログラミング指導法の設計が期待される。

視線分析の結果を踏まえ、データ依存グラフを意識することはコードリーディングを適切かつ効率的に行うために有効かどうかを調査した。実験の結果、データ依存グラフを意識した場合、意識しない場合と比較して、有意に早く読解できるようになることが示された。加えて、主観評価結果からも、データ依存グラフを意識することは読解を容易とし、読解に対する負担を減らし、集中力を高めるといった効果があることが示唆された。以上から、データ依存グラフを意識することは非本質的な認知負荷の抑制につながり、結果として、コードリーディングを適切かつ効率的に行うために有効に働いていた可能性が高いと考えられる。今後、読解学習は特にプログラミングの学習を困難としている初学者により有効であることを示すことが課題となる。すなわち、「プログラミングに苦手意識を感じる学習者はうまくプログラムスライスできていない」といった仮説の検証が課題となる。プログラミングに苦手意識を感じる学習者の場合、通常の被験者よりも改善の幅が大きければ、その仮説はより確かなものであると考えることができる。その他、教材として利用するためのソースコードをより「自然に」するための自動生成方式の採用を検討する。

5. 主な発表論文等

[雑誌論文] (計 7 件)

- ① S. Matsumoto, R. Hanafusa, Y. Hayashi, and T. Hirashima, Analyzing Reading Pattern of Simple C Source Code Consisting of Only Assignment and Arithmetic Operations Based on Data Dependency Relationship by Using Eye Movement, Sakae Yamamoto and Hirohiko Mori (Eds.): Human Interface and the Management of Information. Interaction, Visualization, and Analytics, Springer, pp.545-561 (2018). 【査読有】
- ② 松本慎平, 林雄介, 平嶋宗, 部分間の関係を考えることに焦点を当てたカード操作によるプログラミング学習システムの開発, 電気学会論文誌 C(電子・情報・システム部門誌), Vol.138, No.8, pp.999-1010 (2018). 【査読有】
- ③ 花房亮, 松本慎平, 林雄介, 平嶋宗, 視線運動を用いたプログラム読解パターンへのデータ依存関係に基づく分析-代入演算と算術演算で構成されるプログラムを対象として-, 教育システム情報学会誌, Vol.35, No.2, pp.192-203 (2018). 【査読有】
- ④ S. Matsumoto, S. Yamagishi, and T. Kashima, Difficulty Estimation of Knowledge Elements in Source Codes for Reading-Based Programming Learning by Using Neural Test Theory, Proc. of The First IEEE International Symposium on Artificial Intelligence for ASEAN Development, pp.39-44 (2018). 【査読有】
- ⑤ S. Matsumoto, S. Yamagishi, T. Kashima, Y. Hayashi, and T. Hirashima, A Survey of Analyzing Understanding Process of Programming by Using Eye Tracking, Proceedings of The Twenty-Third International Symposium on Artificial Life and Robotics 2018 (AROB 23rd 2018), GS6-1, pp.116-120 (2018). 【査読有】
- ⑥ K. Okimoto, S. Matsumoto, S. Yamagishi, T. Kashima, Developing a source code reading tutorial system and analyzing its learning log data with multiple classification analysis, Journal of Artificial Life and Robotics, Vol. 22, Issue 2, pp 227-237 (2017). 【査読有】
- ⑦ S. Matsumoto, K. Okimoto, S. Yamagishi, and Tomoko Kashima, Examining Effect and Expectation of Reading Simple Source Codes for Improving Programming Skill, Proceedings of the Fifth Asian Conference on Information Systems, pp.188-191 (2016). 【査読有】

[学会発表] (計 12 件)

- ① 倉本隼, 藤山翔太, 三宝帝斗, 平野拓真, 松本慎平, ソースコード読解におけるチャンク方略に関する基礎研究, 日本経営システム学会イノベーション指向データ分析研究部会 2018 年度第 3 回研究会講演論文集, pp.5-6 (2018). 【査読無】
- ② 森永実子, 松本慎平, 基本的なデータ依存構造を意識したソースコード読解の効果とそれを支援するためのプログラミング学習用システム, 第 75 回情報システム研究会, IS-18-039, pp.13-18 (2018). 【査読無】
- ③ 赤瀬直也, 伊豆田夏帆, 松本慎平, ソースコード読解学習のためのデータ構造に基づいた問題生成に関する研究, 2017 年度教育システム情報学会学生研究発表会講演論文集, 中国地区, p20, pp.205-206 (2018). 【査読無】

- ④ 平野拓真, 伊豆田夏帆, 花房亮, 松本慎平, 林雄介, 平嶋宗, ソースコード読解学習のための視線分析に関する研究, 日本経営システム学会イノベーション指向データ分析研究部会 2017 年度第 2 回研究会講演論文集, pp.25-26 (2018). 【査読無】
- ⑤ 松本慎平, 石井元規, 林雄介, 平嶋宗, 学習者による思考及び操作を中心としたプログラミング学習支援システムの設計と開発, JSAI2017 2017 年度 人工知能学会全国大会(第 31 回)講演論文集, 1F3-3 (2017). 【査読無】
- ⑥ 松本慎平, 林雄介, 平嶋宗, 考えることに主眼を置いたプログラミング教育支援の検討とラーニングアナリティクスの事例, 第 69 回情報システム研究会, IS-17-021, pp107-112 (2017). 【査読無】
- ⑦ 田中貴士, 松本慎平, プログラム読解学習支援システムの開発, 2016 年度教育システム情報学会学生研究発表会(中国地区)講演論文集, P16, pp.215-216 (2017). 【査読無】
- ⑧ 石井元規, 松本慎平, 林雄介, 平嶋宗, プログラミングを苦手とする学習者のための学習支援システムの検討, 第 18 回 IEEE 広島支部学生シンポジウム論文集, 410.情報その他, A4-59, pp.202-205 (2016). 【査読有】
- ⑨ K. Okimoto, S. Matsumoto, S. Yamagishi, T. Kashima, Analyzing Programming Learners and Learning Materials Based on the Learning Activities of Mainly Reading Source Codes, 2016 Annual Conference of Electronics, Information and Systems Society, I.E.E. of Japan, SS4-6, pp.1446-1447 (2016). 【査読無】
- ⑩ 松本慎平, 沖本恒輝, 花房亮, 林雄介, 平嶋宗, プログラム読解戦略とデータ依存関係の基礎的分析, 平成 28 年 電気学会 電子・情報・システム部門大会, 2016 Annual Conference of Electronics, Information and Systems Society, I.E.E. of Japan, TC5-6, pp.121-122 (2016). 【査読無】
- ⑪ 沖本恒輝, 松本慎平, 山岸秀一, 加島智子, 潜在ランク理論によるプログラミング学習者のソースコード読解学習のログ分析と比較, 電気学会研究会資料, 第 68 回情報システム研究会, IS-16-035, pp.43-44 (2016). 【査読無】
- ⑫ 花房亮, 松本慎平, 平嶋宗, 林雄介, 初学者のソースコード読解における視線運動とデータ依存構造の関係分析, 人工知能学会研究会資料 SIG-ALST-B504-02, pp.5-10 (2016). 【査読無】

〔解説記事〕(計 1 件)

- ① 松本慎平, プログラミングが不得意な人でも大丈夫 - 視線運動の定量評価手法で学習支援システム, Fuji Sankei Business i, pp.11 (October 2016)

〔招待講演〕(計 1 件)

- ① 松本慎平, プログラミング学習のための情報システム, 一般社団法人教育システム情報学会中国支部研究発表会特別講演, 広島女学院大学(2017 年 7 月 29 日).

6. 研究組織

(1) 研究分担者

研究分担者氏名：加島 智子
 ローマ字氏名：Tomoko Kashima
 所属研究機関名：近畿大学
 部局名：工学部
 職名：講師
 研究者番号 (8 桁)：30581219

(2) 研究協力者

研究協力者氏名：林 雄介
 ローマ字氏名：Yusuke Hayashi

研究協力者氏名：平嶋 宗
 ローマ字氏名：Tsukasa Hirashima

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。