

令和 5 年 6 月 1 日現在

機関番号：35403

研究種目：基盤研究(C)（一般）

研究期間：2020～2022

課題番号：20K03194

研究課題名（和文）チャンキング技能の獲得を狙いとしたソースコード読解学習システムの開発

研究課題名（英文）Development of a Source Code Reading Learning System to Acquire Chunking Skills

研究代表者

山岸 秀一（Yamagishi, Shuichi）

広島工業大学・情報学部・教授

研究者番号：10609902

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：研究代表者らはこれまで、プログラムスライスの土台であるデータ依存グラフを的確に、かつ効率良く把握することを狙いとして、ソースコード読解に基づいたプログラミング学習システムの開発を進めてきた。また、学習ログを潜在ランク理論により解析し、記述知識単位個々の困難度を段階的に定量化した。実際の講義での利用から統計的に有意に早く読解できるようになるなどその有用性が確認されたが、初学者の支援に留まっていた。そこで本研究では、中級に差し掛かる学習者を支援するため、チャンキング技能の獲得を狙いとしたソースコード読解学習システムを開発する。そして、その有用性を明らかにすることで学術的貢献を果たす。

研究成果の学術的意義や社会的意義

従来システムは、読解時間を短縮できる効率的な学習方法であることが明らかにされた。一方で、従来システムの学習方式は「外的」な知識に一切関与せずプログラムの構造のみに着眼しているため、ソースコードの規模は限定され、それゆえ初学者の支援に留まっている。そこで本研究では、従来よりも規模の大きいC言語のソースコード読解を扱いプログラミング中級程度の学習者の読解学習を支援するため、上級者がソースコード読解の際に適用すると言われているチャンキングに着目する。そして、学術的問い「チャンキング技能が中級に差し掛かるプログラミング学習者のソースコード読解に有効である」ことを明らかにし、学術的貢献を果たす。

研究成果の概要（英文）：Our research team has been working on developing a programming learning support system based on reading source code, with the goal of accurately and efficiently grasping data dependency graphs, which are the foundation of program slices. We have also analyzed learning logs using latent ranking theory, and quantified the difficulty levels of individual descriptive knowledge units in stages. The usefulness of the system has been confirmed through its use in actual lectures, where it has been statistically proven to significantly speed up reading comprehension. However, our support was limited to beginners. In this study, a source code reading learning system aimed at acquiring chunking skills will be developed to support learners approaching the intermediate level. By demonstrating the usefulness of this system, we aim to make an academic contribution.

研究分野：学習工学

キーワード：プログラミング 学習支援 チャンキング技能 ソースコード読解

1. 研究開始当初の背景

プログラミングは高い課題内在性負荷を持つため、その学習の負担は初学者に特に大きい。初学者は認知資源を本質的学習にうまく配分できない場合が多いため、できる限り課題外在性負荷を減らす必要があるとされている。プログラミング力には、経験に裏付けられる創造力や表現力、数学力、国語力、論理的思考力、構造的把握力といった「外的」な技能と、計算機システムの動作原理の理解、言語仕様の知識といった、ソースコード自体の意味を適切に解釈するための「内的」な技能とが内包されていると言われている。学習者の成熟的な学習活動を支援するためには、プログラミングに要求される知識や技能を明確に切り分けただけで、認知資源を本質的学習に容易に配分可能な学習支援システムの導入とそれを用いた実践が求められている。その中でも、プログラムを正確かつ効率的に読むための技能は基本であるため、読解力はプログラミング力の向上に特に重要であると考えられている。

このような背景を踏まえ、研究代表者らはこれまで、プログラミングの「内的」な技能にのみ焦点を当て、プログラムスライスの土台であるデータ依存グラフを的確かつ効率良く把握する力の習得を狙いとして、ソースコード読解に基づいた学習支援システムを開発してきた（以降、従来システム）。具体的には、ソースコード自身の内的な情報であるデータ依存関係のみに基づいた読解学習教材を対象とし、それを用いた学習を支援するための読解学習支援システムの開発を進めてきた。内的情報のみ限定した理由については、初学者を対象とする故、外的な要因を可能な限り排除し課題外在性負荷を減らすためであった。

これまでの研究では、読解学習支援システムの実践的利用を達成できたため、学習ログデータの分析も進められた。学習ログを分析した結果、プログラミング初学者にとって理解を阻害する可能性のある要因として、複合代入演算、インクリメント・デクリメント、整数同士の除算を適切に読解することの困難さ・技量の高さを明らかにし、読解過程での視線の動きからも類似の傾向を確認した。コード読解中の視線運動を分析した結果、視線データは認知過程の推定に有用であることが確認され、同時に「ソースコード読解中、学習者はプログラムスライシングと似た思考を行っている」可能性が導き出された。この知見を踏まえ「データ依存グラフを的確に、かつ効率良く把握することを狙い」として、特定の目的を持たせない形でソースコードを自動生成可能な方式が提案され、学習支援システムの中でその方式が実装された。次に、読解学習の対象を代入文のみから制御文や配列まで拡張させた。まず、学習者適応の教材配信を実現するため、潜在ランク理論によりソースコード記述自体が有する困難度をプログラム依存関係に照合しながら解析した。すなわち、知識単位ごとの知識量の定量化を行うため、潜在ランク理論を用いたソースコード記述の知識単位での段階的困難度定量化を行った。そして、各記述の読解に必要なとされる知識量が段階的に定量化された。以上の知見を活用し、ソースコード読解学習における学習項目・段階のスキル標準を整備した。

従来システムを大学講義に導入した結果、非本質的な認知負荷を相対的に減らしながら、教授者が意図した活動に学習者を集中させることができた。とりわけ、初学者にとってより有効な学習方法であることが示唆された。特に、従来システムは、読解時間の短縮や誤読率の減少が見られ、初学者に有効な学習方法であることが示唆された。一方で、従来システムの学習方式は「外的」な知識に一切関与せずプログラムの構造のみに着眼していたため、その有用性は限定的であった。すなわち、これまでの研究では、プログラミングを不得意とする学習者層を対象とし、「外的」な要因を一切排除したうえでソースコードの記述の知識量を定義し、読解学習の方法を検討していた。それゆえ、ソースコードの規模は限定され、初学者の支援に留まっていた。

中級者以上の読解学習の支援に取り組むため、関連研究の調査を行った。その結果、プログラミング初学者が中・上級者に差し掛かる段階においては「他者が書いたある程度の規模のソースコードを読む作業が求められる」とされている一方、読解に対するツールや方法論による支援が十分ではないことを確認した。次に、学習者がコードの各行についての関係性を正しく理解することを目的とした研究が行われていることを確認した。具体的には、金森らが提案する「プログラミングプロセスの中のプログラムの読解と意味理解」の「プロセス」に位置付けられる読解学習支援に関する様々な研究が試みられていることを確認した。これらは、学習者は知識によって各行を理解することはできるものの、複数行になった際に理解が困難になることに着目した取り組みであった。その中に、機械的にソースコード読解を支援する取り組みとして、学習者が読解すべき文を強調表示する(以下ハイライト表示方式とする)手法を用いた読解作業の支援があった。ハイライト表示方式は課題外在性認知負荷を低減する方法として有効であると考えられるが、ある程度自分の力で要領良く読解できる力も必要であると考えられる。この点については、言語学習においてチャンク方略が広く知られており、その有効性は示唆されている。

以上より、初学者の支援には内的構造の学習が有用である一方、中級に差し掛かる学習者を支援するためには、従来の基盤を活用しつつ、読解技能のひとつを司る外的知識であるチャンクの活用が不可欠であると考えた。そこで本研究では、現在学習者が読解すべき文以外の文を段階的抽象化で定義される「部品」単位で強調し(あるいは伏せて)、チャンキング技能の学習を狙いとした読解学習支援システムの開発を行うこととした。

2. 研究の目的

本研究では、従来よりも規模の大きい C 言語のソースコード読解教材を扱い、プログラミング中級程度の学習者の読解学習を支援するため、上級者がソースコード読解の際に適用されているチャンキングに着目する。本研究の目的は、チャンキング技能の獲得を狙いとしたソースコード読解学習支援システムの開発であり、内容理解の精度向上、速読、バグ発見に対する観点からその有用性を明らかにすることである。そして、学術的問い「チャンキング技能が中級に差し掛かるプログラミング学習者のソースコード読解に有効である」ことを明らかにし、学術的貢献を果たすことを目的とした。

本研究では、これらに対する提案システムの効果の検証だけではなく、以下のそれぞれの点において学術的独創性及び創造性を有するものである。

- ① チャンキング技能の獲得に有効な課題提示方式を検討する際、「注目させたい箇所以外を非表示にする方式」を適用する。課題外在性認知負荷を軽減する手段として、注目させたい箇所を強調表示する方法は一般的である一方、本研究が行う「注目させたい箇所以外を非表示にする方式」は現在までに取り組みられていない手法である。
- ② 田中らによって、「読解の技法は一部の熟練者とその周辺のみに関与されがちである」と指摘されている点の解明を視野に入れ、読解の際に一時的に行われる紙面への簡易記録から知識獲得を試みる。具体的には、読解学習の際に筆記帳を補助的に利用できることとし、この時の記録から構造化技能を評価する。そして「読解力の高い学習者に共通した記録技法の存在」の有無を調査する。読解力の高い学習者に共通した技法が明確になれば、読解を促進するトレース技法として指導法を構築する。読解技能の向上に伴った記録方法の変化などについても分析する。学習課題遂行に直接的に関係する学習活動履歴の分析・評価のみならず、本質的学習に関連した学習活動全体の中に有益な知見が宿っているものと仮定している。
- ③ 本研究では、チャンク単位を段階学習の考え方に基づいて決定する。これは、段階的詳細化プロセスに基づきプログラムの読解と意味理解そのものに着目したものである。
- ④ 研究代表者らの先行研究により、潜在ランク理論を用いてプログラム記述文の困難度が段階定義されている。困難度が定義されたソースコードから「+=」や「++」など記述単位ごとに「知識」を抽出し、各知識の困難度が定義されている。この困難度情報を一次元、段階学習の粒度を二次元として、二次元の学力進捗表の構築を視野に入れ、学習課題を構成する。そして、得られた学力進捗表に基づき学習者適応の学習支援を実現可能な環境を整備することで教育的貢献を果たせるようにする。

本研究では、ある程度読解力を身に付けた学習者を対象とするため、「熟達したプログラマが一連のコードをそのコードが果たす機能として理解している」ことの背景にあるチャンキング技能に着目し、この獲得を学習の狙いとした。これは、プログラミングの理解において、単一の要素を複合して 1 つの大きな要素として理解することを繰り返し、構造的な理解に至ることの重要性を主張した Shneiderman の知見に倣ったものである。

3. 研究の方法

本研究は、教育学的課題と共に、それを検証するための情報工学的な技術課題を含むものである。教育学的課題においては、学習効果の定量評価に加え、認知負荷の計測(チャンキング技能の課題外在性認知負荷低減への効果の分析)、学習者の主観評価を行う。

- ① チャンキング技能の獲得に有効な課題提示方式の検討: 本研究では、チャンキング技能の獲得に有効な課題提示方式を検討する。具体的には、学習者が読解すべき文を強調表示する方法、逆に注目すべき箇所以外の文を非表示とする方法の 2 種類の方法を検証する。ここではそのための 2 種類の読解学習システム構築が技術課題となる。システムは Web アプリケーションとし、表示制御は JavaScript で行う。研究代表者らのこれまでの取り組みにより百問程度の実践的なソースコードが保管されているため、これら全てに対して部品を段階的に定義しチャンク単位を明確にする。その後、研究代表者・協力者らの所属機関で実践的利用を行う。先に述べた二種類の表示方式を A/B テスト形式で適用し、内容理解やバグ発見を題材とした学習課題から、正解率、読解速度の観点に基づき提案法の効果を検証する。
- ② 段階学習の考え方に基づいたチャンク単位決定方式の構築と実装: 古池らは、機能的に有意義な一連のコードを「部品」と定義した。本研究では、この部品をチャンク単位と等価と位置付けている。本研究ではまた、新開らの段階的詳細化プロセスに基づき、部品と同様に多段階のチャンク単位を定義しシステム側でその設定を利活用可能とする。段階的詳細化プロセスとは、目的からプログラムを徐々に細分化していく手法で、アルゴリズム学習の支援を指向したものである。この段階的詳細化プロセスの各水準に倣うことで、学ばせたい粒度に応じたチャンク読解課題が適用可能となる。古池らのシステムは、段階的抽象化プロセスに基づき部品の構築を行わせる学習である。一方、本研究は段階的詳細化プロセスに基づき

プログラムの読解と意味理解自体に着目している。読解と意味理解は学習において有意義であるとした点で古池らの考えと共通しているため、本研究のアプローチは適切であると考えられる。

- ③ 可読性の定量評価：認知的負荷を考慮したソースコード記述の必要性が指摘されているが、提案システムを用いればソースコードの可読性を定量的に明らかにできると考えられる。すなわち、リーダブルなソースコードはチャンク単位が明瞭であると言えるので、目的が同一で記述が異なる複数のソースコードを提案システムで提示し理解に至るまでの時間を測定することで、ソースコードの可読性を量的に定義できると考えられる。提案システムのこのような副次的な活用法についても検討し、更なる学術的貢献を試みる。

4. 研究成果

チャンク単位でのコード読解技能を習得するための学習支援システムを Google Code Prettify を用いて開発した。システムの外観を図 1 に示す。なお、動作環境は Google Chrome、仕様言語は html, css, JavaScript, 使用したオープンソースは Google Code Prettify である。

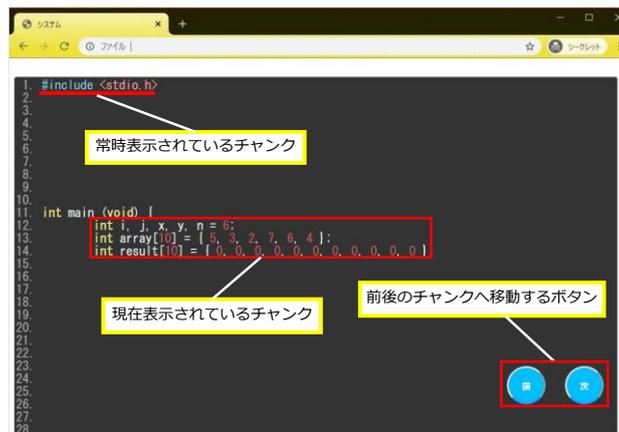


図 1 提案システムの外観

提案手法を用いたシステムは、コードをチャンク単位で区切り、学習者に現在学習者が読解すべきチャンクを順にひとつずつ表示する。チャンクを順に表示する機能については、学習者が次へボタンを押すことで次のチャンクを表示し、かつ、現在表示されていたチャンクを非表示にする。これにより、表示される情報量を必要最低限にすることで、学習者の認知的負荷を軽減することを試みている。

チャンク単位は教授者により決定される。Wang らにより、空白行を用いてコードを意味のあるブロックで分けることを自動化するシステムが開発されている。同様のコードに対して、教授者が与えた空白行によるコードのセグメンテーションは模範と位置付けられているため、彼らの主張に倣い、注目すべきチャンクを順に表示するように設計した。なお、読み込んだライブラリと main 関数の開始・終了のカーリーブラケットは常時表示するようにしている。読解対象箇所と main 関数との関係を把握できるようにすることで、非表示に起因する負担を抑制できるように配慮している。

従来手法を用いたシステム、すなわち、ハイライト表示による読解学習システムは、学習者が着目(選択)する任意の一行に対し、依存関係のある箇所(行)をハイライト表示する。依存関係のある命令とは、入力に依存せず静的スライディングで依存関係がみられた全ての命令を意味する。従来手法を用いたシステムの外観を図 2 に示す。

提案システムは HTML5 に準拠した Web ブラウザで動作する。学習環境をブラウザ上にすることで、学習者に環境構築の負担の軽減を狙っている。学習者は、注目すべきチャンク単位を前後させるための二つボタン操作のみ可能である。提案システムは読解技能の学習を支援するものであるため、学習者をコードの読解に集中させ、課題外在性認知負荷の発生を可能な限り軽減するため簡素なデザインと仕様になっている。

事前に用意されたコードを読解学習する際、提案方式とハイライト方式で学習効果の比較実験を行い、提案法の有用性を確かめる。本研究では、情報学を専攻し中級に差し掛かる段階のプログラミング学習者 20 名を被験者に採用した。C 言語で書かれたコードの読解力を問う事前試験の点数と読解時間に基づき、1 点あたりの秒数を「効率性」と定義し、その値が均一になるよう被験者を実験群と統制群に分け、統制群が従来のハイライト表示方式、実験群が提案手法を用いて本システムを利用した学習を行い、それぞれの方式で最大 45 分の学習を行った。

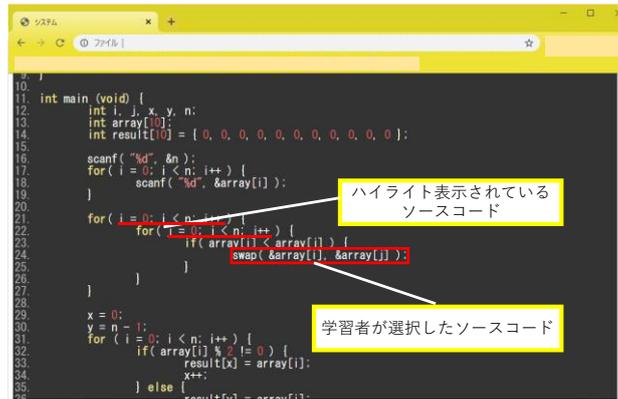


図2 統制群の実験に用いたハイライト表示による読解学習システム

学習後、効率性の観点で学習効果を評価した。試験はコードを読解しある入力を与えられた時のある変数の値を問う形式とした。プレテスト、ポストテストの内容は、C言語のコードを三つ提示し、それぞれに対して三問出題した。読解速度の計測のため、それぞれの問いに対して被験者の解答時間を計測した。本研究では、内容理解に対する読解速度を計測するため、コードを読解し、それに対応した解答を行うまでの時間を計測した。特定の入力に対しての出力を問う問題、プログラムが動いている最中の指定した変数が特定の値になった時のある変数の値を問う問題、出題したコードが行っている動作を与えた選択肢から選ばせる問題とした。

事前試験及び事後試験共に効率性の算出の際にゼロ除算を回避するため最低得点を1とした。事後試験終了後にアンケートを行った。アンケートは、提案手法の有用性を問う項目、提案システムの有用性を問う項目、認知負荷を問う項目とした。提案システムの有用性を問う項目については富士通のウェブユーザビリティ評価スケールを利用した。

表1 実験結果

	事前試験	事後試験
実験群	877	551
統制群	884	534

実験結果を表1に示す。両群共に有意 ($p < .01$, 両側) に学習効果が示された。実験群では提示される情報量が統制群よりも少なく、情報量といった観点では両群は不平等な条件であるにも関わらず、両群は同等の学習効果が示された。提案法は学習者の内容理解を促進するものでない可能性があるが、この点で門田らの知見と一致が見られる。一方、実験群の回答時間の平均は統制群のそれよりも統計的に有意に短かったため、門田らの知見と同様に読解速度の向上に役立つ可能性が示唆された。アンケートからは、提案システムのUIや仕様に関して問題ないことが確認された。使いやすさやプレテストとポストテストの難易度に関して変わった傾向などは見られなかった。これにより、本研究の提案手法を用いた本システムを用いることで、学習者の読解速度の向上に有用である可能性が示唆された。次に、認知負荷アンケートの結果から、提案法は従来法よりも統計的に有意に課題外在性負荷を低減可能な方式であることが確認された。主観評価からも、学習に集中して取り組んでいることが示唆された。このことから、提案法の有用性を明らかにできた。

5. まとめ

本研究では、プログラミング中級程度の学習者の読解学習を支援するため、チャンク単位でのコード読解技能の必要性を明らかにし、そのための学習支援システムの開発を行った。特に、内容理解に費やす速度の向上を狙いとしてチャンキング技能の獲得を狙いとしたコード読解学習支援システムを提案し、システム開発を達成した。提案システムを用いて学習者の読解力向上を確かめる実験を行った結果、チャンク単位でのコード読解は内容理解及び読解速度を促進可能な有用な技能であることが示唆された。これは門田らの主張を拡張し英文読解とコード読解との相違を明らかにした新たな知見であり、意義ある結果と言える。今後、提案システムを用いたコードの可読性の定量評価などに取り組む予定である。

5. 主な発表論文等

〔雑誌論文〕 計6件（うち査読付論文 6件 / うち国際共著 0件 / うちオープンアクセス 2件）

1. 著者名 Shimpei Matsumoto, Shuichi Yamagishi, Kosuke Kaida	4. 巻 Vol.8, No.1
2. 論文標題 Improving User Experience of C Programming Language Learning System for Beginners in Error Correction Learning	5. 発行年 2022年
3. 雑誌名 Information Engineering Express	6. 最初と最後の頁 1-10
掲載論文のDOI (デジタルオブジェクト識別子) 10.52731/iee.v8.i1.527	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Shimpei Matsumoto, Akifumi Ohno, Kosuke Kaida	4. 巻 Vol.8, No.1
2. 論文標題 Evaluating the Usefulness of C Language Learning Support System as a Learning Analytics Tool	5. 発行年 2022年
3. 雑誌名 Information Engineering Express	6. 最初と最後の頁 1-12
掲載論文のDOI (デジタルオブジェクト識別子) 10.52731/iee.v8.i1.622	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Tomoya Iwamoto, Shimpei Matsumoto, Shuichi Yamagishi, Tomoko kashima	4. 巻 1
2. 論文標題 Pre- and Post-survey of the Achievement Result of Novice Programming Learners - On the Basis of the Scores of Puzzle-Like Programming Game and Exams After Learning the Basic of Programming	5. 発行年 2020年
3. 雑誌名 Transactions on Engineering Technologies	6. 最初と最後の頁 130-142
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 松本慎平, 大下昌紀, 買田康介	4. 巻 Vol.140, No.9
2. 論文標題 C言語初学者及びその教授者のためのサーバ・クライアントに基づくプログラミング学習支援システムの開発	5. 発行年 2020年
3. 雑誌名 電気学会論文誌C(電子・情報・システム部門誌)	6. 最初と最後の頁 1096-1109
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 加島智子, 松本慎平	4. 巻 Vol.2, No.1
2. 論文標題 初学者向けCSアンプラグドを用いたプログラミング教育の実践と評価	5. 発行年 2020年
3. 雑誌名 情報教育	6. 最初と最後の頁 1-8
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 松本慎平, 加島智子, 山岸秀一	4. 巻 Vol.2, No.1
2. 論文標題 プログラミング学習前に行われたプログラミングゲームの理解度とその学習後の到達度との関係分析 - アルゴリズムの利用とその学習データの活用を通じて -	5. 発行年 2020年
3. 雑誌名 情報教育	6. 最初と最後の頁 18-27
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計1件 (うち招待講演 0件 / うち国際学会 0件)

1. 発表者名 前田暉正, 松本慎平
2. 発表標題 プログラミング読解学習におけるリファクタリング原則の適用可能性調査 メソッド分割の分析
3. 学会等名 第87回情報システム研究会
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	松本 慎平 (Matsumoto Shimpei) (30455183)	広島工業大学・情報学部・教授 (35403)	

6. 研究組織（つづき）

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分 担 者	加島 智子 (Kashima Tomoko) (30581219)	近畿大学・工学部・准教授 (34419)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関