

アニメーションを活用した 文字列検索アルゴリズム学習のための課題分析

下川 輝^{*1}, 仲林 清^{*2}

^{*1} 千葉工業大学大学院情報科学研究科, ^{*2} 千葉工業大学情報科学部

Learning Task Analysis for String Matching Algorithm Using Animation

Hikaru Shimokawa^{*1}, Kiyoshi Nakabayashi^{*2}

^{*1}Information Science Research of Graduate School of Chiba Institute of Technology

^{*2}Information Science of Chiba Institute of Technology

アルゴリズムアニメーションはアルゴリズムの動的な振舞いを学習者に伝えやすいというメリットがある。しかし、抽象的な概念である時間計算量はアニメーションのみで理解させることは困難な場合がある。本研究ではボイヤー・ムーア法などの文字列検索アルゴリズムを対象としアニメーションの開発を行い、さらに理解を促すための学習課題を取り入れた。今回の報告では、前回報告した時間計算量の理解のつまずきを取り除くために、文字列検索アルゴリズムごとに再度課題階層分析を行い、学習内容や流れを新たに設計した。

キーワード: アルゴリズムアニメーション, アルゴリズム学習, 特徴理解, 文字列検索

1. はじめに

アルゴリズムアニメーションとは、データ構造などを視覚的に表現して、それらを動的に変化させることにより、アルゴリズムの動作を直感的に学習者に伝えようとするものである。アニメーションのメリットとして教科書中の文章による解説や静的な図形による解説では理解することが難しいアルゴリズムの動的な振舞いを学習者に伝えやすい点が挙げられる⁽¹⁾。

しかし、抽象的な概念である時間計算量をアニメーションのみで理解させることは困難な場合があると考えられる。本研究ではボイヤー・ムーア法（以下 BM 法）を含めた 3 種類の文字列検索アルゴリズムを対象として、アルゴリズムアニメーションを開発し、さらに理解を促す学習課題を取り入れる。今回の報告では、前回の報告⁽²⁾を踏まえ、学習者の時間計算量に対するつまずきを取り除くために、再度課題階層分析を行い、新たに学習内容や流れの設計を行った。以下、第 2 章では前回の報告⁽²⁾の概要について示し、第 3 章では今

回実施した課題階層分析について述べる。第 4 章では追加したアニメーション部分について述べ、第 5 章で今後の予定について述べる。

2. 前回の報告

本章では対象とした文字列検索アルゴリズムの説明を行い、前回設計した課題階層分析図及び学習課題について述べる。また前回の実験結果および課題について述べる。

2.1 対象とした文字列検索アルゴリズム

2.1.1 線形探索

線形探索⁽³⁾とは対象とするテキストとパターンを先頭から順に比較していく探索方法である。線形探索の場合、最も早いときは 1 回の比較で見つかり、最悪の場合にはデータ数だけ比較しなければならない。したがってデータ数を M としたとき最大時間計算量は $O(M)$ となる。図 1 に線形探索の照合例を示す。

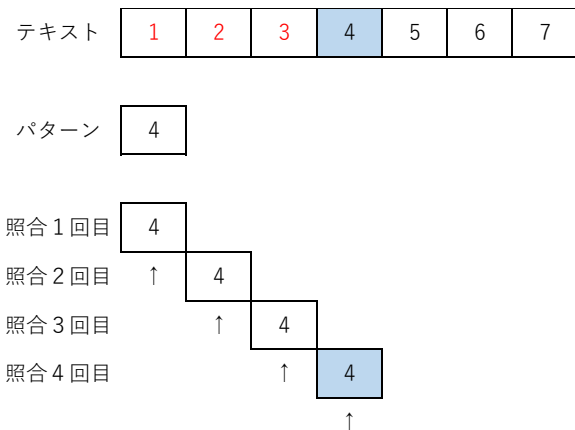


図1 線形探索の照合例

2.1.2 力任せ法

力任せ法^③はテキストとパターンを先頭から順に比較するアルゴリズムである。パターンの終わりに到達する前に、テキストとパターンの対応する文字が一致していなかった場合、パターンを1つ右にずらし、パターンの先頭から同様のことを繰り返し、比較を進めるアルゴリズムである。最悪の場合はテキストの長さ×パターンの長さだけ比較しなければならない。したがってテキスト長を M およびパターン長を N としたとき最大時間計算量は $O(MN)$ となる。図2に力任せ法の照合例を示す。

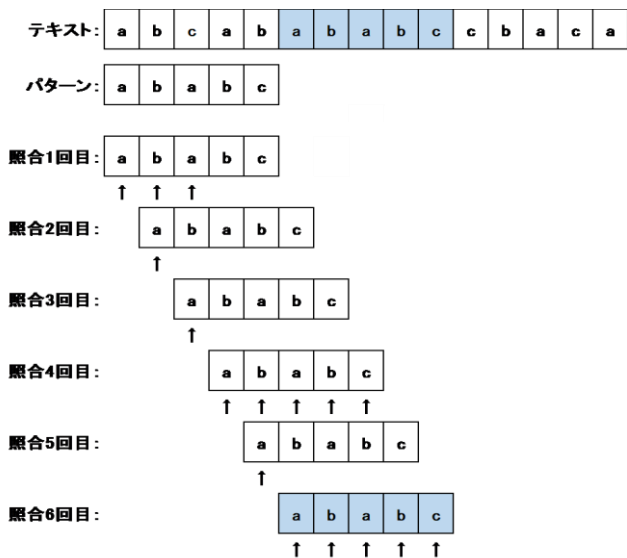


図2 力任せ法の照合例

2.1.3 クヌース・モリス・プラット法

クヌース・モリス・プラット法^③ (以下 **KMP 法**) とは力任せ法と同様に先頭から順に比較を行うが、ずらし表をあらかじめ作成し、それを用いて比較の再開位置を決めるアルゴリズムである。ずらし表はパターン同士を比較し、不一致するまでに一致した回数を用

いて作成される。図3に **KMP 法** の照合例を示す。図3の照合の場合、照合3回目において5文字目で不一致が起こる。ずらし表を参照すると5文字目の **C** で不一致した場合は次の比較再開位置を3文字目からにすることができる。(1,2文字目の **a,b** は一致することが分かるため、3文字目から再開できる)

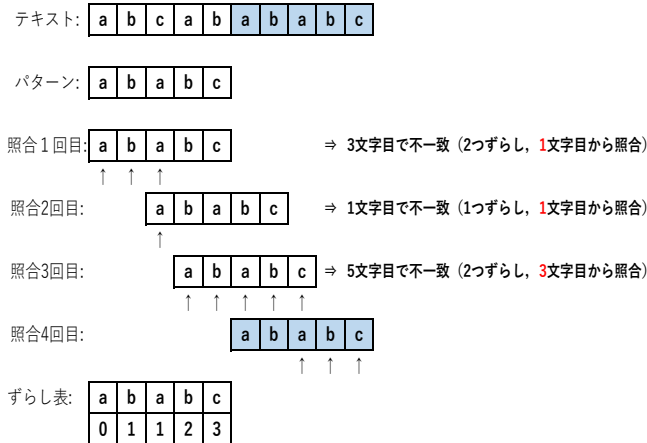


図3 KMP 法の照合例

2.1.4 BM 法

BM 法^③は力任せ法や **KMP 法** がパターンの先頭から比較を始めるのに対し、パターンの末尾から比較を行うアルゴリズムである。また **KMP 法** と同様にずらし表をあらかじめ作成しておき比較を行う。ずらし表はパターンに含まれている文字の末尾からの最短距離を用いて作成される。例えば「ababc」というパターンの場合、**a** の末尾からの最短距離は2であるため、ずらす量は2になる。

BM 法 の特徴である末尾から比較することの利点として、実際の文書処理の場合、文字の種類が多いため、パターンに現れる文字よりも現れない文字のほうが多くなる。つまりテキストとパターンの不一致になる確率が高いことになり、さらに不一致となったテキストの文字がパターンに含まれない確率も高いことになる。したがって不一致となった場合、パターンの長さに近い文字数分だけ照合位置を移動することができるため、力任せ法や **KMP 法** よりも実用的である。また平均時間計算量はテキスト長を M およびパターン長を N とすると $O(M/N)$ となる。最悪の場合はテキストとパターンの一致する確率が高い場合であり、テキストの長さ×パターンの長さだけ比較しなければならない。したがって最大時間計算量は $O(MN)$ となる。図4に **BM 法** の照合例を示す。図4の場合、照合3回目でテキス

トの a とパターン c で不一致が起こる。そこで次の照合ではテキストの不一致した文字がずらし表に含まれている場合は、その数だけ右にパターンをずらすことができる。つまり照合 4 回目ではパターンを 2 文字分ずらして照合を行うことができる。

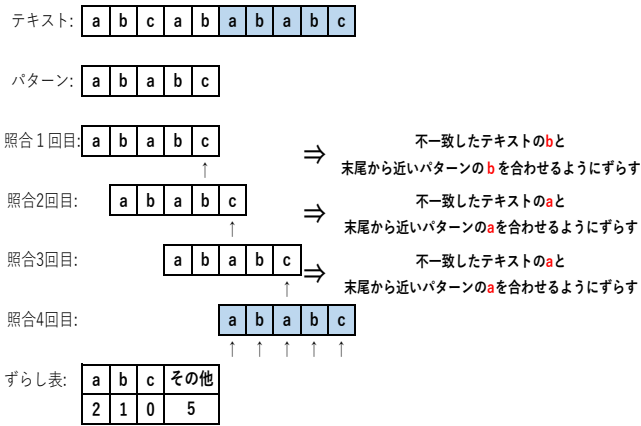


図 4 BM 法の照合例

2.2 前回の学習課題

2.2.1 他の文字列検索アルゴリズムとの比較

BM 法の他に力任せ法, KMP 法のアニメーションを実装した。これらのアルゴリズムと BM 法との違いはパターンの先頭から照合することや不一致情報を利用しないことである。学習者は指定された文字列に対して、それぞれの文字列検索アルゴリズムのアニメーションを確認することで BM 法の特徴を理解しやすくなると考えた。

2.2.2 特定の文字列の動作確認

BM 法の最大時間計算量, 平均的な時間計算量を理解させるためにそれらが顕著にあらわれる文字列を用意し, 学習者にアニメーションを確認することを促した。学習者はこの文字列のアニメーションを確認することによって BM 法の特徴や時間計算量の理解がしやすくなると考えた。

2.3 前回の課題階層分析について

2.3.1 課題階層分析法

課題階層分析法とは, 学習目標を達成するために必要な要素, 知識にどのようなものがあるかを明らかにし, どの順序で学ぶのかを明確化する方法をいう⁽⁴⁾。学習目標とは, 学習者が習得しようとしている知識, 技能のことである。課題階層分析は, もっとも上位にある学習目標から, その目標がなにかを検討すること

を繰り返して分析を行う。そのため, 学習目標より下位にある目標はすべて学習目標を習得するために必要な前提条件となる。洗い出した学習目標を階層に分けて配置した図式を課題階層分析図という。学習者は階層分析図に沿って, もっとも下位に位置する目標から, 学習目標に向かって学習を行う。

2.3.2 課題階層分析図と学習の流れ

表 1 に理解させる BM 法の特徴⁽³⁾を示す。本研究では表 1 で示した特徴の理解度の向上を目標とする。BM 法の動作や仕組みなどの知識と時間計算量の 2 つの要素に分類し, 課題階層分析を行った。設計した課題階層分析図 5 に示す。この課題階層分析図をもとに学習課題の流れを設計した。図 5 の赤枠は学習者が一度に学習する内容である。学習者はまずアルゴリズムの時間計算量やオーダ記法などの基本的事項を線形探索法から学習する。その後, 各アルゴリズムの動作や時間計算量を学習することになる。各アルゴリズムの学習では 2.2.1 項や 2.2.2 項で述べた学習課題を取り入れる。学習者は上位のアルゴリズムを学習する際, 下位のアルゴリズムと比較して動作を見ることになるため, 上位のアルゴリズムの動作が理解しやすくなると考えた。また, それぞれのアルゴリズムの特徴が顕著にあらわれる文字列を用意し, 学習者はその文字列のアニメーションを確認することで動作や時間計算量が学習しやすくなると考えた。表 2 に用意した文字列を示す。例えば「abc1abc2abc3abc4abc5」「abc4」の文字列はパターンの末尾で不一致が起こるため, BM 法にとって効率の良い文字列となる。またパターンの長さだけ比較位置が進むため平均時間計算量である $O(M/N)$ の理解を促すことができると考えた。

表 1 理解させる BM 法の特徴

パターンの末尾から照合
不一致情報をもとに, 何文字ずらすかを決定 実際の文書処理において, 他の文字列検索アルゴリズムよりも実用的である。
最大時間計算量 $O(MN)$ 平均時間計算量 $O(M/N)$ M:テキストの長さ, N:パターンの長さ

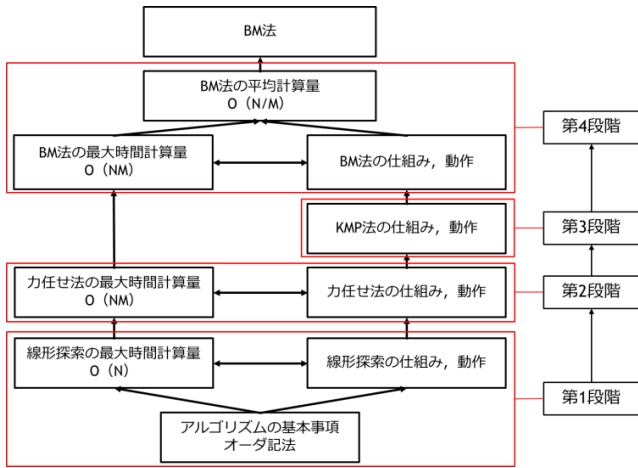


図5 BM法の理解の課題階層分析図

表2 用意した文字列

テキスト	パターン
abcabababc	ababc
123456789	56
abc1abc2abc3abc4abc5	abc4

2.4 前回のアニメーション教材について

本研究で開発したアニメーション教材は Web ブラウザで動作するものである。また学習者の入力文字列などを記録するためにデータベースと連携を行った。アニメーション教材の文字列検索部分を図6に、アニメーションの結果の例を図7に示す。アニメーションは「進む」ボタンを押すごとに処理が進み、テキストとパターンが一致している場合はテキストが緑色に変化し、不一致の場合は赤色に変化し、次の照合位置にパターンをずらす動作を行う。またずらし表部分を図8に示す。ずらし表部分では選択したアルゴリズムに応じて、ずらし表のアニメーションがボタンを押すごとに動作し、ずらし表の作成過程を確認することができる。

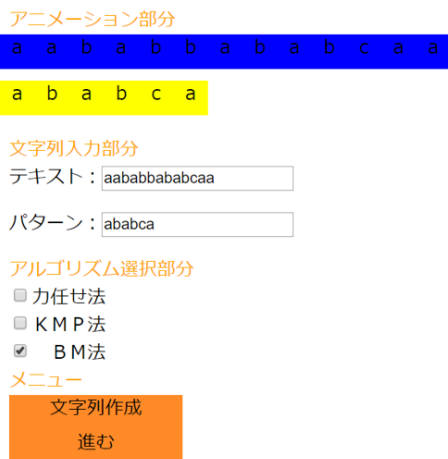


図6 文字列検索アニメーション部分

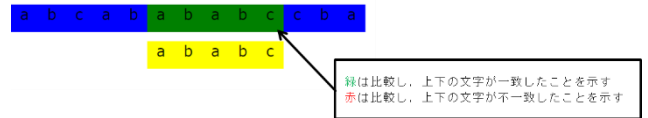


図7 アニメーション結果の例

ずらし表部分



KMP法ずらし表 :

BM法ずらし表 : a:0 b:2 a:0 b:2 c:1 a:0

ずらし表メニュー

ずらし表準備

ずらし表進む

図8 ずらし表部分

2.5 前回の実験結果と考察

前回の実験ではアルゴリズムの動作などの表面的な知識や効率に関する知識については、内容が不足していたり、アルゴリズムの特徴を踏まえていない解答が多いため、十分に理解させることができなかったと判断した。また時間計算量の理解は力任せ法の段階でつまづいている結果となった。そのため再度課題階層分析を行い、学習内容や組み立てを考え直すことや出題内容を改善する必要があると考えた。

3. 課題階層分析

2.5節を踏まえて、各文字列検索アルゴリズムについて、再度課題階層分析を行った。本章では各文字列検索アルゴリズムについて設計した課題階層分析図を示し、さらに各段階でどのような学習を行うのかを説明する。

3.1 力任せ法の課題階層分析図と学習の流れ

力任せ法の動作と最大時間計算量について、課題階層分析を行った。設計した課題階層分析図を図9、図10に示す。これらの課題階層分析図をもとに学習の流れを設定した。最初に学習者は力任せ法の動作の学習を行う。まずアニメーション教材を自由に操作することで、力任せの動作を確認することができる。そして学習者に検索手順の説明を行わせることにより、動作の理解に必要な「一致時の動作」「不一致時の動作」が理解できていることの確認を行う、説明することがで

できれば力任せ法の動作は理解できていると判断する。

次に学習者は時間計算量の学習を行う。前提知識として動作の理解の他に最大時間計算量、オーダ記法の理解が必要であるため、学習者は知識付けの資料を用いて学習を行う。続いて最大時間計算量を理解させる前提として「最大時間計算量となる動作」と「最大時間計算量となる文字列の特徴」が分かることが必要である。そのため最大時間計算量となる文字列を含めた4種類の文字列のアニメーションを確認することを促し、動作や文字列の特徴を意識させた。確認させた文字列を表3に示す。アニメーションを確認する際、比較回数と文字列の特徴、検索時の特徴を意識させることにより、最大時間計算量がテキストの長さ×パターンの長さであることの足掛かりになると考えた。また文字列および検索時の特徴を意識させることにより、最大時間計算量になる場合、テキスト1文字に対してパターンの文字数分比較する動作となることを学習者に気づかせることができると考えた。これらのことを学習者に意識させることにより力任せ法の最大時間計算量である $O(MN)$ を理解させることができると考えた。

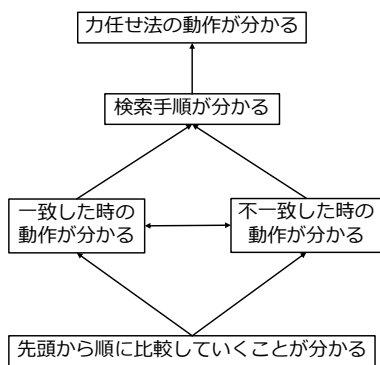


図9 力任せ法の動作の課題階層分析図

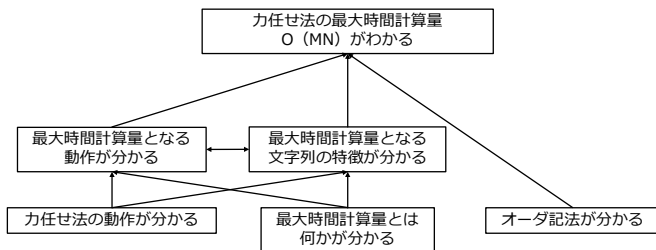


図10 力任せ法の最大時間計算量の課題階層分析図

表3 確認させる文字列

テキスト	パターン
aaaaaaaaab	aab
abcdefghijk	ijk
abc1abc2abc3abc4abc5	abc4
This is an apple	apple

3.2 KMP法の課題階層分析図と学習の流れ

KMP法の動作について、課題階層分析を行った。設計した課題階層分析図を図11に示す。この課題階層分析図をもとに学習の流れを設定した。学習者はアニメーション教材を自由に操作することでKMP法の動作を確認することができる。また「ずらし表の使い方」を理解するための前提として「ずらし表の作られ方」が分かることが必要であるため、ずらし表のアニメーションを意識させる問題を出題する。そして「一致・不一致した時の動作」を理解するための前提として「ずらし表の使い方」が分かる必要があるため、学習者に検索手順の説明を行わせることにより、検索時のずらし表の使われ方を理解しているかの確認を行う。また説明することができれば下位の学習目標である「一致した時の動作」「不一致した時の動作」が分かることを達成しているため、KMP法の動作は理解できていると判断する。また力任せ法と同様に4種類の文字列のアニメーションを確認させることにより、他のアルゴリズムとの効率の比較を行わせる。

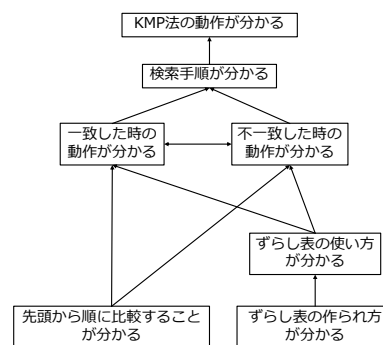


図11 KMP法の動作の課題階層分析図

3.3 BM法の課題階層分析図と学習の流れ

BM法の動作と最大時間計算量、平均時間計算量について課題階層分析を行った。設計した課題階層分析図を図12、図13、図14に示す。これらをもとに学習の流れを設定した。学習者はKMP法と同様にアニメーションの自由操作を行うことにより、BM法の動作

の確認を促し、アニメーションからずらし表の作り方が分かると考えた。そして説明を行わせることにより、ずらし表の使い方を理解しているか確認する。次に時間計算量の学習を行う。最初に学習者は力任せ法と同様にオーダ記法などの知識付けを行う。続いて、学習者に最大時間計算量、平均時間計算量となる文字列を含めた4種類の文字列について比較回数と文字列の特徴、検索時の特徴を意識させ、アニメーションの確認を促す。最大時間計算量は力任せ法と同様の値であるため、3.1節に示した学習の流れを通して学習者に理解させる。平均時間計算量を理解させるための前提は「平均時間計算量となる文字列の特徴」と「平均時間計算量となる動作」が分かることであるため、これを意識させる。具体的には文字の種類が多い文字列のアニメーションを確認させることで、実際の文字列検索処理では末尾で不一致する確率が高く、パターンの文字数分ずらせることを学習者に気付かせることができると考えた。これらのことから平均時間計算量である $O(M/N)$ を理解させることができると考えた。

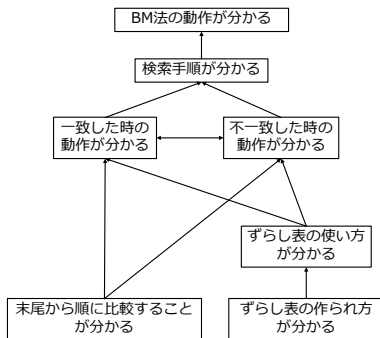


図 12 BM法の動作の課題階層分析図

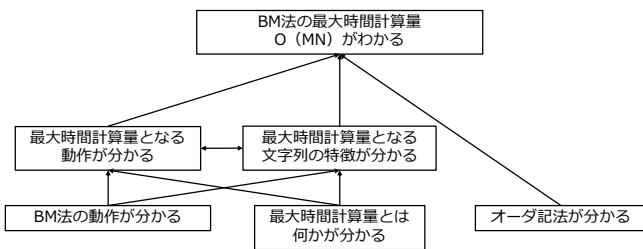


図 13 BM法の最大時間計算量の課題階層分析図

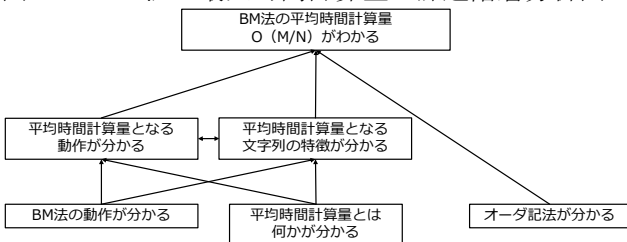


図 14 BM法の平均時間計算量の課題階層分析図

4. アニメーション教材の追加箇所

今回設計した学習内容において比較回数を意識させるための仕組みとして、新たに比較過程を表すアニメーションを追加した。比較過程のアニメーション部分を図 15 に示す。比較が1段階進むごとに、その段階に応じた比較回数、比較位置などを表示する。赤文字は不一致、緑文字は一致したことを示す。

比較過程	
1回目の比較ではテキストのbとパターンのaを比較し、	不一致しました
2回目の比較ではテキストのbとパターンのaを比較し、	不一致しました
3回目の比較ではテキストのbとパターンのaを比較し、	不一致しました
4回目の比較ではテキストのaとパターンのaを比較し、	一致しました
5回目の比較ではテキストのcとパターンのcを比較し、	一致しました
6回目の比較ではテキストのbとパターンのbを比較し、	一致しました
7回目の比較ではテキストのaとパターンのaを比較し、	一致しました
8回目の比較ではテキストのbとパターンのbを比較し、	一致しました
9回目の比較ではテキストのaとパターンのaを比較し、	一致しました

図 15 比較過程

5. 今後の予定

今回設計した学習内容をもとに現在情報系学生数名に対して予備実験を実施している。予備実験の結果が分かり次第、さらに学習課題に改善を加え、本実験を行う予定である。

参考文献

- (1) 古川勝康, 井上勝行, 魚井宏高, 首藤勝: “制御の流れに重点をおいてアルゴリズム学習を支援するシステムの構想”, 電子情報通信学会技術報告, pp159-166(1997-03-15)
- (2) 下川輝, 仲林清: “アニメーションを用いた文字列検索アルゴリズム学習のための課題設計”, 教育システム情報学会第1回研究会論文集, pp33-38(2019)
- (3) 大森克志, 木村春彦, 広瀬貞樹: “アルゴリズムの基礎”, 共立出版株式会社(1997)
- (4) 稲垣忠, 鈴木克明: “授業設計マニュアル ver.2 教師のためのインストラクショナルデザイン”, 北大路書房(2015)