

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

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 Design 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、学習者はアニメーションの速度などを注視してしまい、アニメーションのみで特徴を理解させることは困難な場合がある。本研究では文字列検索アルゴリズムのボイヤー・ムーア法のアニメーションを開発し、さらに特徴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、前回に報告した他のアルゴリズムの動作との比較および特徴が顕著にあらわれる文字列の動作の確認の2種類の学習課題を改善し、新たに課題階層分析によって学習の流れを設計した。実験の結果、動作などの表面的な知識を理解させることはできるが、時間計算量の理解については更なる改善が必要となった。

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。アニメーションの利点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、学習者はアニメーションの速度などを注視してしまい、アニメーションのみで特徴を理解させることは困難であると考えた。本研究では文字列検索アルゴリズムのボイヤー・ムーア法（以下 BM 法）を対象として、アルゴリズムアニメーションを開発し、さらに特徴理解を促す学習課題を取り入れる。具体的には他の文字列検索アルゴリズムの動作と比較させることおよび入力機能を用いて特徴が顕著にあらわれる文字列の動作を確認させることを課題とする。今回の報

告では、前回の報告⁽²⁾を踏まえ、学習者の BM 法の特徴の理解度を向上させるための課題設計を行った。以下、第2章では学習課題について示し、第3章でアニメーション教材について述べる。第4章では実験についてまとめ、考察を行い、第5章ではまとめと今後の課題について述べる。

2. 学習課題

BM 法の特徴理解を促す学習課題について、対象としたアルゴリズムについて説明を行い、前回の報告を踏まえて今回の学習課題について説明する。

2.1 対象として選択したアルゴリズム

本研究で対象としたアルゴリズムについて説明する。

2.1.1 線形探索

線形探索とは最も単純な探索方法である。対象とす

るデータと検索文字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。

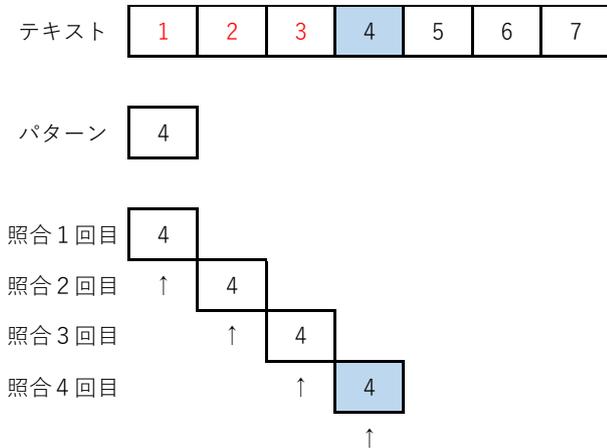


図1 線形探索の照合例

2.1.2 力任せ法

力任せ法とはテキストとパターンを揃え、左から順にテキストとパターンの対応する文字が一致しているかどうかを比較する。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、照合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。

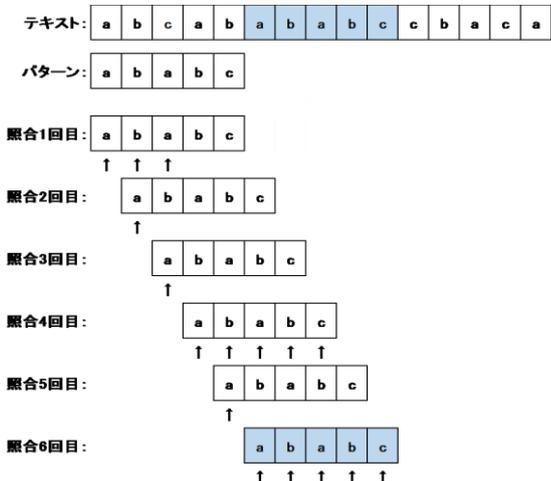


図2 力任せ法の照合例

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（以下 **KMP 法**）とは力任せ法と同様にテキストとパターンの先頭を揃え、左から順にテキストとパターンの対応する文字が一致しているかどうかを比較するが、パターンの不一致位置に応じて、比較の再開位置を決めるアルゴリズムである。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ため、3, 4文字目の **ab** まで一致していることがわかる、したがって照合4回目では1, 2文字目の **ab** は一致することがわかっているため、比較の再開位置をずらし3文字目から比較することができる。

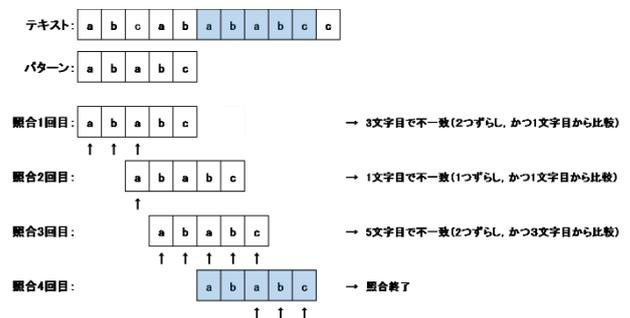


図3 KMP法の照合例

2.1.4 BM法

BM 法は他の力任せ法と **KMP 法**と同じようにテキストとパターンの先頭を揃えて比較を開始するが、両アルゴリズムがパターンの先頭から比較を始めるが **BM 法**はパターンの末尾から比較を行うアルゴリズムである。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列のほうが多くなる。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、検索文字列の長さに近い文字数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。また平均的な時間計算量はテキスト長を M およびパターンの長を N とすると $O(M/N)$ となる。

最悪の場合にはテキストとパターンが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。

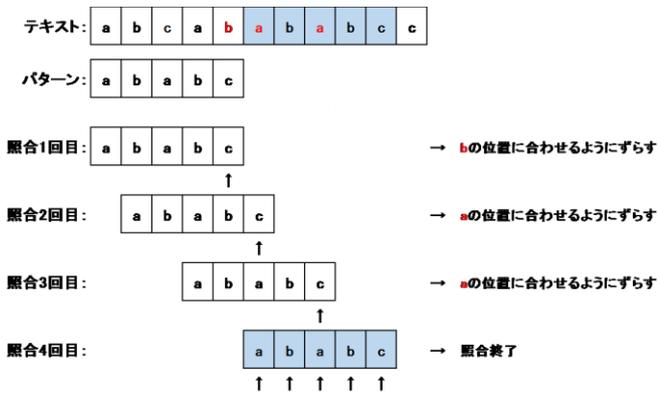


図 4 BM 法の照合例

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。これらの文字列検索アルゴリズムと BM 法との違いはパターン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 前回の実験結果と考察

前回の実験では, 大半の学習者の解答に改善が見られたことから, アルゴリズムアニメーションはアルゴリズムの動作や特徴を理解させることに有効であると考察した。しかし, 時間計算量は学習課題の不十分な内容や学習者の操作の不足, 既有知識を考慮しなかったことなどが曖昧な理解に繋がってしまったのではないかと考えた。そのため学習課題の内容や組み立てをどのようにするのかを考える必要があると考えた。

2.4 今回の学習課題

2.3 節の考察を踏まえて, 課題階層分析^②を行い, 新たに学習の流れを設定した。

2.4.1 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⁽³⁾。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。

2.4.2 課題階層分析図の設計

表 1 に本研究で理解させる BM 法の特徴⁽⁴⁾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 法の理解に必要な要素を考察し, アルゴリズム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を図 5 に示す。この課題階層分析図を基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を例にして学習する。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検索が進むため平均的な計算量である $O(M/N)$ 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 $O(MN)$ 平均的な時間計算量 $O(M/N)$ M:テキストの長さ, N:パターンの長さ

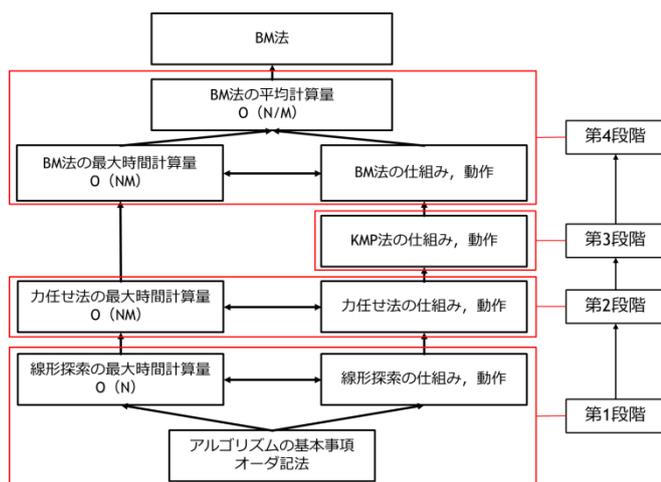


図 5 BM 法の理解の課題階層分析図

表 2 用意した文字列

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

3. アニメーション教材について

本研究で用いるアニメーション教材を Web ブラウザで動作するものを開発した。また、学習者の操作したテキストや検索方法などの操作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 8 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ような動作を行う。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。

アニメーション部分



文字列入力部分

テキスト:

パターン:

アルゴリズム選択部分

- 力任せ法
- KMP法
- BM法

メニュー

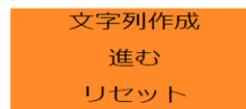


図 6 文字列検索アニメーション部分

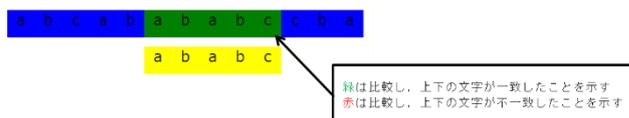
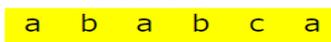


図 7 アニメーション結果

ずらし表部分



KMP法ずらし表:

BM法ずらし表:

ずらし表メニュー



図 8 ずらし表部分

4. 実験について

4.1 実験概要

本研究では、学習者に 2.4.2 でも述べた各アルゴリズムの学習後に動作や時間計算量についての問題を出題し、アルゴリズムの理解度を測る。対象とした学生は情報系学科 3 年生 5 名である。

4.2 実験の流れ

実験は図 9 に示した手順で行った。学習者はまず事

前知識の確認と線形探索を用いてアルゴリズムの基本的事項やオーダ記法についての学習を資料にて行う。続いて各文字列検索アルゴリズムはアニメーション教材および学習課題の操作を行い、問題に取り組む。最後に事後の理解度を調査した。



図9 実験の流れ

4.3 出題内容

各アルゴリズムの理解度を確かめるため、確認問題を出題した。表3に各アルゴリズムでどのような問題を出題したのかを示す。動作は検索する順番などの動作について自由記述をする。効率の問題はそのアルゴリズムにとって、効率良いあるいは悪いテキストやパターンの特徴を記述する。効率の問題はそのアルゴリズムの特徴を理解していないと解答できないと考えたため行った。ずらし表はずらし表の文字と数字の意味を記述する。ずらし表の問題はずらし表の数字によって検索位置が変更されることを理解しているかを確かめるために行った。時間計算量の問題はそのアルゴリズムの平均的あるいは最大時間計算量はどのような値となるか、またその理由を記述させた。

表3 出題内容

アルゴリズム 出題内容	力任せ	KMP	BM
動作	出題	出題	出題
効率の良い場合	出題	出題	出題
効率悪い場合	出題	出題	出題
ずらし表		出題	出題
平均的な時間計算量			出題
最大時間計算量	出題		出題

4.4 実験結果と考察

この章では文字列検索アルゴリズムごとに結果をまとめ、考察を行う。各問題の評価は3段階とし、未解答や的外れな解答は×、抽象的な解答や不足している解答は△、具体的な解答は○とする。

4.4.1 力任せ法の解答結果と考察

力任せ法の問題の結果を表4に示す。力任せ法の動

作の結果ではすべての学習者が「テキストとパターンの最初の文字を比べて同じ文字ならば2文字目も照合するが違う場合は一文字パターンを右にずらして再度最初から照合していく」のような解答をしていたため、アルゴリズムの動作などの表面的な知識の理解度は向上したといえる。

効率に関する問題は「テキスト ababaabbacab パターン cab の様にテキスト内で出現数が少ない文字がパターンの先頭にくるもの」などの効率に関わる点に触れている解答をしていた学習者を○と判断した。学習者Cのみ「宝くじのような紙に書かれている文字や数字を照合して抽選するときなどに時間短縮としての効果がある」のような的外れな解答をしていたため、×とした。学習者Cのみ不正解だったが他の学習者がすべて○であったため、理解度は向上したといえる。また問2、問3で○と評価した学習者の大半が効率について説明する際に、用意した文字列を例にして説明していたため、この文字列は効率の理解度向上に効果があると判断した。

時間計算量は大半の学習者が未回答であったため、理解させることができなかったと判断した。

表4 力任せ法の解答結果

問題 \ 学習者	A	B	C	D	E
動作	○	○	○	○	○
効率の良い	○	○	×	○	○
効率の悪い	○	○	×	○	○
最大時間計算量	○	×	×	×	×

4.4.2 KMP法の解答結果と考察

KMP法の問題の結果を表5に示す。KMP法の動作の結果ではすべての学習者の解答を△と判定した。すべての学習者が2.1.3で述べたKMP法の特徴であるテキスト内の照合開始位置が変更される点に触れた解答をしていなかったため△とした。また、効率の良い場合の問題でもKMP法の特徴を踏まえた解答をした学習者がいない結果となった。原因としてアニメーションにおいて、テキスト内の照合開始位置が変更されることよりもパターンが大きく動くことを注視してしまっただけのため、学習者に特徴が伝わらなかったと考察した。一方効率の悪い場合の問題では「パターン内の文字列がテキスト中に何度も出現しないとき、KMP法

の探索の省略が行えないため効果が低いと考えられる」のような○の解答をした学習者が4人いた。正答が多い理由として効率の悪い場合では力任せ法と同じアニメーションをするため、効率の悪い場合のほうが理解しやすいと考察した。

ずらし表については「パターン文字列との比較時点で直前までに何文字一致していたかという意味」のようなずらし表の求め方を理解している学習者は見られたが、ずらし表に応じて次の比較再開位置を決めるといふ解答をしていた学習者がいなかったため、アニメーションの見せ方や学習内容を改善する必要があると考察した。

表4 KMP法の解答結果

問題 \ 学習者	A	B	C	D	E
動作	△	△	△	△	△
効率の良い	△	△	×	△	△
効率の悪い	△	○	×	○	○
ずらし表	△	×	△	×	△

4.4.3 BM法の解答結果

BM法の解答結果を表5に示す。BM法の動作の結果では「後ろから照合をすること」という点はすべての学習者が記述していたが、「不一致の時にずらし表の数字に応じてパターンをずらすこと」という点を記述している学習者はいなかったため、問題の表現の仕方や学習課題の内容を考える必要があると考察した。

効率に関する問題は「文字の種類が多い場合、不一致が起こる確率が高いため、パターンの文字数分ずらせる」という文字の種類数によって効率に変化する点を記述している学習者がEのみだったこと。また「パターンの文字列を比較的長くする」などの解答を△とし、パターンの長さに着目した解答をする学習者が多くいたため、改善が必要であると考えた。

ずらし表の問題については「パターンの末尾からの距離に応じて決まる」というずらし表の求め方を解答している学習者を○とした。しかし、学習者B、Eについては未回答であったため、ずらし表を考えさせる学習課題が必要であると考えた。

時間計算量は力任せ法の段階で理解させることができていないため、未回答が多い結果となった。学習者Aは平均的な時間計算量は○だったが事前知識を確

認した際に時間計算量について答えていたため、正答したと考えた。学習者Cは時間計算量を正しく答えていたがアニメーションの操作ログから学習課題の操作が足りない点があったこと、事前知識は無かったことを考慮し、正しく学習できていないと判断したため△とした。

表5 BM法の解答結果

問題 \ 学習者	A	B	C	D	E
動作	△	△	△	△	△
効率の良い	△	△	×	△	○
効率の悪い	○	△	○	△	○
ずらし表	○	×	○	○	×
平均的な時間計算量	×	×	△	×	×
最大時間計算量	○	×	△	×	×

5. まとめと今後の課題

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は不足している解答やアルゴリズムの特徴を踏まえた内容で解答している学習者が少ないため、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躓いている結果となった。そのため再度階層分析を行い、学習内容や組み立てを考え直すことや出題する問題の内容をより具体的にするなどの改善する必要がある。

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: “アルゴリズムの特徴理解を促すアルゴリズムアニメーションと学習課題の分析”, 教育システム情報学会第43回全国大会論文集, pp45-46(2018)
- (3) 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)
- (4) 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)