

# Moodle上でのCindyJSを用いた探究学習における ログデータのアナリティクスから学習支援へ（続）

東邦大学・薬学部 金子 真隆

Masataka Kaneko, Faculty of Pharmaceutical Sciences, Toho University

合同会社三玄舎 中原 敬広

Takahiro Nakahara, Sangensha LLC

東邦大学・理学部 野田 健夫

Takeo Noda, Faculty of Science, Toho University

## 1 はじめに

昨年の本共同研究において、フルオンラインの環境下で大学教養課程の数学の授業を実施することが迫られる中、Webブラウザ上で数理モデルを操作することを可能にする動的幾何システムCindyJS(<https://cindyjs.org/>)により生成されるタンジブルなコンテンツを学習管理システムMoodle上で提示し、学習者にこれを用いて探究学習を行わせた上で、操作のログデータをMoodle経由で取得し分析した事例について紹介した[1]。この事例では、テーマとなった関数の多項式近似に関する数式表現とグラフィックスによる図的表現との連関を意識させるべく作成した解説動画による教育的介入が、学習者の操作プロセスに及ぼす効果について、定量的に評価できる可能性を示唆した点がポイントであった。具体的には、Productive Failure [2]と理解しうるような近似の「失敗」が、動画を視聴した群の学生について、そうでない群の学生と比べて高い比率で確認され、該当の「失敗」を伴う探索を通して無限小の次数の概念の獲得に結びつく可能性が示唆されていた。本稿で紹介する事例は、昨年度と同様のフローに従いフルオンラインの環境下で実施したものであるが、昨年度は授業形態がオンデマンド形式であったのに対し、本年度はZoomを用いた同時双方向形式であったという大きな違いがある。とかくこうした違いは形式的なもののように考えられがちであるが、授業実施と学習データ解析の両面について、かなり大きな差異を生んだ。また、操作後に実施した近似の範囲に関する問いかけを内容とするアンケートに対する学習者の回答状況から、上記の介入が操作プロセスに影響をもたらすメカニズムについて、より詳細な知見も得られつつある。本稿では、これらの新たに得られた知見について紹介したい。

## 2 教育工学的背景

近年のラーニング・アナリティクスでは、学習管理システムなどを經由して大量の学習データが得られるようになってきていることを背景に、学習者が教材のファイルを閲覧したり解説動画を視聴したりした際のパターンや、オンラインのチャットでのやりとりを含めた学習者間の対話のプロセスに関する情報を該当のログデータから抽出し、これと

最終的なテストパフォーマンスとを統計的に比較して、後者に関する予測モデルを確立する試みが数多く行われている。こうした分析は、問題を抱えた学習者を早い段階でいかに発見するかといった点で有効な示唆を与えることが期待される一方、教授者による有効な教育的介入に結びつけるためには、学習データから析出されるシグナルが最終的なパフォーマンスに結びつくメカニズムについて理解する必要がある、学習データからより深い情報を取り出すことが求められる。いわゆるマルチモーダルなアナリティクスは、こうした問題意識に応えるべく、学習者の行動の深層部分を解明することを目指したものであると考えられる。たとえば、数学の学習者によるジェスチャーに関するログデータを解析した先行研究 [3] では、センシング技術から得られる大量のデータを扱う際に、行動パターンの時間遷移に基づいて、一連の行動が発生した背景をいかに説明できるかという点に特に注意を払う必要があることが強調されている。また、複数の学習者が数学の協調学習を行った際の、対話の質と最終的なアウトカムの連関を調べた先行研究 [4] でも、いくつかの対話パターンが相互に関連する可能性や、学習者が事前に身につけている数学的な素養によって対話の質そのものが影響を受ける可能性に注目し、アウトカムレベルの分析と並行してプロセスレベルの分析、さらにはこれらのデータを双方向に行き来する分析が必要であることを主張している。この主張の背景には、同じ研究者による別の先行研究 [5] の中で、学習者の事前の学力レベルによって、協調学習を行った際の議論の質に差が出る事例が見出されたことがあると考えられる。動的コンテンツの操作は学習者間の対話とは性質が異なるものではあるが、操作自体を学習者とPCとの対話であると解すれば、操作パターンに関しても学習者の事前の学力レベルによる同様の影響があり得ることとなる。実際、筆者の先行研究 [6] でも、数理的な素養が高いと想定される高校生の操作パターンが、一般的な大学生のそれと比べて、単項式関数のグラフの特徴をより踏まえたものになっていることを強く示唆する結果が得られている（稿末の追記参照）。これは、昨年度の事例 [1] で、操作前の介入が行われた学習者につき、そうでない学習者と比較して多くみられるパターンであることも指摘した。本稿では、前節にふれた近似の「失敗」が発生する前後における操作プロセスを特徴づけるシグナルの時間遷移を追跡すると同時に、アンケート調査に対する回答状況をこれと照合し、解説動画による介入が操作パターンに差異をもたらすメカニズムについて追究する。

### 3 事例研究の方法と結果

実施の基本的なフローは昨年度と同様であるので、コンテンツの詳細や操作ログを取得する手続きについては筆者の過去の講義録 [1][6] に譲るが、前述の通りオンデマンド型から Zoom を用いた即時配信型への講義形態の切り替えが大きな影響を及ぼした。実際、昨年度 [1] のオンデマンド型の状況では、学習者がすべて動画を視聴したか確認することが技術的に難しかったのに対し、今年度は実験群の学生に対して、操作の直前に介入用の動画を見せたことにより、事実上、視聴を強制・徹底した形となった。また、操作時間を 10 分から 20 分程度と指定したことで、各群の学生がほぼ同時並行の形で操作を行う形となり、取得した操作ログデータの処理が非常にやりやすくなった。結果と

して、実験群学生 58 名と統制群学生 60 名の操作ログデータにつき、ほぼ取りこぼしなく取得・分析することができた。また、操作終了後にアンケート調査を行ったのも新たな試みであるが、操作自体が無理関数  $y = \sqrt{1+x}$  の  $x = 0$  の近傍での 3 次関数による近似であったのに対し、「多項式の次数が高くなっていった場合に近似はどのようになっていくと想定されるか」という問いかけが主たる内容であった。これは半ば故意に曖昧な問いかけにしている部分があり、特異点の存在によって次数を上げたとしても多項式で近似できる範囲に限界があることを意識できていたか、学習者自身の記載によって確認することを目的としたものである。

昨年度の事例 [1] と同様に、各学習者の操作における sequential なパターンを可視化したのが図 1 である。ここで赤色は 0 次の係数の操作、黄色は 1 次の係数の操作を示す。

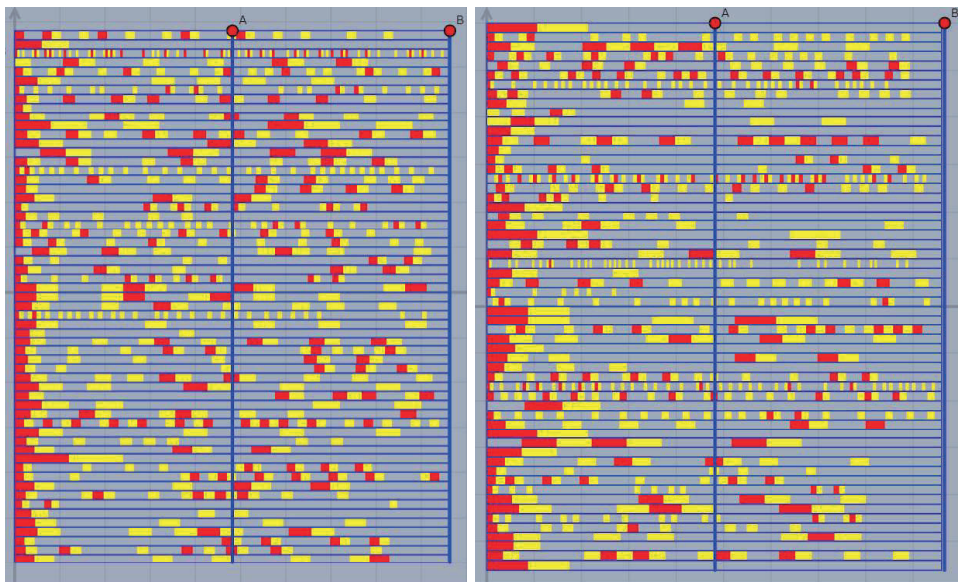


図 1 低次係数の操作パターンの可視化（左：統制群・右：実験群）

この結果を見ると、昨年度の事例と同様に、操作プロセスの後半において、低次（0 次および 1 次）の係数の操作比率に関し、実験群で統制群よりも低いことが観察される。実際、後半の操作全体の中で、これらの低次の係数の操作が占める比率を各学習者について計算し、2 群における比率の分布に対してウィルコクソンの順位和検定を行ったところ、5%水準で有意な結果となった ( $p = 0.0475$ )。先行研究 [7][8][9] において、こうした低次の係数の操作比率の低下が、大域的近似から局所的近似へと学習者の意識が遷移している状況を示すものであることが示唆されていることをふまえると、操作の途中段階で大域的な近似を探すべく低次の係数を動かしてみたものの、1 次近似の部分が接線を与える形と異なる場合には高次の係数をどのように調整しても良い結果が得られないという「失敗」を通じ、高次の無限小の概念を直感的に把握して行っているプロセスが実験群において顕著に多くみられるという解釈が成り立つ。

このような「失敗」の発生を示唆するシグナルとして、 $x$  の区間ごとの近似誤差を想定した。この誤差を追跡することにより、学習者が各時点でどの範囲の近似を目指して

いるか把握できると考えられるからである．具体的に

$$\sup_{x \in [x_0, x_1]} |\sqrt{1+x} - a - bx - cx^2 - dx^3|$$

を考えたときに， $[x_0, x_1] = [-0.5, 0.5]$ における誤差を小さく保ちながら， $[x_0, x_1] = [-1, -0.5]$ における誤差を小さくしてみるトライアルの存在を指摘した [1]．このような場合，区間  $[0.5, 2]$ における誤差は大きくなるというトレードオフの関係が成り立ち，上記の「失敗」につながることになる．そこで，前者が0.01以下であり，後者がいくつか設定した閾値 (0.20, 0.25, 0.30, 0.35, 0.40) 以下になるようなトライアルを経験している学生の比率を2群についてそれぞれ計算し可視化すると，図2のようになる．

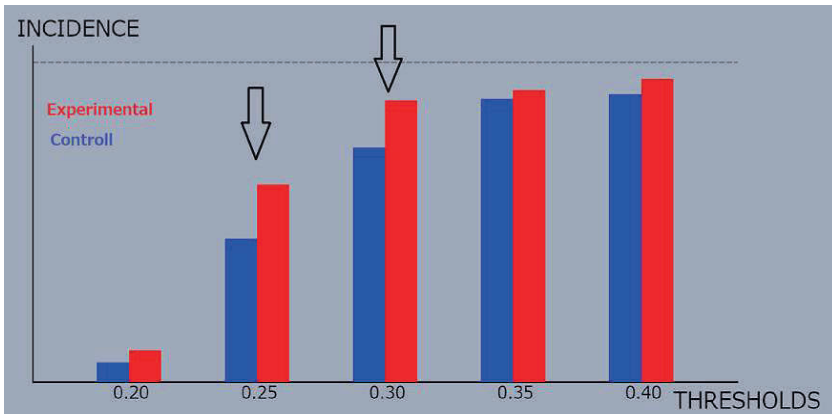


図2 Productive Failure の発生率 (青左：統制群・赤右：実験群)

図中に矢印で示した通り，これらの比率が等しいか否かについて，群間比較の検定を行うと，閾値が0.25と0.30のケースで2群の比率に有意差が確認された (閾値が0.30の場合，統制群では60名中44名，実験群では58名中51名が該当し， $p = 0.0385$ )．

以上から想定されるのは，実験群におけるこれらの「失敗」の発生率の高さが，操作プロセス後半における低次の係数の操作比率の低さにつながったのではないかということである．一般論として，本研究における学習活動のように，学習者がある時点で行った操作の結果を観察し，その結果をそれ以後の操作に反映させているような場合には，操作の sequential なパターンを分析しただけでは不十分であり，よりイベントの中身に密着した分析が求められることが指摘されている [10]．そこで，操作プロセス中に該当の「失敗」の存在が確認された学生について，3つの区間  $[-1, -0.5]$ ， $[-0.5, 0.5]$ ， $[0.5, 2]$ における誤差の時間遷移を可視化した上で，最初に「失敗」が発生した瞬間をプロットすることを試みた．図3は，実験群のある学生について該当の可視化を行ったものである．ここで，横軸は左から右に向けて時間の経過を表し，1目盛が1分である．縦軸は1目盛が $\frac{1}{6}$ である．また，上記の3区間における誤差の遷移をそれぞれ黒色，赤色，青色の線で表示している．さらに，上記の「失敗」が最初に発生した時点を黄色の線で表示している．当然，この時点が全操作プロセスの中でどこに位置するかは，学習者個々に異なることになる．

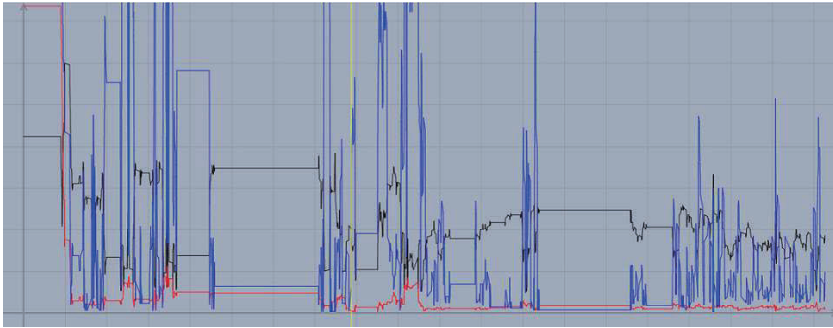


図3 誤差の時間遷移と Productive Failure の発生時点

Productive Failure が操作傾向に変化をもたらした因果関係を探るためには、「失敗」の前後で区間  $[-0.5, 0.5]$  における誤差の推移がどのように変化したか追跡することが自然である。ここで問題になるのが、図3の事例にも見られるように、操作が全時間帯にわたり絶え間なく行われているわけではなく、いったん操作を止めて観察している時間帯が存在するという点である。もとより、操作を止めて観察している時間帯にも学習者の思考が進展していることを想定すると、誤差が変動している時間帯だけを抽出したのでは、思考プロセスの評価として不十分になる危険性がある。ここでは発想を逆転させて、「操作」に関する評価ではなく「観察」に関する評価と考えて分析することとした。具体的には、最初に「失敗」が確認された時点の前後で、それぞれ区間  $[-0.5, 0.5]$  における誤差の AUC (Area Under Curve) を計算し、その時間平均を計算して比較することを試みた。「失敗」が確認できない学習者については、やむを得ないので、全時間帯にわたる AUC の平均をもって代替することとした。図4は、2群に分けて各学習者の操作ログから計算した該当の数値をプロットしたものである。左軸が失敗前、右軸が失敗後の値を表しており、同一の学習者のデータを線分で結んで表示している。「失敗」が確認できない学生のデータは便宜的に黄色い水平線で表示されている。

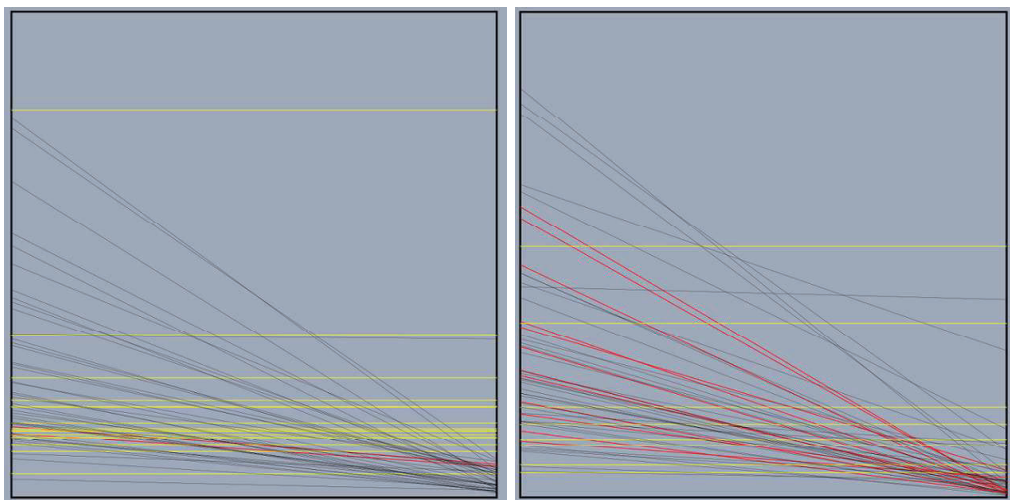


図4 Productive Failure の前後での AUC の時間平均の推移 (左: 統制群・右: 実験群)

ここで、図中の赤線は、学習者のアンケート調査での回答に関連している。問いかけが曖昧であったため、回答には様々な観点のものが含まれるが、その中で、 $x$ の範囲に制限なく近似が可能になっていくという趣旨の回答をしたものが、統制群で7名、実験群で9名あった。これに対し、限定された範囲でのみ近似が可能であるという趣旨の回答をした者は、統制群で6名だったのに対し、実験群で16名となった。仮に近似範囲が限定されるという意識を持っていた学生であっても、それが必ずしもアンケートの回答として現れるとは限らないので、回答数を単純に比較するのは意味がないが、実験群に偏って多いことは指摘できるであろう。さらに問題なのは、これらの近似範囲の限定について言及した学生のうち、操作プロセス中に失敗を含む者がどの程度含まれたかである。結果として、実験群では該当の回答が16名中12名と大多数だったのに対し、統制群では6名中わずか2名のみであった。これら計14名分のデータが図4の中で赤色の表示がされている。2群の間でこれらのデータの分布を比較すると、特に失敗後の分布について、実験群で低い水準に集中していることが確認できる。実際、実験群で失敗を経験した51名を、近似範囲の限定に言及した12名とそれ以外の39名に分け、該当者の操作ログデータから失敗後のAUCの時間平均値の分布を計算し、その分布を対象にしてウィルコクソンの順位和検定を行うと、これら2つのグループの分布に有意な差が認められる( $p = 0.0305$ )。

## 4 総合論議と今後の研究の方向性

第2節に触れたラーニング・アナリティクスの立場から本研究の事例をみた場合、操作ログデータから比較的容易に取得できる、図1にあるようなsequentialなパターンをもとに学習者の困難を把握し、これに対する教育的介入を行うために思考の状況を適切に解釈する道筋を確立することが求められる。ここでは、介入用の動画の内容、操作後のアンケートへの回答状況、操作プロセス中での「失敗」の発生状況、「失敗」前後での近似誤差の推移、といった複数の要素を照合し、アナリティクスの文脈から何を結論できるか考察したい。

動画では、3次関数のグラフの形状を単項式関数の重ね合わせという観点から理解した上で、 $|x|$ が1より小さい範囲か、それ以外の範囲かに応じ、単項式関数の間のパワーバランスが逆転する点に触れられている[1]。従って、動画を操作前に視聴した実験群の学生の間で、アンケートへの回答の中に近似範囲の限界に言及するものが多かったのは自然なことと考えられる。図1で確認されたように、実験群の学生について、操作プロセス後半で低次係数の操作比率が低下しているが、こうしたsequentialなパターンの変化が $x = 0$ の近傍での局所的な近似への意識の集中を意味するものだとすれば、結果として $x = 0$ の近傍での近似誤差の低下に反映されることが想定できる。図4の結果をふまえると、操作の途中でProductive Failureを経験したことがこうした近似誤差の低下を引き起こし、「失敗」の発生率が高い実験群でその傾向が強くなったという因果関係が強く示唆される。以上を総合すると、動画の視聴によって学生の意識の中に近似できる範囲に関する関心が生じ、これがProductive Failureの発生に結びつき、結果として大域的な近似への「諦め」を伴いつつ、図1にあるような実験群特有の操作パターンを

生んだというシナリオが成立する。

仮に分析の目的が解説動画の効果の検証であれば、以上の知見だけでもそれなりの価値を持つと考えられるが、学習者の操作プロセスを把握しながら、どのように足場架けを行うかという次元での知見を求めるとすると、事情はとたんに複雑になる。その理由は、動的コンテンツの操作を通じて「理解が進んだ」とひとくちに言っても、その深度や内容は学習者個々の状況に依存する部分が大きく、高度に重層的だという点である。実際、図4の実験群のデータに見られる通り、全体として近似誤差の低下が認められる一方で、近似範囲の限界についてアンケートに言及するまでの明確な意識を持っている学習者と、認識がそこまでには至っていない学習者とで、操作方略の遷移に小さくない差異があることが明らかである。一方、統制群のデータに関しても、近似範囲の限界にアンケートで言及し、操作中の失敗も経験している学習者が、失敗後の近似精度の点で必ずしも良いパフォーマンスを残しておらず、操作方略に少なからずゆらぎを残していた可能性が読み取れる。この結果からは、解説動画で示された単項式関数のパワーバランスに関する説明が Productive Failure の発生を促したという方向性だけでなく、逆に、Productive Failure を操作中に経験することで、単項式関数の増減のパターンを想起するという方向性の思考プロセスを経た学習者の存在が示唆される。こうしてみると、操作を通じた理解の重層性に対応し、教授者側の診断モデルとしても、階層的なものを準備しておく必要が想定される。一例として、次のような診断リストが考えられる。

1. 探索は  $x = 0$  の近傍における誤差の点で収束性を示しているか？
  - $x = 0$  の近傍での誤差が操作終盤までゆらぎを見せていないか？
  - $x = 0$  の近傍における AUC は低下しているか？
2. 探索は exhaustive なもので、偶発的に最終的な近似を見つけたものではないか？
  - 右側の区間で誤差を大きくしてみるトライアルがあるか？
  - 右側と左側の誤差についてトレードオフの関係を認識できているか？
3. 近似できる範囲の限界を意識するに至っているか？

もとより、ICT を活用した探究学習のプロセスについては、偶発的な要因による影響が排除できず、動的コンテンツの操作のログデータのみを用いて学習活動を十分に把握することは困難な場合が多いことは否定できない。しかし、本研究の結果が示す通り、教授者側が学習目的や利用するコンテンツの特性に応じた工夫を少しだけ行い、学習プロセスに関する他種のデータを取得した上で、操作ログデータとの照合に基づくマルチモーダルなアナリティクスを実施することによって、学習者の思考を解明できる可能性は小さくない。筆者が現在取得する必要を感じているのは、動的コンテンツを用いた協調学習を行った際に、学習者が見せる体動、特に手の動きに関するログである。本研究で用いている CindyJS のコンテンツは、オンラインでの利用とは別に、iPad に実装して2人ないし3人の学生に協力して操作してもらおう形でも利用してきている。そうした際の経験では、操作方略を話し合う際に、学習者間の対話だけでなく、画面上を指さす行動が多く見られ、これがグループの思考プロセスを追跡する上で重要な手掛かりとなる場合が少なくない。図5は、山口大学の北本卓也氏が開発されたシステム [11] を利用し、記述統計における相関係数の性質について学習するために作成した、動的コンテンツと数式表示とを統合させた HTML 形式の教材の画面である。

$A = (a_x, a_y), B = (b_x, b_y), C = (c_x, c_y), D = (d_x, d_y)$  に対して  

$$m_x = \frac{a_x + b_x + c_x + d_x}{4}, m_y = \frac{a_y + b_y + c_y + d_y}{4}$$
 とし  

$$v_A = a_x - m_x, v_B = b_x - m_x, v_C = c_x - m_x, v_D = d_x - m_x$$

$$w_A = a_y - m_y, w_B = b_y - m_y, w_C = c_y - m_y, w_D = d_y - m_y$$
 とおくと、相関係数  $cor$  は  

$$\frac{v_A w_A + v_B w_B + v_C w_C + v_D w_D}{\sqrt{v_A^2 + v_B^2 + v_C^2 + v_D^2} \sqrt{w_A^2 + w_B^2 + w_C^2 + w_D^2}}$$
 で定義される。  
 次のコンテンツは、6つのデータをいろいろ動かした時に、相関係数がどのように変化するか観察できるようにしたものである。

図5 数式表現と動的コンテンツを一体化させたHTML教材の事例

図6は、2人の学習者がこのコンテンツを用いて協調学習をした際の様子であるが、左側の写真では動的コンテンツの部分の手前で手を動かす動作が見られるのに対し、右側の写真では数式表現の部分を同時に指さす動作が見られる。



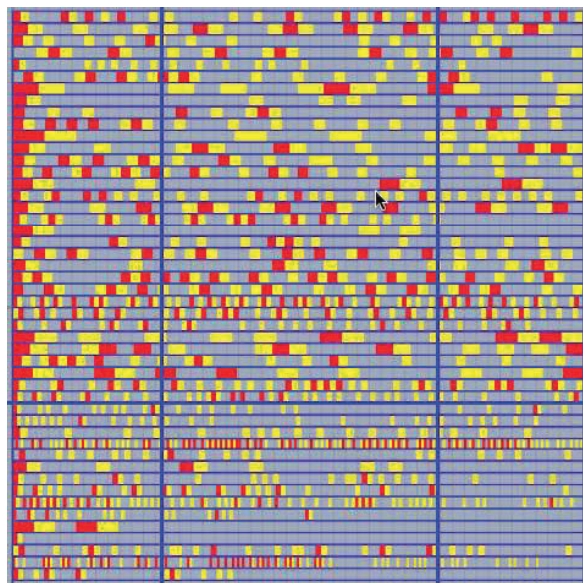
図6 コンテンツを利用した協調学習の様子



教材の内容や利用法については紹介記事 [12] に譲るが、同じく手を動かす動作であっても、2つの部分のいずれを対象としているかによって、その時点での活動内容が大きく異なることが推定され、グループの思考プロセスを時系列的に追跡する上で、これらの動作は大きな手掛かりになる。現在、深度センサーを用いて学習者の行動を上方から撮影し、手指の位置情報を取得して、操作ログや対話記録と対照した分析ができないか、検討を行っているところである。

### 追記

筆者の過去の講究録 [6] に収録した試行について、その後対象事例の追加があったことから、この際、結果の図解を更新する。該当の図解は図4であるが、図の下半分にある高校生の事例数が8例となっているところ、その後15例まで増え、その結果が可視化できていることから、これらを追加したものを以下に掲載する。これを見ると、高校生群の操作プロセス後半における低次の係数の操作頻度低下が、図1の実験群のデータにおける傾向と同一であることがわかるであろう。



講究録 2142・27 ページ 図4の差し替え

### 謝辞

本研究は JSPS 科研費 19K03175, 21K02752 の助成を受けている。

### 参考文献

- [1] 金子真隆, 中原敬広, 野田健夫: Moodle 上での CindyJS を用いた探究学習におけるログデータのアナリティクスから学習支援へ, 京都大学数理解析研究所講究録 2178, pp.37-46, 2021

- [2] M. Kapur: Productive failure, *Cognition and Instruction*, 26, pp.379–424, 2008
- [3] A. Andrade, J. A. Danish, A. V. Maltese: A measurement model of gestures in an embodied learning environment: Accounting for temporal dependencies, *Journal of Learning Analytics*, 4-3, pp.18-45, 2017
- [4] M. M. Chiu: Statistically modelling effects of dynamic processes on outcomes: An example of discourse sequences and group solutions, *Journal of Learning Analytics*, 5-1, pp.75-91, 2018
- [5] M. M. Chiu, B. W. Y. Chow: Classmate characteristics and student achievements in 33 countries, *Journal of Educational Psychology*, 107-1, pp.152-169, 2015
- [6] 金子真隆, 中原敬広, 野田健夫: CindyJS によるコンテンツの web 上での操作ログの解析, *京都大学数理解析研究所講究録 2142*, pp.23–32, 2020
- [7] 金子真隆: CindyJS によるコンテンツを用いた協調学習における操作と対話との関連の追跡 (I), *日本教育工学会第 34 回全国大会講演論文集*, pp.279–280, 2018
- [8] 金子真隆, 野田健夫: CindyJS によるコンテンツを用いた協調学習における操作と対話との関連の追跡 (II), *日本教育工学会研究報告集*, 18-5, pp.45–52, 2018
- [9] 金子真隆, 江木啓訓, 中原敬広, 野田健夫: CindyJS によるコンテンツを用いた協調学習における操作と対話との関連の追跡 (III), *日本教育工学会 2019 年度秋季全国大会講演論文集*, pp.289–290, 2019
- [10] P. Reimann: Time is precious: Variable-and-event-centered approaches to process analysis in CSCL research, *International Journal of Computer-Supported Collaborative Learning*, 4, pp.239-257, 2009
- [11] 北本卓也: HTML5 を用いた授業支援ツールについて, *京都大学数理解析研究所講究録 2178*, pp.39–47, 2020
- [12] 金子真隆, 北本卓也, 野田健夫: オンラインでの動的な探究学習における数学ソフトウェア活用について, *数式処理*, 掲載予定