

学習項目間の因果・依存関係計量による 学習者個人の学習状況適合カリキュラム生成方式

小西 亮介[†] 高橋 雄介[‡] 清木 康^{‡‡}

[†]慶應義塾大学総合政策学部 〒525-8520 神奈川県藤沢市遠藤 5322

[‡]慶應義塾大学大学院 政策・メディア研究科 〒525-8520 神奈川県藤沢市遠藤 5322

^{‡‡}慶應義塾大学環境情報学部 〒525-8520 神奈川県藤沢市遠藤 5322

E-mail: {†rkonishi, ‡yt, ‡‡kiyoki}@mdbl.sfc.keio.ac.jp

あらまし 本稿では、学習項目間の因果・依存関係計量による学習者個人の学習状況に応じたカリキュラムの動的生成方式を示す。本方式は、学習項目間の静的な順序・依存関係、および、特定の集団内および個人のテスト結果を用いて、その集団のテスト結果あるいは集団の学力傾向を比較することにより、既知の学習項目間の順序・依存関係の強さを計量し、さらに、その依存関係のルールを抽出する。本方式により、既知の学習項目間の順序・依存関係の強さとルールを用いて、個人のテスト結果から、個人別に重要な学習項目とその学習の順序を発見することが可能となる。本方式の実現により、学習項目と学習者個人との関係の計算が可能となり、動的に変化する個人の学習状況と学習項目間との関係計量による個人に特化したカリキュラムを生成し、個人の学習を効果的に支援することが可能となる。

キーワード eラーニング, データマイニング, 個人化, 学習管理システム

A Dynamic and Personalized Curriculum Generation Method by Analyzing Causality and Dependency between Learning Objects

Ryosuke Konishi[†] Yusuke Takahashi[‡] and Yasushi Kiyoki^{‡‡}

[†] Faculty of Policy Management, Keio University 5322 Endo, Fujisawa, Kanagawa, 252-8520 Japan

[‡] Graduate School of Media and Governance, Keio University 5322 Endo, Fujisawa, Kanagawa, 252-8520 Japan

^{‡‡} Faculty of Environmental Information, Keio University 5322 Endo, Fujisawa, Kanagawa, 252-8520 Japan

E-mail: {†rkonishi, ‡yt, ‡‡kiyoki}@mdbl.sfc.keio.ac.jp

Abstract In this paper, we show a dynamic and personalized curriculum generation method by analyzing causality and dependency the learning objects. This method calculates the strength and extracts the rules of causality and dependency between individual learning objects, which are described beforehand, by using results of testing and trends of understanding in a particular community and the definitions of learning objects with the order of learning. This method enables discovering important learning objects for an individual user with the order of effective learning, by analyzing an individual result of testing and that of a particular community. This method realizes computation of relationships between leaning objects and the learning situation of individual users, that is changing in a dynamic way, and enables generation of personalized curriculum for supporting the users learn effectively.

Keyword e-Learning, Data Mining, Personalization, Learning Management System

1. はじめに

データベースの高度応用領域として、21世紀知識情報化社会における教育および学習管理の実現への期待が高まっている。個人差を取り扱うことが効果的な学

習において重要である教育および学習分野に対して、情報検索、情報統合機能群を応用することによって、多様な専門分野における教育に関する知識情報源群から、個人の学習に関して、時間的、空間的、意味的に

統合可能にする機能を実現することが有効である [1~4].

本稿では、学習項目間の因果・依存関係計量による学習者個人の学習状況に応じたカリキュラムの動的生成方式について示す。本方式は、学習項目間の静的な順序・依存関係、および、特定の集団内および個人のテスト結果を用いて、その集団のテスト結果あるいは集団の学力傾向を比較することにより、既知の学習項目間の順序・依存関係の強さを計量し、さらに、その依存関係のルールを抽出する。本方式により、既知の学習項目間の順序・依存関係の強さとルールを用いて、個人のテスト結果から、個人別に重要な学習項目とその学習の順序を発見することが可能となる。本方式の実現により、学習項目と学習者個人との関係の計算が可能となり、動的に変化する個人の学習状況と学習項目間との関係計量による個人に特化したカリキュラムを生成し、個人の学習を効果的に支援することが可能となる。

本方式は、具体的に、以下の2つの特徴を有する。

- 集団のテスト結果を分析することにより、学習項目間の順序・依存関係について、関係の強さを確率的に定義し、同時に、その依存関係のルールを定義する。
- 抽出した学習項目間の依存関係の強さ及びルールを用いて、個人のテスト結果を分析し、個人の重要学習項目と優先順位を計量し、個人に特化した学習カリキュラムを生成する。

本方式は、前提として、学習項目間の静的な順序・依存関係、および、テストの設問と学習項目間の関係が既知であるものとする。その上で、同一のコミュニティ内では、学習項目毎の理解の仕方、理解度、項目間の依存関係の強さに、同一の傾向があるという前提を設定している。ここで、特定のコミュニティ内におけるテスト結果を分析することで、既知の学習項目間の順序・依存関係の強さおよびルール抽出し、個人のテスト結果からの個人の重要学習項目および優先順位の計量を実現する。

従来方式の教材提案システムでは、誤答した設問に関連のある教材が出力されるのみで学習者の学習到達度を学習項目間の依存・因果関係を考慮した教材の提案ができていない [5~7]. 事象間の関係やメタデータ間の関係性計量と、ある事象との関係性を考慮した情報検索手法の研究もある [8~10]. これらの手法は、ある事象における原因および結果に関するベクトル間の計量は可能であるが、個々の事象間の因果関係の強さを抽出し、そこから個人に特化した学習カリキュラムを生成するといった使用は想定されていないため、教育および学習状況への適用は困難である。特に、学習項

目間の依存・因果関係性を計量し、その関係量と学習状況間と尾関係を計算する個人に特化した学習カリキュラムを生成する方式として応用することには適していない。協調フィルタリング、アソシエーション・ルールは、単純に同様な傾向を持つ依存関係は抽出してきても、それらの依存関係のルールの分析までは対象としていない [11~13]. 本方式は、正誤判定や正答率、問題解答時間を考慮した教材提案する従来方式とは異なり、学習者の学習状況の知識構造化によりその学習段階、理解状況に合致し、さらにその次の段階の最重要学習項目を提示する新しいカリキュラム生成方式として位置づけることができる。

2. 実現方式

学習項目間の因果・依存関係計量による学習者個人の学習状況適合カリキュラム生成方式を、次の3つの手順によって実現する。

【Procedure-1】学習項目の表現

【Procedure-2】学習状況の把握

【Procedure-3】個人に応じたカリキュラムの作成

2.1. 学習項目の表現

1. 知識体系を学習項目として整理、細分化する

数学の学習項目 (Learning Object) を X について、各項目を整理・細分化し、下記のとおり定義する。

(実験のところで記述した) 文部科学省が策定した学習指導要領 [14] および学習指導要領に基づいた教科書 [15] を参照した。

$$X = \{X_1, X_2, X_3, \dots, X_n\}$$

2. それらの項目を階層構造で表現する

整理・細分化した学習項目を、学習対象として適切な粒度に階層化 (図1) し、下記のとおりに定義する。

$$X_{ijk} \in X_{ij} \in X_i \in X$$

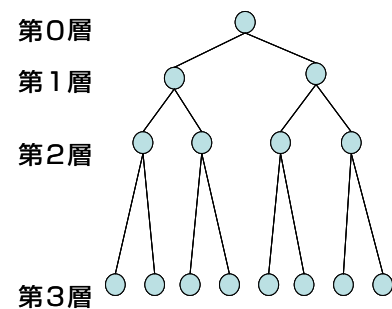


図1 学習項目を粒度に応じて階層化する

3. 学習項目の順序・依存関係を表現する

専門家によって予め定義された知識を用いて、数学の学習項目 (Learning Object) を X について、各項目を整理・細分化し下記のとおり定義する (図2)。

	X_1	X_2	X_3	X_4	...	X_{n-1}	X_n
X_1	0	1	1	0	0	1	1
X_2	0	0	1	0	0	0	0
X_3	0	0	0	1	0	0	1
X_4	0	0	0	0	0	1	0
\vdots	0	0	0	0	0	0	0
X_{n-1}	0	0	0	0	0	0	0
X_n	0	0	0	0	0	0	0

図2 学習項目の順序関係を表現した隣接行列

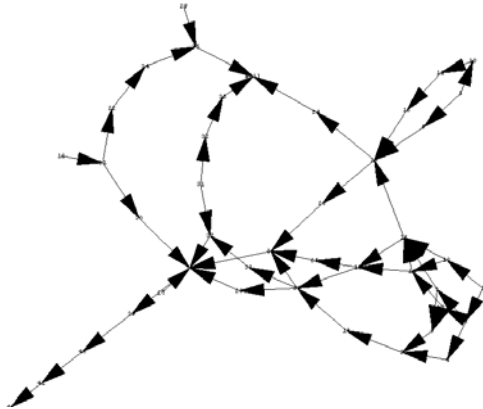


図3 学習項目の順序関係

2.2. 学習状況の把握

1. 学習項目と試験問題との関係の定義

試験問題を作成し、関連する知識を定義する。その関連する知識においてのみ学力を評価する。具体的には、設問ごとに関連する知識領域との関係を定義し、関係があれば1、そうでなければ0で表現する。

question / metadata	X_1	X_2	X_3	X_4	X_5
1	1	1	1	0	0
2	1	1	1	1	0
3	1	0	0	1	1
4	0	1	0	1	0
5	0	1	1	0	0

学習者が与えられた設問に対して、正答したことを1で表現する。そうでない場合は、0で表現する。

$$Q_i \longrightarrow \begin{cases} Q_i = 1 & (\text{正答が得られた場合}) \\ Q_i = 0 & (\text{誤答が得られた場合}) \end{cases}$$

2. 設問ごとに、その設問の正答者と誤答者の比較

試験 ID と関連づけられた学習項目を元に、試験の結果を学習項目の評価値(E)を、次式によって求める。ただし、E はk番目にある学習項目の評価値で、R は評価する学習項目に関連付けられた設問であり、Q は正答した設問数を表している。i は、設問の識別子である。それぞれの除法が、当該学習項目の評価値となる。正答のみを記録する。

$$E_k = \frac{\sum_{i=1}^n Q_i}{\sum_{i=1}^n R_i} \quad (\text{where. } E_k \leq 100)$$

3. 学習者個人の学習状況をベクトル形式での表現

学習者個人の学習項目ごとの正答率を学習状況ベクトルと呼び、下記のとおり学習者個人の学習状況を表現する。

$$PS = \begin{bmatrix} P(X_1) \\ \vdots \\ P(X_n) \end{bmatrix}$$

4. 学習状況ベクトルのうち大きな学習課題(正答率の低い学習項目)から上位3つを重要な学習課題として抽出

3で計量した学習状況ベクトルの学習項目ごとの大きさで、学習課題を個別に取り出す。

$$PS = \{X_1, X_2, \dots, X_n\} \mapsto \max \{X_i, X_{i-1}, X\}$$

2.3. 個人に応じたカリキュラムの作成

1. 学習項目間の順序・依存関係性を計量する。学習項目間の関係性計量により、前手順で発見した最も大きな学習課題に関連の深い学習項目を探索し、重要な学習課題を発見する

Procedure-1で定義した学習項目の順序・依存関係の関係性の強さを定量的に表現し、個人の学習状況との積で重要な学習課題を発見する。

学習項目Xに対して、順序関係で下位にある学習項目 X^- とし、学習者ごとに前後関係のある学習項目の正答状況の関係を計算し、その値を比較することで学習項目間の依存・因果関係を計算する。

[Step-1] 任意の一組の依存関係のある学習項目間の正誤同時生起確率を計算する。正誤が生起する組み合わせは、4通りあり、それぞれを下記のように条件付確率を計算する

(1) $\text{if } X^- = 0 \text{ then } P(X | \text{value} = 1)$

下位の学習項目を誤答した場合、正答する確率 $P(k)$ とする。この確率が高いと、依存・因果関係が弱いと認める。

(2) $\text{if } X^- = 0 \text{ then } P(X | \text{value} = 0)$

下位の学習項目を誤答した場合、正答する確率 $P(e)$ とする。この確率が高いと、依存・因果関係が強いと認める。

(3) $\text{if } X^- = 1 \text{ then } P(X | \text{value} = 1)$

下位の学習項目を誤答した場合、正答する確率

$P(i)$ とする。この確率が高いと、依存・因果関係が強いと認める。

(4) $if X^- = 1 then P(X | value = 0)$

下位の学習項目を誤答した場合、正答する確率 $P(o)$ とする。この確率が高いと、依存・因果関係が弱いと認める。

[Step-2] Step-1の計算結果に複数の計量系(Metrics)を適用し、それぞれに意味的な解釈を加える

$$M = \{m_1, m_2, \dots, m_n\}$$

さまざまな計量系を適用し、解釈を加えることが可能であるが、本稿では、個人の学習状況との関係量を計算するために必要な計量系を適用することを目的とする。特に、下記の計量系を適用し、解釈を加えた。Step-1で計算した確率に対してヒューリスティクスによる解釈を加え、下記のような計量系を適用した。

$$(I) P(e) - P(k) \mapsto \begin{cases} + \\ - \end{cases}$$

$$(II) P(i) - P(o) \mapsto \begin{cases} + \\ - \end{cases}$$

上記の計量系により、依存・因果関係量を計算し、下記のように評価する。

(i) +の場合、依存・因果関係が強い

(ii) -の場合、依存・因果関係が弱い

次に、2つの値を座標軸上に表現することで、4象限による分類し、有効な値を判別することができる。2つの学習項目間に依存・因果関係がある場合、上記の計量系の結果が正の値をとることがヒューリスティクスにより当然と考える。ゆえに、本稿では、第1象限にマッピングされる学習項目の組み合わせを有効な組み合わせと認め、下記のように原点からの距離を関係量として用いる。

$$dependence\ and\ causality(x, y) = \sqrt{(x + y)^2}$$

- (1) 発見された学習課題を、学習順序や関係量に基づき並べ替え、個人に特化したカリキュラムを作成する。学習課題に関連のある学習項目の関係量の平均値をとり、学習者個人の学習課題の関係量とする。
- (2) 学習項目間の依存・因果関係性計量の計量系ごとに解釈を加えずに評価し、学習項目と学習者個人との関係を解釈することができる。個人の学習状況を(1-正答率)で学習課題を表現することで、重み

ベクトルとして学習課題として出力した値との距離計算が可能となる。

2. 学習者個人に特化したカリキュラムを作成する

学習者個人の学習状況に応じて動的にカリキュラムを生成する。大きな学習課題(正答率の低い)とそれらに関連のある学習項目を提示することを特徴とし、下記のとおり学習カリキュラムを生成する。

- ① 大きな学習課題の提示する。
- ② それらに関連のある学習項目を提示する。
 - (a) 学習項目間の依存・因果関係量で並べ替える
 - (b) 学習項目による学習の順序で並べ替える
 - (c) 重要な学習課題で並べ替える

上記の手順により、学習者個人に特化したカリキュラムを動的に生成する。

3. 実験

学習項目間の因果・依存関係計量による学習者個人の学習状況に応じたカリキュラムの動的生成方式について実現可能性を検証する為に以下の実験を行なう。

3.1. 実験環境

実験用データとして、次のとおり、擬似的な履歴データ群を生成した。まず、学習項目群について、中学数学における数学知識領域、特に中学2年生の学習範囲を整理・細分化した学習項目群を用いた[11]。学習指導要領に基づいた数学教科書[12]におけるそれぞれの章につけられた名称を基本データとして用い、数学学習項目を表現した。

試験設問群としては、擬似的な試験設問群を設定し、上記で生成した学習項目と関連付けた。試験結果については、擬似的な学習者で10人分定義した。

3.2. 評価手法

従来方式による学習課題の発見および教材等を提示した結果と、本方式による学習項目と学習者個人との関係性計量の結果を比較することにより、本方式の実現可能性を示す。以下に示す手順によって実際に生成した個人に特化したカリキュラムと、既存方式によるカリキュラムを比較する。

1) 学習項目間の依存・因果関係量を計量する

本実験では、学習項目間の依存・因果関係量を計算した結果で、全部の組み合わせ(図4)と第一象限にマッピングされた学習項目の組み合わせ(図5)を計量し、カリキュラムの生成には、第一象限のみにマッピングされた学習項目の組み合わせのみを採用した。

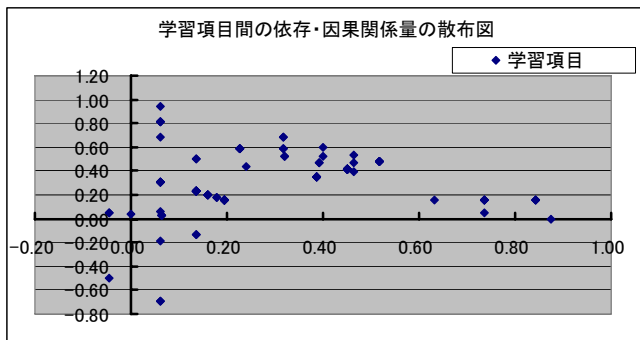


図4 学習項目間の依存・因果関係量の散布図

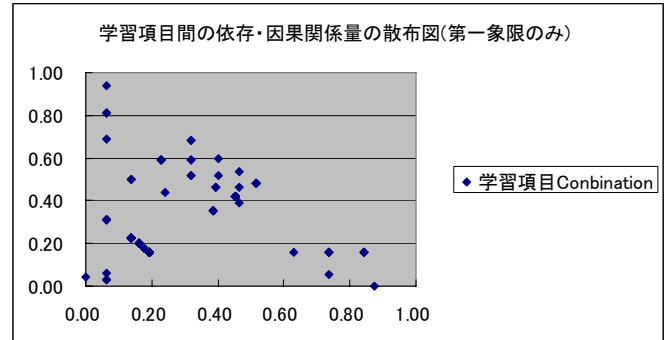


図5 学習項目の依存・因果関係量の散布図(第一象限のみ)

2) 学習者個人の学習状況を計量する

擬似的な学習者個人(L)の学習状況ベクトル(表4)を生成し、その状況ベクトルのうち大きな学習課題(正答率の低い)となっている学習項目を抽出した(表5)。また、学習者Lの学習課題に関連の強い学習課題との関係の強い学習項目を抽出した(表6)。

3) 学習者個人の学習課題を発見する

個人の学習課題と学習項目との関係計量による学習課題の発見するために、関連の強いと認められた学習項目群に対して、学習者個人(L)の学習状況ベクトルから学習課題ベクトル(学習状況の重みベクトル)との距離計算により、学習課題ベクトルを生成した(表7)

4) 学習者個人に応じた学習カリキュラムを生成する

上記の結果から、個人に特化した学習カリキュラムを生成した(表8)。今回は、大きな学習課題(表5)に加え、依存・因果関係の強い学習項目(表7)のうち課題の大きさに上位3つの学習項目をカリキュラムに取り入れた。

3.3. 結果と考察

実験結果を表4~8に示す。ここから次のことが云える。因果・依存関係を考慮した適切な学習項目の優先順位の決定が可能となった。つまり、学習課題間の因果量を決定し学習項目の学習順序を決定することができた。しかし、本実験の課題は、因果・依存関係を考慮する上で有効な値として、第一象限にマッピングされる学習項目間を採用し、原点からの距離を因果量として用いたが、今後、実証研究で今回複数用意した計量系を適用し、経験知から現実に適切な計量系を適用し、本方式を評価、改善していきたい。

学習到達度を考慮した学習課題の選択が可能となった。従来方式による学習課題の提示は、学習課題に関連付けられた教材および学習課題を提示する(表5)とことにとどまり、学習課題の背景や本質的な解決すべき課題を提示することが困難であった。従来方式では、表5にある学習項目のみ独立で取り扱い、それらに定義された教材あるいは学習課題を提示していた。一方、提案手法は、それらの学習課題(表5)に対して依存・

因果関係を考慮したその他学習項目も、学習カリキュラムに組み込むことが可能である。

表3. 中学数学の学習項目

ID	学習項目
1	正の数・負の数
2	文字の式
3	方程式
4	比例と反比例
5	平面図形
6	空間図形
7	式の計算
8	連立方程式
9	一次関数
10	図形の調べ方
11	図形と合同
12	確率
13	平方根
14	式の計算
15	二次方程式

表4. 学習者個人(学習者L)の学習状況ベクトル

ID	学習項目	正答率
1	正の数・負の数	0.733
2	文字の式	0.692
3	方程式	0.667
4	比例と反比例	0.667
5	平面図形	1.000
6	空間図形	1.000
7	式の計算	0.500
8	連立方程式	0.500
9	一次関数	0.600
10	図形の調べ方	1.000
11	図形と合同	1.000
12	確率	1.000
13	平方根	0.500
14	式の計算	0.500

15	二次方程式	0.500
----	-------	-------

表5. 学習者Lの重要な学習課題

ID	重要な学習課題	正答率
13	平方根	0.500
14	式の計算	0.500
15	二次方程式	0.500

表6. 学習課題に関係の強い学習項目

ID	学習項目	関係量
10	図形の調べ方	0.840
9	一次関数	0.763
11	図形と合同	0.763
2	文字の式	0.703
8	連立方程式	0.703
3	方程式	0.279
1	正の数・負の数	0.125

表7. 学習者Lの学習課題に関係の強い学習項目

ID	学習項目	重み	個人	課題
8	連立方程式	0.168	0.5000	0.0840
9	一次関数	0.183	0.4000	0.0732
2	文字の式	0.168	0.3077	0.0517
3	方程式	0.067	0.3333	0.0223
1	正の数・負の数	0.072	0.2667	0.0192
10	図形の調べ方	0.201	0.0000	0.0000
11	図形と合同	0.183	0.0000	0.0000

表8. 個人に特化した学習カリキュラム

学習の順序	学習項目
1	文字の式
2	一次関数
3	連立方程式
4	平方根
5	式の計算
6	二次方程式

4. 結論と展望

本稿では、学習項目間の因果・依存関係計量による学習者個人の学習状況に応じたカリキュラムの動的生成方式について示した。本方式を擬似的なデータに対して適用した実験を行なうことにより、学習項目間の順序・依存関係の強さとルールを用いて、個人のテスト結果から、個人別に重要な学習項目とその学習の順序の発見について、実現可能性を評価した。

本稿において示した方式では、現在の段階では時系列データを考慮しておらず、今後、学習履歴を用いた時系列学習状況分析を導入して行きたい[16]。今後の発展として、実際の学習者を被験者としてデータを集積、

分析をする実証実験に取り組む予定である。

文 献

- [1] Raghavan, V. V. and Wong, S. K. M. A critical analysis of vector space model for information retrieval. *Journal of the American Society for Information Science*, Vol.37 (5), p. 279-87, 1986.
- [2] Perl, J.: "Probabilistic Reasoning in Intelligent Systems," Morgan Kaufmann, San Francisco, CA, 1988.
- [3] Salton, G.: "The S M A R T Retrieval System Experiments in Automatic Document Processing," Prentice Hall Inc., Englewood Cliffs, NJ, 1971
- [4] Salton, G., Wong, A. and Yang, C. S.: "A vector space model for automatic indexing," *Communications of the ACM*, Vol. 18, No. 11, pp.613-620, 1975.
- [5] 三田泰正, 藤岡健司, 萩野哲男, 高田秀行, 上林弥彦, "学習履歴を利用した動的な問題提示を行う学習システムの提案", 電子情報通信学会第15回データベースワークショップ, March.2004
- [6] 佐野洋, "個人適合の学習教材自動生成を実現する語学教育システム", 電子情報通信学会第13回データベース工学ワークショップ, March.2002
- [7] 清水智公, 中村純一, 吉田尚史, 服部隆志, 萩野達也, "個人の知識に応じた教材の自動構成システムの実現," 情報処理学会研究報告, 2002-CE-65, 情報処理学会コンピュータと教育研究会, pp.41-48, 2002.
- [8] 鷹野孝典, 清木康: "異分野データベース群を対象とした意味的検索空間統合プロセスの実現," *DBSJ Letter*, Vol.1, No.1, pp.55-58, 2002.
- [9] 鷹野孝典, 関子泰三, 清木康, "事象間の因果関係を扱う動的な文脈解釈機能を有する意味的連想検索方式の実現" 情報処理学会論文誌: データベース, Vol.46, (TOD25), 2005
- [10] 関子泰三, 清木康, 鷹野孝典, 波内みさ, 但田育直: 事象データ間の因果関連性計量機能をともなったベクトル空間検索方式, 情報処理学会論文誌: データベース, Vol.45, No.SIG 7(TOD22), pp.124-136, 2004.
- [11] Sarwar, Badrul and Karypis, George and Konstan, Joseph and Riedl, John, "Item-based Collaborative Filtering Recommendation Algorithms." In *Proceedings International WWW Conference* (10), Hong-Kong, 2001
- [12] Herlocker, J., Konstan, J., Borchers, A., and Riedl, J. (1999). An Algorithmic Framework for Performing Collaborative Filtering. In *Proceedings of ACM SIGIR'99*. ACM press.
- [13] George Karypis, "Evaluation of Item-Based Top-N Recommendation Algorithms", 10th Conference of Information and Knowledge Management (CIKM), pp. 247-254, 2001
- [14] 文部科学省, "中学校学習指導要領", http://www.mext.go.jp/b_menu/shuppan/sonota/990301c.htm
- [15] 岡本和夫, 小関照純, 森杉馨, 佐々木武ほか 39名, "未来へひろがる 数学1, 2", 啓林館, 東京, 平成17年.
- [16] 小西亮介, 高橋雄介, 清木康, "個人に特化した数学教育環境実現のための学習状況分析システム", データベースワークショップ 2006 (DBWS2006), p27-32, July.2006