

教育用 Web システムにおける 多次元的データマイニングシステムの開発

中富 竜一郎[†] 光吉 和哉[†] 堀江 隆太[†] 杉本 典子[‡]
 崇城大学大学院工学研究科[†], 崇城大学情報学部情報学科[‡]

1. はじめに

現在, 多くの大学では, e ラーニングシステムを導入し, ネットワークを使った教材の提供, 共有, 成績管理, 課題提出, 小テスト等に活用している. それに伴い, 学生による学習活動履歴, 解答履歴, 正答・誤答, 及び教員による指導履歴, 学習用問題等, 教育効果を向上するために有用であると考えられる様々なデータが蓄積されている. これに着目し, e ラーニングに関するデータを対象としたデータマイニングの研究が現在も活発に行われている[1]. しかし, 一方では, 効果的なデータの蓄積方法が確立されていない等の課題が残されているため, 実際の教育活動に有効に活用されているとは言いがたい状況となっている. 本来, 教育効果を向上させるためには, 単に学習時間のみを考慮するだけでは不十分であり, 学生の興味を引く課題の提供, 満足度を向上させるための段階的な学習問題の難易度設定等, 学習用問題の内容や教員による指導・コメントに関するデータまでを考慮することが必要であると考えられる.

本研究の目的は, e ラーニングシステムを活用した教育活動で得られるあらゆるデータを対象としたデータマイニングの手法を開発することである. そのために, 九州大学及び崇城大学で活用されている教育用 Web システム Web Drill[2, 3]を対象としたデータマイニングシステムを構築する. Web Drill では, Web ブラウザを使って, 教員による学習問題を作成, 及び学生による解答が行われる. その際, システム内部に蓄積される学習問題や解答履歴は, 全て XML 形式で保存される. これらを統合し, 学習問題とそれに対する解答履歴を対応付けた XML データを構築する. さらに, 決定木を使って, XML 形式データの分類を行うシステムを構築し, その有効性を検証する.

Development of a Multi-dimensional Data Mining System for Web Based Instruction.

[†]R. Nakatomi, K. Kazuya, R. Horie, · Faculty of Engineering, Sojo University.

[‡]N. Sugimoto · Department of Computer and Information Science, Sojo University.

2. 教育用 Web システム Web Drill

図1に教育用 Web システム Web Drill の構成を示す. Web Drill は, アカウント情報, 学生による解答履歴, 教員によるヒントやコメント等の様々なデータを XML 形式で蓄積している. 本研究では, これらのデータを統合して活用し, 以下の3つの課題を明らかにする.

- (1) e ラーニングシステムにおける効果的なデータ蓄積手法の提案.
- (2) 多次元的データマイニングアルゴリズムの開発.
- (3) 実稼動システムによるアルゴリズムの有効性の検証.

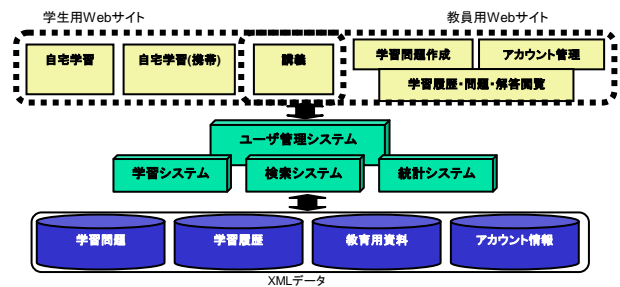


図1. Web Drill の構成

図2と図3に, Web Drill の中で使われている学習問題用 XML データ, 及び学習履歴 XML データの例を示す. これらのデータは, 統合されて蓄積される.

```
<question>↓
<qid>0</qid>↓
<type>text</type>↓
<query>Everyone laughed at me.</query>↓
<hint>みんなは私を笑った。→私は…</hint>↓
<hint2nd></hint2nd>↓
<referenceU></referenceU>↓
<size>32</size>↓
<answers>↓
<ans>↓
<text>I was laughed at by everyone.</text>↓
<feedback></feedback>↓
<correct>1</correct>↓
</ans>↓
↓
<mistake>↓
<text>am I'm</text>↓
<comment>元の文が過去形なので, 受動文も過去形になります。</comment>↓
</mistake>↓
<mistake>↓
<text>-(at)</text>↓
<comment>laugh atで「～を笑う」という句動詞なので, atは省</comment>↓
</mistake>↓
</answers>↓
```

図2. 学習問題 XML データ

```

<log>↓
<log-data>↓
<user>sugimoto</user>↓
<date>2010/11/24</date>↓
<time>13:27:01</time>↓
<answer>I was laughed by everyone.</answer>↓
</log-data>↓
<log-data>↓
<user>sugimoto</user>↓
<date>2010/11/24</date>↓
<time>13:32:08</time>↓
<answer>I was laughed at by everyone.</answer>↓
</log-data>↓
<log-data>↓
<user>sugimoto</user>↓
<date>2010/11/24</date>↓
<time>13:36:01</time>↓
<answer>I was laughed at by everyone.</answer>↓
</log-data>↓
</log>↓
    
```

図 3. 学習履歴 XML データ

3. XML データからのデータマイニング

本研究では、Web Drill に蓄積される XML データの中で使われている任意の 2 つの項目に対して、それらの間の相関関係を見つける手法を提案する。まず、ユーザは、XPath を使って対象とする項目を指定する。さらに、どちらか一方の項目に対する分類条件を指定する。システムでは、前処理として、二つのタグを含む部分木を抽出する。Web Drill における問題のタイプ（選択問題や記述問題等）と、学生による解答数の関係を例としたときに得られる部分木を図 4 に示す。得られた部分木の集合から、指定された分類条件を満たす項目 XML データの特徴を決定木で表現する。学生による問題への解答が 5 回以上行われたかどうかを分類条件としたときに得られる決定木を図 5 に示す。

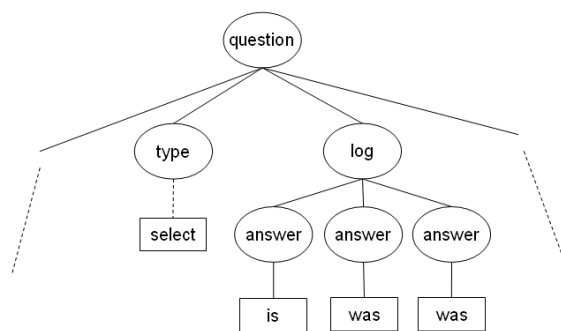


図 4. 部分木

4. まとめと今後の課題

XML データを対象とした 2 次元データマイニング手法を提案し、教育用 Web システム Web Drill を活用して、その有効性を検証した。e ラーニングシステムにおけるデータマイニングにおいて、これまでは、学習履歴や学生の成績等の限定さ

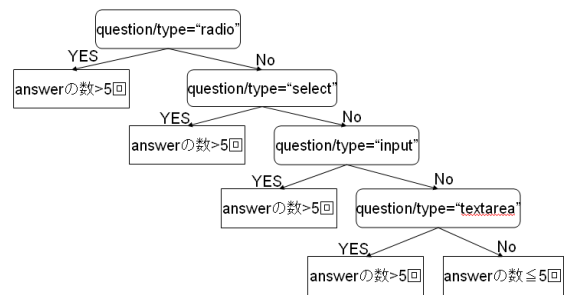


図 5. 決定木による XML データの特徴付け

れたデータのみを扱うに留まっていたが、本研究では、学習用問題そのものが XML 形式で蓄積され、活用されている Web Drill を対象とすることにより、与えられる学習問題の形式やその難易度と、学習到達度の関係など、教員が学習者の理解度を向上させるために有益な情報を抽出することが可能となった。

今後の課題としては、多様な属性を持つ XML データから、学習効果に有益と考えられる二つの属性を自動的に見つけ出す手法の開発が挙げられる。本研究で提案した手法は、データマイニングの対象となる属性を教員が選ぶしくみになっており、教員がある程度の予測をもってデータマイニングを行うこととなるため、教員自身が認識していない教育上の問題点を見落とす可能性がある。そこで、データの中から頻出する属性の組を前処理で見つけ出すしくみが必要となることが考えられる。さらに、文書やスライド等の形式で活用されている教育用教材から Web Drill の形式へ自動変換する機能を実装し、システムの利便性を向上させることが必要と考える。

参考文献

[1] 植野正臣：e ラーニングにおけるデータマイニング，日本教育工学会論文誌，No. 31，Vol. 3，pp. 271-283，2007.
 [2] 田畑義之，杉本典子：外国語学習教材作成・管理システム Web Drill の構築，ドイツ語情報処理研究(2007).
 [3] 田畑義之：オンライン外国語学習システム Web Exercise-可能性と課題-，言語科学，第 37 号，九州大学大学院言語文化研究院言語研究会，2002.