

シラバスから行うカリキュラム分析

The Curriculum Analysis Using the Syllabus

三好 善彦

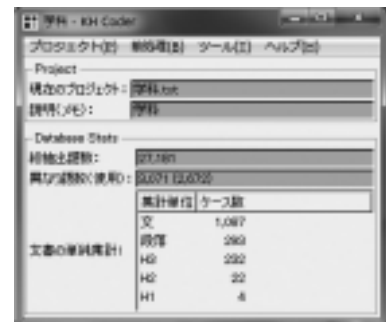
MIYOSHI Yoshihiko

In all universities and colleges, we can see all class contents, such as lectures, exercises and practical skills, in their syllabi now. These class contents are decided based on the curriculum. In this paper, I analyzed the curriculum by using summaries of all classes in our syllabus of Saitama W. J. College.

1. はじめに

埼玉女子短期大学には、2011年度現在、商学科と国際コミュニケーション学科があり、各学科にはそれぞれ4つのコースと両学科に共通した3つのコースの合計11のコース¹が設置されている。そして、カリキュラムは、一般教育科目としてキャリア選択科目と教養選択科目、各学科の専門科目として学科基本科目と英語科目（国際コミュニケーション学科のみ）と各コース科目がある。また、コースが違って共通するテーマや目標もあるため、同じ科目が複数のコース科目になっている場合もある。例えば、『身体表現とコミュニケーション』は『ビューティー・キャリアコース』『心理と健康コース』『スポーツ・パフォーマンスコース』の3つのコース科目となっているなどである。

今回、カリキュラムとコースの関連性やカリキュラムの分類をするために、科目名称のみではなく、授業内容からそれらを分析することとした。これらの分析は、シラバスに記載されている各授業の概要に対してテキストマイニング²を行い、各コース間や各科目間の関連性を分析するため対応分析を行った。また、これらの分析にはフリーソフ



統計単位: クラス数	
文	1,087
単語	280
H0	200
H2	20
H1	4

トウェアである「KH Coder³」を用いた。

2. 分析データ

分析するデータはシラバスに記載されている「授業概要」を用いることとし、これらは全て本学 Web シラバス⁴から取得した。データの記述は、「一般教育科目」「商学科科目」「国際コミュニケーション学科科目」「共通コース科目」を<見出し1>、「キャリア基礎選択科目」、一般教育科目の各分野、各学科基本科目と各コース科目を<見出し2>、各授業科目名称を<見出し3>としている。各見出しは、<H1><H2><H3>を用いて html 文章と同様のタグを付ける。以下に、データの一部を記載する。

<H1>一般教育科目</H1>

<H2>キャリア基礎選択科目</H2>

<H3>コンピュータ A</H3>

コンピュータは単なる情報処理のためのツールから新しいコミュニケーションツールとして、またビジネスの現場においても必要不可欠なものになっている。この授業ではコンピュータの持っている可能性を理解するためにコンピュータの仕組みや利用方法について解説する。さらに、ビジネスの現場において必要となるワープロを使用した文書の作成方法およびスライドを使用したプレゼンテーションの方法を習得することも目標とする。

…中略…

<H1>商学科科目</H1>

<H2>商学科学科基本科目</H2>

<H3>現代商学</H3>

生産、流通、消費に分けられている経済活動において、商学とは財やサービス、アイデアなどの商品が生産者から流通業者を通じて消費者に行きわたる商業全般を学ぶ学問である。この授業では、商品の取引および流通システム、流通を担っている卸・小売業者の機能・構造などを主に学習する。商学の学習は「マーケティング」

関連科目の理解を高めると共に、経済活動の理解につながる。

…以下略…

分析において、以下の図のように集計単位を「H1」にすれば学科に対する関連性、「H2」にすればコースに対する関連性、「H3」にすれば授業科目に対する関連性を調べることができる。



しかしながら、授業数（＜見出し3＞）が232科目にもものぼるこの分析データを用いた場合、結果として表示される項目数があまりにも多すぎるため授業科目についての関連性を調べることが容易でなかった。そのため、授業科目についての関連性は、コースごとに分析データを別途作成し、授業科目名称を＜見出し1＞として分析を行い、コース内での授業科目の関連性を調べることとした。

3. 学科の特色

はじめに、学科の特色を分析するために「一般教育科目」「商学科科目」「国際コミュニケーション学科科目」「共通コース科目」についての特徴的な言葉を「関連語探索」を用いて行ったところ次のような結果が得られた。

関連語探索

Search Entry
 コード選択
 辞書選択
 検索範囲入力
 検索条件
 検索結果表示
 国際コミュニケーション学科

コード入力
 辞書名
 検索範囲
 検索条件
 検索結果表示
 国際コミュニケーション学科

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名

辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名
 辞書名</

各学科を特徴づける語

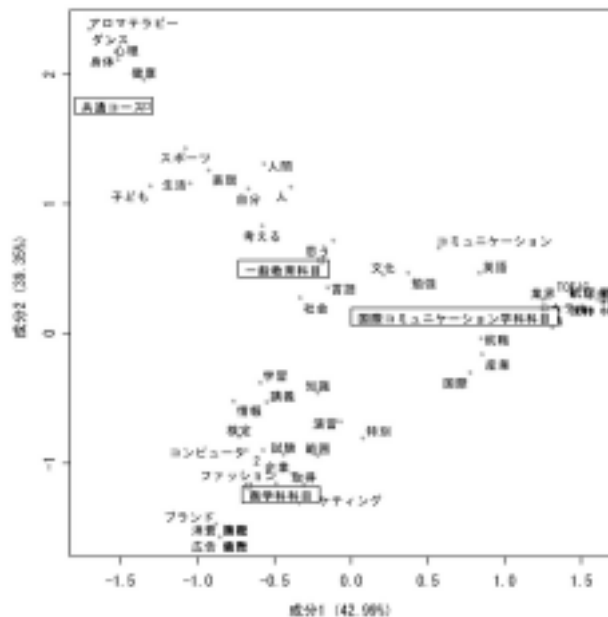
N	一般教育科目		商学科科目		国際コミュニケーション学科科目		共通コース科目	
1	文化	0.1272	講義	0.1303	英語	0.1329	健康	0.1212
2	授業	0.1065	ファッション	0.1291	観光	0.1226	身体	0.1146
3	社会	0.0922	簿記	0.1277	ホテル	0.1034	心理	0.1111
4	英語	0.0915	知識	0.1198	コミュニケーション	0.0831	ダンス	0.1064
5	思う	0.0701	医療	0.1184	業界	0.071	アロマセラピー	0.1064
6	言語	0.0532	企業	0.1118	旅行	0.0692	学ぶ	0.1033
7	日本	0.0519	学ぶ	0.0964	国際	0.0646	表現	0.0971
8	勉強	0.0448	学習	0.092	日本	0.0616	生活	0.0966
9	必要	0.0447	検定	0.0909	目指す	0.0605	考える	0.0841
10	目標	0.0441	基礎	0.0874	理解	0.0604	行う	0.0605
11	人	0.0399	ブランド	0.0807	ブライダル	0.0564	子ども	0.0567
12	問題	0.0397	理解	0.0776	就職	0.0545	人間	0.055
13	学生	0.0385	試験	0.0723	エアライン	0.0535	自分	0.0493
14	フランス語	0.0385	販売	0.0717	身	0.0525	人	0.0488
15	生活	0.0382	演習	0.0678	学生	0.0503	心	0.0474

数値は Jaccard の類似性測度

この結果をもとに、各学科の特徴をまとめると以下のように考察することができる。

- 「一般教育科目」では【文化】【社会】【英語】【言語】などの語が特徴的に使われており『一般教養としての特徴』がよくあらわされていることがわかる。
- 「商学科科目」では【ファッション】【簿記】【医療】【企業】などの語が特徴的に使われており『より実務的な特徴』がよくあらわされていることがわかる。
- 「国際コミュニケーション学科科目」では【英語】【観光】【ホテル】【コミュニケーション】【業界】【旅行】【国際】などの語が特徴的に使われており『より国際的な業界という特徴』がよくあらわされていることがわかる。
- 「共通コース科目」では【健康】【身体】【心理】【ダンス】【アロマセラピー】などの語が特徴的に使われており『より人間的な部分に対する特徴』がよくあらわされていることがわかる。

また、より視覚的に全体の特徴をつかむために「集計単位を<H1>」として「対応分析」を行うと以下の通りとなった。



この結果から、「商学科科目」「国際コミュニケーション学科科目」「共通コース科目」が3方向に離れた位置に分布しており、「一般教育科目」がほぼ中心に位置していることがわかる。「関連語検索」の時と同様に、それぞれについての特徴を考察すると以下のようにまとめることができる。

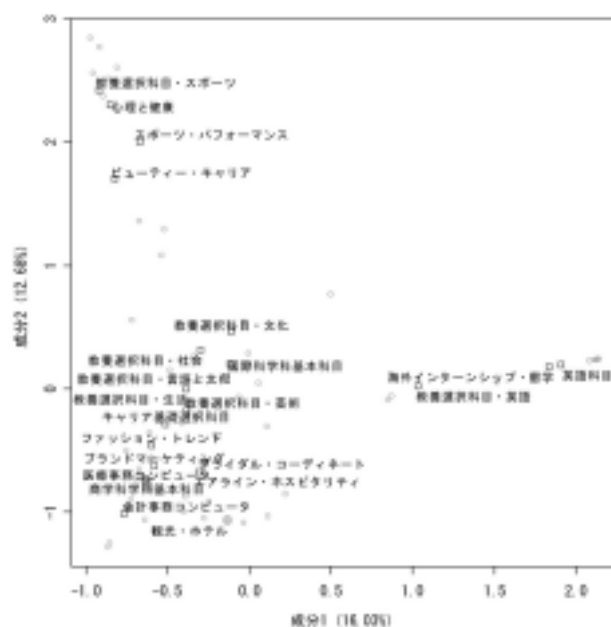
- 中心に位置している「一般教育科目」は全体の中で目立った特徴がないと判断できる。これは、一般教育科目の性質上当然のことであると考えられる。
- 下方方向に位置している「商学科科目」に対しては、関連語探索で出てきた語以外にも【検定】【販売】【会計】などの語も特徴的な語として見られ、『より実務的な特徴』を示している。
- 右方向に位置している「国際コミュニケーション学科科目」に対しては、同様に【航空】【TOEIC】などの語も特徴的な語として見られ、『より国際的な業界という特徴』を示している。
- 上方方向に位置している「共通コース科目」に対しては、同様に【子ども】【スポーツ】などの語も特徴的な語として見られ、『より身体的な部分に対する特徴』を示している。

以上の分析と考察からそれぞれの学科の特徴をシラバスから導き出すことができた。これらの特徴は、学校紹介において示されている各学科の特徴⁵と同じであることがわかり、本学のカリ

キュラムは全体的に学科設置の目的に合致していると判断できるであろう。

4. コースの特色

次に、コースの特色を分析するために、「集計単位を<H2>」として対応分析を行うと以下の通りとなった。ここでは、コース間の特徴を分かりやすくするため「変数のみ」の表示としている。



全体の特徴として、「教養選択科目」「キャリア基礎選択科目」といった『一般教育科目』は中心に配置されていて基本的な部分を担っていることがわかる。また、「海外インターンシップ・留学」「英語科目」「教養選択科目・英語」といった『英語』に関する部分が右の方に配置されていて【TOEIC】【英語】という特徴を表している語が見られた。さらに、「教養選択科目・スポーツ」「心理と健康」「スポーツ・パフォーマンス」「ビューティー・キャリア」といった『共通コース』に関する部分が上の方に配置されていて【心理】【健康】【身体】などより人間的な部分の特徴を表している語が見られた。

商学科の「ファッション・トレンド」「ブランドマーケティング」「会計事務コンピュータ」「医療事務コンピュータ」および国際コミュニケーション学科の「観光・ホテル」「ブライダル・コ

ーディネート」「エアライン・ホスピタリティ」の各コースについては、同じような位置にありコースによる大きな特徴が表れることがなかった。あえて特徴をあげるとすれば、それら全体としては左下の方に配置されていて【医療】【会計】【ホテル】【航空】など職業に関する特徴を表している語が見られた。その中で、「商学科学科基本科目」は商学科の4つのコースと同じような位置に配置されているので、コースの専門性を取り入れた形の学科基本科目となっていると判断でき、全体として商学科カリキュラムという一つ枠に収まっていることがわかる。それに対し、「国際コミュニケーション学科学科基本科目」（図中は、「国際科学科基本科目」）は『一般教育科目』群の中に位置していることから、どちらかという、コースの専門性よりも一般教養的な部分を取り入れた形の学科基本科目となっていると考えられる。

また、成分1（横方向）について着目すると、商学科の各コースは左側に位置し、国際コミュニケーション学科の各コースは右側に位置していることがわかる。これらに対し、キーワードとなる語を見ると、商学科は【試験】【検定】【講義】などの語が見られ、国際コミュニケーション学科は【業務】【就職】【業界】などの語が見られている。このことから、各学科のコースの特徴として、商学科は『卒業後の進路としての職業を見据えたうえで資格取得も目標としているコース設定』となっており、国際コミュニケーション学科は『卒業後の進路を最初から意識したコース設定』となっていると考えることができる。

以上の考察より、商学科の学科基本科目およびコースについては、全体的にまとも商学科の特徴がよく表れていると考えられる。しかし、国際コミュニケーション学科の学科基本科目および海外インターンシップ・留学コースについての科目設定は、他のコースと大きく離れていて学科としてのバランスを崩していると考えられる。特に、『海外インターンシップ・留学コース』⁶については検討する必要があると考えられる。

5. 科目の関連性

これまでに、学科とコースの分析を行い、それぞれにおける特徴を見出したり問題点を指摘したりすることができた。シラバス分析の最後として、科目の関連性の考察を行うこととする。ここでは、コース内の科目設定についての分析を行って科目設定のバランスや類似性を検討することになる。

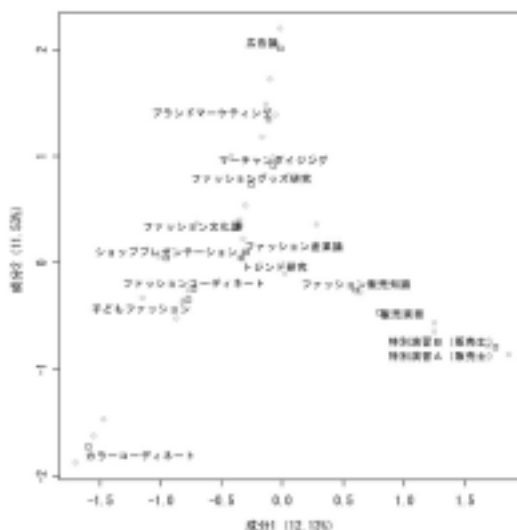
この分析に関しては、コースごとに分析データを作成し、授業科目名称を〈見出し1〉として

分析を行った。以下に、各コースにおける対応分析の結果と考察を示す。また、全体を通して比較した結果から、成分1（横軸）と成分2（縦軸）の値が ± 2 の範囲内に入っているものについては科目ごとの関連性があり、 ± 2 を超えるものについては関連性が薄く、 ± 3 を超えるものについては関連がないと考えられると判断した。

5.1 ファッション・トレンドコース

ファッション・トレンドコースについては、全体的にまとまっておりコースの目的に合わせた科目設定がされていると考えることができる。中でも、「特別演習 A（販売士）」「特別演習 B（販売士）」「販売実務」「ファッション販売知識」といった科目で、【資格】【販売】【検定】などの語がキーワードとして見られ『資格取得』に関するグループとなっていて、「広告論」「ブランドマーケティング」「マーチャダイジング」「ファッショングッズ研究」といった科目で、【消費】【製品】

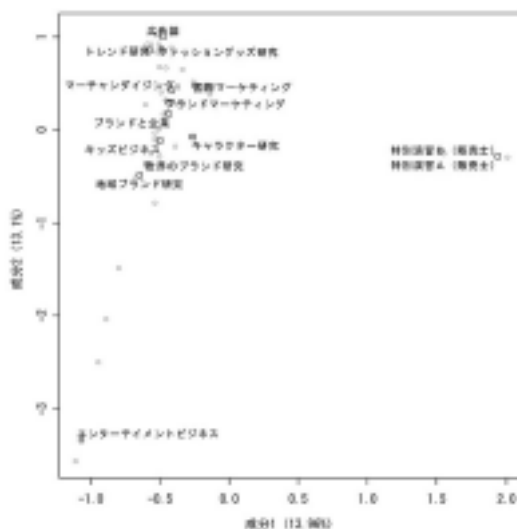
【企業】といった語がキーワードとして見られ『ビジネス』に関するグループとなっていると考えることができる。また、ファッションに関する科目が中心的に配置されており、全体として3つのグループがバランスよく配置されていると考えることができる。「カラーコーディネート」が他の科目と離れた左下に配置されているが、成分1と成分2とも ± 2 の範囲内に入っているため、コースにおける科目の関連性は保たれていると考えることができる。



5.2 ブランドマーケティングコース

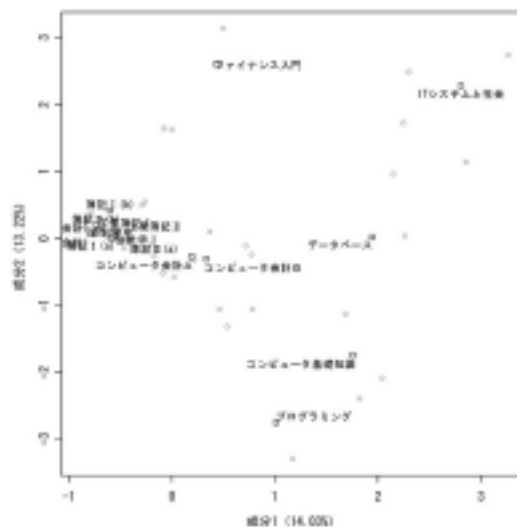
ブランドマーケティングコースについては、【マーケティング】【ブランド】といった語がキーワードとして見られほとんどの科目が中心に位置している。また、特徴ある科目としては、ファッション・トレンドコースと共通科目である「特別演習 A（販売士）」「特別演習 B（販売士）」は、【販売】【資格】などの語がキーワードとして見られ『資格取得』を明確に示していると考えられる。

左下に位置する「エンターテインメントビジネス」は、【楽しむ】【映画】【娯楽】といった語がキーワードとして見られ他のコース科目との関連性が見られないと考えることができる。



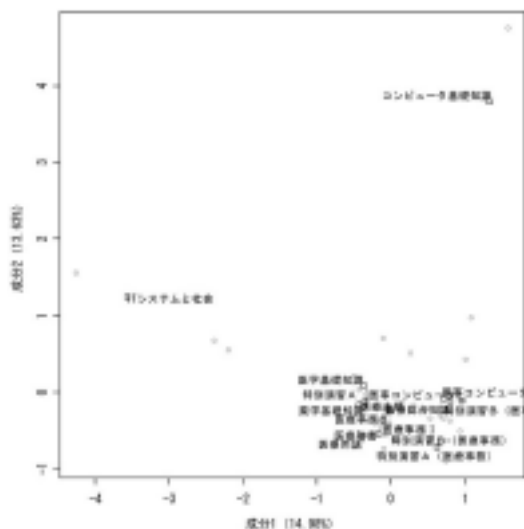
5.3 会計事務コンピュータコース

会計事務コンピュータコースについては、コースの目的となる簿記や会計に関する科目が中心に多く位置している。中心から離れた位置に「プログラミング」「コンピュータ基礎知識」「データベース」「IT システムと社会」といった簿記や会計との関連性が薄いコンピュータなど IT に関係する科目が配置されている。また、【金融】という語がキーワードとして見られる「ファイナンス入門」も離れた位置に配置されており関連性が薄いと考えることができる。



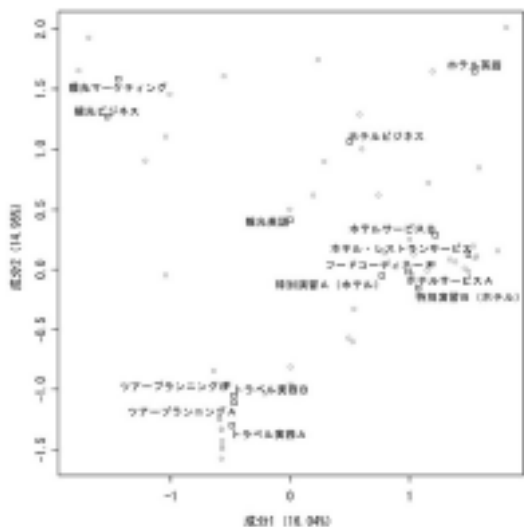
5.4 医療事務コンピュータコース

医療事務コンピュータコースについては、会計事務コンピュータコースと同様にコースの目的となる医療に関する科目が中心に多く配置されている。中心から離れた位置には、会計事務コンピュータコースと共通科目となっている「IT システムと社会」「コンピュータ基礎知識」といった IT に関係する科目が配置されており、コースにおける科目の関連性がないと考えることができる。



5.5 観光・ホテルコース

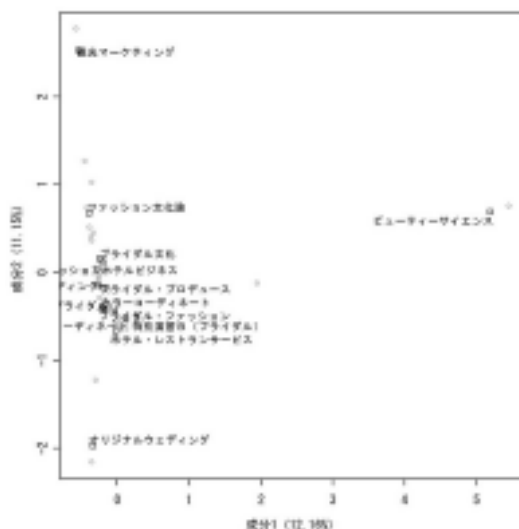
観光・ホテルコースのついては、全体的にまとまっておりコースの目的に合わせた科目設定がされていると考えることができる。その中でも、【ホテル】【レストラン】【サービス】【接客】などの語がキーワードして見られるホテル関係、【旅行】【宿泊】などの語がキーワードとして見られる旅行関係、【マーケティング】【ビジネス】などの語がキーワードとして見られるビジネス関係の3つの科目群がバランスよく配置されていると考えることができる。



5.6 ブライダル・コーディネートコース

ブライダル・コーディネートコースについては、【ブライダル】【婚礼】【ドレス】【ウェディング】などの語がキーワードとして見られ、コースの目的であるブライダルに関する科目が中心に多く配置されている。

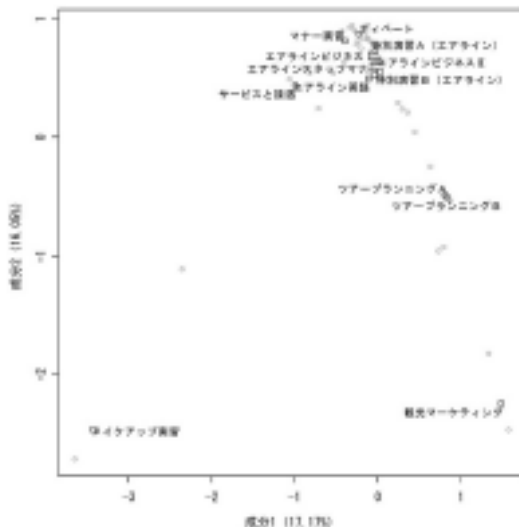
また、左下に位置している「オリジナルウェディング」は成分1と成分2の両方とも±2の範囲内に入っているためコースにおける科目の関連性は保たれていると考えられるが、左上に位置している「観光マーケティング」は【観光】【マーケティング】といった語がキーワードとして見られるようにコースにおける科目の関連性が薄いと考えることができる。さらに、ビューティー・キャリアコースとの共通科目となっている「ビューティーサイエンス」は、他の科目とは大きくかけ離れた右に位置しておりコースにおける科目の関連性がないと考えることができる。



5.7 エアライン・ホスピタリティコース

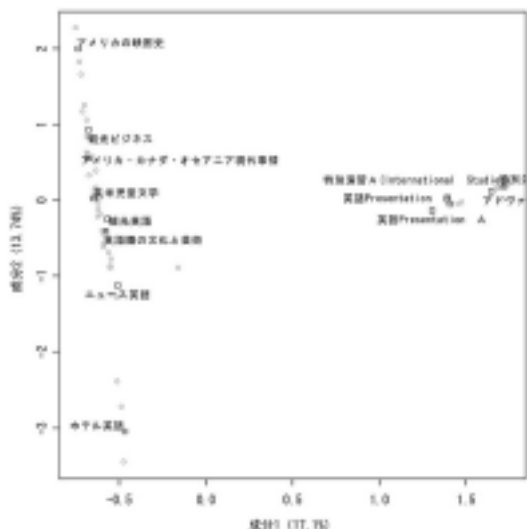
エアライン・ホスピタリティコースについては、【ホスピタリティ】【マナー】【エアライン】【コミュニケーション】といった語がキーワードとして見られるなどコースの目的であるエアラインに関する科目が中心に多く配置されている。

また、【旅行】【ツアー】などの語がキーワードとして見られる「ツアープランニング A」「ツアープランニング B」については、少し離れた位置にあるがコースに関連する科目として考えることができる。しかし、【観光】【マーケティング】という語がキーワードとして見られる「観光マーケティング」は関連性が薄いと考えられるし、【化粧】【飾る】という語がキーワードとして見られ左下に位置しているビューティー・キャリアコースとの共通科目である「メイクアップ実習」はコースにおける科目の関連性がないと考えることができる。



5.8 海外インターンシップ・留学コース

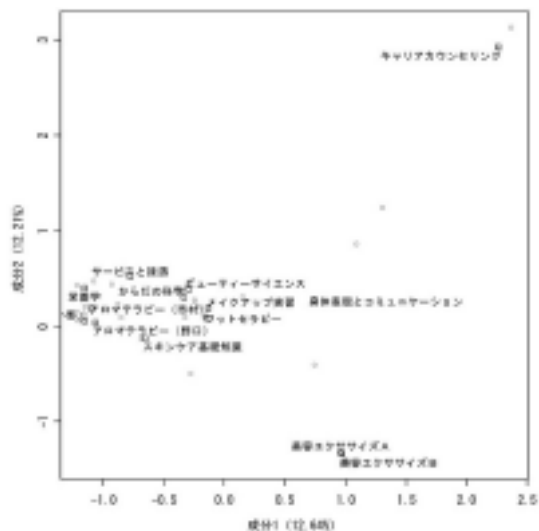
海外インターンシップ・留学コースについては、「特別演習 A (International Studies)」 「特別演習 B (American Studies)」 「英語 Presentation A」 「英語 Presentation B」 「アドヴァンスド・オーラルイングリッシュ」のシラバスが英語で書かれているため、正確な関連性の分析を行うことができない。これを踏まえたうえで科目の関連性を分析すると、【ホテル】【英会話】【フレーズ】といった語がキーワードとして見られる観光・ホテルコースとの共通科目である「ホテル英語」は関連性が薄いと考えられる。



5.9 ビューティー・キャリアコース

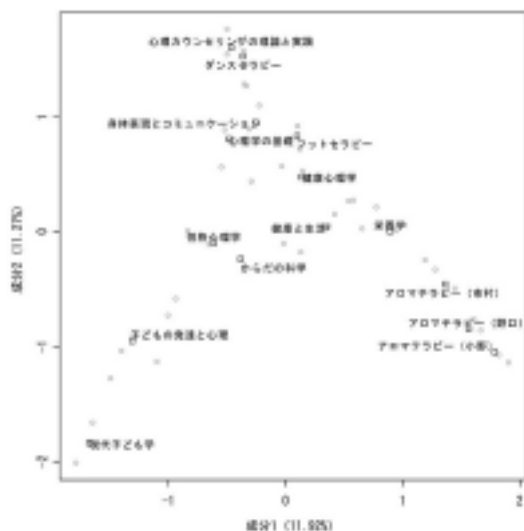
ビューティー・キャリアコースについては、【美しい】【健康】【考える】【変化】といった語がキーワードとして見られるなどコースの目的である美容に関する科目が中心に多く配置されている。

また、【ヨガ】【ボディ】などの語がキーワードとして見られる「美容エクササイズ A」 「美容エクササイズ B」については、右下の少し離れた位置にあるがコースに関連する科目として考えることができる。しかし、【ゲーム】【カウンセリング】という語がキーワードとして見られる「キャリアカウンセリング」は右上の大きく離れた場所に位置しておりコースにおける科目の関連性がないと考えることができる。



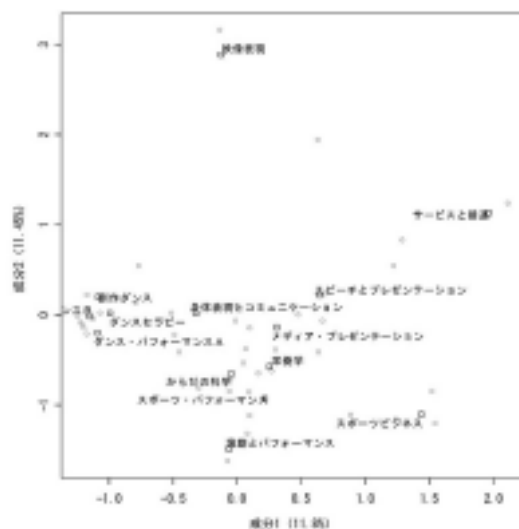
5.10 心理と健康コース

心理と健康コースについては、全体的にまとまっておりコースの目的に合わせた科目設定がされていると考えることができる。その中でも、【カウンセリング】【セラピー】【コミュニケーション】などの語がキーワードして見られる上部に位置する『心理関係』、【植物】【食品】などの語がキーワードとして見られる右下に位置する『健康関係』、【成長】【発達】などの語がキーワードとして見られる左下に位置する『子供関係』の3つの科目群がバランスよく配置されていると考えることができる。



5.11 スポーツ・パフォーマンスコース

スポーツ・パフォーマンスコースについては、【ダンス】【身体】【パフォーマンス】【プレゼンテーション】などの語がキーワードとして見られなどコースの目的である『スポーツ』および『表現』に関する科目が中心に多く配置されている。【映画】【子供】などの語がキーワードとして見られる「映像表現」は、他の科目とは大きくかけ離れた上部に位置しておりコースにおける科目の関連性がないと考えることができる。



6. まとめ

今回、これまで示したとおり、KH Coder を利用した対応分析によりシラバスからカリキュラムの分析を行うことができた。ここでは、『KH Coder チュートリアル』⁷⁾に基づき一連の分析を行っている。そこでは、「2段階での計量テキスト分析」として以下のように2段階での分析が紹介されている。

- 段階1 データ中から言葉を自動的に取り出して計量的分析を行うことで、分析者の予断を極力含めない形でデータを探索・提示する。
- 段階2 コーディングルールを作成して、分析者が明示的にデータ中の概念（コンセプト）を取り出すことで、分析を深めることを目指す。

「KH Coder チュートリアル」2ページ

今回のカリキュラム分析においては、この2つの段階の中でも最初の「段階1」を中心に行った。今後、このカリキュラム分析を発展させ、シラバス以外にも様々な資料をもとに行うようになれば、「段階2」にあるように明示的にカリキュラムのコンセプトを取り出すような分析に進むことが可能であると考えている。

この「段階1」の分析のみでも、各学科については、『より実務的な特徴』を持つ商学科と『より国際的な業界という特徴』を持つ国際コミュニケーション学科であると分析することができ、各コースについては、『卒業後の進路としての職業を見据えたうえで資格取得も目標としているコース設定』をしている商学科と『卒業後の進路を最初から意識したコース設定』をしている国際コミュニケーション学科というように学科間では多少目標設定が異なっていると分析できた。また、各科目においてはコースの目的に合わせてうまく配置されていることが分析できたが、いくつかの科目がコースにおける科目の関連性がないことも分析できた。ただし、ここで関連性がないと分析できた科目でも、資格取得など目標のために設置している場合などがあることも忘れてはならない。そして、これらの分析結果を参考にすることにより、今後のカリキュラム作成や募集活動のヒントを得ることができるであろう。

注

1. 2011年度は、商学科には『ファッション・トレンド』『ブランドマーケティング』『会計事務コンピュータ』『医療事務コンピュータ』の4コース、国際コミュニケーション学科には『観光・ホテル』『ブライダル・コーディネート』『エアライン・ホスピタリティ』『海外インターンシップ・留学』の4コース、両学科共通として『ビューティー・キャリア』『心理と健康』『スポーツ・パフォーマンス』の3コースがある。
2. テキストマイニングとは、『定型化されていない文章の集まりを自然言語解析の手法を使って単語やフレーズに分割し、それらの出現頻度や相関関係を分析して有用な情報を抽出する手法やシステム。(IT用語辞典 e-Words, URL <http://e-words.jp/>より。)』である。また、Yahoo!や Google などで検索すると多くの Web ページがヒットし、様々な文献や資料を閲覧することが可能である。
3. KH Coder Ver. 2.beta.27, 2011/11/26, URL <http://khc.sourceforge.net/>, 樋口耕一, ko-ichi@ss.ritsumei.ac.jp
4. http://sws.saijo.ac.jp/public/web/Syllabus/WebSyllabusKensaku/UI/WSL_SyllabusKensaku.aspx から検索できる。
5. 「2011年度埼玉女子短期大学ガイドブック」によれば、商学科は「情報通信分野の発展、海外企業との提携の増加など社会が新たな変貌を遂げていくなか、豊かな知識と確かな実務能力を養い、多彩なビジネスシーンで活躍できる女性を育成します。」(25ページ)、国際コミュニケーション学科は「国際化、情報化社会で活躍するため、英語力とコンピュータの運用能力やビジネスマナーを修得、次代の異文化交流の担い手となり、グローバルな視野で社会に貢献する人材を育成します。」(33ページ)、両学科共通コースは「実践的な知識を学びながら、仕事を進めていくうえで必要なコミュニケーションのあり方を美容、心、スポーツそれぞれの視点から考えます。女性が最も魅力的に輝く学びを探ります。」(41ページ)、一般教育科目は「日本や海外の言語と文化、社会、生活、芸術、スポーツなどに種別された科目を幅広く設定、社会人としての基礎となる教養を身につけて、豊かな人間性を育みます。」(47ページ)と紹介されている。
6. 2011年度は国際コミュニケーション学科に海外インターンシップ・留学コースを設置しているが、2012年からはコース改変を行い、同コースを廃止する予定である。
7. KH Coder チュートリアル, 2011/11/26, <http://khc.sourceforge.net/> よりダウンロード可

平成 18 年度
プロジェクト研究報告書

シラバス可視化のための シラバス情報抽出支援システム

1070383 刈谷 悠

指導教員 妻鳥 貴彦

2007 年 3 月 9 日

高知工科大学 情報システム工学科

要 旨

シラバス可視化のための シラバス情報抽出支援システム

刈谷 悠

近年，多くの大学でシラバスが導入されている．シラバスとは，講義内容，講義計画，評価方法などの講義に関する情報を事前に学生に示す冊子である．最近では Web 上でシラバスを公開する形が増加してきている．

学生はシラバスを読むことで講義内容を知り，履修計画を立てる．しかし，現在のシラバスは科目間に存在する関係が読み取りづらいことや，各大学のシラバス同士の比較が困難であるという問題点が指摘されている．これらの問題の解決方法として，各大学のシラバスの科目間の相互関係を可視化する方法が考えられている．そのためにはシラバスの情報を抽出し，その情報を体系的にまとめる必要があるが，シラバスから手動で情報抽出を行うには多大なコストが掛かってしまう．

そこで本研究では，ユーザが Web 上で公開されているシラバスからシラバス情報を抽出するためのシステムを構築し，抽出したシラバス情報を体系的に記述してファイル出力する方法を提案する．これにより，抽出したシラバス情報をシラバスの可視化に活用できるようになる．

キーワード シラバス，体系的記述，シラバスの可視化，シラバス情報の抽出

目次

第 1 章	はじめに	1
第 2 章	研究背景	2
2.1	シラバスとは	2
2.2	シラバスの問題点	3
2.2.1	シラバス上の暗黙情報	3
2.2.2	シラバスの構成要素の相違	4
2.3	問題の解決方法	8
2.4	シラバス可視化システム	9
2.4.1	シラバスの可視化方法	9
2.4.2	システム構成	9
2.4.3	システム概要	10
2.5	シラバス情報の抽出	11
2.6	研究の目的	11
第 3 章	オントロジーによるシラバスの体系的記述	12
3.1	オントロジーとは	12
3.2	本研究におけるオントロジーの役割	13
3.3	オントロジーを導入する利点	13
3.4	データ記述要素	14
第 4 章	シラバス情報抽出支援システムの設計	16
4.1	対象とするシラバスの形式	16
4.2	シラバス情報の抽出方法	17

目次

第 5 章	シラバス情報抽出支援システムの実装	20
5.1	システムの構成	20
5.2	システムの概要	21
5.3	システムの実行方法	21
5.3.1	シラバスのトップページのアドレス入力	21
5.3.2	シラバスページへのリンクの選定	22
5.3.3	テキストと構成要素の関連付け	22
5.3.4	シラバス情報抽出結果の確認	23
5.3.5	出力ファイル表示	25
5.3.6	出力ファイルの記述結果	25
第 6 章	評価	27
6.1	評価方法	27
6.2	結果	27
6.3	考察	28
第 7 章	まとめ	31
謝辞		32
参考文献		33

目次

2.1	科目間に存在する関係	4
2.2	プレレキジット表	5
2.3	高知工科大学のシラバス	6
2.4	山形大学のシラバス	7
2.5	シラバス可視化システム：起動時	10
2.6	シラバス可視化システム：可視化例	11
4.1	情報抽出の方法	18
5.1	トップページのアドレス入力	22
5.2	情報抽出対象の選定	23
5.3	テキストと構成要素の関連付け	24
5.4	情報抽出結果の確認	24
5.5	出力ファイルの表示	25
5.6	XML ファイルの記述結果	26
5.7	抽出元のシラバスページ	26
6.1	情報抽出に成功したシラバス	29
6.2	情報抽出に失敗したシラバス	30

表目次

4.1 シラバス形式の調査結果	17
---------------------------	----

第 1 章

はじめに

近年，多くの大学でシラバスが導入されている．シラバスとは，講義内容，講義計画などの講義に関する情報を事前に学生に示す冊子であり，最近では Web 上でシラバスを公開する形も多くなっている．学生はシラバスを読むことで講義内容を知り，履修計画を立てる．

しかし，現在のシラバスにはいくつかの問題点が指摘されている．まず，科目間に存在する関係が読み取りづらい点である．多くのシラバスでは，各科目に関しての情報しか記述されていない．しかし，実際には各科目はそれぞれが独立して存在しているのではなく，他の科目の基盤となる科目や同じ知識を用いる科目など，他の科目と何らかの関係を持っている．これらの関係は学生が履修計画を立てる上で必要な情報であるため，これらの関係が読み取りづらいことが問題となっている．次に，シラバス同士の比較が困難であるという点である．シラバスには標準的な記述項目・記述形式といった記述方法が明確に定められていないため，各大学が独自に記述方法を決定し，シラバスを作成している．シラバスにはその大学のカリキュラムの特色などの情報が多く含まれている．そのため，他大学のシラバス同士を比較することにより，その大学のカリキュラムの特色などを相対的な形で明らかにすることができ，学生の学習意欲の向上に繋がれると考えられる．

これらのシラバスの問題点を解決する方法として，各大学のシラバスを可視化するという方法が考えられている．シラバスの可視化は，シラバスのデータ項目から暗黙的な情報である科目間に存在する関係を可視化する．しかし，そのためにはシラバスの情報を抽出し，その情報を体系的にまとめる必要がある．

そこで本研究では，シラバスの情報を抽出するシステムを構築し，抽出したシラバス情報を体系的に記述してファイル出力し，シラバスの可視化に活用できるようにする．

第 2 章

研究背景

2.1 シラバスとは

シラバスは、大学で開講される講義の内容を科目ごとに文章で簡潔にまとめた講義要綱である。現在シラバスは、Web 上で閲覧できる形や冊子として学生に配布するといった形で公開されることが多い。

シラバスには、大学で開講される講義についての概要やスケジュールなどが記載されている。学生はこれを読むことで講義の内容を理解した上で履修計画を行うことができる。また、大学はシラバスを導入することで、講義の内容や形式を予め学生に理解してもらうことができる。それにより、学生が講義全体の見通しを持って学習を進めることができるため、現在では多くの大学がシラバスを導入している。さらに、シラバスを外部にも公開することにより、外部の人にその大学ではどのような内容、質の講義をしているのかを知ってもらうという役割もある。これにより、大学の掲げるカリキュラムの特色を多くの人に理解してもらうことができる。

しかし、現在のシラバスにはいくつかの問題点が指摘されている。利用上の問題としては、シラバス上に読み取りづらい暗黙的な情報が含まれているという問題が指摘されている。また、シラバスの構成上の問題としては、シラバスが各大学ごとに独自の形式で作成されているためにその記述方法、内容に差が大きく、シラバス同士の比較が困難であることが指摘されている [1]。

2.2 シラバスの問題点

本節では、前項で述べた現在のシラバスの問題点とその原因について考察する。

2.2.1 シラバス上の暗黙情報

現在のシラバスには、履修前提・被履修前提の関係や、複数の科目で用いられる知識やキーワードなどといった科目間に存在する関係は明記されておらず、暗黙的に存在している。これらの科目間に存在する関係は、学生が系統立った学習をする上で不可欠なものであるため、これらの情報が暗黙的にしか存在していないことが学生の履修科目の決定を困難にしている。例として、高知工科大学 情報システム工学科のシラバスから「データ構造とアルゴリズム」、「人工知能」、「データベースシステム」、「離散数学」の4教科の繋がりを挙げる。まず、「データ構造とアルゴリズム」は「データベースシステム」を事前履修が望ましい科目としており、「データベースシステム」は「人工知能」を履修前提科目としている。さらに、「データ構造とアルゴリズム」、「人工知能」、「離散数学」の3教科のシラバスには共通して「木構造」というキーワードが記述されていることから、これらの教科は何らかの繋がりを持っていると考えられる(図 2.1)。しかし、これらの関係を学生が理解するためには全ての科目のシラバスページを読み、同じキーワードの有無や履修前提科目などの情報を調べる必要がある。

この問題の原因として、シラバスでの科目情報の記載方法と、実際の科目の存在の仕方との相違が挙げられる。現在のシラバスでは、各科目の情報が科目ごとに区切られた形でまとめられて記載されている。だが、科目は他の科目から独立して存在しているのではなく、カリキュラム上で何らかの役割を持ち、他の科目との繋がりを持って存在している。この相違から、科目間に存在する関係が読み取りづらくなり、暗黙的な情報となってしまっていると考えられる。これらの暗黙的な情報は、厳密にはシラバス上に存在しないわけではなく、複数の科目間に渡って記載されている。よって、科目の役割やカリキュラム全体を理解している教員がシラバスを読んだ場合には、科目間の関係が読み取りづらいと感じることはない

2.2 シラバスの問題点

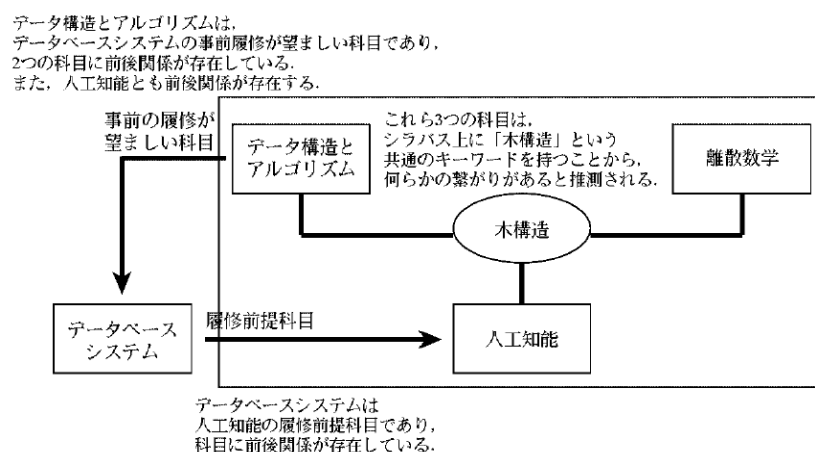


図 2.1 科目間に存在する関係

考えられる。しかし、主にシラバスを利用するのは履修計画を立てる学生であり、学生にシラバス全体の理解を求めることは困難である。そのため、これらの暗黙的な情報を理解できる形にすることは学生の履修支援を行う上で非常に重要である。

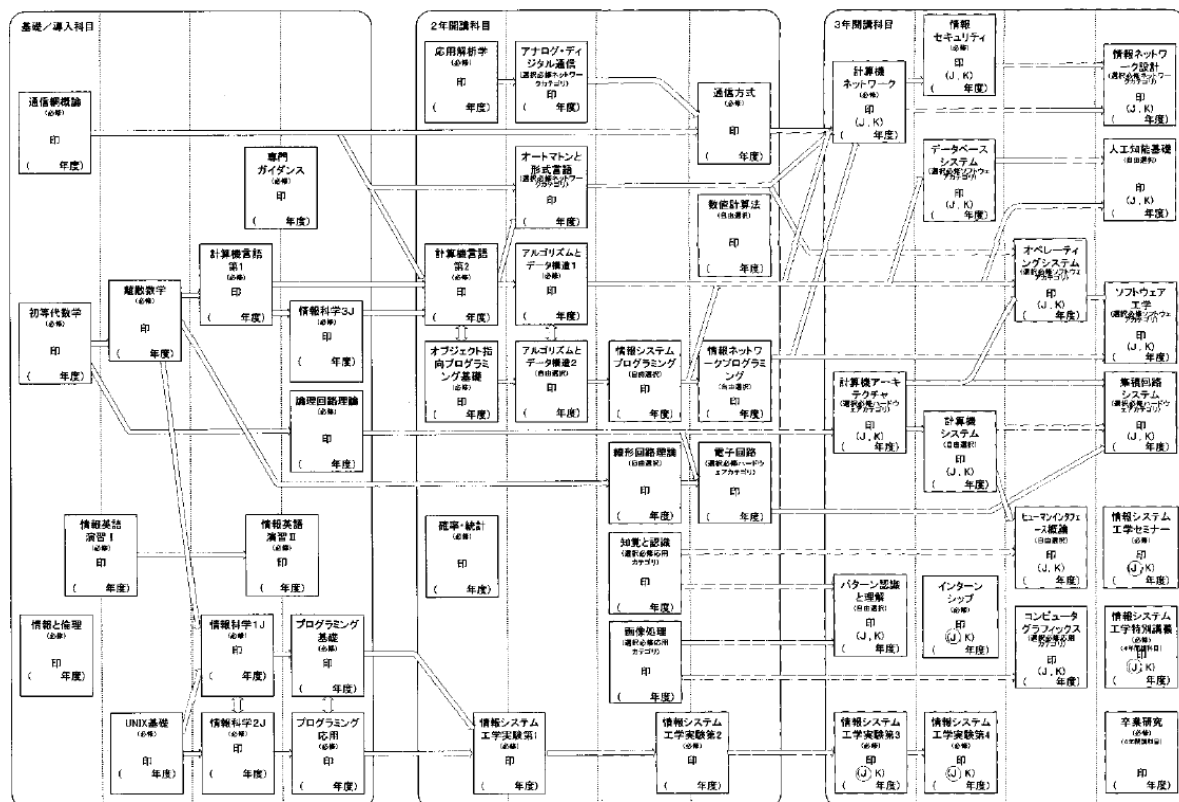
現在、この問題を解決するために、高知工科大学情報システム工学科ではプレレキジット表と呼ばれる表がシラバスと併せて用いられている（図 2.2）。これは、カリキュラム上にある履修前提科目の一覧を表にまとめたものであり、開講される講義に合わせて毎年作成されている。これを用いることで、上記の問題の内の、履修前提・被履修前提の関係を視覚的に理解することができる。しかし、現在はカリキュラムが変更された場合に手作業で表を作成しなおす必要がある。また、これにより可視化された関係は科目の前後関係だけであり、科目間に存在する関係が全て可視化されているわけではないため、学生の履修科目の決定を困難にしているという問題を解決できるとは限らない。

2.2.2 シラバスの構成要素の相違

また、シラバスの構成上の問題点として、大学ごとにシラバスの記述方法が異なっている点が挙げられる。現在のシラバスは、各大学が独自に記述形式や記述項目を決定して作成している。例として、高知工科大学 情報システム工学科のシラバス [2] と、同じく情報系学科

2.2 シラバスの問題点

情報ネットワークシステムコース 専門科目のプレレキジット



注)このプレレキジット票に記された履修前提科目(群)の単位を全く取得していない場合、当該科目の単位は取得できません。

図 2.2 プレレキジット表

である山形大学工学部 情報科学科のシラバス [3] を取り上げる。

高知工科大学 情報システム工学科のシラバスは、以下の構成要素により構成されている。

- ・ 科目名
- ・ 担当教員名
- ・ 開講年次
- ・ 開講クォータ
- ・ 単位数
- ・ 区分
- ・ 講義目標
- ・ 講義計画
- ・ 履修前提科目
- ・ 履修前提が望ましい科目
- ・ 同時履修すべき科目
- ・ 成績評価方法
- ・ 教科書/参考書
- ・ 備考

2.2 シラバスの問題点

高知工科大学では、全ての科目は人文科目や語学、各学科の専門科目といったカテゴリに分類されている。科目はカテゴリの1つに属することになっており、区分はその科目がどのカテゴリに属しているかを示している。履修前提科目は、事前に必ず履修しておかなければならない科目を示している。履修前提が望ましい科目は、必ずしもではないが、事前に履修しておくことが望ましい科目を示している。同時履修すべき科目は、その科目に関連した実験科目など、同時に履修することが前提となっている科目を示している。図 2.3 は、Web 上で公開されている高知工科大学のシラバスの例である。科目名、開講年次、開講クォータ、単位数、区分のみが名称または数値で記述されており、それ以外の内容は文章で記述されている。

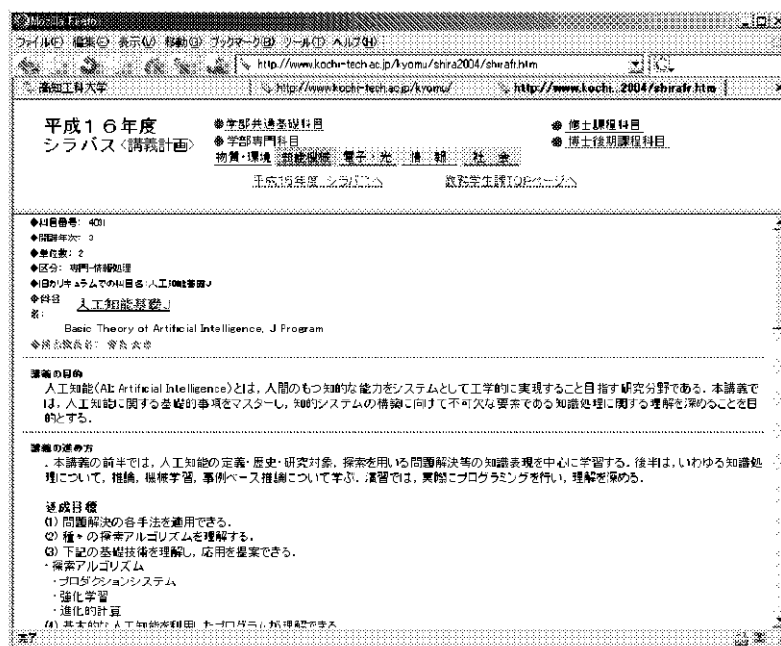


図 2.3 高知工科大学のシラバス

山形大学工学部情報化学科のシラバスは、以下の構成要素により構成されている。

2.2 シラバスの問題点

- ・ 科目名
- ・ 担当教員
- ・ 担当教員の所属
- ・ 開講学年
- ・ 開講時期
- ・ 単位数
- ・ 開講形態
- ・ 開講対象
- ・ 科目区分
- ・ 授業概要
- ・ 授業計画
- ・ 学習の方法
- ・ 成績評価の方法
- ・ テキスト
- ・ 参考書
- ・ 科目の位置付け
- ・ その他

授業概要では科目のテーマ，ねらい，キーワードが示されている．授業計画では講義の実施方法と日程が示されている．科目の位置付けではどの履修コースのどの教育目標に対応しているかが示されている．図 2.4 は，Web 上で公開されている山形大学のシラバスの例である．山形大学のシラバスも，科目名，開講学年，開講時期，単位数のみが名称または数値で記述されており，他は基本的に文章で記述されている．

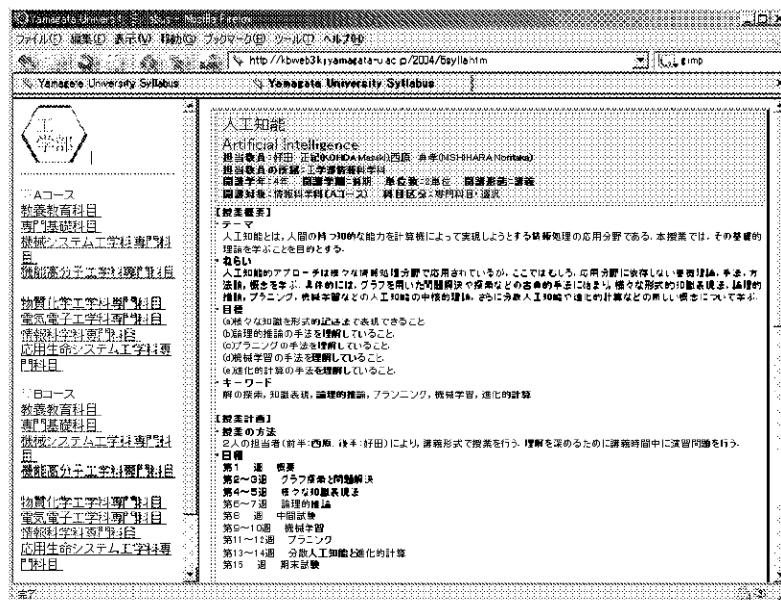


図 2.4 山形大学のシラバス

2.3 問題の解決方法

これら 2 大学のシラバスの構成要素を比較すると、科目名のように共通する要素も存在するものの、講義計画と授業計画のように記述内容が同じ要素であっても、要素名が異なっているものが存在する。また、高知工科大学に存在する履修前提科目、同時履修すべき科目などの情報は山形大学では存在しないなど、各大学のシラバスでは構成要素が大きく異なっていることが分かる。

記述内容について比較した結果、シラバス上の科目名は各大学のカリキュラムに依存していることが分かった。例えば、高知工科大学においての「アルゴリズムとデータ構造Ⅰ」、
「アルゴリズムとデータ構造Ⅱ」の科目が、山形大学では「データ構造とアルゴリズム」とされているように、内容は同じでも科目名が異なっている場合がある。また、同一の科目名であっても学習する範囲が異なるなどの場合も考えられる。

以上のことから、大学ごとにシラバスの構成要素が異なる上に、同じ構成要素同士でも記述されている内容まで同じであるとは限らないため、シラバス同士を単純に比較することは困難である。

2.3 問題の解決方法

これらの問題の解決方法として、各大学のシラバスを可視化する方法が考えられている[4]。可視化する情報は、科目間の関係や科目の属する分野などの学生にとって分かりづらい情報である。科目の前後関係や複数の科目で同じ知識やキーワードを用いるなどの関係は、各科目のシラバス情報を分析することにより抽出することができる。また、科目の属する分野の特定も、各科目のシラバスに記述されている語彙を各分野の専門用語辞典などと照らし合わせるにより可能である。抽出されたシラバス情報は、シラバスの構成要素を定めて体系的に記述する。これにより、大学ごとに異なっているシラバスの構成要素を比較できる形にまとめられる。

2.4 シラバス可視化システム

2.4 シラバス可視化システム

本節では，シラバス可視化システムによるシラバスの可視化方法と，システムの構成・概要を述べる．

2.4.1 シラバスの可視化方法

シラバス可視化システムで取り扱う科目間の関係は以下の 7 つの関係である．

- ・ 共通の修得項目を持つ
- ・ 共通の知識項目を持つ
- ・ 共通の関連項目を持つ
- ・ 共通の講義項目を持つ
- ・ 先修科目
- ・ 後修科目
- ・ 関係科目

これらの関係を，科目をノード，科目間の関係をパスとして表示することにより，有向グラフとして表すことができる．可視化システムでは，科目名を表示したノード同士を有向の矢印で表したパスで繋ぐ形式をとり，暗黙情報の可視化を行う．科目間の関係の区別は，パスの色を変更することで区別することにする．

2.4.2 システム構成

現在はシラバスを Web 上で公開する形が多くなってきている．可視化システムは，将来的に Web 上のシラバスの補助資料として使うことを想定しているため，開発環境に Macromedia Flash MX 2004 を用い，Web アプリケーションとして実装した．クライアントは Web コンテンツを閲覧するためのプラグインをインストールした Web ブラウザを用い，可視化システムの置かれた Web ページにアクセスすることでシステムを使用することができる．

2.4 シラバス可視化システム

2.4.3 システム概要

クライアントは Web ブラウザを用いて可視化システムの設置された Web ページにアクセスする．Flash で作成されたシラバス可視化システムのコンテンツが読み込まれると，図 2.5 の画面が表示される．

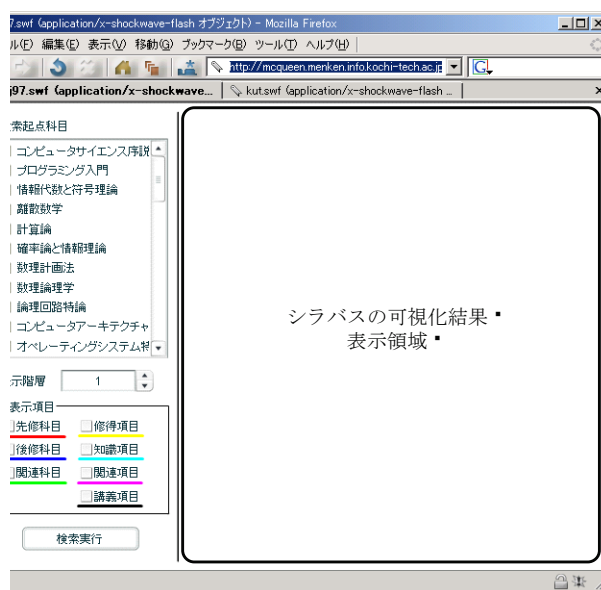


図 2.5 シラバス可視化システム：起動時

画面左側には，検索基点科目，表示項目，表示階層の設定入力と検索開始ボタンが表示される．検索基点科目は，シラバス上に存在する科目の一覧が表示されており，ここで他の科目との関連を調べたい科目を選択する．表示項目では，科目間に存在する関係の内，どれを表示するかを設定することができる．表示階層の設定は，表示する項目をどの程度の深さまで検索して関係を表示するかを設定することができる．設定完了後，検索開始ボタンを押すことにより検索と結果表示が行われる．画面中央から右側には，検索基点の科目とその科目と関係を持っている科目が設定内容に応じて表示される．実際にシラバス可視化システムにより科目間の関係を可視化した例を図 2.6 に示す．

2.5 シラバス情報の抽出

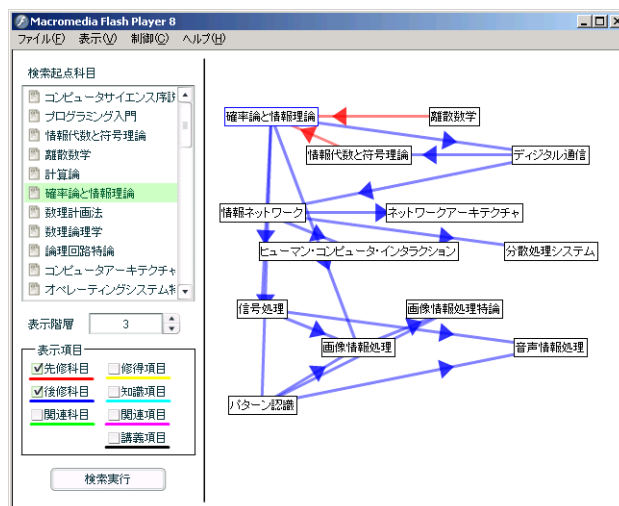


図 2.6 シラバス可視化システム：可視化例

2.5 シラバス情報の抽出

シラバス可視化システムでシラバス情報を可視化するためには、事前に対象のシラバスのテキストを抽出し、体系的に記述しておく必要がある。しかし、手動でシラバスから情報を抽出し、その情報を体系的にまとめるには、シラバスの規模にもよるが数時間～数十時間掛かってしまう。さらに、その際の語彙の抽出方法によっては、科目間で同じ知識を用いる関係が抽出できない場合もある。

このことから、シラバス情報を抽出するには専用のシステムにより効率的に情報収集することが望ましいと考えられる。

2.6 研究の目的

そこで本研究では、シラバスの可視化を行うためのシラバス情報抽出支援システムを構築し、シラバス情報を体系的に記述することにより、シラバスの可視化に用いることのできるデータの作成を行うことを目的とする。対象とするシラバスは、Web 上で公開されているシラバスとする。また、シラバスの標準化やデータの共有・再利用の助けとするために、本研究ではシラバス情報を体系的に記述する際に工学的オントロジーの考えを取り入れる。

第 3 章

オントロジーによるシラバスの体系的記述

3.1 オントロジーとは

オントロジーとは、本来は哲学の存在論を指す。存在論とは、世界に存在する全ての事柄を説明し、それらの体系と関係を明らかにしようとする理論である。しかし、近年のインターネットに代表されるネットワークや各種メディアの急速な進歩により、世の中で大量の情報が氾濫するようになった。それに伴い、哲学ではなく知識工学の分野でオントロジーが注目されるようになってきた。

大量の情報の中から必要な情報を効率的に検索するにはどのようにすればよいのか。または、非明示的な知識を情報として明示・共有するにはどのようにすればよいのかという観点から、この工学的オントロジーまたはオントロジー工学という形でオントロジーが利用されるようになった。オントロジーを工学の立場から見た場合、中核を成す定義は「存在一般という基本的なものだけではなく、人工物を含めた具体的なもの、たとえば企業活動、熱力学、問題解決構造などを対象としたものであって、そこに現れる概念と関係を明示的に示し、明確な意味定義を与えたもの」とされており、その目標は「実世界、厳密には情報科学が対象とし得る全てのモデル構築の基盤を与えること」とされている。このことから、知識工学の分野ではオントロジーは情報の共有や再利用ができ、人間やコンピュータを問わず理解できるものとして定義される。また、人工知能の立場においては「概念化 (Conceptualization) の明示的な仕様」と定義され、知識ベースの立場では「人工システムを構築する際のビル

3.2 本研究におけるオントロジーの役割

ディングブロックとして用いられる基本概念語彙の体系（理論）」と定義されている。

3.2 本研究におけるオントロジーの役割

上記 3 つの立場におけるオントロジーの定義から考えられる本研究におけるオントロジーの役割を以下に示す。

- 暗黙情報の明示化

シラバス上で暗黙的な情報として記述されている科目間の関係をデータ記述により明示化する。

- 記述方法の標準化

現在、各大学ごとに定められているシラバスの記述方法に対し、客観的な立場で標準の記述方法を定める。これにより、シラバスの標準化を行えると考えられる。

- 知識の体系化

シラバス上の暗黙的な情報を明示化し、共通の記述方法を定義することにより、コンピュータ上での知識の体系化を図る。

- 知識の共有と再利用

シラバスの記述方法を客観的に標準化するため、記述内容の共有や再利用を行うことができる。

- メタモデルの提供

オントロジーの利用の際には、オントロジーをクラス定義と見なしてそのインスタンスを作成しながらモデル構築を行う。そのため、シラバスのモデルとなる概念や構成要素といったものがメタモデルとして提供される。

3.3 オントロジーを導入する利点

2.2 節で述べたように、各大学のシラバスを比較した際に記述方法の相違があるにも関わらず、科目間にある関係が読み取りづらいという問題はどの大学のシラバス同士を比較した

3.4 データ記述要素

際にも共通して起こる。また、同じ分野の大学のシラバスでは記述方法に相違があるにも関わらず学習する内容はほぼ同じであることが多い。このことから、分野にはそこに含まれる知識や項目に関して暗黙の合意があり、シラバスはそこから学習すべき知識や項目を抽出することで作成されているといえる。

各大学のシラバスの記述項目に相違がある原因は、この分野に対する認識が異なっているためであり、これは分野というものが明確な定義を与られていないために起こる問題であると考えられる。従って、分野を端的に表すことのできるオントロジーを定義することにより、大学が持つ分野に対する相違を解消することができると考えられる。

3.4 データ記述要素

本研究では、先行研究で提案されたオントロジーを考慮したシラバスの構成要素を基にシラバスの構成要素を決定した [1]。文章で記述する要素については該当する内容を文章で記述し、キーワードで記述する要素については該当するキーワードを列挙する形で記述する。構成要素と記述方法の一覧を以下に示す。

- キーワードで記述する要素

- － 科目名：その科目の名称を記述する。
- － 修得項目：その科目中で学生に修得させたい項目を記述する。
- － 知識項目：修得科目ほどではないが、その科目中で触れられたことを学生に覚えておいてほしい項目を列挙する。
- － 関連項目：その科目中で触れられる項目、またはそれに関連した項目を列挙する。
- － 講義項目：その科目中の各講義で用いられる項目をグループ化し、より粗なレベルの項目名を付けたものを列挙する。
- － 先修科目：事前に履修しておくことが望ましい科目の名称を列挙する。ただし、先修科目と後修科目が必ずしも対応している必要はなく、ある科目の先修科目がある科目を後修科目にしているとは限らない。

3.4 データ記述要素

- － 後修科目：その科目を履修した後，続けて履修することが望ましい科目の名称を列挙する．ただし，後修科目と先修科目が必ずしも対応している必要はない．
 - － 関連科目：先修科目・後修科目のように科目の前後をつけることはできないが，その科目と関連している科目の名称を列挙する．ただし，対象の関連科目が必ずしもその科目を関連科目にしている必要はない．
- 文章で記述する要素
 - － 概要：科目の内容や，科目内で修得する範囲・到達点を記述する．
 - － 実験/演習：その科目で行う実験や演習の内容や進め方について記述する．
 - － 実施方法：その科目中で行う講義に関して，内容や進め方を記述する．
 - － 教科書/参考書：その科目中で用いる教科書や参考書の名称，出版社，ISBN 番号やコメントを記述する．
 - － 必修/選択：その科目が必修科目であるか，選択科目であるかの情報を記述する．
 - － 備考：他の項目には該当しないが，特に記述しておくべき内容やコメントがあれば，その内容を記述する．

第 4 章

シラバス情報抽出支援システムの設計

本章では，第 3 章で定めた体系的記述方法を用いたシラバス情報抽出支援システムの設計について述べる．

4.1 対象とするシラバスの形式

シラバス情報抽出支援システムを構築する上で，どのような形式で公開されているシラバスを対象にすることがより有効であるかを調べるために，平成 19 年 1 月現在のシラバスの形式調査を行った．対象としたシラバスは，文部科学省にリンクの張られている国立，公立および私立の四年制大学で公開されているシラバスである．また，本研究で扱うシラバスとは，科目についての情報がある程度詳しく記載されているものとしており，単純に科目一覧を表にまとめたものなどはシラバスに含めないことにした．さらに，Web サイト上でシラバスが公開されているとする基準は，科目間の関係をシラバスから抽出するという目的上，少なくとも学科単位でシラバスがまとめられているものとした．よって，本研究では講義の担当教員などが独自に作成したシラバスなどはシラバスに含めないことにした．調査結果を表 4.1 に示す．

調査結果から，本研究では公開されているシラバスの約 77 %を占めている HTML 形式のシラバスを対象とすることにした．

HTML 形式のシラバスが Web サイト上でどのようなファイル構成で設置されているか

4.2 シラバス情報の抽出方法

表 4.1 シラバス形式の調査結果

	調査数（公開数）	HTML	PDF	その他（形式）
国立大学	87（73）	59	13	1（FLASH）
公立大学	76（35）	18	17	0
私立大学	544（215）	171	42	2（DOC）
合計	707（323）	248	72	2（DOC），1（FLASH）

を調べたところ，大半のシラバスでは統一された記述方法で各科目ごとのシラバス（以下，シラバスページ）が作成されており，さらにそのシラバスページへのリンクを大学または学科単位の一覧でまとめたページ（以下，一覧ページ）が作成されていることが分かった．また，シラバスページには各大学ごとないし各学科ごとに記述方法が定まっており，それを基に作成されているか，CGI などで機械的に作成されている場合が多いことも分かった．このことから，シラバスページの各項目を囲む HTML のタグパターンは全てのシラバスページで統一されていることになる．そこで本研究では，シラバスの記述項目を分解するために HTML のタグパターンに注目する．

4.2 シラバス情報の抽出方法

HTML 形式はテキスト処理で情報が抽出できるため，HTML Parser を用いることで対象のシラバスからテキストと HTML タグを抽出し，HTML タグのパターンによってシラバス情報を記述項目ごとに分解する．しかし，分解された項目がシラバス上のどの記述項目にあたるのかをシステムが自動的に判別することは困難である．構文解析により，シラバスページの文書構造を自動的に判別することはある程度可能であると考えられるが，多くのシラバスに適用することは困難である．

そこで本研究では，対象とするシラバスの記述方法をユーザが指定することにより，対象のシラバスからの情報抽出を自動で行うシステムを構築する．これにより，HTML 形式で

4.2 シラバス情報の抽出方法

公開されているシラバスであればどの大学のシラバスにでも使用することができる．情報の抽出方法について図 4.1 に示す．

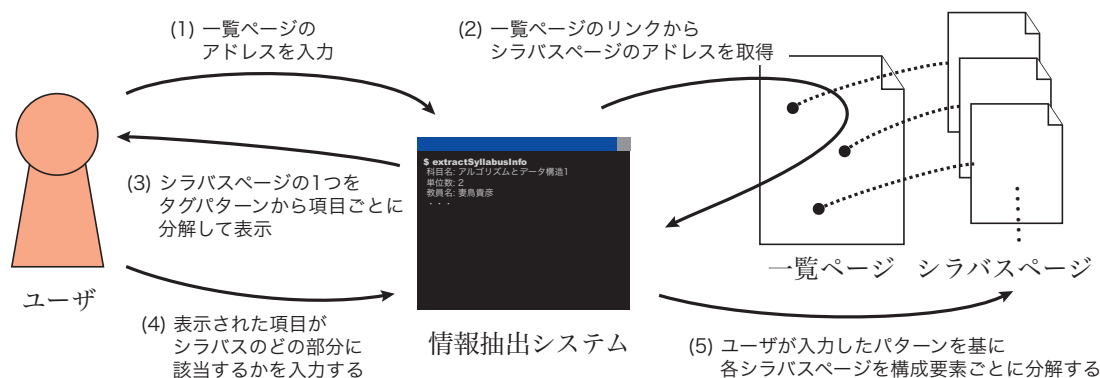


図 4.1 情報抽出の方法

また、タグパターンから記述項目を分解する際、強調表示などのシラバスページごとに使用場所や回数が異なるタグまで分解の際のタグ要素にしてしまうと、記述項目の分解がうまくできなくなってしまう恐れがある．そこで、項目上の区切りのためには用いられないと考えられる文字装飾やリンクなどのタグに関しては、タグパターンとして認識しないことにする．

構成要素を体系的に記述する際、キーワードで記述する修得項目、知識項目、講義項目、関連項目の4項目に関しては各テキストの形態素解析を行い、その結果からキーワードを作成する．キーワード作成の例として、「自然言語処理を行う」という文章があった場合、形態素解析を行うとそれぞれ「自然（副詞）」、「言語（名詞）」、「処理（名詞）」、「を（助詞）」、「行う（動詞）」に分けられ、それぞれの形態素の品詞が分かる．本研究では、キーワードとして名詞と、名詞を修飾する品詞と、それらの品詞の集合から考えられる組み合わせ全てをキーワードとして抽出することにした．この例の場合、まず品詞3つの集合から「自然言語処理」というキーワードが作成される．また、品詞2つの組み合わせとして「自然言語」、「言語処理」というキーワードが作成され、最後に品詞1つの「自然」、「言語」、「処理」もキーワードとする．

作成されたキーワードが意味を持ったものであるかどうかを調べるために辞典参照を行

4.2 シラバス情報の抽出方法

い、一致したキーワードをその科目内容を示す語彙とする。その際、キーワードの抽出を正しく行うためには情報学科のシラバスには情報分野の辞典を、教育学科のシラバスには教育分野の辞典といったように各学科の分野の辞典を用意し、その辞典によりキーワードの一致をとることが望ましい。本システムでは、辞典参照の方法として品詞数の多いものから順に参照していき、一致する語彙が見つかった場合は、そのキーワードの品詞数以下のキーワード群に関しては辞典参照しないことにした。これは、より多くの品詞を持つキーワードはそれよりも品詞数の少ないキーワードに比べ、より深い意味を持つものであるという仮定に基づいている。例えば「言語処理」というキーワードが抽出された場合、このキーワードは「言語」の分野と「処理」の分野のどちらにも属している上に「言語」と「処理」よりも、より限定的な意味を持っていると考えられる。本システムは、語彙の抽出と分野の特定を目的としたものであるため、キーワードを抽出する際にはできるだけ深い意味を持ったキーワードを抽出することが望ましい。

第 5 章

シラバス情報抽出支援システムの実装

本章では、第 4 章で行った設計に基いたシラバス情報抽出支援システムの実装について述べる。

5.1 システムの構成

本システムの開発環境には Java 言語を用いた。本システムは、シラバスから抽出した HTML の開始タグと終了タグのタグパターンによってシラバスの記述項目を分解する。その際、文字装飾やリンクなどのタグ (a, b, big, center, div, font, link, s, span, strike, sub, sup, u) は扱わないことにした。また、開始タグ・終了タグを必要としないタグであるシンプルタグも、タグパターンとして扱わないことにした。本システムでは、シラバス情報としてキーワードの抽出を行う。そのためのキーワード作成の際に、テキストの形態素解析ツール JUMAN を用いてテキストの形態素解析を行う。JUMAN を用いた理由は、本システムは今後、構文解析ツール KNP を用いた情報抽出を行うシステムへの改良を予定しており、KNP は JUMAN を用いることにより構文解析を行うためである。

JUMAN とは、計算機による日本語解析の研究を目指す多くの研究者たちに共通して利用できるシステムとして京都大学言語メディア研究室で開発された形態素解析ツールである。このシステムは、隠れマルコフモデルを使った解析で高い精度を誇り、形態素解析ツール ChaSen の元となっている。また、このシステムは学校文法が計算機向きではないという

5.2 システムの概要

問題を考慮し，使用者によって文法の定義，単語間の接続関係の定義などを容易に変更できる．現在は JUMAN Ver 5.1 が最新版である．

5.2 システムの概要

本システムは，まず，ユーザが対象のシラバスのトップページのアドレスを入力することにより，そのシラバスのシラバスページの情報を抽出する．次にユーザが，抽出された各テキストにシラバスの構成要素を関連付けすることにより，そのシラバスの記述方法を指定する．その後，システムが関連付けされた記述方法に従って各シラバスページから情報抽出を行い，シラバスの構成要素ごとに抽出された情報を記述した XML ファイルを出力する．

5.3 システムの実行方法

本節では，シラバスからの情報抽出の実行方法について述べる．例として，山形大学工学部 情報科学科のシラバスを用いる．また，キーワードの抽出の際に用いる辞典には「改訂 電子情報通信用語辞典」を事前にデータベース化したものを用いた．本システムでは，キーワードを正しく抽出するために，対象とするシラバスが属する分野の辞典を用意し，その辞典を用いてキーワード抽出を行うことが望ましい．

5.3.1 シラバスのトップページのアドレス入力

本システムは，まず，ユーザがシステムに情報を抽出したいシラバスのトップページのアドレスを入力する．ここでのシラバスのトップページとは，各シラバスページへのリンク一覧が記述されているページを指す．トップページの入力を行う際のインタフェースを図 5.1 に示す．アドレスの入力が完了すると，「次へ」ボタンと押すことにより情報抽出を行う対象のシラバスページへのリンクを選択する画面へ進むことができる．システムを終了したい場合は，「終了」ボタンを押すことによりシステムを終了することができる．

5.3 システムの実行方法

図 5.1 トップページのアドレス入力

5.3.2 シラバスページへのリンクの選定

トップページのアドレス入力画面で「次へ」ボタンが押されると、シラバスのトップページに記述されているリンク情報を全て抽出し、表示する。また、抽出された各リンクそれぞれに対応したチェックボックスも表示する。シラバスページ以外へのリンクが含まれている場合は、ユーザがそのリンクをチェックボックスで指定することによって情報抽出を行う対象から除外することができる。リンクの選定を行う際のインタフェースを図 5.2 に示す。リンクの選定が完了すると、「次へ」ボタンを押すことによりテキストと構成要素との関連付けを行う画面へ進むことができる。トップページのアドレス入力をやり直したい場合は、「戻る」ボタンを押すことによりトップページのアドレス入力画面に戻ることができる。その際、以前入力されたトップページのアドレスはそのままになっている。

5.3.3 テキストと構成要素の関連付け

リンクの選定画面で「次へ」ボタンが押されると、まず、情報抽出を行う対象のリンクの先頭にあるアドレスのシラバスページ（以下、サンプルページ）を、このシラバスの記述方法を指定するためのページとしてサンプルページからテキストと HTML タグ情報を抽出する。さらに、サンプルページに記述されているテキストを HTML タグごとに分解して表示

5.3 システムの実行方法

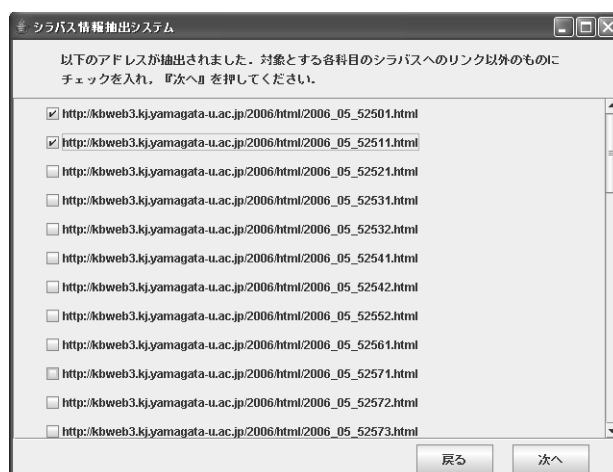


図 5.2 情報抽出対象の選定

し、それぞれのテキストの下に構成要素が入っているコンボボックスを表示する。ここでユーザが、抽出された各テキストにシラバスの構成要素をそれぞれ対応させることにより、そのシラバスがどのような構成要素を持ち、どのような形式で記述されているかを指定することができる。テキストと構成要素の関連付けを行う際のインタフェースを図 5.3 に示す。テキストと構成要素の関連付けが完了すると、「次へ」ボタンを押すことによりサンプルページからの情報抽出の結果を確認する画面へ進むことができる。シラバスページへのリンクの選定をやり直したい場合は、「戻る」ボタンを押すことによりリンクの選定画面に戻ることができる。その際、以前選択されたリンクのチェックボタンはそのままになっている。

5.3.4 シラバス情報抽出結果の確認

テキストと構成要素の関連付けの画面で「次へ」ボタンが押されると、まず、5.3.3 節でユーザが行った関連付けを基にサンプルページの各テキストからキーワード抽出を行い、その結果の辞典参照を行う。さらに、辞典の語彙と一致したキーワードとそのキーワードが含まれていたテキストを、ユーザが関連付けした構成要素ごとに表示する。これにより、ユーザは自身が行った関連付けの確認と、サンプルページからどのようなキーワードが抽出されたかを確認することができる。情報抽出結果の確認の際のインタフェースを図 5.4 に示す。

5.3 システムの実行方法

シラバス情報抽出システム

各項目内容に当て嵌まる構成要素をそれぞれ選び、『次へ』を押してください。
項目内容は、各シラバスページ内の『教養1』や『2単位』などが記述されている欄を指します。
項目名である『講義名』や『単位数』などが記述されている欄については、構成要素を選択する必要はありません。

・ねらい
指定なし

ワークステーションを用いて、コンピュータの基本的な操作方法ならびにC言語によるプログラミング
取得項目

・目標
指定なし

1) UNIXの操作の基礎を習得すること (学習教育目標(B)-1-(1))
知識項目

2) プログラムを作成する際の考え方の基礎を習得すること (学習教育目標(D)-1-(1))
知識項目

戻る 次へ

図 5.3 テキストと構成要素の関連付け

情報抽出結果の確認後に、「実行」ボタンを押すことにより全てのシラバスページからの情報抽出を行い、その結果出力されたファイル名の一覧を表示する画面へ進むことができる。テキストと構成要素の関連付けをやり直したい場合は、「戻る」ボタンを押すことによりテキストと構成要素の関連付けを行う画面に戻ることができる。その際、以前行われた構成要素の関連付けはそのままになっている。

シラバス情報抽出システム

選択した項目内容と構成要素により、先程のページからキーワード抽出した結果です。
『実行』を押すとすべての各科目のシラバスページに対してキーワード抽出を行います。
結果をXMLファイルで出力します。

科目名
プログラミング演習 (A情報)
Computer Programming I

情報
取得項目

概要
「C言語によるプログラミングの初歩」を学ぶこと

言語
C言語

取得項目
ワークステーションを用いて、コンピュータの基本的な操作方法

戻る 実行

図 5.4 情報抽出結果の確認

5.3 システムの実行方法

5.3.5 出力ファイル表示

情報抽出結果の確認画面で「実行」ボタンが押されると、まず、ユーザが行った関連付けを基に全てのシラバスページに対して情報抽出を行い、その結果抽出された情報を体系的に XML ファイルに記述し、ファイル出力する。さらに、出力された XML ファイルのファイル名の一覧を表示する。実行が完了し、出力ファイルの一覧表示を行う際のインタフェースを図 5.5 に示す。実行が完了すると、「完了」ボタンを押すことによりシステムを終了することができる。



図 5.5 出力ファイルの表示

5.3.6 出力ファイルの記述結果

出力された XML ファイルの記述形式を図 5.6 に、情報の抽出元である山形大学工学部 情報科学科のシラバスを図 5.7 に示す。キーワードで記述する項目には抽出されたキーワードを記述し、文章で記述する項目には関連付けされたテキストをそのまま記述している。これにより、文章で記述する項目である概要には『「C 言語によるプログラミングの初歩」を学ぶこと』という文章が抽出され、記述されていることが確認できる。また、キーワードで記述する項目である修得項目に関連付けされたテキストからは『言語』というキーワードが抽出され、記述されていることが確認できる。

5.3 システムの実行方法

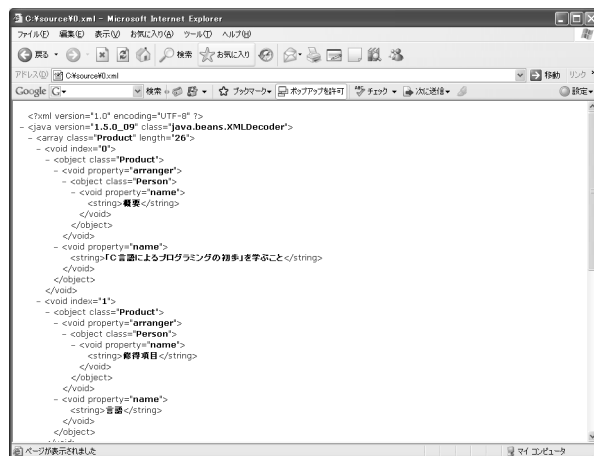


図 5.6 XML ファイルの記述結果

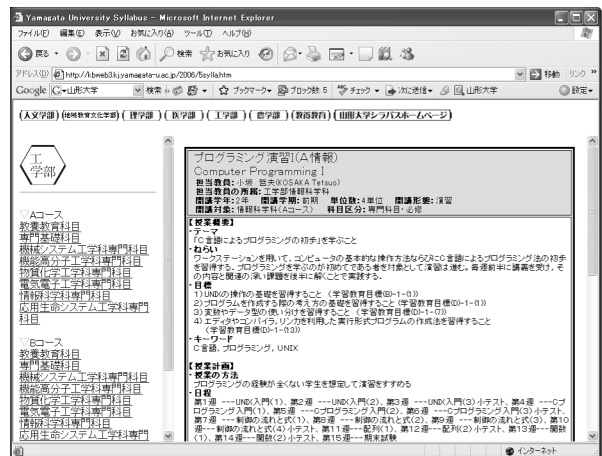


図 5.7 抽出元のシラバスページ

第 6 章

評価

本章では，第 5 章で実装した情報抽出支援システムが実際に公開されているシラバスから手動で情報を抽出し，その結果を体系的に記述することによりシラバスの可視化に活用できるようにできたかの評価を行うために実施した検証内容について述べる．

6.1 評価方法

本システムを文部科学省にリンクの張られている大学 10 校のシラバスに対して使用し，その結果と実際に各シラバスから情報抽出した結果とを比較した．本システムの辞典参照には「改訂 電子情報通信用語辞典」をデータベース化したものを用いているため，対象とするシラバスは電子情報通信分野に関連していると考えられる学校または学科のシラバスとした．今回の検証では，対象のシラバスの全てのシラバスページから，ユーザが行った関連付けの通りに情報抽出できていれば成功，1 科目でも関連付けした通りに情報抽出できていなければ失敗したものとして扱うことにした．

6.2 結果

本システムによる情報抽出を実施した各シラバスの大学名と結果を以下に示す．

- 成功
 - － 武蔵工業大学 環境エネルギー工学科
 - － 北海学園大学 工学部 電子情報工学科

6.3 考察

- 山口東京理科大学 電子・情報工学科
- 東北大学 工学部 電気情報・物理工学科
- 大阪工業大学 情報システム学科
- 西日本工業大学 電気電子情報工学科
- 新潟工科大学 情報電子工学科
- 失敗
 - 高知工科大学 情報システム工学科
 - 東海大学 情報理工学部 情報科学科
 - 山形大学 工学部 情報科学科

10 大学のシラバスに対して本システムを使用して情報抽出を行った結果，7 校で情報抽出に成功し，3 校で情報抽出に失敗した．

6.3 考察

情報抽出に成功した 7 校のシラバスは，全てのシラバスページの記述方法が完全に統一されていたため，情報を抽出することができた．記述方法が統一されているシラバスの例として，東北大学工学部 電気情報・物理工学科のシラバスを例に挙げる．このシラバスでは全てのシラバスページで，授業科目名，開講系，開講セメスター，担当教官名・所属，授業科目の目的・概要及び達成目標等，他の授業科目との関連及び履修上の注意，授業計画，成績評価の方法及び基準，教科書・参考書，関連ホームページの順に項目が記述されている．このうち，担当教官名・所属の項目では，図 6.1 に示した 2 科目のシラバスページのように記述されている教員数が異なる場合もあった．しかし，これは BR タグによる改行が行われているだけであり，本システムは BR タグなどのシンプルタグをタグパターンとして扱わないことにしているため，タグパターン上ではこのような場合でもずれは生じず，正しく情報抽出できた．また，その他の項目で BR タグなどのシンプルタグが記述されていた場合にも同様に情報抽出できた．

6.3 考察

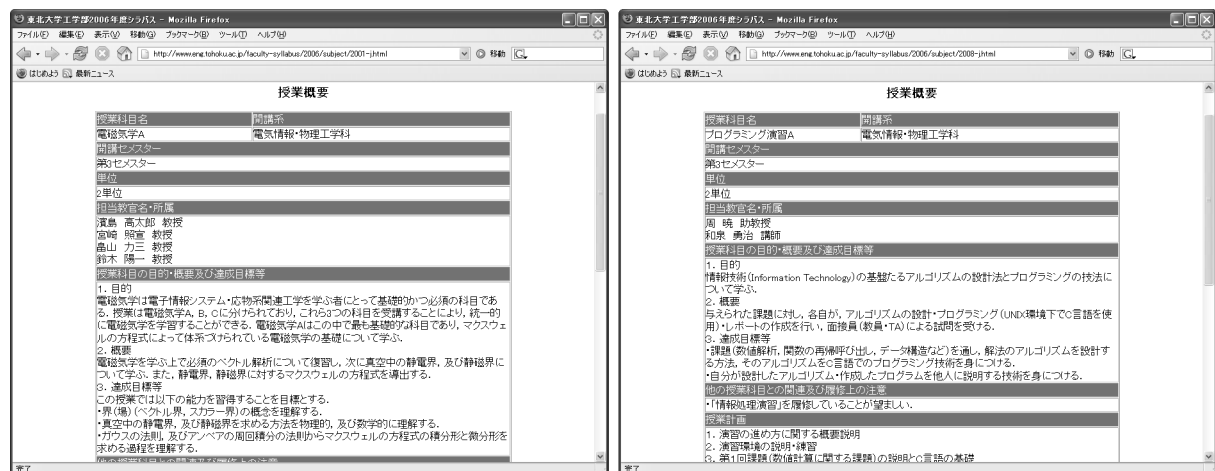


図 6.1 情報抽出に成功したシラバス

失敗した 3 校のシラバスに関しては，全てのシラバスページからの情報抽出が失敗した訳ではないが，科目や教員ごとに記述方法が異なる場合があったためにユーザが指定した関連付けにずれが生じ，正しく情報抽出できないシラバスページがあった．確認された情報抽出の失敗例として，タグパターンのずれによって抽出されたキーワードや文章が関連付けされた項目ではない項目に誤分類されてしまう場合や，ファイルに記述する際に一致するタグパターンがないためにファイル記述されない場合があった．また，シラバスページごとに項目の記述順序が異なっている場合もあり，この場合も正しく情報抽出できなかった．最も多い失敗例である項目内容の記述方法の相違によりタグパターンのずれが生じる例を，東海大学情報理工学部 情報科学科のシラバスページを用いて示す．東海大学情報理工学部 情報科学科の「情報と職業」と「プログラミング実習」のシラバスページを図 6.2 に示す．この 2 つのシラバスページでは授業の目標項目のテキスト数が異なっているが，この項目からは図 6.1 で示したシラバスと同様に情報を抽出することができた．問題となるのは，必修/選択の別の項目である．図 6.2 の左側に示した「情報と職業」のシラバスページをサンプルページとして関連付けを行った場合，必修/選択の別の項目には 3 つの学科名と，その学科の学生の履修の必要性が関連付けされる．この 3 学科分の情報を記述している部分の HTML タグは TR タグと TD タグであり，本システムがタグパターンとして扱うことにしているタグである．そのため，システムはこのシラバスの全てのシラバスページの必修/選択の別の項目

6.3 考察

から、3 学科分の情報を抽出しようとする。しかし、図 6.2 の右側のように 1 学科分の情報しか記述されていないシラバスページの場合はタグパターンにずれが生じ、抽出した情報を関連付けした通りの構成要素に記述することができないため、正しく情報抽出することができなかった。

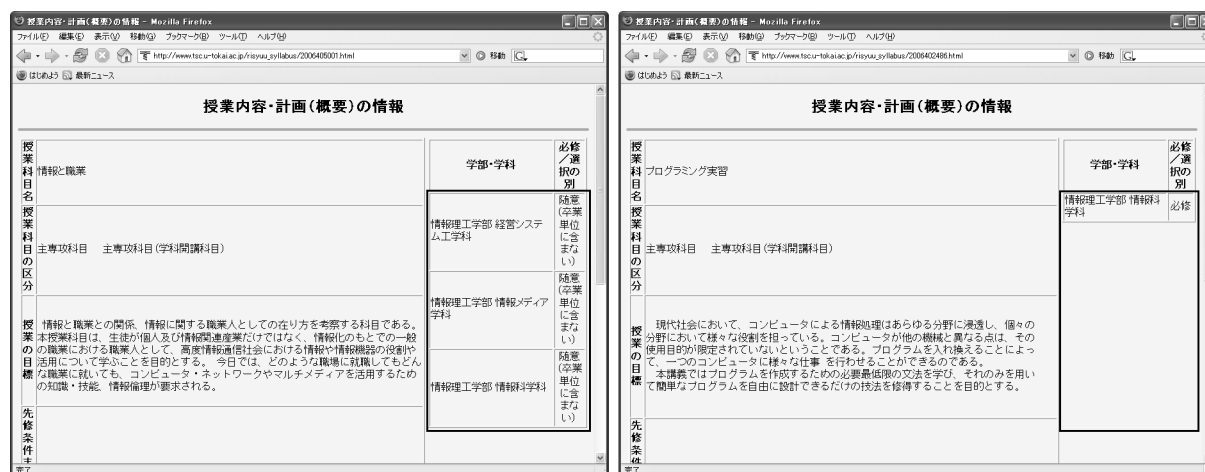


図 6.2 情報抽出に失敗したシラバス

本システムは同一シラバス内で記述形式や記述項目が統一されていない場合を考慮していない。そのため、シラバスページごとに記述項目までのタグパターンが異なっている場合には、抽出したキーワードや文章がユーザが指定した項目とは別の項目に分類されてしまうことが分かった。

今回構築したシステムでは、同一のシラバス内で統一された記述方法がとられているシラバスから情報を抽出できる。しかし、同一のシラバス内で異なった記述方法がとられている場合も少なくはないため、今後はそれらのシラバスからも情報抽出を行うことのできるシステムに改良する必要がある。現在考えられる方法として、構造解析や構文解析などを用いることにより、システムがサンプルページでユーザに指定された情報からシラバスに記述されている内容を取得し、それにより他のシラバスページの記述項目の判別・抽出を行えるようにすることが挙げられる。

第 7 章

まとめ

本研究では，シラバスの可視化に必要なシラバス情報の抽出を行い，その情報を体系的に記述するシステムを構築した．本システムは，ユーザが対象とするシラバスの記述方法を指定することにより，自動でシラバス情報の抽出を行うシステムとして実装した．また，開発環境には Java 言語を用い，シラバス情報の記述は XML 形式で行った．

評価を行うために，本システムを文部科学省にリンクの張られている 10 大学のシラバスを対象として使用し，情報抽出を行った．また，その結果と実際にその 10 大学のシラバスから手動で情報抽出した結果とを比較することにより，本システムの情報抽出が正しく行えているかを検証した．その結果，大学内ないし学科内でシラバスの記述方法が統一されている場合では情報抽出が正しく行えることを確認した．

現在のシステムの問題点は，HTML 形式で公開されているシラバスのみの対応である点，記述方法の統一されていないシラバスからは情報抽出できない点．ユーザがシラバスの記述方法を指定する必要があるため，ユーザに多少の負担が掛かる点が挙げられる．

今後の課題として，上記の問題点の解決に加え，語彙の抽出方法の検討・実装などを予定している．また，構造解析や構文解析によるシラバス情報の抽出に関しても検討・実装する予定である．

謝辞

本研究の全てにおいて，多大なるご指導を賜りました高知工科大学 情報システム工学科 妻鳥貴彦 講師に心より感謝致します．

本論文及び本研究において，貴重なご意見を賜りました高知工科大学 情報システム工学科 篠森敬三 教授に厚くお礼を申し上げます．

また，本研究を進めるにあたって，貴重なご意見を賜りました高知工科大学大学院 工学研究科 基盤工学専攻 情報システム工学コース 1 回生の木下聡氏，高木翔平氏，川島真一郎氏及び，高知工科大学 情報システム工学科の 4 回生の寒川剛志氏，大黒隆弘氏，松本康祐氏，3 回生の太田和也氏，加集広希氏，藤山翔太氏，橋田味加子氏，畠山博和氏，福田将行氏，藤原健太郎氏，山崎雄大氏に深く感謝致します．

参考文献

- [1] 石川 大地, “オントロジー理論に基づくシラバス可視化のための情報システム工学領域の体系的記述”, 高知工科大学 平成 16 年度学士卒業研究論文, 2005.
- [2] 高知工科大学工学部情報システム工学科シラバス
- [3] 山形大学工学部情報科学科シラバス,
<http://kbweb3.kj.yamagata-u.ac.jp/2006/home.htm>
- [4] 木下 聡, “シラバス可視化システムの構築”, 高知工科大学 平成 17 年度学士卒業研究論文, 2006.