

シラバスデータを活用した書籍検索手法の検討

越智 洋司

近畿大学理工学部

ochi@ele.kindai.ac.jp

概要 本研究ではシラバスにおける科目間の関連性を抽出し、関連科目を見つけ出すことで、シラバスに書かれている書籍検索をするシステムを開発している。本稿では、システムの概要とシステムで検索される関連科目ならびに書籍の妥当性を評価するための評価アンケートについて報告する。最後にその結果から有用性を検討する。

キーワード シラバス分析, 書籍検索, データベース

1 はじめに

講義内容を理解する方法として、書籍などを利用した自学自習を行う方法が挙げられる。原田⁽¹⁾は、図書館の実際の貸出記録約 90 万件を用いて、協調フィルタリングの手法を用いて図書のおすすめに関する研究を行っている。しかしこのアプローチでは、ユーザの行動に着目したものであり着目した科目に一致する書籍が推薦されるとは限らない。その際に、適切な参考書を見つけ出すことが重要となる。そこで、本研究では講義シラバスに着目した。シラバスを利用した研究としては、シラバスの内容解析による LOM の自動生成の研究⁽²⁾、LDA と Isomap を利用した科目のマッピング⁽³⁾等、シラバスから科目間の関係や内容を抽出する試みがある。

2 システム概要

2.1 書籍検索のアプローチ

シラバスに掲載されている科目は、大学にカリキュラムに則ったものがあるため、カリキュラムにおける順序性が開講年度や関連科目といった項目で表現されている。我々はその関係性に着目し、下記のアプローチによる書籍システムシステムを構築する。

(1) シラバスの記述からの関連科目検索

(2) 科目間の関連性を利用した書籍検索

本研究では、シラバスに明示的に関連性が記載されている科目を関連科目とし、そこから逆方向の関係性を持つ科目を逆関連科目と呼び、書籍検索に利用する。

2.2 書籍検索機能

(1) 明示的関連科目による書籍推薦

明示的関連科目とは、対象科目の関連科目の項目に記載されている科目である。関連科目は、対象科目の専門家である先生が対象科目を理解するために必要な科目を記載、あるいは関係性の深い科目を記載してい

るため、明示的に書かれた関連科目に記載されている教科書や参考文献を推薦する手段が有効と考える。

(1) 類似度計算による関連科目と書籍検索

シラバスの内容を名詞で形態素解析を行い、コサイン類似度による科目間の類似度計算することで、類似度の高い科目(以下、類似度関連科目)を抽出する。類似度関連科目による書籍推薦では、対象科目を入力することで、対象科目の教科書や参考文献を提示し、類似度関連科目の教科書や参考文献も提示する。

(3) 逆関連科目による書籍検索

逆関連科目とは、対象科目を明示的な関連科目として記載している科目である。仮に、上記で示している 2 のように関連科目の書籍が有効であれば、逆関連科目に記載されている書籍も有効ではないかと考える。

(4) セメスターを用いたレベル別書籍検索

講義には開講学年が設定されており、セメスターが低ければ簡単な科目であり、セメスターが高ければ難しい科目であると仮定できる。本機能では、その順序性を活用する。

ユーザインタフェースを図 1 に示す。



図1 システムインタフェース

表 1 類似度関連科目の評価

順位	対象科目											平均
	A	b	c	d	e	f	g	h	i	j	k	
1	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00
2	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00
3	5.00	5.00	5.00	4.67	5.00	4.67	5.00	5.00	5.00	4.67	4.67	4.88
4	4.75	5.00	5.00	4.75	5.00	4.25	5.00	4.75	4.75	4.00	4.75	4.73
5	4.60	4.80	4.80	4.60	4.80	4.00	5.00	4.60	4.80	4.20	4.60	4.62

3 評価実験

3.1 評価目的

本研究では関連科目によって書籍推薦を行っているため、まずは明示的な関連科目以外の関連科目が妥当であるかを検証する必要がある。検証項目は、

- ・類似度関連科目の妥当な検索順位
- ・逆関連科目の妥当性
- ・関連科目からの書籍の妥当性

である。類似度関連科目は対象科目を除く全ての科目との類似度を表す。このため、システムとして導入する際、類似度関連科目の何位までが関連科目として妥当であるかを評価する必要がある。

逆関連科目については、元々自明な関連科目を逆説的に捉えていることになるので自明のようであるが、逆関連科目には双方向で関連科目を挙げているものと片方向のみ関連科目をあげているものの 2 種類があるため、検証する必要がある。

3.2 評価対象と方法

本学電気電子工学科と情報学科の 2 学科の専門科目を対象に解析を行った。科目数は電気電子工学科が 141 科目、情報学科が 66 科目の合計 207 科目である。

科目を担当している先生(5 名)を被験者とし、担当されている科目の中から 11 科目を対象として、システムが提示する関連科目の妥当性を 5 段階評価する。科目評価のアンケートでは関連科目を「対象科目を理解する助けとなる科目」と定義付けする。アンケート項目数は関連科目 15 科目とし、この 15 科目は「明示的な関連科目を除いた逆関連科目と、明示的な関連科目を除いた類似度関連科目のランクが高い順に抽出した科目」である。書籍の妥当性については、「対象科目を理解する助けとなる書籍」と定義付けてアンケートを行った。

3.3 評価結果

類似度関連科目の評価結果をまとめたものを表 1 に示す。縦軸に類似度関連科目の順位を、横軸に対象科目を示し、それぞれに対応する 5 段階評価の平均値を表に示す。ただし、表内の平均値の数値は、該当順位までの平均値である。つまり、科目 f において順位 4

の数値は 4.2 であるが、これは類似度関連科目順位 1～4 の評価値の平均値を示している。また、明示的な関連科目については評価値 5 とした。また明示的関連科目が出現した箇所の背景色を灰色とした。

逆関連科目の妥当性については、半数の科目が関連科目として明示的に記述されていた。関連科目として明示されていない逆関連科目(28 科目)については、評価平均が 2.5 と高いものではなかった。逆関連科目を関連科目として使用することは難しいと考える。

書籍の妥当性については、類似度関連科目の書籍については、各科目の平均値を平均すると、4.205 となり、類似度関連科目の順位を反映した書籍検索は有効であると示唆された。一方、明示的関連科目の書籍は平均が 3.415 と低くなった。明示的関連科目は先生方が示している関連科目であるため評価値が高いと予想していたが、類似度関連科目よりも低い値を示した。

4 おわりに

本研究ではシラバスデータを用いた書籍推薦システムを提案した。シラバスを用いた書籍推薦システムで有効な書籍推薦手法は類似度関連科目を用いた書籍推薦手法であることが示唆された。類似度関連科目による書籍推薦手法が有効であるとわかったことから、シラバスの持つ可能性の一つ確認できた。また、類似度関連科目を求めることにより、学科間を越えた関連科目が導き出せることが分かった。このことから、シラバスの関連科目を自動生成するシステムの可能性が考えられる。

参考文献

- [1]原田 隆史, 図書館の貸出履歴を用いた図書の推薦システム, デジタル図書館「デジタル図書館」編集委員会 [編], pp22-31, 2009
- [2]辻 靖彦, 森本 容介, LOM の自動生成を目的としたシラバス文書の情報抽出, 電子情報通信学会技術研究 報告. ET, 教育工学 109(335), 131-136, 2009
- [3]関谷 貴之, 松田 源立, 山口 和紀, Isomap を用いたカリキュラムの可視化, 人工知能学会 インタラクティブ情報アクセスと可視化マイニング研究会(第 5 回), pp.33-40, 2013