

学生の得意不得意を考慮した時間割の自動作成

西森友省 †

堀幸雄 †

今井慈郎 ‡

† 香川大学大学院工学研究科

† 香川大学総合情報基盤センター

‡ 香川大学工学部

1 はじめに

大学において学生は大学から配布されるシラバス等を用いて時間割を作成しなければならない。しかし、シラバスを閲覧するだけでは開講されている授業科目の内容を把握しきれず、学生の能力に適していない授業科目を履修してしまい、単位を取得できなかつたり、授業科目の内容を十分に理解できないことで留年したり、進級や就職活動などに不利になると考えられる。大学における留年者数の全学生に占める割合である留年率は、1988年度から増加傾向を示し [1]、留年学生に対するサポートが必要であると考えられる。

我々は、学生の得意不得意を考慮して学生の能力に適した時間割を作成するシステムを考案、開発した。これにより、学生の得意不得意分野を反映した時間割を作成し、学生に効率良く学習させることで学生の留年者率の低下や学生の授業科目の内容の理解に貢献することを目的とする。

2 関連研究

高橋らは、フィルタリング技術を活用したシステム *Airs* を構築している [2]。Airs では学習履歴に協調フィルタリングを活用することで、その学習者の特徴を推測している。学習者の学習履歴を用いる点では本研究と同じであるが、本研究では時間割を提供することができ、大学が開講している授業科目に対して支援できる。

由谷らは、電子シラバスに現れる用語などを用いた教養教育のカリキュラム分析を行っている [3]。科目間の類似度をコサイン類似度を用いることで表現している。科目間の類似度を算出している点では本研究と同じであるが、本研究では科目間の類似度を用いて学生に時間割を提供し、カリキュラムを把握できていない学生にも科目間の繋がりを反映した時間割を提供できる。

3 提案方式

3.1 Active Syllabus[4]

本研究室では、Active Syllabus という時間割自動作成システムを提案、開発した。Active Syllabus とは活性伝

播モデル [5] を利用した時間割自動作成システムである。

3.2 カリキュラム分析

大学で開講されている授業科目のシラバスデータを WEB 上で公開されているシラバスから取得する。各シラバスに出現する単語を形態素解析を用いて抽出し、その出現頻度により、授業科目 c_i のシラバスの文書ベクトルを作成する。1つの授業科目 c_i に出現する単語を tf-idf により重み付けを行い、 c_i を次の様に表わした。

$$c_i = (\omega_{i1}, \omega_{i2}, \dots, \omega_{in})$$

$$\omega_{i,t_j} = tf(t_j, c_i)idf(t_j)$$

$$idf(t_j) = \log(N/df(t_j))$$

ただし、 $tf(t_j, c_i)$ は授業科目シラバス c_i における単語 t_j の出現頻度であり、 N は総科目数、 $df(t_j)$ は単語 t_j が 1 回以上出現する科目数である。また、名詞の数、接続、記号、及び形態素解析に失敗しているであろうひらがな、カタカナ、アルファベット 1 文字の単語は除外した。

本システムでは授業科目間の内容の類似度を用いて学生の得意不得意分野を推定する。授業科目間の内容の類似度を算出するためにベクトル空間におけるコサイン類似度を用いた。コサイン類似度を用いて授業科目間の類似度は式 (1) より算出される。

$$L_{ij} = \frac{c_i c_j}{\|c_i\| \|c_j\|} = \frac{\sum_{k=1}^N \omega_{ik} \omega_{jk}}{\sqrt{\sum_{k=1}^N \omega_{ik}^2} \sqrt{\sum_{k=1}^N \omega_{jk}^2}} \quad (1)$$

3.3 授業科目の難易度と適性値

過去に授業科目が開講された時の不合格者数 *failed*、履修者数 *total* から授業科目の難易度 *difficult* を推定する。授業科目の難易度は式 (2) により定義される。

$$difficult = failed/total \quad (2)$$

授業科目の難易度、授業科目間の類似度を用いて、授業科目が学生に対してどれだけ向いているかを表わす適性値 *apt* を定義する。授業科目 a の適性値は式 (3)、(4) により算出される。

$$apt_a = \frac{\sum_{i=1}^n sim(a, is) \times difficult_{is}}{n} - \alpha_a \quad (3)$$

$$\alpha_a = \frac{\sum_{i=1}^m sim(a, if) \times difficult_{if}}{m} + difficult_a \quad (4)$$

ここで、 $sim(a, i)$ とは授業科目 i と授業科目 a の内容の類似度、 $difficult_a$ とは授業科目 a の難易度である。ま

A System for Generating University Time Schedules based on subject grade score points

†Yusei Nishimori †Yukio Hori ‡Yoshiro Imai

た, $difficult_{is}$ は単位を取得している授業科目の難易度, n はその総数. $difficult_{if}$ は単位を落とした授業科目の難易度, m はその総数である. 時間割作成時には式 (3), (4) により算出された授業科目の適性値が高い授業科目から履修する授業科目として時間割を作成する.

4 評価と考察

本システムの有効性を調べるため学生が実際に履修した時間割を用いて, 精度 $precision$, 誤認率 $misrecognition$ を算出した. 精度, 誤認率は式 (5), (6) により算出される.

$$precision = \frac{\text{不可科目回避数}}{\text{不可の科目総数}} \quad (5)$$

$$misrecognition = \frac{\text{取得単位を外した数}}{\text{取得単位総数}} \quad (6)$$

本提案 (apt) と他に作成した時間割作成手法とを精度, 誤認率で比較する. 今回比較に用いた他の時間割作成方法とは, rand, kw, as, ds, apt+kw, apt+as の 6 つの手法である. rand とは, ランダムに時間割を作成する手法である. kw とは, 学生の興味のあるキーワードに関連の強い授業科目を優先して選択する手法である. as とは, Active Syllabus による時間割の作成手法である. ds とは, 授業科目の難易度が低いものから順に選択する手法である. apt+kw とは, apt と kw を合わせたものである. apt における適性値と kw におけるキーワードとの関連度の合計が高い科目から履修科目に選択する. apt+as とは, apt と as を合わせたものである. as で時間割を作成する場合, 初期値として kw で作成される時間割を設定するが apt+as では apt を初期値とする.

実験に用いたシラバスは 2009 年度のシラバスである. 実験には香川大学工学部信頼性情報システム工学科の学生 10 名の 1 年後期から 3 年後期までの時間割を用いた. また, rand では精度と誤認率は 10 回試行を行い, その平均を取った. また, キーワードは学生が過去に履修した科目内で出現する単語の出現回数が多いもの上位 20 件を学生のキーワードとした. 精度と誤認率の実験結果は表 1 のようになった. 今回の考察からは, rand の結果は無視するものとする. これは, rand は実行毎に結果が異なるため, 最も良い場合の時間割を学生に提供できるとは限らないためである.

表 1 から精度では apt が最も良く, as, apt+as が最も悪いことが分かった. また, 次に精度が良い手法が ds であることから授業科目の難易度を学生の履修科目に考慮すると学生は単位を落としづらくなること考えられる. また誤認率では as, apt+as が最も良く, apt が最も悪いことが分かった. 精度の結果と対照的であることから現在の学生は授業科目の難易度や授業科目が自分に

表 1: 精度と誤認率

	精度の全体平均	誤認率の全体平均
rand	0.33	0.55
kw	0.20	0.52
as	0.18	0.42
ds	0.32	0.54
apt	0.33	0.56
apt+kw	0.25	0.53
apt+as	0.18	0.42

合っているかなどを考慮せずに時間割を作成していると考えられる.

5 まとめ

本稿では, 学生の得意不得意を考慮した時間割の自動作成について述べた. 評価実験により, 提案手法は自分に合った授業科目を知りたい学生や単位の取得が重要な学生にとって有効であると考えられる. 今後の課題として, 本システムを全学部対応にすることで実験人数を増やし, より正確な精度, 誤認率を算出すること, 及び精度, 誤認率をより良くするために提案手法の見直しを行う予定である.

参考文献

- [1] 内田千代子, 大学における休・退学, 留年学生に関する調査 第 27 報, 全国大学メンタルヘルス研究会報告書 第 28 回; 13-25, 2007.
- [2] 高橋泰樹; 松澤俊典; 山口未来; 土肥紳一; 和田雄次, 履歴データを活用した教育支援システム Airs, 情報科学技術フォーラム一般講演論文集 5(4), 367-368, 2006-08-21.
- [3] 由谷真之; 森幹彦; 喜多一, 電子シラバスを用いた大学教養教育のカリキュラム分析, 情報科学技術フォーラム一般講演論文集 4(4), 315-316, 2005-08-22
- [4] 堀 幸雄; 中山 堯; 今井 慈郎, 科目ネットワーク上の活性伝播を用いた時間割の自動生成システム, 情報処理学会論文誌, Vol.52, No.7, pp.2332-2342, 2011.
- [5] J.R. Anderson, A spreading activation theory of memory, Journal of Verbal Learning and Verbal Behavior, 22, pp.261-295, 1983.