

様式VI

博士学位論文調査報告書

論文題目 Studies on Content Analysis and Ordering of Courses from
a Knowledge-Based Perspective
(知識に基づく科目の内容分析と順序付けに関する研究)

申請者氏名 戴 憶菱

最終学歴 平成27年3月
立教大学大学院経営学研究科経営学専攻修士課程 修了

令和3年3月
京都大学大学院情報学研究科社会情報学専攻博士後期課程
研究指導認定見込

学識確認 令和 年 月 日 (論文博士のみ)

論文調査委員 京都大学大学院情報学研究科
(調査委員長) 教授 吉川 正俊

論文調査委員 京都大学大学院情報学研究科
教授 田島 敬史

論文調査委員 京都大学大学院情報学研究科
教授 緒方 広明

(続紙 1)

京都大学	博士 (情報学)	氏名	戴 憶菱
論文題目	Studies on Content Analysis and Ordering of Courses from a Knowledge-Based Perspective (知識に基づく科目の内容分析と順序付けに関する研究)		
(論文内容の要旨)			
<p>With the development of open and online education, students have easier access to a larger amount of learning materials. One of such resources, Massive Open Online Courses (MOOCs) make it possible for students to select courses based on their own learning goals. On the other hand, the one-fits-all curriculum policy in traditional higher education cannot prepare students for the competitive and segmented job market. To this end, adapting learning experience according to the rapidly-changing job market is essential to achieve fruitful learning and successful career development. This thesis envisions a picture of higher education in which the students are encouraged to select courses and design a personalized curriculum for various learning goals. Facilitating the integration and adoption of online courses motivates this thesis.</p> <p>With diversified learning goals and heterogeneous courses, it is challenging for students to discover appropriate courses. In this thesis, the author tackles the problems of course content analysis and course ordering from a knowledge-based perspective. In other words, this thesis puts emphasis on the matching and analysis of learning goals and courses at the level of knowledge. Considering different types of learning goals and scales of courses, this thesis addresses the following three tasks:</p> <ol style="list-style-type: none">1. Course content modeling. This task models the knowledge taught in a single course, helping understand and evaluate the course content. A curriculum guideline is utilized as the domain knowledge categorization to indicate what categories of knowledge are contained in the course. Specifically, the thesis proposes a Wikipedia structure-based method to capture the implicit relatedness of a course syllabus and the descriptions of the domain knowledge categories.2. Knowledge coverage estimation. In this task, the author assumes that the student has a specific learning goal, and estimate how much of the target knowledge is covered in a single course. The thesis defines a concept of knowledge category coverage as the proportion of the knowledge taught in the course to the knowledge required in the category. The thesis then models the knowledge category and the course as a set of concepts, and utilize the centrality upon a taxonomy to quantify the importance of concepts for the computation of the coverage.3. Course ordering. This task identifies the optimal order to take related courses for a			

given learning goal. This thesis focuses on “technical terminologies” which are frequently required in the job market. Given a technical terminology, the author aims at identifying an order of courses which contributes to the acquisition of the terminology and also follows the prerequisite relationships among courses. The author developed a two-step approach in which course-terminology relatedness is firstly estimated and then courses are ordered based on the prerequisite relationships and the estimated relatedness. In addition to an information retrieval-oriented evaluation metric, the author explored whether the order is effective from pedagogical perspectives.

This thesis has advantages over existing works as it a) supports direct learning goals represented as domain knowledge categories and job opportunities, b) provides a systematic view on course content by utilizing the domain knowledge categorization, c) quantifies knowledge coverage to provide a helpful criterion in course selection, and d) models a rank-sensitive ordering problem where the information gain of every position is optimized. Furthermore, promising results were obtained to show that the proposed methods are effective.

(論文審査の結果の要旨)

大規模公開オンライン講座の推進により、学習者は時間や空間の制限なく学習することができるようになりつつある。学習者は、知識への好奇心、スキルアップ、就職の準備など多様な学習目標を持つと考えられる。その一方、公開されたオンライン科目の教育目的、教育内容や知識の表現は多岐にわたる。したがって、学習者が自身の目標に合う科目を識別することは極めて困難である。本論文は、こういった状況に着目し、科目の内容分析と順序付けに取り組むことで学習者の科目選択を支援する手法を研究した結果として、以下の三つの成果をまとめている。

1. 知識階層を用いた科目内容の分析

一つの科目の内容を理解する、あるいは複数の科目の内容を比較するには共通のドメイン知識階層が有用である。そこで、この知識階層を用い、科目内容とそれぞれの知識カテゴリとの関連度を推定することを目標とした。Wikipedia といった外部の知識ベースを利用し、科目内容と知識階層の深層的な関連度を推定することを提案した。実験では、提案手法が、Labeled Latent Dirichlet Allocation というベースライン手法より良い精度が得られることを示した。

2. 知識カバレッジの定義と推定

目標とする知識のカテゴリが科目の受講によって習得できる知識の割合を知識カバレッジと定義した。そして、知識のカテゴリが概念により構成されることから、概念の依存関係によるカテゴリへの重要度を推定する手法を提案した。実験では実データを収集し、手法の有効性を示した。

3. 複数の科目の順序付け手法の開発

職に就くための学習目標を達成するには計画的に科目の受講プランを立てなければならない。この研究では、学習目標として与えられた専門用語に対し、科目の依存関係を満たす科目の受講順序を与える手法を開発した。既存研究と異なり、順序中の各ポジションにある科目がもたらす情報量が最適化される点が優れている。実験では、手法が生成した順序の学習体験への影響も評価し、より学習者の個人的なニーズに応えられるモデルの開発に繋がることを示唆した。

以上、本論文は、主に科目の内容分析と順序付けに関する三つの問題に取り組んだ。その研究成果は、大規模公開オンライン講座の科目選択の発展に資するもので、学術上、實際上寄与するところが少なくない。よって、本論文は博士（情報学）の学位論文として価値あるものと認める。また、令和3年2月18日、論文内容とそれに関連した事項について試問を行った結果、合格と認めた。また、本論文のインターネットでの全文公表についても支障がないことを確認した。