

標準カリキュラムに基づくオンラインコース間の順序関係分析

蔵野晴基[†] 高橋幸雄[†] 河村奨^{††} 一島力男[†] 布田徹[†] 卯木輝彦[‡]
[†] 国士舘大学 ^{††} リブライズ [‡] フォトロン

1. はじめに

オンライン上で様々な知識、技術を学ぶことのできるオンラインコースや学習コンテンツが増えている。このような学習コンテンツは教育の機会をオープンにすることで多くのユーザに対して知識獲得の機会を提供するものである。オンライン学習コンテンツには前提となる知識や後続する内容といった情報が体系化されておらず、学習者はオンライン学習コンテンツをどう順序で学ぶかということが把握しにくいという問題がある。本研究はコンピュータサイエンス分野の標準カリキュラムの知識体系を基準として、オンライン学習コンテンツ間の順序関係の抽出方法を提案する。

2. 関連研究

中島らはシラバスデータに含まれる名詞の包含関係を用いて、科目間の学習順序を推定する仕組みを提案している[1]。この手法により科目間に関して学ぶ順序情報を機械的に得ることができるが、この手法ではシラバスにこれから学ぶ学習分野とそれを学ぶ上で必要な前提分野が記載されていることが前提となっている。本研究では学習分野、前提分野の学習コンテンツから順序関係を得るために学習内容テキストから Wikipedia を用いて上位語、下位語を抽出し[2]、科目の順序関係の判定を行なう。疋田らはコンピュータサイエンス分野の標準カリキュラム[3]に出現するトピックをどの程度カバーしているのかという特徴を用いてカリキュラムの特徴を分析する方法を提案している[4]。この研究ではカリキュラムの特徴は分析できるが、どう順序で学ぶことが望ましいのかということについては考慮していない。

3. 提案手法

提案手法は学習コンテンツ c_i について Wikipedia から得られる上位語 $\text{hypernym}(c_i)$ 、下位語 $\text{sub}(c_i)$ を用いて標準カリキュラムから得られる技術的キーワードを考慮する。具体的には学習コンテンツ c_i, c_j 間における順序関係の強さを示すスコアを S_{ij} とし、下記のように定義する。 α は予備実験により最も精度の高かった値 (0 - 1) を用いるものとする。

$$S_{ij} = \alpha \frac{c_i \text{ hypernym}(c_j)}{\|c_i\| \|\text{hypernym}(c_j)\|} + (1 - \alpha) \frac{\text{sub}(c_i) c_j}{\|\text{sub}(c_i)\| \|c_j\|}$$

具体的に図 1 に Coursera にあるコース Master of Computer and Information Technology における順序関係の例を示す。

Data Structures & Software Design

This course focuses on data structures, software design, and advanced Java; starting with an introduction to data structures and basics of algorithm analysis. Important data structures covered include arrays, lists, stacks, queues, trees, hash maps, and graphs. The course also focuses on software design and advanced Java topics such as software architectures, design patterns, and concurrency.

Computer Systems Programming

This course is a continuation of Data Structures & Software Design and introduces students to fundamental concepts in computing systems. Divided into two parts, the first half of the course introduces important concepts in modern operating systems: processes, scheduling, caching, and virtual memory. The second half of the course provides an introduction to fundamental concepts in the design and implementation of networked systems, their protocols, and applications. The course will use the C program language, and will develop your knowledge on C system calls, and libraries for process/thread creation and manipulation, synchronization, and network communication.

図1. 順序関係のある学習コンテンツとその説明

このコースでは前者を前提として、後者のコースを履修することが指定されている。各コースの概要から標準カリキュラムに含まれるキーワードに下線を示した。学習内容には知識内容だけでなく具体的技術である C Program, System Call 等のキーワードが書かれている場合(二重線)も存在する。標準カリキュラムのみを用いる方法では、このような具体的技術を表すキーワードを捉えることができないため Wikipedia から得られる上位語、下位語を用いる。

4. おわりに

本研究はオンライン学習コンテンツ間の学習順序関係の抽出方法として、学習コンテンツの説明内容に含まれる内容と Wikipedia から得られる上位語、下位語を用いて順序関係を計算する方法について述べた。今後の課題として順序関係の分かっている学習コンテンツを調査し、提案手法の効果を検証することが挙げられる。

謝辞

本研究の一部はトランスコスモス財団の助成を受けたものです。

参考文献

- [1] 中島克也, 小林亜樹: テキストメタデータを用いたオープン教材のための学習順序推定手法の提案, 情報処理学会第 14 回情報科学技術フォーラム, RK-006, 2015.
- [2] Jong-Hoon Oh, Ichiro Yamada, Kentaro Torisawa and Stijn De Saeger, "Co-STAR: A Co-training Style Algorithm for Hyponymy Relation Acquisition from Structured and Unstructured Text," In Proceedings of COLING-2010, pp.842-850, 2010.
- [3] カリキュラム標準コンピュータ科学領域(CS) https://www.ipsj.or.jp/annai/committee/education/j07/ed_j17-CS.html (2020 年 1 月 30 日参照).
- [4] 疋田輝雄, 石畑清: シラバスに基づく理工系情報学科のカリキュラム調査, 情報処理, Vol.52, No.8, pp.1020-1025, 2011.