

# 標準カリキュラムを用いた求人情報からの対応知識の抽出

戴 憶菱<sup>†</sup> 浅野 泰仁<sup>‡</sup> 吉川 正俊<sup>‡</sup>

<sup>†</sup> 京都大学大学院情報学研究科 〒606-8501 京都市左京区吉田本町

E-mail: <sup>†</sup> daiyiling@db.soc.i.kyoto-u.ac.jp, <sup>‡</sup> {asano, yoshikawa}@i.kyoto-u.ac.jp

**あらまし** 大量な電子求人情報から業界のニーズを把握することが期待できる。学生にとって就きたい職業を意識して学習の計画を立てることが望まれる。また、教育者が科目の提供において業界のニーズへの対応策に関心がある。ゆえに、高度教育の知識体系から職業に求められる知識を捉えることが学習計画の作成や科目の設計などに役立つと考える。本研究は標準カリキュラムを照合知識体系として用い、求人情報から大学で取り扱う知識が対応すると対応しない知識を抽出することを目的とする。求人情報には当該ポジションに必要な知識を表す文以外に、企業宣伝、選考プロセスなどの情報も含まれている。また、知識を表す文に絞っても、広い概念と狭い概念、具体的な技術と抽象的な知識を表す文が混在している。本研究はこれらの難点を克服できる情報検索手法を模索する。

**キーワード** 標準カリキュラム 求人情報 知識体系

## 1. はじめに

学生の興味と長所に合う仕事に必要な知識を授けることは高度教育の一つの役割である。学生が、早い段階で将来のキャリアを意識しながら自分に適切な科目を選択することが望ましいと考える。Jones ら[1]の工学部学生のモチベーションについての調査から、工学への価値認識が将来の工学関連のキャリアへの興味と繋がっていることが分かる。ゆえに、大学新入生のような初心者にとって、科目やプログラムと職業との対応関係が分かれば、当該分野への認識が深まり、学習計画の作成が容易になると考えられる。

電子求人情報から業界に求められる知識やスキルが覗える。そのポジションに必要な知識とある科目が取り扱う知識の対応関係からそのポジションとその科目の関連度を計算することができる。しかし、こういうアプローチはアドホックで、ポジション間、科目間、業界全体、プログラム全体の状況が見えない。また、知識の体系の視点からみると、業界のポジションの配置から反映される知識体系と教育界の科目の設置から反映される知識体系には大きな違いが存在する。業界において、一つひとつのポジションは特化した業務を執行するために設けられたもので、知識の応用シーンに応じた知識体系が埋め込まれている。一方、教育界では、それぞれの科目がその分野の知識を理解するための学習の効率や便宜を考慮したもので、知識の学習シーンに応じた知識体系が埋め込まれている。図1の例をもって説明すると、あるポジションが「オブジェクト指向プログラミング」、「MVCモデル」と「アジャイルソフトウェア開発手法」三つの知識を要する。このポジションにとって、三つの概念が同じ重要な位置を占めていることが覗える。しかし、学術的な視点からみると、この三つの概念は異なる広さと深さを持ち、同一に扱われるのは望ましくない。具体的に、計算機

科学分野の権威たる標準カリキュラム Computer Science Curricula 2013[2] (下記 CS2013) の知識体系に参照すれば、「オブジェクト指向プログラミング」は「プログラミング言語」という知識カテゴリに属し、「MVCモデル」は「ヒューマンコンピュータインタラクション」に属し、「ジャイルソフトウェア開発手法」は「ソフトウェア工学」に属することになる。計算機科学の知識を身に付けるには、「ソフトウェア工学」と「プログラミング言語」はより基礎的な知識であり、色んな職業に必要とされる。一方、「ヒューマンコンピュータインタラクション」は人間と計算機とのインタラクションを中心とする知識であり、計算機科学の一つの分科である。図1の示すポジションを目指す学生が学習の計画を立てるとすると、「プログラミング言語」、「ソフトウェア工学」、「ヒューマンコンピュータインタラクション」という学習の順序が進められる。

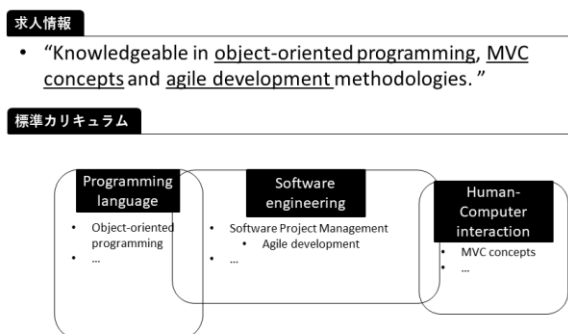


図1 求人情報と標準カリキュラムの知識体系の違い

上述のように、職と科目との関連性を抽出するには共用の知識参照体系が必要である。また、学習計画の作成、科目の設計などの教育的な利用シーンに支援するため、当該分野の知識の基礎度や抽象度を示せる知

識体系が望ましい。したがって、本研究は CS2013 に定義された知識体系を用い、求人情報と標準カリキュラムとの対応知識の抽出を目的とする。

## 2. 関連研究

### 2.1. 教育界と業界とのギャップについて

計算機科学の分野では、業界に求められる知識や能力と学術プログラムで教える知識や能力とのギャップに関して数多くの研究が行われてきた[3]-[7]。これらの研究は業界の管理者や現職者、大学の講師や学生を対象に、インタビューやアンケート調査を行い、業界と教育界が重視する知識や能力のギャップを考察した。先ず、これらの研究は業界が求めている知識や能力と学術プログラムで扱う知識や能力の差をマクロな視点から分析した。つまり、学術プログラム全体や計算機科学に関連する業界の全体が計算機科学分野の全体に対する認識を分析したため、1) それぞれの会社のポジションとそれぞれの教育機関の科目との対応関係が分からない、2) 分析の対象となる知識の分類が粗雑であるという二つの欠点がある。

もう一つのグループの研究は電子求人情報から業界に必要とされる知識をマイニングした[8], [9]。これらの研究は主に「HTML」、「Java」のような具体的な技術に着目し、求人情報に出現頻度の高い技術を業界に求められる計算機科学分野の知識と見なす。しかし、計算機科学分野の知識を身に付けるには単なるそれぞれ具体的な技術を学ぶことは効率的ではない。各技術に埋め込まれている抽象的な概念や理論を見極め、体系化してから学ぶのは望ましい。本研究は標準カリキュラムの知識体系で求人情報に必要とされる知識を捉えることでポジションと科目のマッチングに支援する。

### 2.2. 求人情報の推薦について

企業の採用プロセスの電子化によって求人側と求職側の求人サイトの利用が増えている。これを背景に、利用者のプロフィールに応じた職を推薦することが求人サイトの一つのタスクである。

Agrawal ら[10]は求人サイトの利用者プロフィール、職業プロフィールと科目内容に表れる概念の類似度から利用者に適応する職業や科目を推薦するフレームワークを提案した。しかし、実践結果の欠如により提案手法は求人情報と科目とどの程度の精度でマッチングできるかは未知である。また、あるポジションとある科目の類似度という出力だけは求職者の学習計画の作成には十分な情報ではない。本研究は標準カリキュラムの知識体系を経由した求人情報と科目との知識対応状況を提供することで、求職者あるいは学習者の知識

のマネジメントに支援する。

### 2.3. 職場における学習資料の推薦について

職場で任された業務を遂行するため常に新しい知識を学ばなければならない。Saraswat ら[11]は職場の環境における学習に着目し、従業員の背景知識と任されたタスクに必要とされる知識に適応する学習資料の推薦システムを提案した。彼らはある図書館の分類システムに定められたキーワードを用い、従業員の背景知識、タスクに必要な知識、学習資料がカバーする知識を手で注釈した。しかし、こういうアプローチは大量な求人情報と科目の関連付けには不向きである。また、自由記述の求人情報から専門分野知識を自動的に識別することは容易ではない。本研究はそういう難点を克服できる情報検索の手法を模索する。

## 3. 問題設定

求人情報には、当該ポジションに必要な知識と能力以外に、多様な情報が混在している。Chopra と Shivangi [12]によると、求人情報の内容を「専門知識」、「教養知識」、「役得」、「選考プロセス」、「企業宣伝」、「学歴」などのカテゴリに分類することができる。本研究は求人情報と標準カリキュラムとの知識対応の抽出を目的とするため、「専門知識」を表す文章に絞る。

一方、CS2013 は ACM と IEEE との連携特別委員会により編集され、大学学部計算機科学プログラムで扱うべき知識をまとめた指導要領である。CS2013 では、18 個の Knowledge Area (下記 KA) から成り立つ。各 KA がさらに幾つの Knowledge Unit (下記 KU) を含み、各 KU は複数の Topic を含む。表 1 が示すように、「Information Management」という KA が「Database Systems」などの KU を含み、「Components of database systems」などの Topic を含む。

表 1 CS2013 知識体系の例

KA	KU	Topic
Information Management	Database Systems	• Approaches to and evolution of database systems
		• Components of database systems
	Data Modeling	• Data Modeling
..	..	• ...

本研究は図 2 の示すように、与えられた求人情報から専門分野知識を表す文章を識別し、それと対応する CS2013 における知識のカテゴリを抽出することを問題とする。

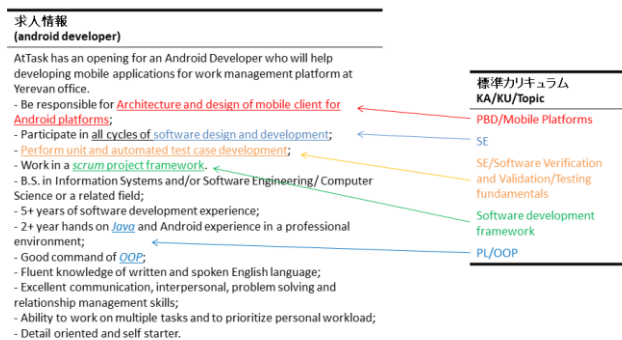


図2 問題設定の例

#### 4. 予想される困難な点

図1と図2の例から分かるように、求人情報の中に表す専門知識には異なる粒度を持っている。例えば、「software design and development」という上位概念もあれば、「automated test case development」という下位概念もある。また、求人情報の中に具体的な技術と抽象的な概念を表す文章が混在している。例えば、「Java」と「OOP」が同時に書かれているケースがある。これらの違う粒度や特徴を持つ知識を表す文章を正確に区別する必要がある。一つの案としては、知識のタイプを整理した上で、それぞれのタイプに応じた手法を通じて標準カリキュラの知識体系と対応する知識を抽出することが考えられる。

#### 参考文献

[1] B. D. Jones, M. C. Paretti, S. F. Hein, and T. W. Knott, "An Analysis of Motivation Constructs with First-Year Engineering Students: Relationships Among Expectancies, Values, Achievement, and Career Plans," *Journal of Engineering Education*, vol. 99, no. 4, pp. 319–336, 2010.

[2] I. C. Society, "Computer Science Curricula 2013: Curriculum Guidelines for Undergraduate Degree Programs in Computer Science," ACM, 2013.

[3] S. Surakka, "What Subjects and Skills Are Important for Software Developers?," *Commun. ACM*, vol. 50, no. 1, pp. 73–78, Jan. 2007.

[4] C. Aasheim, L. Li, and S. Williams, "Knowledge and Skill Requirements for Information Technology Graduates: A Comparison of Industry and Academia," *Journal of Information Systems Education*, pp. 349–356, Jan. 2009.

[5] A. Radermacher, G. Walia, and D. Knudson, "Investigating the Skill Gap Between Graduating Students and Industry Expectations," in *Companion Proceedings of the 36th International Conference on Software Engineering*, New York, NY, USA, 2014, pp.

291–300.

[6] T. C. Lethbridge, "What Knowledge Is Important to a Software Professional?," *Computer*, vol. 33, no. 5, pp. 44–50, May 2000.

[7] M. Exter, "Comparing Educational Experiences and On-the-job Needs of Educational Software Designers," in *Proceedings of the 45th ACM Technical Symposium on Computer Science Education*, New York, NY, USA, 2014, pp. 355–360.

[8] P. J. Kovacs and G. A. Davis, "Determining Critical Skills and Knowledge Requirements of IT professionals by Analyzing Keywords in Job Postings," *Issues in Information Systems*, vol. 1, no. 9, pp. 95–100, 2008.

[9] K. S. Koong, L. C. Liu, and X. Liu, "A Study of the Demand for Information Technology Professionals in Selected Internet Job Portals," *Journal of Information Systems Education*, vol. 13, no. 1, pp. 21–28, Jan. 2002.

[10] B. Agrawal *et al.*, "4C: Continuous Cognitive Career Companions," in *Artificial Intelligence in Education*, 2017, pp. 623–629.

[11] N. Saraswat, H. Ghosh, M. Agrawal, and U. Narayanan, "Contextual Recommendation of Educational Contents," in *Artificial Intelligence in Education*, 2015, pp. 439–448.

[12] Chopra, Shivangi, "Job description mining to understand undergraduate co-operative placements," Master's Thesis, UWSpace, 2017.