

電子化されたシラバスに基づく 学位授与事業のための科目分類支援システムの試作

宮崎 和光[†] 井田 正明[†] 芳 鐘 冬 樹[†]
野 澤 孝 之[†] 喜 多 一^{††}

大学評価・学位授与機構では、短期大学・高等専門学校卒業生および専門学校修了者等を対象に、単位累積加算を基にした学士の学位授与事業を行っている。この制度を利用し学士の学位授与を希望する者は、各専門分野ごとに定められた所定の単位を修得しなければならない。申請者は、自らの判断で修得した科目を分類・整理し申告する。それに対し、大学評価・学位授与機構では、申請者による分類の正しさを、各専門分野の専門委員が申告された科目のシラバスを読むことで検討している。しかしながら、近年の申請者数の増大から、この作業には膨大な時間と労力を要しており、情報技術を活用した科目の分類支援が望まれている。本稿では、電子化されたシラバスを利用して、申告された科目の分類を支援する科目分類支援システムの試作を行う。例として、R大学の39科目に適用することで、専門委員の作業量の軽減に寄与することを示す。

Development of a Course Classification Support System for the Awarding of Degrees Using Syllabus Data

KAZUTERU MIYAZAKI,[†] MASAOKI IDA,[†] FUYUKI YOSHIKANE,[†]
TAKAYUKI NOZAWA[†] and HAJIME KITA^{††}

The National Institution for Academic Degrees and University Evaluation is engaged in the awarding of academic degrees based on the accumulation of credits. These credits must be classified according to pre-determined criteria for the chosen disciplinary field. This work is carried out by the sub-committees within the Committee of Validation and Examination for Degrees, whose members are well-versed in the syllabus of each course. The number of applicants is increasing yearly so that a course classification system supported by information technology is strongly desired. In this work, we have developed a prototype of *Course Classification Support system* using syllabus data. We have shown its effectiveness in 39 courses of R university.

1. はじめに

大学評価・学位授与機構 (National Institution for Academic Degrees and University Evaluation, 以下、NIAD-UE と略す) では、短期大学・高等専門学校卒業生および専門学校修了者等を対象に学士の学位を授与する事業を行っている。この制度を利用し学士の学位授与を希望する者 (申請者) は、大学の科目等履修生制度を利用した学修や、短期大学や高等専門学校に置かれる NIAD-UE が認定した専攻科にお

ける学修等により、各専門分野ごとに定められた所定の単位を修得しなければならない。

NIAD-UE では、専門分野として、「文学」「工学」等全部で 26 の専攻分野、およびその下位区分として、「国語国文学」「情報工学」等全部で 53 の専攻区分を用意している (平成 14 年度時点)。申請者は、専攻区分の中から 1 つの区分を選択し、修得した単位を申請者の判断で、その選択した区分に合致するように分類・整理し申告する。それに対し NIAD-UE では、申請者が所定の単位を適切に修得しているかどうかを、当該科目のシラバス (授業計画書) に基づき審査する。特に、専門性の高い大量の科目をシラバスのみを考慮して短期間で集中的に分類することが要求されている。

シラバスは学生に対する授業の選択と履修時のガイドのための情報提供を本来の目的とするが⁽¹⁾、授業内

[†] 独立行政法人大学評価・学位授与機構
National Institution for Academic Degrees and University Evaluation

^{††} 京都大学
Kyoto University

容を最も詳しく示す資料として、大学評価の支援^{2),3)}や大学における教育情報の活用⁴⁾等にも積極的に活用されている。さらに、単位レベルでの教育の質の保証は、高等教育の多様化、国際化に際しての重要事項となっており、その判断材料の1つとして、シラバスは国内外を問わず、近年、注目を集めている。

現在、WWWでの学生への情報提供等を目的として、各大学ではシラバスの電子化が進んでいる。この電子化されたシラバスを扱った研究として、電子化されたシラバスの自動収集とそこからの情報抽出を目指したものが知られている⁴⁾⁻⁶⁾。それに対し、我々の主眼は、学位授与事業の支援のためにシラバスをその内容に基づき分類することにある。現在、この作業は、大学等の教員で構成される専門委員会で行われているが、近年の申請者数の増大から、膨大な時間と労力を要している。特に、年間の総申請者数は平成14年度には2,457名に達し、その後も、年約5%の割合で増加しつつある。そのため、情報技術を活用した分類支援が強く望まれている。

ところで、NIAD-UE同様、直接には教育機能を持たない機関が、学位を授与する例として、英国の学位授与評議会(CNAA)やアイルランド共和国の国立学位評議会(NCEA)が知られている^{7),8)}。しかし、そこでの学位授与事業は、機関認定を基本とするものであり、NIAD-UEの個人申請とは趣が異なる。一方、米国には、当該機関での履修を要件とせず評価のみによって学位を授与するいくつかの大学が存在する^{9),10)}。しかし、そこでは、その各大学に属する学生により作成されたポートフォリオや面接等による総合的な評価が主であり、教育機関としての機能を持たないNIAD-UEの制度とは根本的に異なる。これらのことから、NIAD-UEの学位授与事業は我が国独自の制度であり、そこで求められている「専門性の高い大量の科目をシラバスのみを考慮して短期間で集中的に分類すること」はほかに類をみない要求であると考えられる。

本稿では、電子化されたシラバスを利用して、申告された科目の分類を支援する科目分類支援システムの試作を行う。以下、2章では、問題の所在を明確にするために、NIAD-UEの学位授与制度、ならびにそこでの科目分類の現状および改善のための基本方針について述べる。3章では、科目の分類を支援するシステムである科目分類支援システムを試作する。続く4章

では、科目分類支援システムの活用例を示すことで、システムの評価を行う。5章は結論であり、今後の課題をとりまとめる。

2. 学位授与制度と科目分類

2.1 単位修得に着目した申請方法の概略

図1に、学士の学位授与制度の概要を示す。学位授与事業の制度・申請方法等の詳細は、NIAD-UEが毎年度発行している「新しい学士への途」に詳しく記されている。申請者は、まず、第1に、短期大学や高等専門学校を卒業する等の基礎資格を有する必要がある。NIAD-UEでは、基礎資格をその修業年限等によって、3つの申請区分に分類している。各申請区分には、それぞれ修得すべき単位数が基本基準として示されている。申請者は、自らの学修に適する申請区分を1つ選択し、その申請区分に課せられている基本基準を満足する必要がある。

さらに申請者は、自らの学修に最も適する専攻区分を1つ選択し、その選択した専攻区分に課せられている単位の修得基準である専攻基準も同時に満足させなければならない。例として、図2に、平成14年度版の「新しい学士への途」¹¹⁾から転載した「情報工学」区分の専攻基準表を示す。専攻基準表には、修得すべき単位数が明記された「専門科目」、「関連科目」の区分欄、および「専門科目の例」が記載されている。「専門科目」、「関連科目」の区分欄は、さらに、「科目区

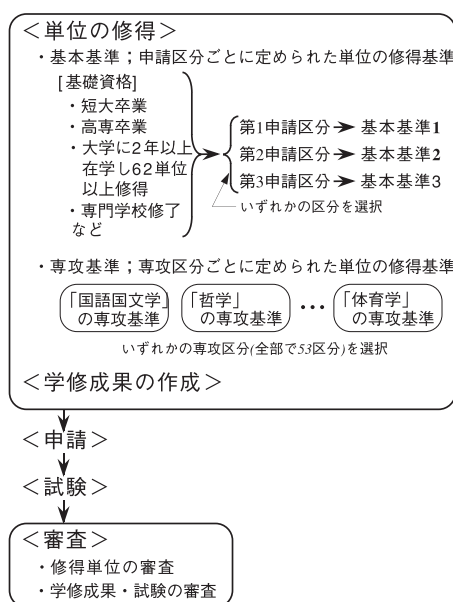


図1 学士の学位授与制度の概要

Fig. 1 Illustration of the awarding of a bachelor's degree in NIAD-UE.

なお、現在、これらの機関はともに廃止され、その機能は、各々、オープン・ユニバーシティ評定サービス(OUVS)および高等教育訓練資格カウンシル(HETAC)に引き継がれている。

整理番号	専攻区分	専門・関連科目の区分及び修得すべき単位数 (62単位以上)	専攻分野
41	情報工学	<p>専門科目 (40単位以上)</p> <p>【A群 (講義科目)】 (30単位以上) ○情報工学基礎理論に関する科目 (4単位以上) ○計算機システムに関する科目 (6単位以上) ○情報処理に関する科目 (6単位以上) ○情報に関連する科目 【B群 (演習・実験・実習科目)】 (6単位以上) ○情報工学に関する演習・実験・実習科目</p> <p>関連科目 (4単位以上)</p> <p>◇工学の基礎となる科目 ◇工学及び周辺技術等に関する科目</p>	工学

■専門科目の例■

【A群 (講義科目)】
 ○情報工学基礎理論に関する科目
 データ構造, 言語理論とオートマトン, アルゴリズム, 計算理論, 符号理論, スイッチング回路理論, 信号処理論, 論理学, 情報工学, 情報数学, 離散数学 (グラフ理論, 組合せ論), 計画数理 (OR, 待行列) など
 ○計算機システムに関する科目
 プログラミング, プログラム言語, コンパイラ, オペレーティングシステム, デジタル回路, 計算機構成, 計算機アーキテクチャ, ネットワーク, 情報セキュリティ, 分散処理, ソフトウェア工学, データベースシステム, 性能評価など
 ○情報処理に関する科目
 数値解析, 人工知能, 知識工学, エキスパートシステム, 自然言語処理, 音声処理, 画像処理, 図形処理, パターン認識, シミュレーション, 通信処理 (プロトコル), グループウェア, マルチメディア, ヒューマンインターフェース, コンピュータグラフィクス, ロボティクスなど
 ○情報に関連する科目
 乗積回路, 電子回路, システム制御理論, 通信網, 最適化論, 計測工学, 生体情報処理など
 【B群 (演習・実験・実習科目)】
 ○情報工学に関する演習・実験・実習科目

図 2 「情報工学」区分の専攻基準表
 Fig. 2 The criteria for the information engineering sub-field.

表 1 「情報工学」区分の科目区分番号
 Table 1 Item numbers of the information engineering sub-field.

科目区分 1	情報工学基礎理論に関する科目
科目区分 2	計算機システムに関する科目
科目区分 3	情報処理に関する科目
科目区分 4	情報に関連する科目
科目区分 5	情報工学に関する演習・実験・実習科目
科目区分 6	関連科目
科目区分 7	専攻外科目

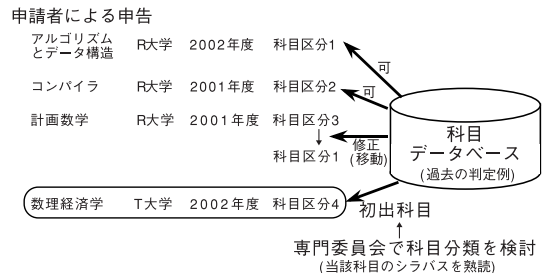


図 3 現状の科目審査プロセスの例
 Fig. 3 An example of an actual course classification.

分」と呼ばれるいくつかの「～に関する科目」に分類され、それぞれの「科目区分」ごとに必要な単位数が示されている。「専門科目の例」は、「専門科目」について、それぞれの「科目区分」に該当する具体的な科目名の例が示されている。

申請者は、修得した科目を、その内容により、各専攻区分に係る「専門科目」、「関連科目」およびそれ以外の「専攻外科目」に自ら分類・整理し申告しなければならない。これは、上述の「情報工学」区分の場合は、表 1 の 1~7 のいずれかの科目区分に修得した科目を分類することを意味する。このように申請が自己申告を基本としているため、場合によっては、不適切な科目分類が行われる可能性があり、NIAD-UE において、その適切性の確認等が行われている。

図 2 のような表は、はじめに述べた米国の大学にお

さらに、学修成果と呼ばれるレポートを提出し、そのうえで、修得単位の審査および、学修成果の内容に基づいて実施される試験に合格した後に、学士の学位が授与される (図 1 参照)。

いても作成され学生の評価に利用されている⁹⁾。しかしここでは、主として教養レベルの学修の評価を目的に行われており、NIAD-UE が行っているような専門科目の詳細な分類は要求されていない。

2.2 現状の科目審査プロセスの概要

専門科目の判定は、各分野に精通した専門化による判定が必要である。NIAD-UE では、そのような専門性の高い科目分類の確認を支援するために、平成 12 年度から、過去の科目分類の判定結果をデータベース化し、専攻基準の判定に活用している。このデータベースは科目データベースと呼ばれており、現状では、当該シラバスと分類判定結果が「1 対 1」に対応するように作られている。図 3 に、科目データベースを用いた専攻基準判定のための科目審査プロセスの一例を示す。まず申請者による分類と科目データベースとのマッチングをとり、一致するものは「可」とする。図 3 の例では、「アルゴリズムとデータ構造」「コンパイラ」

がこれに相当する。一方、不一致となった場合には、機械的に申請者による科目分類を、科目データベースの分類に修正（移動）する。図3の例では、「計画数学」がこれに相当する。

それに対し、過去に判定例が存在しない科目（初出科目）については、専門委員会において、申請者による科目分類の正しさを、その科目のシラバスを読むことで検討している。図3の例では、「数理経済学」がこれに相当する。初出科目に対する判定作業は、基本的には対応するシラバスをすべて読むことで行われており、近年の申請者数の増大から、その作業には膨大な労力と時間を要している。

2.3 科目審査プロセス改善のための基本方針

ところで、決定木や Support Vector Machine 等のテキスト分類手法¹²⁾を用いれば、初出科目を自動分類することも可能である。しかし、本制度の性格上、判定結果に誤りが存在することは絶対に許されない。そのため、科目判定そのものの自動化は考えない。むしろ、判定すべき科目数の削減を通じて、専門委員の負担軽減を目指す。すなわち、我々の主眼は、情報システムのみによる自動分類の実現にはなく、人を積極的に取り込んだ形での「人による分類の情報システムによる支援」にある。

判定すべき科目数の削減という意味では、専攻基準を満たすために必要な科目以外の科目を判定する必要はない。もちろん、すべての科目を判定しなければ、専攻基準を満たさないケースも考えられるが、そうでない場合には、各科目区分ごとに、その科目区分に分類されることが有望であると考えられる科目を優先的に提示することができれば、判定作業の軽減につながるものとする。

次章では、このような科目分類に係る判定作業を支援するシステムとして、科目分類支援システムを試作する。

3. 科目分類支援システム

3.1 システムの概要

本稿では、電子化されたシラバスを入力とし、「新しい学士への途」に即した科目分類候補（群）を出力するシステムの試作を行う。ここで「科目を分類する」とは、たとえば、「情報工学」区分の場合、申請者が大学等で修得し NIAD-UE に申告した各科目を表1の1~7のいずれかに分類することを指し、「科

目分類」は、その分類された科目区分を指す。科目分類の候補が得られれば、2.3節で述べたような形での専門委員会の判定支援に寄与する利用が可能になるものとする。以下では、「情報工学」区分に話題を限定するが、他の専攻区分に対しても、同様の作業は可能である。

科目分類の要は、科目データベースにある。しかし現実には、科目データベース中に判定例が存在しない科目が多々申告されてくる。また、現在の科目データベースは、シラバスと科目区分が「1対多」に対応していない等十分に整備されているとはいえない。そこで本研究では、まず初めに、実際の審査業務に活用されているものとは独立した研究用の科目データベースである科目DBの作成を行う。そしてその後、科目DBから各科目区分ごとに特徴的な語句を抽出し、その語句と分類支援を必要とする科目のシラバスとのマッチングをとるシステムとして科目分類支援システムを提案する。

3.2 科目分類支援システムの試作

3.2.1 ステップ1：科目DBの作成

本項では、科目DBの作成を行う。具体的には、まず初めに、平成14年度に、インターネット上に公開されている比較的充実した内容を持つ13大学の情報系学科の電子化されたシラバスデータを収集した。これらすべてのシラバスを筆者が熟読し、「新しい学士への途」¹¹⁾に記載されている例示科目を参考に、表1の科目区分1~7までの番号付けを行った。この際、1つの科目が、表1の複数の科目区分に分類されることを許す。これは、当該科目のシラバスと表1が「1対多」に対応付けられたDBの作成を意味する。また、表1の科目区分1~5に分類困難な専門科目は科目区分0として分類した。

対象科目数は全部で1,010科目であるが、以下では、結果の検証用に用いたR大学の49科目を除いた961科目の分類結果を科目DBに登録した。この961科目を次に示す分類項目に従って分類したときの各分類項目に含まれる科目数を表2に示す。

- c_0 ; 専門科目ではあるが、科目区分1~5には分類困難な科目。
- c_i ; 科目区分 i ($i = 1, 2, \dots, 7$) に分類される科目。
- ci_0 ; 便宜上、科目区分 i ($i = 1, 2, \dots, 7$) に分類したが、それ以外の区分に分類される可能性も考えられる科目。
- c_{ij} ; 科目区分 i ($i = 1, 2, \dots, 7$) と科目区分 j ($j = 1, 2, \dots, 7$; $j \neq i$) の両方の内容を含む

¹¹⁾「新しい学士への途」は毎年度改定されているが、本稿では、平成14年度版¹¹⁾をもとに検討を進めた。

表 2 各分類項目に含まれている科目数および専門用語数
Table 2 The number of courses and technical terms for each item.

分類項目	c0	c1	c2	c3	c4
科目数	68	108	104	85	54
専門用語数	2496	2419	2746	2637	2055
分類項目	c5	c6	c7	c10	c12
科目数	86	351	48	7	21
専門用語数	2395	7075	1249	294	929
分類項目	c13	c14	c16	c23	c24
科目数	3	2	6	8	4
専門用語数	186	238	305	539	270
分類項目	c35	c36	c234	c236	c1234
科目数	1	1	1	2	1
専門用語数	40	61	86	93	121

科目。

また、c234, c236, c1234 は、それぞれ、科目区分「2, 3, 4」、「2, 3, 6」および「1, 2, 3, 4」の内容を含む科目である。

3.2.2 ステップ 2: 専門用語集合の作成

科目 DB に登録されている科目のシラバスから科目の分類を支援する情報を得ることを考える。まず、大学によって書式の異なるシラバスから所望の属性を取り出すために、各シラバスを、あらかじめ井田ら¹³⁾が提案する XML 形式のフォーマット に変換する。そしてその後、中川らの「接続頻度に基づく手法」¹⁴⁾を利用し、変換された各シラバスデータから専門用語を抽出する。先のフォーマットには、「科目名」等全部で 16 のフィールドが存在するが、ここでは「科目名」、「授業概要」、「授業計画」、「履修により達成される目標」フィールドのみを抽出の対象とした。

抽出後、算用数字を含むもの、および、重複して抽出されていた用語を除外した。前者は「(1) 電気通信」のように専門用語に数字 (シラバス中の列挙番号等) が混ざったものを除外するために行ったが、「3D モデル」等数字を含む専門用語も同時に除外されてしまう問題も残されている。この時点での各分類項目ごとの専門用語数を表 2 に示す。また、以下では、分類項目 c_i ($i=0, 1, 2, 3, 4, 5, 6, 7, 10, 12, 13, 14, 16, 23, 24, 35, 36, 234, 236, 1234$) に含まれる専門用語集合を C_i と書く。

次に、この結果を利用し、表 1 に即した分類を支援する情報の提示を考える。そのために「科目区分 i ($i=0, 1, 2, 3, 4, 6, 7$) であると判定された科目のシラバスのみ」に含まれている専門用語集合 C'_i を作

成する。これにより、その科目区分のみに内包されている特徴的な語句の抽出を狙っている。ここで、科目区分 5 は、情報工学全般にわたる語句を含んでいる可能性があるため、あえて C'_5 は作成しないものとする。本稿では、対象とする「情報工学」区分の各科目区分の特性を考慮し、以下のように各科目区分を 3 つのグループに分類し、各 C'_i を作成した。

- グループ A (科目区分 1, 2, 3, 4)

$$C'_1 = C_1 \setminus \cup_{i \in \{2,3,4,6,7,23,24,35,36,234,236\}} C_i$$

$$C'_2 = C_2 \setminus \cup_{i \in \{1,3,4,6,7,13,14,16,35,36\}} C_i$$

$$C'_3 = C_3 \setminus \cup_{i \in \{1,2,4,6,7,12,14,16,24\}} C_i$$

$$C'_4 = C_4 \setminus \cup_{i \in \{1,2,3,6,7,12,13,16,23,35,36,236\}} C_i$$

- グループ B (科目区分 0)

$$C'_0 = C_0 \setminus \cup_{i \in \{1,2,3,4,6,7,12,13,14,16,23,24,35,36,234,236,1234\}} C_i$$

- グループ C (科目区分 6, 7)

$$C'_6 = C_6 \setminus \cup_{i \in \{1,2,3,4,0,7,10,12,13,14,23,24,35,36,234,1234,5\}} C_i$$

$$C'_7 = C_7 \setminus \cup_{i \in \{1,2,3,4,0,6,10,12,13,14,16,23,24,35,36,234,236,1234,5\}} C_i$$

C'_i の作成の基本は、科目区分 i ($i=0, 1, 2, 3, 4, 6, 7$) に分類されている専門用語集合 C_i から、その科目区分 i に関係しない分類項目に分類されている専門用語を取り除くことにある。ただし、グループ A では、科目区分 5 または 0 に関係する専門用語集合との差集合はとらない。これは、実際はそれらの科目区分に分類されていたとしても、科目区分 1~4 のいずれかの要素を含むことが考えられるためである。同様に、グループ B では、科目区分 5 に関係する専門用語集合との差集合はとらない。一方、グループ C では、科目区分 5 および 0 に関係する専門用語集合との差集合をとる。これは、科目区分 5 や 0 は少なくとも専門科目ではあるので、専門科目ではない科目区分 6 や 7 からは除外可能なためである。

各科目区分 i ($i=0, 1, 2, 3, 4, 6, 7$) に対応する専門用語集合の語数、すなわち $|C'_i|$ は、それぞれ 682, 954, 1318, 1127, 754, 4154, 430 である。以下では、これらの語句をまとめたものを myDB と呼ぶ。すなわち、myDB = $\{C'_i; i=0, 1, 2, 3, 4, 6, 7\}$ である。

3.2.3 科目分類のアルゴリズム

myDB と分類すべき科目のシラバスに出現する語とのマッチングをとることで、その科目の分類を支援

次の URL より入手可能。

<http://svrrd2.niad.ac.jp/syllabus/10/syllabus10.xsd>

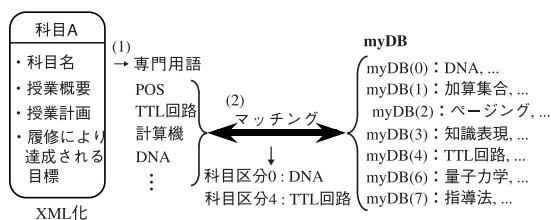


図4 科目分類支援システムの動作例

Fig. 4 A sample result of CCS.

するシステムを考える。本システムに、あるシラバスが入力されたときの動作は次のとおりである。

- (1) 分類すべき科目のシラバスから中川らの手法¹⁴⁾により専門用語を抽出する。この際、各シラバスは、あらかじめ井田らが提案するXML形式のフォーマット¹³⁾に変換した後、「科目名」、「授業概要」、「授業計画」、「履修により達成される目標」フィールドのみを抽出の対象とする。
- (2) (1)で抽出された各専門用語とmyDBとのマッチングをとり、マッチした用語とその用語に対応する科目区分番号を出力する。

以下では、本システムを科目分類支援システム(CCS)(Course Classification Support system)と呼ぶ。

図4にCCSの動作例を示す。CCSに分類すべき科目として科目Aのシラバス(XML変換済み)が入力されると、(1)そのシラバス中の「科目名」、「授業概要」、「授業計画」、「履修により達成される目標」フィールドから中川らの手法¹⁴⁾により専門用語が抽出される。図4の例では、「POS」、「TTL回路」、「計算機」、「DNA」等が抽出されている。(2)そしてこれら抽出された専門用語とmyDBとのマッチングをとる。その結果、科目区分0に相当する専門用語として「DNA」、科目区分4に相当する専門用語として「TTL回路」をシステムは出力する。

このようにCCSの出力は、「マッチした用語とその用語に対応する科目区分番号」である。有効な分類支援を実現するためには、適切な科目区分とともに適切な用語を出力することが望ましい。しかし現時点では、出力された用語に対する検討は十分ではない。したがって本稿では、提示された科目区分の適切さのみに焦点をあてCCSを評価する。出力された用語に対する検討は、今後の課題の1つである。

3.3 科目分類支援システムの利用方法

現状では、専門委員には、原則として、申請者から申告されたすべての科目の判定を依頼している。しかし学位授与の審査という観点からは、専攻基準を満た

した時点でそれ以上の判定が行われる必要はない。もちろん、すべての科目を判定しなければ専攻基準を満たさないケースも考えられるが、そうでない場合には、各科目区分ごとに、その科目区分に分類されることが有望であると考えられる科目から優先的に提示してゆけば、判定作業の軽減につながるものとする。

具体的には、科目区分 i ($i=1, 2, 3, 4$)に分類されるべき科目を探しているときには、まず最初に、申請者、CCSともに科目区分 i に分類されるとしている科目を優先的に提示し、それらの科目の判定を専門委員に依頼する。そしてその結果、その区分に関する基準を満たした場合は、その申請者のその科目区分の判定を終える。それに対し、基準を満たさない場合には、さらに、(申請者は科目区分 i には分類していないが)CCSが科目区分 i に分類されると示唆する科目を次に有望な候補として専門委員に提示する。これにより、すべての科目の判定が必要とされる場合以外は、判定作業の軽減が期待できる。

3.4 科目分類支援システムの特徴

更新可能性: CCSで用いられているmyDBは、科目DBが更新されるごとにつねに更新可能である。NIAD-UEでは、年2回学士の学位授与と申請を受け付けているが、各期の判定を終えるごとに科目DBを更新し、その結果をmyDBに反映させられることで、次の期には、より良い分類支援が期待できる。

専攻区分への依存性の局在化: 本研究では、「情報工学」区分を対象にCCSを構築したが、他の専攻区分であっても科目DBを差し替えるのみで、同様のシステムが構築可能である。

大学での利活用が可能: myDBを参照することで、その科目区分での特徴的なトピックスを知ることができる。これをうまく活用すれば、教師が授業を組み立てる際の材料として利用可能である。

4. 科目分類支援システムの評価

4.1 R大学情報系学科

本章では、CCSの動作評価を目的に、充実したシラバスの内容を持つR大学情報系学科を取り上げ、科目の分類支援の具体例を示す。R大学情報系学科には49の科目があり、その中で科目名から「演習・実験・実習科目(科目区分5)」であることが明らかな科目(7科目)や「卒業研究(科目区分7)」、ならびにシラバスにほとんど情報がない科目(2科目)を除いた39科目をCCSへ入力する。

CCSの有効性を議論するための比較対象として、3.2.3項で述べたCCSのアルゴリズム中の(2)で用い

ている専門用語集合 myDB を「新しい学士への途」¹¹⁾ に記載されている例示科目のみをデータベース化したものに置き換えたものを考える．これを以下では，CCS' と呼ぶ．「新しい学士への途」¹¹⁾ に記載されている例示科目は科目区分 1~4 のみなので，CCS' の出力も科目区分 1~4 に対応するもののみとなる．また，以下では，CCS' の各科目区分 i ($i=1, 2, 3, 4$) に含まれる専門用語集合を $icDB(i)$ と書き，それらをまとめたものを $icDB$ (Illustration Courses DB) と呼ぶ．すなわち， $icDB = \{icDB(i); i=1, 2, 3, 4\}$ である．

表 3 に，R 大学の 39 科目を CCS および CCS' に入力したときの myDB(i) および icDB(i) ($i=1, 2, 3, 4$) とマッチしていた用語の数を示す (科目名は適宜省略してある)．表 3 の「筆者分類」には，3.2.1 項で筆者が直接シラバスを読んで分類した結果を示す．この「筆者分類」内には科目区分 1, 2, 3, 4 に関連する科目としては，表中でゴチック体で示した「情数 a 演」から「基礎情数」まで全部で 28 科目ある．これら 28 科目の CCS の結果を見ると，該当する科目区分 i ($i=1, 2, 3, 4$) に関しては，すべてそれに相当する myDB(i) といくつかの専門用語がマッチしている．それに対し，CCS' では，アンダーラインを付した「応用情科」「論理数 1」等全部で 10 科目で，ある科目区分 i に相当する icDB(i) とマッチする専門用語が存在していない．

4.2 科目の分類支援結果

表 3 の結果をふまえ，科目の分類支援について検討する．ここでは先の R 大学の 39 科目すべてが初出科目である場合を想定する．その場合，現状では，専門委員には，原則すべての科目の判定を依頼する．それに対し，本研究では，すでに述べたように，各科目区分ごとに，その区分に関する基準を満たした時点で，その科目区分の判定を終える．このような方法で，判定作業の軽減を目指す場合には，本来，科目区分 i に分類されるべき科目を CCS が見落とすことがあってはならない．いい換えれば，本来，科目区分 i に分類されるべき科目に対しては，myDB(i) とマッチする専門用語数が 0 であってはならない．これが満たされたうえで，CCS が科目区分 i であると示唆する科目数が少なければ少ないほど，より効果的な分類支援が実現される．

この観点から，表 3 を集計すると，表 4 のようになる．該当数は，先の 39 科目に対する 3.2.1 項ステップ 1 の結果である．これは，39 科目中，各科目区分に分類されると筆者が判断した科目数を意味する．こ

表 3 R 大学情報系学科の 39 科目を CCS および CCS' に入力した結果

Table 3 The results of CCS and CCS' in R university.

科目名	筆者分類	CCS の結果				CCS' の結果			
		1	2	3	4	1	2	3	4
情数 a 演	c16	3				1			
Prog	c2		14	2			3		
Prog2	c2		8	1			2		
情報構造	c1	4	5	1		2			
コンプロ	c2	4	6			1	2		
データ解析	c1	3	1			1			
人工知能	c3	1		8			2	3	
計算機	c2		13			1	1		
DB	c2	1	14	1			1		
計算理 1	c1	7	1			2			
計算理 2	c1	7	1			2			
確率 1 演	c16	3		1		1			
情数 b	c16	3				2			
シスプロ	c2		7	1			3		
計算機概	c124	3	15		2	2	1	2	1
オトマタ	c1	7	1			2	1		
情数 1 演	c1	1				1			
計算機演	c2	1	11	1			3		
応用情科	c3	1	1	2					
論理数 1	c1	1	1	3	1		1		
論理数 2	c12	5	5						
生体情報	c34	1		2	1			2	
確率 2 演	c16	1		1	1				
統計 3	c16	2		1					
計画 1 演	c134	21		1	1	2	1		
計画 2	c14	5			1	1	1	1	
情理論演	c1	6							
基礎情数	c16	1							
統計演	c6	1							
多変量	c6	2		1					
情数 3 演	c6			1		1			
幾何	c7								
光通信理	c0							1	
複雑理	c0					1			
数理経済	c0							1	
遺伝情報	c0	1							
解析演	c7	1							
代数演	c7	1			1				
物理	c7								

ゴチック体は，科目区分 1, 2, 3, 4 に関連する科目を表し，アンダーラインは，(CCS' において) 専門用語集合 icDB とのマッチが十分ではない科目を表す．

れは，表 3 の「筆者分類」の部分で科目区分 i ($i=1, 2, 3, 4$) に関係する科目数を各々数え上げたものと一致する．たとえば，39 科目中，科目区分 1 に関係すると判断されていた科目は「情数 a 演」等全部で 18 科目である．

それに対し，39 科目すべてを CCS および CCS' に入力したときの結果が示唆数と見落数である．示唆数は，各手法がその科目区分に分類されると示唆した科目数である．これは単純に，myDB(i) または icDB(i) とマッチした科目数を数え上げたものである．たとえ

表 4 科目移動支援の結果
Table 4 The result of the course movement.

	該当数	示唆数		見落数	
		CCS	CCS'	CCS	CCS'
科目区分 1	18	29	16	0	6
科目区分 2	9	16	13	0	0
科目区分 3	4	16	6	0	2
科目区分 4	4	7	0	0	4

ば、CCS が科目区分 1 に分類される可能性があるとして示唆した科目は 39 科目中 29 科目、CCS' の場合は 16 科目であった。参考までに、myDB を「科目区分 i であると判定されたシラバスに含まれている専門用語」集合により構築した場合、すべての科目区分において示唆する科目数が 39 になった。

見落数は、その科目区分に分類されるべき科目を見落とした数である。これは、該当数に数え上げられた科目のうち示唆数に含まれなかったもの数である。たとえば、科目区分 1 に対しては、CCS が科目区分 1 である科目を見落とすことはないが、CCS' では、科目区分 1 に分類されるべき 18 科目中 6 科目を見落としていた。

4.3 結果の解釈

表 4 より、科目区分 2 に関しては、CCS、CCS' とともに、その科目区分に分類されるべき科目を見落とすことはなかった。しかし、それ以外の科目区分 1, 3, 4 の場合には、CCS' は、その各科目区分に分類されるべき科目のうち数科目を見落としている。特に、科目区分 4 に関しては、示唆数が 0 となっており、その区分に分類されるべき 4 科目すべてを見落としてしまっている。科目区分 4 は広範囲な科目を対象としているにもかかわらず、「新しい学士への途」¹¹⁾ の例示科目数は他の区分に比べ少ない。このことが、CCS' の性能悪化につながったものと考えられる。

それに対し、CCS では、各科目区分に分類されるべき科目を見落とす例は皆無であった。これは、本研究の目的である分類支援の信頼性を主張する意味で非常に重要である。しかしながら、示唆数に関しては、CCS は、つねに CCS' よりも多くなっている。これはある意味、見落数をゼロにするうえでやむをえないことではあるが、示唆数の削減は今後の重要な課題である。

ちなみに CCS の利用においては、3.3 節で述べたように、専攻基準の満足に必要な科目区分に分類されることが有望な科目を優先的に専門委員へ提示することが基本となる。この具体的な手順は、各専攻区分ごとに、その専攻基準に応じてカスタマイズする必要がある。

表 5 myDB(1) とマッチした科目の各分類項目ごとの科目数
Table 5 The number of courses that match with myDB(1) for each item.

分類項目	c0	c1/c16	c2	c3	c4	c12
科目数	1	14	3	2	0	1
分類項目	c14	c124	c134	c34	c6	c7
科目数	1	1	1	1	2	2

る。しかしここでは、CCS の有効性を手短かに示すために、「上の R 大学の 39 科目を履修した申請者が情報工学区分の専攻基準（図 2 参照）を満たしているかを判定する」という事例に対し、「科目区分 1 から順に専門委員に提示し判定を依頼する」という CCS の利用方法を採用した場合を考える。この場合、表 3 で myDB(1) とマッチしている科目の数を各分類項目ごとに集計すると表 5 のようになる。今、科目区分 5 は対象外としているので、専門科目としては 40 - 6 の 34 単位必要である。R 大学の各科目は 2 単位なので、表 5 に示した 29 科目の判定のみで十分この基準を満たすことが可能となる。これは、これまでの 39 科目の判定に比べ、およそ $\frac{3}{4}$ の時間で判定作業を終えられることを意味する（各専攻区分ごとに異なる専攻基準表の特徴を加味すれば、さらなる時間短縮も期待できる）。

5. おわりに

本稿では、NIAD-UE における学位授与事業のための科目分類支援システム（CCS）試作の現状について述べた。R 大学情報系学科の 39 科目に適用することで有効性を確認した。

5.1 実用化に向けて

現在、我々は、本稿の結果をふまえ、CCS の実用化に取り組んでいる。本格的な運用に至るには、以下に述べる 3 つの課題を解決する必要があると考えている。

第 1 の課題は、申告された大量の科目のシラバスをどのように処理するかというものである。現在、年間、約 2,000 件にものぼる申請があり、その結果、全体では、数万ページに達するシラバスを、毎年、取り扱う必要がある。これらすべてのシラバスからの用語抽出は非現実的な作業をとまなう。そこで、我々は、シラバスを直接 CCS に入力する方法を、現在、検討中である。これが実現されれば、本稿で対象とした「情報工学」区分に関しては、実用化に大きく前進すると考える。

次に、「情報工学」区分以外の他の専攻区分での動作確認も本格運用に際しては重要である。この場合、

まず第 1 に、その専攻区分に合致した科目 DB を作成しなければならない。科目 DB の作成はある程度の専門性を要するので、今後は、本稿で行った方法以外にも、既存の科目データベースの活用を積極的に考えていきたい。具体的には、科目データベースには、本来、不十分さが含まれているという前提に立ち、その不十分さを吸収するようなシステム作りを検討していきたいと考えている。また、4 章の最後で述べた、各専攻区分ごとの特徴を加味したシステム作りも実運用に際しては重要になると考える。

最後に、「科目区分」や「例示科目」に大幅な変更が生じた場合への対応も重要である。この場合、過去の判定例である科目 DB が無意味になる可能性がある。したがって、そのような場合には、システムの再構築が必要となる。NIAD-UE では、毎年度、「新しい学士への途」を更新しているので、このような変化は例外的ではない。本研究は平成 14 年度版の「新しい学士への途」¹¹⁾ を参考にシステムを構築したが、専攻区分によっては、平成 15 年度版の「新しい学士への途」では、例示科目に若干の変更が生じている。このような変更にどの程度まで追従可能であるかを調べるとともに、大幅な変更が生じた場合には、変更前後で myDB の活用方法に差異をつける等の工夫が必要であると考えられる。

5.2 さらに改善に向けて

上記の実用化に向けた課題とともに、CCS のさらなる改善も重要である。この件に関しては、現在、CCS が提示する情報をより豊かにすることを考えている。具体的には、myDB に統計処理を施すことで提示される情報に信頼度を付加することを、現在、検討している。また、それと同時に、CCS が提示する用語の積極的な活用も検討中である。

さらに、専門用語抽出手法に関する考察もより深く行う必要がある。専門用語を抽出する手法には、本稿で採用した方法以外にも、テキスト中に出現した回数を基本にする手法¹⁵⁾ や、候補語の出現箇所の周辺に共起する語の性質に基づく手法^{15),16)} 等、様々なものが存在する。今後、提示された用語の活用を考える際には、抽出手法の差異による影響が大きくなると思われるので、いくつかの抽出手法を比較検討することが重要であると考えられる。

4 章で述べた CCS の活用例に関しては、該当外の科目区分に分類されると示唆するケースを少しでも減らす努力を行っていきたい。見落数と示唆数の間のトレードオフを考慮しながら、より有効な支援情報の提示を考えていきたい。また、科目区分 4 のような、広

範囲な科目を対象としているにもかかわらず、myDB 中の専門用語数 ($|C'_4|$) が比較的少ない区分に関しては、今後は、専門用語の類義語処理等何らかの工夫が必要になることが予想される。

ところで、NIAD-UE における科目審査では、専門委員の判定が主であり、その責任は非常に重く、かつ、作業負担も大きい。それに対し、情報システムには、専門委員の作業負担を軽減する従の役割が期待されている。本稿では、そのような「人による分類の情報システムによる支援」を重視する立場から、判定作業そのものの自動化には触れず、科目を提示する順番を工夫することで、科目判定に係る作業時間の短縮を目指した。このような形での情報システムと人間・組織・社会との相互関連は、審査等、判定間違いの許されない重責を担う人間の負担軽減を目的として、今後ますます重要になるとと思われる。

謝辞 本研究を遂行するにあたりご協力いただいた大学評価・学位授与機構学位審査研究部の教員をはじめ「大学評価情報の構造解析と評価プロセスへの応用の研究会」参加者の皆様に謝意を表します。

参 考 文 献

- 1) 井下 理：シラバスの意味と機能，大学力を創る：FD ハンドブック，大学セミナーハウス（編），pp.62-72，東信堂（1999）。
- 2) 野澤孝之，井田正明，芳鐘冬樹，宮崎和光，喜多一：シラバスデータのクラスタリングに基づく教育コース分析システムの開発，情報処理学会第 66 回全国大会，pp.377-378（2004）。
- 3) 井田正明，芳鐘冬樹，野澤孝之，宮崎和光，喜多一：シラバスデータベースシステムの検討と応用，第 48 回システム制御情報学会講演会講演論文集，pp.563-564（2004）。
- 4) 山田信太郎，伊東栄典，廣川佐千男：自動収集した Web シラバスデータの分析と考察，情報科学技術フォーラム 2002（FIT2002），一般講演論文集第 4 分冊（N-32），pp.301-302（2002）。
- 5) 山田信太郎，松永吉広，伊東栄典，廣川佐千男：Web シラバス情報収集エージェントの試作，電子情報通信学会和文論文誌 D-I，Vol.J86，No.8，pp.566-574（2003）。
- 6) 伊東栄典，松永吉広，山田信太郎，廣川佐千男：Web シラバスからの DB 構成，2003 年度人工知能学会全国大会（JSAI2003）1D4-08（Jun. 2003）。
- 7) 森 利枝：英国オープン・ユニバーシティにおける単位認定と評定サービス，学位研究，No.17，pp.183-198（2003）。
- 8) シェーマス・パーセル（著），吉川裕美子（訳）：アイルランド高等教育における品質保証—高等教育訓練資格カウンシル（HETAC），学位研究，

No.15, pp.124-140 (2001).

- 9) 森 利枝：リージェント大学における評価のシステム—学習とクレジットの評価を中心に，学位研究，No.10, pp.107-129 (1999).
- 10) 縮 昭：ニュージャージー州学外学位機関・トーマスエジソン州立大学の仕組と意義，学位研究，No.10, pp.73-89 (1999).
- 11) 平成 14 年度版 新しい学士への途．
http://www.niad.ac.jp/sub_gakui/siryu/new/new_gakushiH14.pdf
- 12) 高須淳宏，相原健郎：テキスト分類における訓練データと性能の実験的考察，*NII Journal*，No.6, pp.1-8 (2003).
- 13) 井田正明，宮崎和光，芳鐘冬樹，喜多 一：シラバス XML データベースシステム構築に関する考察，情報処理学会第 65 回全国大会，pp.247-248 (2003).
- 14) 中川裕志，森 辰則，湯本紘彰：出現頻度と接続頻度に基づく専門用語抽出，自然言語処理，Vol.10, No.1, pp.27-45 (2003).
- 15) Mima, H. and Ananiadou, S.: An application and evaluation of the C/NC value approach for the automatic term recognition of multi-word units in Japanese, *Terminology*, Vol.6, No.2, pp.175-194 (2000).
- 16) Hisamitsu, T., Niwa, Y., Nishioka, S., Sakurai, H., Imaichi, O., Iwayama, M. and Takano, A.: Extracting terms by a combination of term frequency and a measure of term representativeness, *Terminology*, Vol.6, No.2, pp.211-232 (2000).

(平成 16 年 6 月 17 日受付)

(平成 17 年 1 月 7 日採録)



宮崎 和光

昭和 42 年生．平成 8 年東京工業大学大学院総合理工学研究科知能科学専攻博士後期課程修了．同年同大学大学院総合理工学研究科助手．平成 10 年から 11 年にかけて同大学大学院総合理工学研究科リサーチアソシエイト．平成 11 年学位授与機構審査研究部助教授を経て平成 12 年より大学評価・学位授与機構学位審査研究部助教授．機械学習，特に強化学習に関する研究に従事．博士（工学）．人工知能学会，計測自動制御学会，日本機械学会，日本高等教育学会の会員．



井田 正明（正会員）

昭和 39 年生．平成 2 年京都大学大学院工学研究科精密工学専攻修士課程修了．同年京都大学工学部助手．平成 14 年より大学評価・学位授与機構評価研究部助教授．システム工学，評価と意思決定，数理計画法，データベースに関する研究に従事．博士（工学）．計測自動制御学会等の会員．



芳鐘 冬樹（正会員）

昭和 44 年生．平成 14 年東京大学大学院教育学研究科総合教育科学専攻博士課程単位取得退学．同年大学評価・学位授与機構評価研究部助手．研究コミュニティの構造分析，専門用語の自動抽出手法に関する研究に従事．修士（教育学）．日本図書館情報学会の会員．



野澤 孝之

昭和 50 年生．平成 14 年東京工業大学大学院総合理工学研究科知能システム科学専攻博士課程修了．平成 14 年から 15 年にかけて京都産業大学客員研究員．平成 15 年より大学評価・学位授与機構評価研究部助手．非線型力学，複雑系の理論とシミュレーション，ネットワーク構造分析に関する研究に従事．博士（理学）．電子情報通信学会の会員．



喜多 一

昭和 34 年生．昭和 57 年京都大学工学部卒業．昭和 62 年同大学大学院工学研究科博士後期課程研究指導認定退学．昭和 62 年京都大学工学部助手．平成 9 年東京工業大学大学院総合理工学研究科助教授．平成 12 年大学評価・学位授与機構教授を経て平成 15 年京都大学学術情報メディアセンター教授．社会システム，ニューラルネットワーク，進化的計算，エージェントシミュレーションの研究に従事．工学博士．電気学会，電子情報通信学会，計測自動制御学会，システム制御情報学会等の会員．