

学位研究 第18号 平成16年3月（論文）

[大学評価・学位授与機構 研究紀要]

電子化シラバスに基づく学位授与のための科目分類支援システムの検討

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

宮崎 和光, 井田 正明, 芳鐘 冬樹, 野澤 孝之, 喜多 一

MIYAZAKI Kazuteru, IDA Masaaki, YOSHIKANE Fuyuki, NOZAWA Takayuki and KITA Hajime

Research in Academic Degrees, No. 18 (March, 2004) [the article]

The Journal on Academic Degrees of National Institution for Academic Degrees and University Evaluation

1. はじめに	135
2. 学位授与事業における科目分類	135
2.1 単位修得に着目した申請方法の概略	135
2.2 現状の科目審査プロセスとその改善	136
3. 科目分類支援システム	137
3.1 システムの概要	137
3.2 科目分類支援システムの開発	138
3.2.1 ステップ1：科目DBの作成	138
3.2.2 ステップ2：専門用語集合の作成	139
3.2.3 科目分類のアルゴリズム	141
3.3 科目分類支援システムの特徴	142
4. R大学情報系学科の分類支援結果	142
5. 科目分類支援システムの活用例	146
5.1 科目の移動支援	146
5.2 科目移動支援の結果	147
5.3 結果の解釈	147
6. おわりに	147
ABSTRACT	150

電子化シラバスに基づく学位授与のための科目分類支援システムの検討

宮崎 和光*, 井田 正明**, 芳鐘 冬樹***, 野澤 孝之***, 喜多 一****

1. はじめに

大学評価・学位授与機構（以下、機構と略す）では、短期大学・高等専門学校卒業生及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機構が認定した専攻科（認定専攻科）における学修等により、機構の定める所定の単位を修得しなければならない。

申請者は修得した単位を個人の判断で分類し機構に申告する。機構では、申請者が所定の単位を適切に修得しているかどうかを、大学等の教員で構成される専門委員会において、申告された科目のシラバスを読むことで検討する。

しかしながら、近年の申請者数の増大から、この検討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分類支援が望まれている。そこで、本論文では、電子化されたシラバスを利用した、この分類を支援する科目分類支援システム試作の現状について述べる。

以下、第2章では、問題の所在を明確にするために、機構の学位授与事業における科目分類の現状について述べる。第3章では、科目分類支援システムの概要を述べるとともに具体的なシステムの詳細を記す。第4章では、科目分類支援システムを利用した科目の分類支援例を示す。さらに第5章では、科目分類支援システムの活用例としての「科目の移動支援」について触れる。第6章は結論であり、本研究の成果を総括し、今後の課題をとりまとめる。

2. 学位授与事業における科目分類

2.1 単位修得に着目した申請方法の概略

本論文が対象とする学位授与事業の制度・申請方法などは、機構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。「新しい学士への途」には、平成14年度時点で、「文学」「工学」など全部で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」など全部で53の専攻区分が記載されている。参考までに、付録に平成14年度版の「新しい学士への途」から「情報工学」区分のページを抜き出したものを示す。

まず、第一に、申請者は、短期大学や高等専門学校を卒業するなどの基礎資格を有する必要がある。その下で、申請者は、いずれかの専攻分野とその専攻分野の中で該当する専攻区分のひとつを選択し、大学の科目等履修生制度を利用した学修や、認定専攻科における学修等で、

* 大学評価・学位授与機構 学位審査研究部 助教授

** 大学評価・学位授与機構 評価研究部 助教授

*** 大学評価・学位授与機構 評価研究部 助手

**** 京都大学 学術情報メディアセンター 教授

所定の単位を修得しなければならない。

機構では、基礎資格の修業年限などによって、3つの申請区分を用意している。申請者は、申請区分ごとに定められた**基本基準**と、専攻区分ごとに定められた**専攻基準**を満たすように単位を修得しなければならない。そして申請者は、これら2つの基準に基づき修得した単位を、その内容により、各専攻区分に係る「**専門科目**」「**関連科目**」及びそれ以外の「**専攻外科目**」に分類・整理し、申告する。

各専攻区分ごとの具体的な「**専門科目**」及び「**関連科目**」については、専攻基準の基準表に「**科目区分**」および、複数の「**科目区分**」をまとめた「**群**」という形で分類・整理され記載されている。さらにこの基準表には、「**専門科目**」,「**関連科目**」,「**群**」,「**科目区分**」ごとに申請上必要な単位数が明記されているとともに、「**専門科目**」については、それぞれの「**科目区分**」ごとに具体的な科目名の例が「**専門科目の例**」として示されている。

申請者は、取得しようとする学位の専攻区分を決定し、各専攻区分ごとに示されているこの「**専門科目の例**」を参考に、自らの判断で修得した科目を分類する。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る。

2.2 現状の科目審査プロセスとその改善

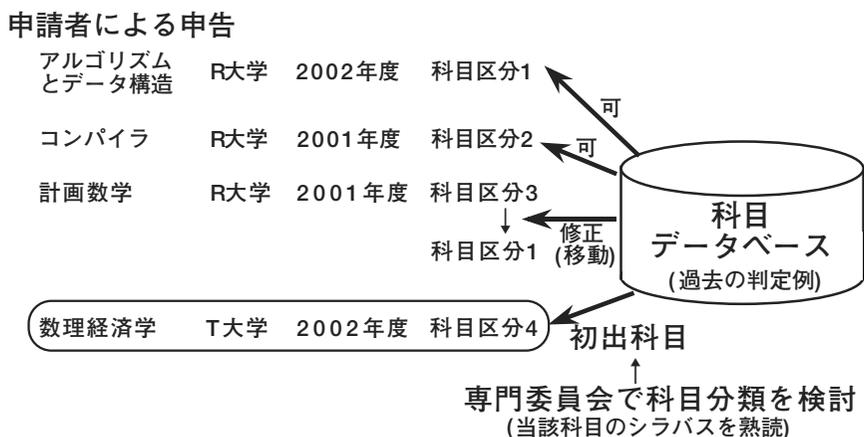


図1 現状の科目審査プロセスの例

それに対し機構では、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 (DB) 化し、科目審査に活用している。このDBは**科目DB**と呼ばれており、現状では、シラバスと判定結果が「1対1」に対応するように作られている。図1に、科目DBを用いた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DBとのマッチングをとり、一致するものは「可」とする。図1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DBの分類に修正 (移動) する。例えば、図1の例では、「計画数学」がこれに相当する。

ところで、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、その科目に対応するシラバスを読むことで検討している。

例えば、図1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、本来、省略できるものではなく、基本的には対応するシラバスをすべて読む必要がある。しかし、個々の申請者の立場からは、専攻基準を満たすために必要な科目以外の科目を判定される必要はない。勿論、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。そこで本研究では、情報技術を利用した科目分類のためのシステムを提案し、このような判定作業の支援を考える。この件については、4章で詳しく述べる。

科目審査の結果、「基本基準」および「専攻基準」を満たした場合、単位に関しては合格とされる¹⁾。それに対し、単位が不足した場合には、専門委員会において慎重な審議を重ね合否が決定される。この際、仮にある同一の科目が複数の科目区分に分類される可能性が考えられたとしても、過去の判定例である科目DBの変更は許していない。しかし、シラバスの内容によっては、シラバスと科目区分が「1対多」に対応するDBの構築を許してもよいはずである。但し、この場合、つねに申告されたすべての科目のシラバスの再分類が要求される。5章では、この作業の支援に提案システムを活用することを考える。具体的には、移動可能な科目を優先的に提示することで作業の効率化を目指す。

次章では、これらのもととなるシステムである**科目分類支援システム**を提案する。

3. 科目分類支援システム

3.1 システムの概要

本論文では、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。ここで科目分類とは、例えば、「情報工学」区分の場合、申請者が大学等で修得し機構に申告した各科目を表1の1～7のいずれかの科目区分に分類することをいう。「新しい学士への途」は毎年度改定されているが、本論文では、「[新し

表1 「情報工学」区分の科目区分番号（[[新しい学士への途]平成14年度版]より抜粋したものに番号を付した）

科目区分1	情報工学基礎理論に関する科目（4単位以上）
科目区分2	計算機システムに関する科目（6単位以上）
科目区分3	情報処理に関する科目（6単位以上）
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（4単位以上）
科目区分7	専攻外科目

注、「情報工学」区分の専門科目はA群（科目区分1～4）とB群（科目区分5）に分かれており、A群から30単位以上、A群とB群の合計が40単位以上となるように修得しなければならない。

¹⁾ さらに学修成果と呼ばれるレポートおよび、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。

い学士への途」平成14年度版]をもとに検討を進めた。また以下では、「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DBにある。しかし現実には、科目DB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DBは分類精度の点で十分とはいえない。そこで本研究では、まず初めに、科目DBの作成を考える。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。

次節の大まかな流れは次の通りである。まず3.2.1節で、科目DBの作成を行う。続く3.2.2節では、3.2.1節で作成した科目DBから表1の各科目区分ごとに特徴的な語句を抽出する。そして3.2.3節では、3.2.2節で求めた各科目区分ごとの特徴的な語句と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開発

3.2.1 ステップ1：科目DBの作成

有効な分類支援を実現するためには、分類精度の高い科目DBが必要となる。機構では、平成12年度から、科目DBの整備を続けているが、そもそも現状の科目DBは、シラバスと科目区分とが「1対多」に対応していないなど十分に整備されているとは言えない。そこで本研究では、まず初めに、この科目DBの作成を行った¹⁾。

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宮崎が熟読し、「[新しい学士への途]平成14年度版」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²⁾。この際、ひとつの科目が、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1010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。

表2 科目DBに登録されている科目の分類結果

分類項目	c0	c1	c2	c3	c4	c5	c6	c7	c10	c12
科目数	68	108	104	85	54	86	351	48	7	21
分類項目	c13	c14	c16	c23	c24	c35	c36	c234	c236	c1234
科目数	3	2	6	8	4	1	1	1	2	1

¹⁾ 現時点でこのDBは、現実の審査業務に活用されている科目DBとは独立したあくまでも研究用のDBである。

²⁾ 実際の運用に際しては、この作業は専門委員会が実施すべきものである。

- c0；科目区分1～5には分類困難な専門科目。
- ci；科目区分*i* (*i*=1,2,...,7) に分類される科目。
- ci0；科目区分*i* (*i*=1,2,...,7) に分類されると思われる科目。
- cij；科目区分*i* (*i*=1,2,...,7) と科目区分*j* (*j*=1,2,...,7. *i* ≠ *j*) の両方の内容を含む科目。

また、c234,c236,c1234は、それぞれ、科目区分2,3,4および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、予め[井田03]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、[湯本01]の「接続頻度に基づく手法」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」など全部で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専門用語を除外した。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」など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は表3の通りである。以下では、分類項目*ci* (*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*Ci*と書く。

表3 表2の各分類項目ごとの専門用語数

分類項目	c0	c1	c2	c3	c4	c5	c6	c7	c10	c12
専門用語数	2496	2419	2746	2637	2055	2395	7075	1249	294	929
分類項目	c13	c14	c16	c23	c24	c35	c36	c234	c236	c1234
専門用語数	186	238	305	539	270	40	61	86	93	121

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」集合というような専門用語集合を作成する。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ので、敢えて専門用語集合は作成しないものとする。

具体的な各専門用語集合の作成方法は次の通りである。まず、科目区分*i* (*i*=0,1,2,3,4,6,7) に

^{††} 次のURLより入手可能。http://svrrd2.niad.ac.jp/syllabus/10/syllabus10.xsd

^{**} その他のフィールドとしては、「授業コード」「開講対象学年」「開講学期」「曜日」「時限」「単位数」「授業形式」「教室」「教官情報」「成績評価方法」「教科書」「参考書」がある。

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。これは、例えば、「科目区分1であると判定されたシラバスのみに含まれている専門用語」集合を作成する際には、 C_{13} や C_{1234} などの科目区分1の内容を含む専門用語集合以外の専門用語集合（例えば、 C_{23} など）との差集合をとることを意味する。本来、分類項目 c_{13} に分類される科目などはシラバスの内容を科目区分1の部分と科目区分3の部分に分け、科目区分3の部分の専門用語のみを除外する処理を行うべきである。しかし、今回は、このようなシラバス内での分割は行っていないため、 C_{13} や C_{1234} などとの差集合をとると、「1番のみに含まれる」という意味での重要な語句が除外されてしまう可能性があり、このような方針とした。

以下では、各科目区分の特性を考慮し、科目区分0,1,2,3,4,6,7を次の3つのグループに分類し具体的な作成方法を説明する。

- グループA；科目区分1,2,3,4
- グループB；科目区分0
- グループC；科目区分6,7

グループAでは、科目区分5または科目区分0に関する専門用語集合との差集合はとらない。これは、実際は科目区分5や科目区分0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および科目区分0に関する専門用語集合との差集合をとる。これは、科目区分5や科目区分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や科目区分7からは除外可能なためである。

これらをまとめると以下のようになる。ここで、「分類項目 c_i ($i=0,1,2,3,4,6,7$) のみに含まれる専門用語集合」を C'_i と書く。

- グループA

$$C'_1 = C_1 \setminus (C_2 \cup C_3 \cup C_4 \cup C_6 \cup C_7 \cup C_{23} \cup C_{24} \cup C_{35} \cup C_{36} \cup C_{234} \cup C_{236})$$

$$C'_2 = C_2 \setminus (C_1 \cup C_3 \cup C_4 \cup C_6 \cup C_7 \cup C_{13} \cup C_{14} \cup C_{16} \cup C_{35} \cup C_{36})$$

$$C'_3 = C_3 \setminus (C_1 \cup C_2 \cup C_4 \cup C_6 \cup C_7 \cup C_{12} \cup C_{14} \cup C_{16} \cup C_{24})$$

$$C'_4 = C_4 \setminus (C_1 \cup C_2 \cup C_3 \cup C_6 \cup C_7 \cup C_{12} \cup C_{13} \cup C_{16} \cup C_{23} \cup C_{35} \cup C_{36} \cup C_{236})$$

- グループB

$$C'_0 = C_0 \setminus (C_1 \cup C_2 \cup C_3 \cup C_4 \cup C_6 \cup C_7 \cup C_{12} \cup C_{13} \cup C_{14} \cup C_{16} \cup C_{23} \cup C_{24} \cup C_{35} \cup C_{36} \cup C_{234} \cup C_{236} \cup C_{1234})$$

- グループC

$$C'_6 = C_6 \setminus (C_1 \cup C_2 \cup C_3 \cup C_4 \cup C_0 \cup C_7 \cup C_{10} \cup C_{12} \cup C_{13} \cup C_{14} \cup C_{23} \cup C_{24} \cup C_{35} \cup C_{234} \cup C_{1234} \cup C_5)$$

$$C'_7 = C_7 \setminus (C_1 \cup C_2 \cup C_3 \cup C_4 \cup C_0 \cup C_6 \cup C_{10} \cup C_{12} \cup C_{13} \cup C_{14} \cup C_{16} \cup C_{23} \cup C_{24} \cup C_{35} \cup C_{36} \cup C_{234} \cup C_{236} \cup C_{1234} \cup C_5)$$

各分類項目 $i = 0, 1, 2, 3, 4, 6, 7$ に対応する専門用語集合の語数、すなわち $|C'_i|$ は、表4の通りである。表4に含まれるすべての語句をまとめたものを **myDB** と呼ぶ。すなわち、 $\text{myDB} = \{C'_i; i=0,1,2,3,4,6,7\}$ である。

表4 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の語数

科目区分 <i>i</i>	0	1	2	3	4	6	7
$ C'_i $	682	954	1318	1127	754	4154	430

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とのマッチングをとることで、その科目を分類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通りである。

- (1) 分類すべき科目のシラバスから [湯本01] の手法により専門用語を抽出する。この際、各シラバスは、予め [井田03] が提案するXML形式のフォーマットに変換した後、「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。

本システムは、その性質上、科目の自動分類も可能であるが、我々の主眼は、機械のみによる自動分類の実現にはない。むしろ人を積極的に取り込んだ支援システムの構築を目指している。すなわち本論文が対象としている課題（科目の分類支援）においては、あくまで専門員会での判定が主であり、本システムの出力結果は、専門委員会の判定を支援するものであることが望ましい。このような理由から以下では、本システムを科目分類支援システム (CCS) (Course Classification Support system) と呼ぶ。

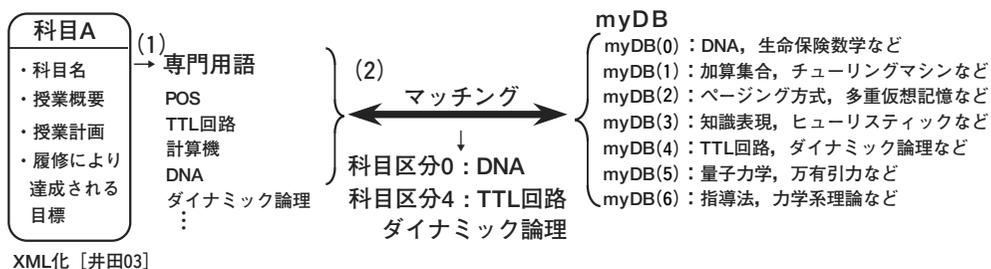


図2 科目分類支援システムの動作例

図2に、CCSの動作例を示す。CCSに科目Aのシラバス (XML変換済み) が入力されると、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 [湯本01] の手法により専門用語が抽出される。図2の例では、「POS」「TTL回路」「計算機」「DNA」「ダイナミック論理」などが抽出されている。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」および「ダイナミック論理」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とは言えない。したがって本論文では、提示された科目区分の適切さのみに焦点をあてたCCSの活用法を検討する。具体的な活用事例については、4章および5章で詳しく述べる。

3.3 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。機構では、年2回学士の学位授与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に反映させることで、次の期には、よりよ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。しかし2.2.1節で行った科目DBの作成は一般には容易ではない。既存の科目DBを利用した科目分類支援システムの構築は今後の重要な課題のひとつである。

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。

「科目区分」や「例示科目」の大幅変更に弱い： 「科目区分」や「例示科目」の大幅な変更が生じた場合は、過去の判定例である科目DBが無意味になる可能性が高い。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。機構では、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。大幅な変更が生じた場合への対応は、今後の重要な課題である。

4. R大学情報系学科の分類支援結果

本章では、CCSの活用例として、R大学情報系学科の49科目の分類支援を考える。ここでは、49科目中、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への入力の対象とし、この39科目すべてが初出科目である場合を考える。CCSの有効性を議論するための比較対象として、3.2.3節で述べたCCSのアルゴリズム中の(2)で用いているmyDBを表5に示す〔「新しい学士への途」平成14年度版〕の例示科目をデータベース化したものに置き換えたもの考える。このシステムを以下では、CCS'と呼ぶ。〔「新しい学士への途」平成14年度版〕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。

以下では、表5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5 「[新しい学士への途]平成14年度版」に記載されている「情報工学」区分の例示科目
※印を付したものは、表記揺らぎを考慮し、独自に追加したものである。

科目区分1	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, オペレーションズリサーチ※, 待ち行列※, 組み合わせ論※, 組み合わせ論※, オペレーションズ・リサーチ※
科目区分2	計算機システム, 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, 情報セキュリティー, 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, OS※, デジタル回路※, 計算機アーキテクチャ※, 情報セキュリティー※
科目区分3	情報処理, 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, 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, CG※, コンピュータグラフィックス※, ヒューマンインタフェース※, ヒューマンインタフェイス※, ヒューマンインターフェイス※
科目区分4	集積回路, 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

図3に、R大学「計算機方式論」のシラバスをCCSおよびCCS'に入力した結果を示す。この場合、まずシラバスから「計算」「計算機」など全部で172語の専門用語が抽出され、それらのうち15語がmyDBとマッチした。具体的には、myDB(0)には「方式論」、myDB(2)には「チャンネルプログラム」「データ形式」など全部で13語、myDB(6)には「保護方式」がマッチしたが、それ以外のmyDB(1)などにはマッチする専門用語はなかった。それに対し、CCS'を用いた場合、「アルゴリズム」がicDB(1)とマッチし、「プログラミング」がicDB(2)とマッチしていた。

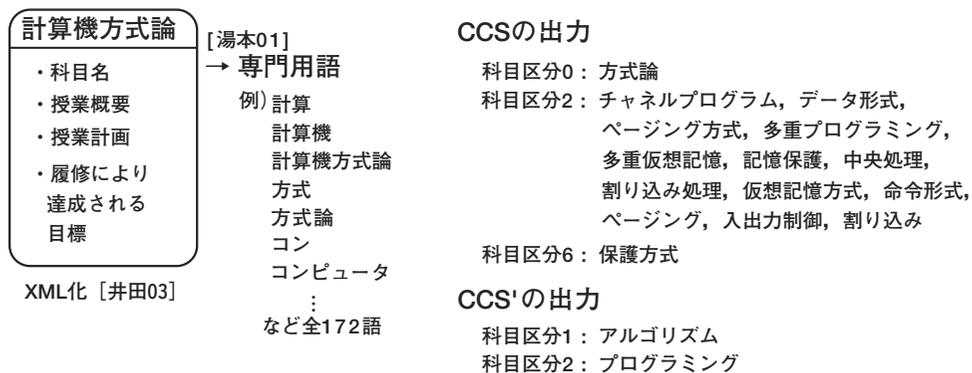


図3 R大学「計算機方式論」の分類支援結果

同様に、「論理数学1」のシラバスをCCSおよびCCS'に入力した結果を図4に示す。この場合は、シラバスから「数学」「論理」など全部で104語の専門用語が抽出され、それらのうち8語がmyDBとマッチした。具体的には、myDB(0)には「解法例」、myDB(1)には「論理数学」、myDB(2)には「断片」、myDB(3)には「多値論理」「Gantzen流」「Hilbert流」、myDB(4)には「ダイナミック論理」、myDB(7)には「時制」がマッチしたが、myDB(6)にはマッチする専門用語はなかった。それに対し、CCS'を用いた場合、「OS」がicDB(2)とマッチしたのみである。

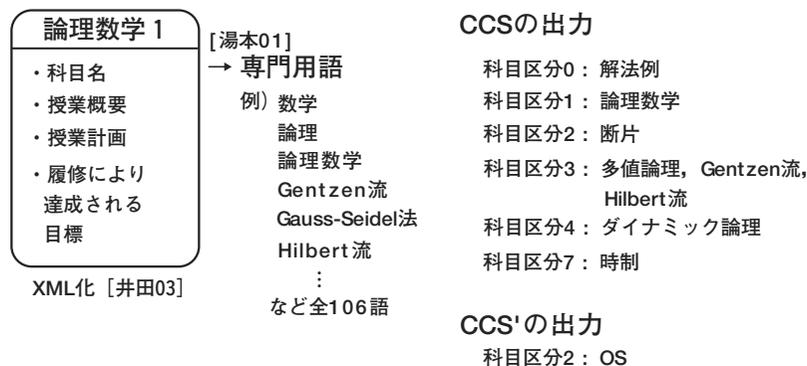


図4 R大学「論理数学1」の分類支援結果

既に述べたように、本章では、ある申請者が申告した科目（39科目）すべてが初出科目である場合を想定している。このような場合には、現状では、専門委員には、原則すべての科目の判定を依頼している。しかし実際には、専攻基準を満たした時点で、それ以上の判定を行う必要はない。勿論、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考えられる。

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 に分類されるとしている科目から優先的に提示してゆく。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。

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。言い換えれば、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。以下では、この観点から、CCSおよびCCS'を評価する。ところで、科目区分6, 7に関しては、「新しい学士への途」に例示科目が存在しないことからわかる通り、現実には、分類上、問題となることが少ない。また科目区分0は「分類困難な専門科目」であり、過去に同様の判定を行っていないことが多く、分類支援は、通常、困難である。従って以下では、科目区分1, 2, 3, 4につい

でのみ、先の観点を検証することとする。

表6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。さらに表6の「3.2.1節での分類結果」には、3.2.1

表6 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

科目名	3.2.1節での 分類結果	CCSの結果				CCS'の結果			
		1	2	3	4	1	2	3	4
情報数学2-A及び演習	c16	3				1			
統計学及び演習	c6	1							
プログラム言語	c2		14	2			3		
プログラム言語2	c2		8	1			2		
情報構造	c1	4	5	1		2			
コンパイラ	c2	4	6			1	2		
多変量解析	c6	2		1					
情報数学3及び演習	c6			1		1			
データ解析	c1	3	1			1			
人工知能	c3	1		8			2	3	
応用情報科学	c3	1	1	2					
論理数学1	c1	1	1	3	1		1		
論理数学2	c12	5	5						
計算機方式論	c2		13			1	1		
データベースシステム	c2	1	14	1			1		
生体情報論	c34	1		2	1			2	
計算の理論1	c1	7	1			2			
計算の理論2	c1	7	1			2			
確率論1及び演習	c16	3		1		1			
確率論2及び演習	c16	1		1	1				
幾何学	c7								
統計学3	c16	2		1					
光通信理論	c0							1	
複雑さの理論	c0					1			
情報数学2-B	c16	3				2			
計画数学1及び演習	c134	21		1	1	2	1		
システムプログラム	c2		7	1			3		
数理経済学	c0							1	
遺伝情報学	c0	1							
計算機概論	c124	3	15		2	2	1	2	1
計画数学2	c14	5			1	1	1	1	
情報理論及び演習	c1	6							
オートマトン	c1	7	1			2	1		
解析学及び演習	c7	1							
基礎情報科学	c16	1							
情報数学1及び演習	c1	1				1			
線形代数及び演習	c7	1			1				
計算機入門及び演習	c2	1	11	1			3		
物理学	c7								

節で宮崎が分類した結果を示す。この「3.2.1節での分類結果」内には科目区分1, 2, 3, 4に関連する科目としては、「情報数学2-A及び演習」から「計算機入門及び演習」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、全てそれに相当する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、「応用情報科学」「論理数学2」など全部で11科目で科目区分 i に相当するicDB(i)とマッチする専門用語数がない。

以上の結果から、CCSでは、 i ($i = 1, 2, 3, 4$) に関する分類を行っている際、その科目区分に含まれるすべての科目を専門委員に提示することができるのに対し、CCS'では、そのような情報を十分には提示しきれない場合があることがわかる。このことから、分類支援情報の提示という意味で、CCSはCCS'に比べ、より有効であると考えられる。

5. 科目分類支援システムの活用例

5.1 科目の移動支援

既に判定された科目の中に単位数が不足している科目区分に移動可能な科目がないかを探することを考える。この場合、基本的には、移動の対象となる科目は、申請者が申告しているすべての科目である。本研究では、CCSを活用し、移動の検討対象となる科目数の削減を目指す。

4章で用いたR大学工学部情報科学科の39科目を対象に、この「科目の移動」支援を検証する。申請者がR大学の科目のみを申告していたとすると、CCSを用いない場合は、この39科目すべてが移動の検討対象となる。すなわち、基本基準および専攻基準を満たす科目分類への移動を調べるためには、39科目すべてのシラバスを読む必要がある。しかしCCSの利用により、ある科目が明らかに移動に適さないということが示されれば、移動に要する時間の短縮が期待される。

具体的な活用方法としては、CCSが移動可能と示唆した科目から順に、専門委員にシラバスの内容を検討して頂くことを想定している。ところで、一般に、機械による自動分類には「分類すべきところに分類していない」という誤り（第1種の誤り）と「分類すべきでないところに分類してしまう」という誤り（第2種の誤り）が考えられる。本論文の立場上、第1種の誤りは存在してはならない。すなわち、移動可能な科目を「移動可能でない」と判定することがあってはならない。これが満たされた上で、第2種の誤り、すなわちCCSが移動可能であると示唆する科目が少なければ少ないほど、より少ない労力で科目の移動が実現されることとなる。

表7 科目移動支援の結果

	該当科目数	示唆する科目数		見落した科目数	
		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

5.2 科目移動支援の結果

科目移動支援の結果を表7に示す。左端（該当科目数）は、先の39科目に対する3.2.1節ステップ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宮崎が判断した科目数を意味する。これは、表6の「3.2.1節での分類結果」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。例えば、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「情報数学2-A及び演習」など全部で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残りの4列である。最初の2列は、各手法がその科目区分、すなわちその科目区分に該当する分類項目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。例えば、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。

最後の2列は、その科目区分に分類されるべき科目を見落としした数である。すなわち、第1種の誤りに相当する。これは、「該当科目数」に数え上げられた科目のうち「示唆する科目数」に含まれなかったものの数である。例えば、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。

5.3 結果の解釈

表7より、科目区分2に関しては、CCS、CCS'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。すなわち、第1種の誤りはゼロである。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、4科目すべてを見落とし、移動の対象として示唆された科目の中に科目区分4に分類されるべき科目が存在しないという最悪の結果を呈している。このようにCCS'では、科目区分によっては、第1種の誤りをゼロにすることができない。

それに対し、CCSでは、本来、その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。すなわち、第1種の誤りはつねにゼロである。これは、本研究の目的である移動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、第2種の誤りに関しては、CCS'に劣る場合がある。これはある意味、第1種の誤りをゼロにする上でやむを得ないことではあるが、第2種の誤りの縮小は今後の重要な課題である。

6. おわりに

本論文では、電子化されたシラバスに基づいた、学位授与事業のための科目分類支援システム試作の現状について述べた。R大学情報系学部の39科目に適用することで有効性を確認した。

今後は、まず第一に、他の「専攻区分」での動作確認を行いたいと考えている。この場合、その「専攻区分」に合致した科目DBの作成が重要となる。科目DBの作成はある程度の専門性

を要するので、今後は、既存の科目DBの活用も積極的に考えていきたい。具体的には、科目DB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。

次に、CCSの出力として提示された用語の有効性に関する検討、ならびに、提示される情報をより豊富にすることも重要である。例えば、提示した情報に信頼度のようなものを付与したいと考えている。これには統計処理や機械学習的な要素をシステムに導入することである程度実現可能であると考えている。

さらに、3.3節で述べたように「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。本研究は平成14年度版の「新しい学士への途」を参考にシステムを構築したが、平成13年度版の「新しい学士への途」とは、「情報工学」区分の「例示科目」に若干の変更が生じている。このような変更によ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などの工夫が必要であると考えられる。

また、本研究では、専門用語を抽出するために[湯本01]を利用しているが、専門用語を抽出する手法には、他にも様々なものが存在する[辻03]。4章ならびに5章で述べた活用事例では、提示された用語の活用には至っておらず、その意味で、現時点では、専門用語抽出手法の差異による影響はさほど大きくないと思われる。しかし、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特に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。

用語抽出に関連した他の改良としては、シラバスをXML化した後の処理における抽出対象から「科目名」を除外することも考えている。通常、「科目名」に含まれる事項は「授業概要」や「授業計画」に含まれることが多く、あえて「科目名」を抽出の対象にする必要がないと思われる。また、現実には、「科目名」が、その講義の内容を正確に反映していないことがあり、そのようなシラバスの不正確さを「科目名」を抽出対象から除外することである程度吸収可能であると考えている。

CCSの活用例として4章で述べた「科目の移動支援」に関しては、本来、移動できない科目を移動可能であ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。「見落とし科目数」(第1種の誤り)と「示唆する科目数」(第2種の誤り)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。

最後に、本研究は、あくまで分類支援を重視したものであり、機械による自動分類の精度向上が主眼ではないが、決定木やSupport Vector Machine (SVM) などのテキスト分類手法[高須03]との比較も今後の重要な課題である。また、本研究同様に、電子化されたシラバスを扱った研究に[山田03, 伊東03]などがある。そこでは、電子化シラバスの自動収集とそこからの情報抽出が主たる課題である。今後はそこで得られている知見を積極的に活用し、CCSの充実をはかる予定である。

謝辞

本研究を遂行するにあたりご協力いただいた当機構学位審査研究部の教官をはじめ「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〕

- [井田03] 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)。
- [湯本01] 湯本紘彰, 森辰則, 中川裕志: 出現頻度と接続頻度に基づく専門用語抽出, 情報処理学会第145回自然言語処理研究会, pp. 111-118 (2001)。
- [「新しい学士への途」平成14年度版] http://www.niad.ac.jp/sub_gakui/siryu/new/new_gakushiH14.pdf
- [辻03] 辻慶太, 芳鐘冬樹: 専門用語として普及しそうな語の自動抽出, 第51回日本図書館情報学会研究大会発表要綱, pp. 105-108 (2003)。
- [高須03] 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)。
- [山田03] 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)。
- [伊東03] 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 (2003)。

付録

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								
42	情報工学	<table border="1"> <tr> <td colspan="2">専門科目 (40単位以上)</td> </tr> <tr> <td> 【A群 (講義科目)】 (30単位以上) <input type="radio"/> 情報工学基礎理論に関する科目 (4単位以上) <input type="radio"/> 計算機システムに関する科目 (6単位以上) <input type="radio"/> 情報処理に関する科目 (6単位以上) <input type="radio"/> 情報に関連する科目 【B群 (演習・実験・実習科目)】 (6単位以上) <input type="radio"/> 情報工学に関する演習・実験・実習科目 </td> <td></td> </tr> <tr> <td colspan="2">関連科目 (4単位以上)</td> </tr> <tr> <td> <input type="checkbox"/> 工学の基礎となる科目 <input type="checkbox"/> 工学及び周辺技術等に関する科目 </td> <td></td> </tr> </table>	専門科目 (40単位以上)		【A群 (講義科目)】 (30単位以上) <input type="radio"/> 情報工学基礎理論に関する科目 (4単位以上) <input type="radio"/> 計算機システムに関する科目 (6単位以上) <input type="radio"/> 情報処理に関する科目 (6単位以上) <input type="radio"/> 情報に関連する科目 【B群 (演習・実験・実習科目)】 (6単位以上) <input type="radio"/> 情報工学に関する演習・実験・実習科目		関連科目 (4単位以上)		<input type="checkbox"/> 工学の基礎となる科目 <input type="checkbox"/> 工学及び周辺技術等に関する科目		工学
専門科目 (40単位以上)											
【A群 (講義科目)】 (30単位以上) <input type="radio"/> 情報工学基礎理論に関する科目 (4単位以上) <input type="radio"/> 計算機システムに関する科目 (6単位以上) <input type="radio"/> 情報処理に関する科目 (6単位以上) <input type="radio"/> 情報に関連する科目 【B群 (演習・実験・実習科目)】 (6単位以上) <input type="radio"/> 情報工学に関する演習・実験・実習科目											
関連科目 (4単位以上)											
<input type="checkbox"/> 工学の基礎となる科目 <input type="checkbox"/> 工学及び周辺技術等に関する科目											
■専門科目の例■		【A群 (講義科目)】 <input type="radio"/> 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), 計画数理論 (OR, 待行列) など <input type="radio"/> 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 <input type="radio"/> 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 <input type="radio"/> 情報に関連する科目 集積回路, 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)】 <input type="radio"/> 情報工学に関する演習・実験・実習科目									

[ABSTRACT]

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

MIYAZAKI Kazuteru*, IDA Masaaki**, YOSHIKANE Fuyuki***,
NOZAWA Takayuki*** and KITA Hajime****

The National Institution for Academic Degrees and University Evaluation (NIAD-UE) 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 Course Classification Support System (CCS) using analysis of syllabus data. We have shown its effectiveness using two examples.

* Associate Professor, Faculty for the Assessment and Research of Degrees, National Institution for Academic Degrees and University Evaluation

** Associate Professor, Faculty of University Evaluation and Research, National Institution for Academic Degrees and University Evaluation

*** Research Fellow, Faculty of University Evaluation and Research, National Institution for Academic Degrees and University Evaluation

**** Professor, Academic Center for Computing and Media Studies, Kyoto University