

シラバスに基づく理工系情報学科の カリキュラム調査

正田 輝雄 石畑 清

明治大学理工学部

調査の目的

全国の大学理工系情報学科の授業内容について、シラバスを資料として用いる調査を一昨年から行っている。これまで49学科を調査した。

日本の大学の情報学科は、電気系、数理系、物理系、経営工学系など異なった専門の学科から派生したところが多い。情報の名称のつく学科は、理工系に限っても、その理念や設立経緯、歴史は多様である。その教育内容は、機械系や電気系と比べても、これらの事情を反映してやはり多様であると言われている。しかしそのカリキュラムの実態を詳しく調査したことはこれまでなかったようである。

情報処理学会は永らく、大学理工系情報学科の学部カリキュラム標準を制定してきた。これらは大学学科の科目を定めるのに広く利用されてきた。前回のJ97まで、標準は科目の内容によって定めるものだったが、情報処理教育委員会はJ07において初めて、CS、IS、CE、SE、IT、GEの6分野に分けて標準をまとめた。また科目ではなく知識体系(Body of Knowledge)をエリア、ユニット、トピックの3レベルからなる構造として定めた。特にコアユニットとは必修のユニットである。

このJ07プロジェクトの一環として、コンピュータ科学教育委員会は大学CS系情報学科で学習する内容の標準モデルJ07-CSをまとめた^{1), 2)}。その主要部分がコンピュータ科学知識体系CS-BOK-J 2007(単にCS-BOK-J)である。米国版^{4), 5)}に対応する。CS-BOK-Jは、15エリア、138ユニット、そして

66コアユニット(255時間)からなる。CS-BOK-Jはソフトウェア工学やヒューマンインタフェースを重視する一方で、数値計算はコアに含めていない。

本稿での調査は、日本の情報学科で、知識体系CS-BOK-J 2007の、各ユニットがどの程度教えられているかの統計データを得ることを目的としている。CS-BOK-Jを基準として、各ユニットが各学科でどの程度学習されているかを測るため、シラバスを調査してデータ化する。各学科の個々のカリキュラムの評価が目的ではなく、理工系情報学科全体としてのカリキュラムの分布と傾向を知ることがねらいとしている。これによってコアユニットのカバー率の分布や、授業内容の傾向によるクラスタ解析の予備的な結果が定量的に得られた。

大学理工系情報学科(CS系とは限らない)の多くが加盟する、理工系情報学科協議会(現在約150学科参加)を通してシラバス提供を依頼した(2009年9月)。シラバスの形は冊子あるいはインターネットである。今回の調査は、49学科についてシラバスを調査した。調査には主に提供された冊子を用いた。

エリア別のコアカバー率と総時間の分布

最終的な集計結果としては、次の2つをCS-BOKの15のエリアごとに算出した。

(1) コアカバー率

(2) エリア総時間

(1)は、コアユニット(必修)だけに注目して、CS-

BOK に示されているコア時間 (コアユニットに割り当てるべき時間) に対して, その学科で実際に教えている時間の割合をパーセントで表示したものである。

(2) は, エリアに相当する内容を教えている時間を単純に合計したものである。コアユニットとそれ以外のユニットを区別せず, すべてのユニットの時間を加えてある。単位は時間 (60 分) である。

図-1 と図-2 は, 15 のエリア (表-1) での学科数分布を, (1) コアカバー率と (2) エリア総時間によってグラフ化したものである。エリアごとに, コアカバー率の 5% きざみでの学科数分布と, エリア総時間の 10 時間きざみでの学科数分布である。各図の表題のところには, カッコ内に平均値と標準偏差を載せている。

コアカバー率は 0% から 100% までの値をとる。エリア総時間は, 200 時間を上限と決め, それを超えた学科は 200 時間として表示している。これは, エリアごとに最大値がさまざまであることの影響を排除するためである。エリア間の比較を可能にするとともに, グラフが小さくなりすぎないようにした。たとえばソフトウェア工学のエリアでは, 200 時間を超える学科が 2 校あることが分かる。AR エリアでは総じてコアカバー率が高い。

図-1 と図-2 のそれぞれの最後の図は, すべてのエリアの合計を示す。エリアごとの学科数の分布の違いが表れている。

□ 学科数分布の解析結果

- DS, PF, AL, AR, OS など「コアユニット時間数の大きいエリア」あるいは「CS で伝統的エリア」では, コアユニットカバー率は平均 60 ~ 70%。分散は小さい。
- NC (ネットワークコンピューティング) ではコアユニットカバー率は 81% で高い。
- SE では 44% で低い。
- コアユニット時間数の小さいエリア HC, MR, GV, IS, SP では, 学科間の差は大きい。

結果は解析しにくい, このあたりにその学科の特徴が現れると考えられる。

ユニット総時間数の分布においては, コアカバー率と似ているが, 違う点もある。総時間数は, 学科個別の事情によることが多いように思われる。

■ クラスタリングとデンドログラム

前章の統計数値結果を用いて, クラスタリングとデンドログラムの作成を行った。これは中間結果で分析は十分ではない。

クラスタ数 $k=3$ の, カバー率によるクラスタリング結果を図-3 に示す。49 対象が 15, 17, 17 のクラスタ A0, A1, A2 に分かれる。SP (68%), MR (55%), SE (38%) の順でエリア差が大きい。またバラツキの小さいエリアは NC, HC, GV であり, 特に NC と IM でクラスタ A1 (65.5%, 35.6%) が, HC でクラスタ A2 (60.5%) が例外的である。

□ コアユニットカバー率のクラスタ分析

- 全般に, コア時間数の少ないエリアや SE で, クラスタの間に差がついている。コア時間の多い CS 伝統的エリアではクラスタ間に差は少ない。
- クラスタ A0 は各エリアに対しておおむね平均的であり, SP エリアが低い (8.3%) ことを除けば平均カバー率 59.0% である。
- クラスタ A1 は, 平均カバー率 53.1% と 3 つの中では低く, 特に NC, MR, SE, IM のカバー率が低い。一方 IS のカバー率はトップである。
- クラスタ A2 は平均カバー率が 70.2% であり, いずれの分野も高いカバー率を示す。いわば CS 的なクラスタだと言える。

図-4 のマトリックスの行がエリア, 列が大学 (学科) を表し, その要素は各科目のコアのカバー率の値を示す。マトリックスの左側に科目, 上側に大学 (学科) に対してクラスタリングを行った結果をデンドログラムで示している。クラスタリングについては, コアの一一致率のベクトルのユークリッド距離を

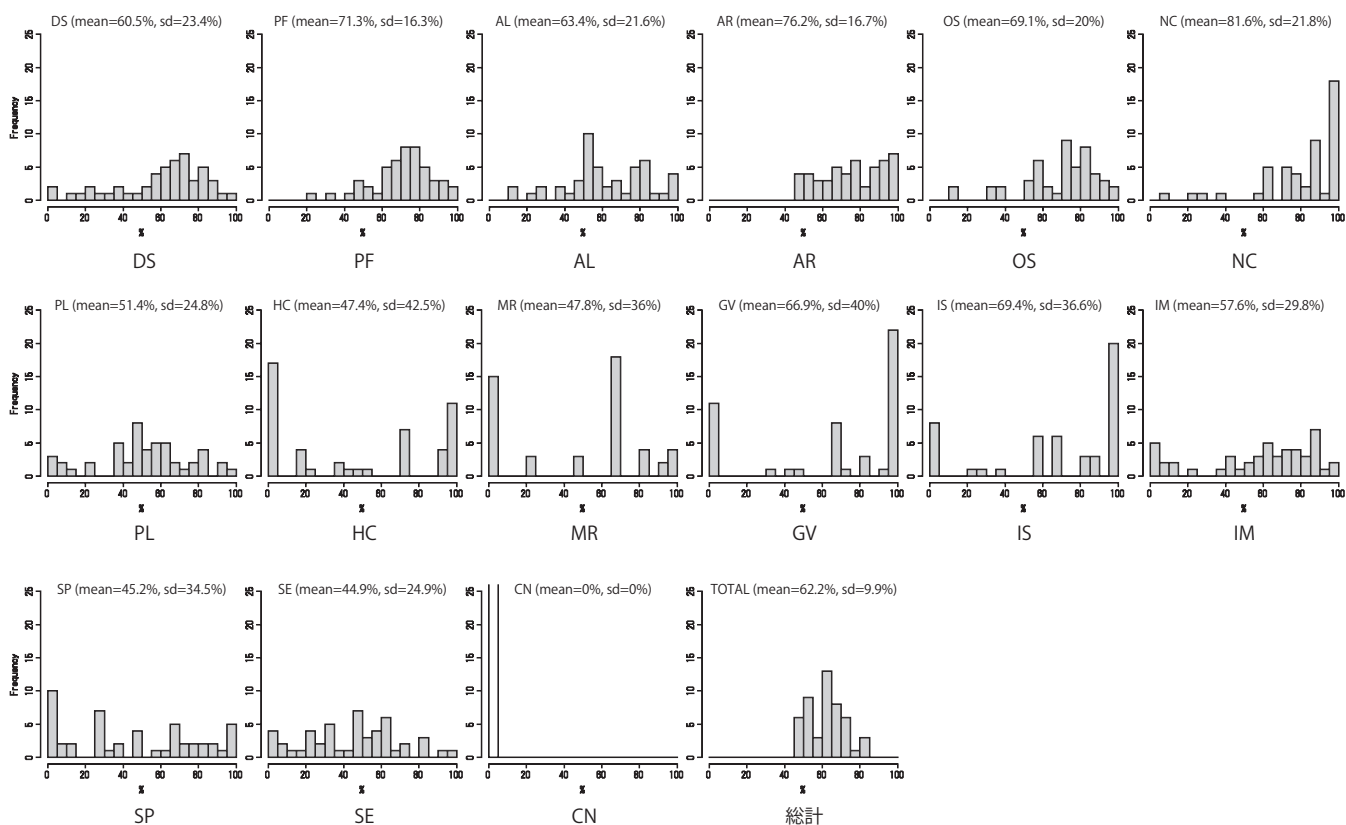


図-1 コアカバー率の分布

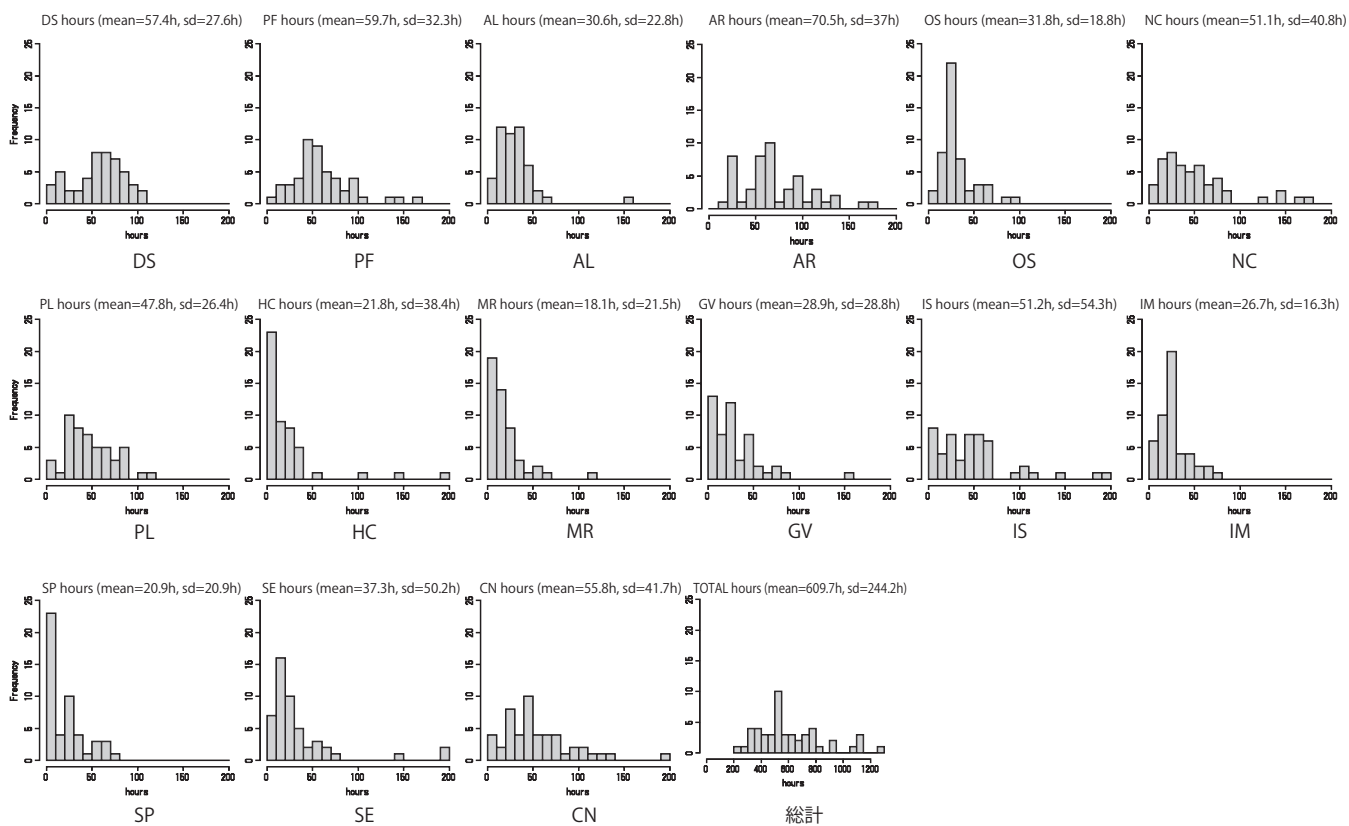


図-2 エリア総時間の分布

DS	離散構造	コア 41 時間
PF	プログラミング基礎	コア 38 時間
AL	アルゴリズム	コア 20 時間
AR	アーキテクチャと構成	コア 32 時間
OS	オペレーティングシステム	コア 17 時間
NC	ネットワークコンピューティング	コア 14 時間
PL	プログラミング言語	コア 17 時間
HC	ヒューマンコンピュータインタラクション	コア 8 時間
MR	マルチメディア表現	コア 3 時間
GV	グラフィックスとビジュアル・コンピューティング	コア 3 時間
IS	インテリジェントシステム	コア 5 時間
IM	情報管理	コア 14 時間
SP	社会的視点と情報倫理	コア 11 時間
SE	ソフトウェア工学	コア 32 時間
CN	計算科学と数値計算	コア 0 時間
	総 計	コア 255 時間

表-1 エリア一覧

距離として、平均距離法による階層型クラスタリングを実施した。デンドログラムは、枝の長さが距離に対応し、短い枝で結ばれている科目（あるいは大学）ほど距離が短い、すなわち関係が強いことを示している。

エリアについては、AR と OS が非常に似た傾向にあり、さらに PF, AL, NC の順に似た傾向のグループを形成する。また、DS と PL, IM と SE, GV と IS, MR と SP がそれぞれ似た傾向にある。DS と PL のグループと AR, OS, PF, AL, NC のグループに分けられるが、それらを合わせたものと、IM と SE のグループに分けられ、さらにそれらを合わせたものが、GV と IS のグループに分けられる。CN は科目の中で特に特異な傾向を示していることが分かる。このような階層関係が左側のデンドログラムから読み取れる。大学(学科)についても同様にデンドログラムから階層関係(系統関係)を見ることができる。

図-5 は、同様にエリア総時間のデンドログラムである。クラスタリングについては、時間数のベクトルのユークリッド距離を距離として、平均距離法による階層型クラスタリングを実施した。

エリアについては、OS と IM, 次に MR と GV が非常に似た傾向にあり、AL, OS, IM, SP のグ

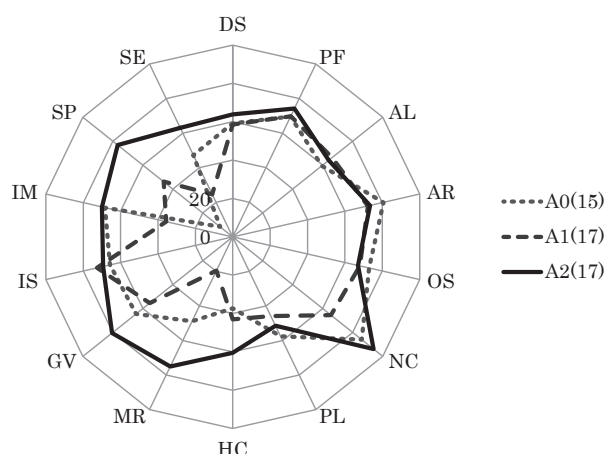


図-3 コアカバー率によるクラスタリング (k=3)

ループ, MR, GV さらに HC を加えたグループが存在するが、それらをまとめたグループが他と比較的離れたグループを形成している。あと、DS と PL, PF と AR のグループが併存している。CN, SE, IS はそれぞれ他と異なる傾向にあることが分かる。大学間の関係もそうだが、コアのカバー率と実施時間数では、その傾向がかなり異なっていることが分かる。

調査の方法

CS-BOK-J の各ユニットの内容つまりトピックが、その学科においてカバーされている時間数を調べる。その結果として、その学科における、15 エリアそれぞれに対する、コアユニットのカバー率と、総時間数が得られる。コアカバー率は CS 系からの視点、総時間数は他系向きと言えるかもしれない。さらにクラスタ分析他によって、エリア間や学科間の類似関係を調べる。

情報系の学科のシラバスを読んで、それぞれの科目の各回に教えている内容が J07 CS-BOK のどのユニットに対応するかを判定した。調査結果を記入する専用 Excel フォームを用意した。1 つの Excel ファイルが 1 つの学科に対応する。ファイルは複数の

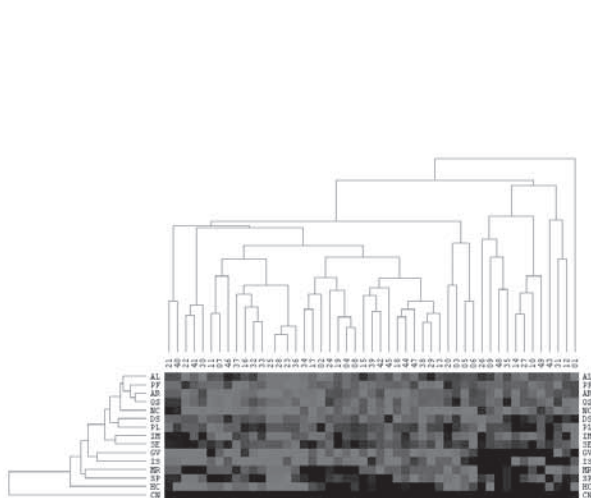


図-4 コアカバー率によるデンドログラム

シートからなり、それぞれのシートが1つの科目に対応する。シートの横軸は授業の回数で、第1回から第15回までのそれぞれに対応する。縦軸はJ07 CS-BOK-Jのユニットである。それぞれの回でどのユニットの内容を教えているかを判定し、該当する欄に印を付けるのが作業の主要部分である。

理工系情報学科協議会の会員となっている学科が146ある。これらの学科を調査対象と決め、CS教育委員長名でシラバスの送付を依頼した。インターネット形式のシラバスも多かったが、冊子の形では57の学科からシラバスの提供を受けることができた。

冊子の形のシラバスを廃止してネット上でのアクセスに限っている大学が最近は多いが、調査の作業は冊子の方がやりやすい。ネット上のシラバスは、科目名をキーとして検索する方式が多く、学科のカリキュラムの全体像を把握しにくい。印付けの作業そのものも冊子の方が短時間で終わることが多い。このような理由から、今回の調査では冊子体のシラバスを入手できた57学科の中から49学科をランダムに選んで調査対象とした。57学科すべてを対象としなかったのはマンパワーの制約のためである。

これらの学科をCS教育委員会の委員（および少数の専門家）にランダムに割り当て、それぞれが分担作業の形で上記Excelフォームへの記入作業を

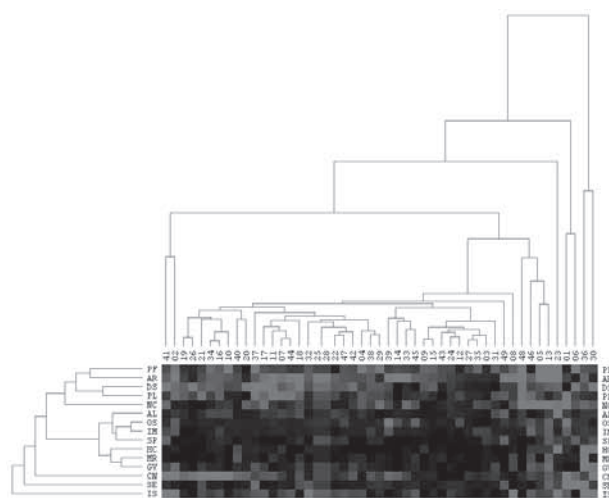


図-5 エリア総時間によるデンドログラム

行った。

学科で教えている科目は、基本的にすべて調査対象とした。ただし、一般教養科目（学部共通科目としている大学が多い）に類する科目は除外した。語学、物理、化学などの科目である。線形代数、統計、離散数学などの数学科目は、専門か一般教養か判断に迷う場合があり、大学によって違う扱いになっていることがある。

今回はそれぞれの科目が学科の中でどのように扱われているかも無視した。具体的には、必修／選択の別は無視したし、コース分けも無視した。後者は、1つの学科の中に複数のコースを設け、特定のコースにしか対応しない科目を設ける場合である。

□ クラスタ分析

比率（コアカバー率、コア時間が0のために対象とならないCNを除く14種）と時間（15エリア総時間）の2種類の調査データについて、k-Means法によりクラスタリングを行った。ここでは前者だけを示す。この手法は対象データを排他的にグループ化を行うための代表的なものであるが、グループ数(kと呼ぶ)はあらかじめアドホックに与える。初期値は、データの中からランダムにk個選ぶ。それぞれのデータがk個の代表値の中のどれに最も近いかを判定し、その代表値のグループに入れる。こうしてできたグループの重心を新たな代表値としてプロセ

スを繰り返し、変化がなくなるまで続ける。

実際には $k=3$ とし、各データは 14 次元のベクトルとして扱い、「近さ」をユークリッド距離で定義する。ここでは 15 エリアは独立であると仮定し、現実に従属関係があっても因果性は想定しない。

得られたクラスタを解釈するために、いくつかの判断用語を設定する。個々のエリアごとに、「クラスタ間に差がある」とするのは、そのクラスタの重心値の差がクラスタ間で大きく違っているときである。そのエリアにおいて「バラつきが小さい」とは、差が大きい標準偏差が小さい、つまり特異なクラスタが存在する場合を示している。差も標準偏差も大きいエリアはまとまりが悪く特徴を抽出しにくい（クラスタを抽出しにくい）ことを意味する。差が小さいときは標準偏差にかかわらず特徴差がないと考えられる。

この分析では、クラスタリングの結果を解釈するために、各クラスタ重心を比較し、エリアのカバー率の最大最小の差が 15% 以内であれば当該エリアで差は生じていないとする。差が大きいとき当該エリア内で 15% 以内の標準偏差であればバラツキが大きく、5% 以内なら小さいとする。

結果と今後

大学学科の教育内容の調査はこれまで少なく、それもアンケートの部分的な結果によるものがほとんどだったと思われる。シラバスによる調査は、実際の授業内容の実地調査には劣るが、アンケートよりは正確である。ただし 1 学科あたりの調査は、大学カリキュラムに関する一般的知識と知識体系 CS-BOK-J に通じた調査者による、4～5 時間程度の作業を要する。

つまり今回の調査の特徴は次の 3 点である。

- 1) アンケートでなくて、シラバスに基づく調査という新手法。定量的。
- 2) 結果は中間結果だが、ある程度予想通りのもの。

- 3) 作業量が多いことと、基準およびシラバス一般に関する知識が必要。

今後の課題として次のことがある。

- CS を基準に用いることには当然だが長所と短所がある。
- データの複数チェックなど、精度や解析をより詳しく正確にする。
- 以下についての区別：科目の必修／選択、講義／実習／実験、一般／専門科目。

本稿は文献 3) からまとめたものである。シラバス提供に協力いただいた学科および調査協力者の方々に厚く謝意を表する。

参考文献

- 1) 情報処理学会カリキュラム標準コンピュータ科学 J07-CS 報告書, 151p. (2009 年 1 月 20 日). http://www.ipsj.or.jp/12kyoiku/J07/20090407/J07_Report-200902/4/J07-CS_report-20090120.pdf
- 2) 正田：情報専門学科カリキュラム標準「J07」：コンピュータ科学領域（J07-CS），情報処理，Vol.49, No.7, pp.728-735 (July 2008). http://www.ipsj.or.jp/12kyoiku/J07/20090407/J07_Report-200902/4/IPSJ-MGN4907_J07-CS-200806.pdf
- 3) 石畑 清，大岩 元，角田博保，清水謙二郎，玉井哲雄，長崎等，中里秀則，中谷多哉子，正田輝雄，三浦孝夫，箕原辰夫，和田耕一，渡辺 治：理工系情報学科の授業内容分布のシラバスによる調査（中間報告），情報処理学会情報教育シンポジウム SSS2010, pp.139-149.
- 4) The Joint Task Force on Computing Curricula : Computing Curricula 2001 : Computer Science, Final Report, IEEE Computer Society, ACM (Dec. 15, 2001). http://www.acm.org/education/education/education/curric_vols/cc2001.pdf
- 5) Computer Science Curriculum 2008 : An Interim Revision of CS 2001, Report from the Interim Review Task Force, ACM, IEEE Computer Society (Dec. 2008). <http://www.acm.org/education/curricula/ComputerScience2008.pdf>

(2011 年 4 月 2 日受付)

正田輝雄（正会員） hikita@cs.meiji.ac.jp

1970 年東京大学卒業。東京大学、東京都立大学を経て、1989 年から明治大学理工学部情報科学科教授。計算理論、Web 技術等に興味を持っている。2003 年から 2010 年まで本会コンピュータ科学教育委員会委員長。著書は「コンパイラの理論と実現」（共立出版）など。「情報科学こんせぶつ」（朝倉書店）編集。

石畑 清（正会員） ishihata@cs.meiji.ac.jp

1974 年東京大学理学部物理学卒業。東京大学を経て、明治大学理工学部教授。言語を中心として、プログラミング全般に興味を持つ。