数理解析研究所講究録
第1801巻 2012年 68-79

情報教育と数学の関わり

一橋大学　山崎秀記 (Hideki Yamasaki) *

Hitotsubashi University

概要

情報・計算機教育に関わる立場から，主に大学初年次において，「教育数学をどう考えるか」についてささやかな論考を試みる．

1　はじめに

情報教育の目的については，文部科学省などが「情報活用能力」の涵養を提案しており，その中には「情報活用の実践力」「情報の科学的な理解」「情報社会に参画する態度」の3つの要素があるとされている [1].

大学初年時における情報教育では，主に

1. How To 的なコンピュータ，とりわけ代表的なアプリケーションソフトの活用能力
2. ネットワークシステムのセキュリティに関する知識
3. 情報倫理，著作権などの社会科学的知識

等の獲得が期待され，「情報の科学的な理解と取扱い」，とりわけその数学的側面が軽視されてきたように思われる．事情は高校の教科「情報」でも同様であり，とくに進学校などでは実質的に数学など他の科目に振り替えられてきた実態もある [2].

筆者は，情報教育および数学教育の両面において，その共通部分に位置する情報科学的・計算機科学的要素がもっと重視されてよいと考えている [3]．本論考では，情報教育との関連の中で立ち現れる数学的内容について，筆者が関心を持ついくつかのトピックを通して概観することで，本研究集会で与えられた課題に応えたいと考える．

*E-mail:yamasaki.hideki@r.hit-u.ac.jp

2 集合論・論理学との関わり

集合論や論理学の基本的素養は，科学一般を学ぶ上での必須事項であるが[1]，コンピュータについて学ぶ際にも，その基礎をなす集合論や論理学の理解は欠かせない．基本論理素子（演算）であるAND（$\wedge$），OR（$\vee$），NOT（$\neg$）の組み合わせで任意の論理回路が構成できる「完全性」の議論はもとより，一種類の電気（電子）的なスイッチング素子があれば十分であることも，集積回路を理解する上で重要である．

> 例．ティンカートイ・コンピュータ [4] では，スイッチング素子を木組みのおもちゃで実現して三目並べをするコンピュータを構築している．

さらに，基礎的な問題としては，濃度に関わる，

> 自然数濃度 (加算無限) ＝記述全体の濃度＜実数濃度＝自然数のベキ集合濃度

がある．コンピュータは，本質的に記述（データ）処理機械であるから，そのメモリー量および計算時間に上限のない抽象的なコンピュータにおいても，そのモデル化能力には限界が生じる．さらに，数学基礎論に関連して，計算可能性の理論も本質的であろう．

> 例．プログラムの停止性判定問題は，筆者の経験では文系の学生でも比較的容易に理解される．

3 計算量の感覚

オーダーのおよび計算量の概念も重要であろう．

> 例．$\displaystyle\sum_{n=1}^{\infty}\frac{1}{n}=\infty$ は数値計算の実行可能性の観点からは2つの非現実的仮定，無限精度の計算と無限の計算時間，を必要とし，計算機実験（数値計算）的には納得することは困難である．実際，単精度計算（有効数字7～8桁）では15.4程度から増えないし，n まで加えて，値は $\log n$ 程度であるから，和の値を k にするのに，e^k に比例する時間がかかる．

2の集合・論理学との関連を含めて，コンピュータを人間の脳の働きの延長としての機械的補助装置として捉えるときに，その形式的限界についての理解とともに，アルゴリズムの実行可能性に関する正しい感覚を養うことが肝要であろう．

---

[1] AならばBを示せ，と言われているのに，BからAを導いて平気でいる学生（受験生）はけっこう多い．

4 離散数学との関わり

情報の科学的な理解における最も基本的要素は，対象の概念モデルの構築であろう．そこでは，

1. 路線図等に代表されるネットワークモデル
2. 組織図等に現れる木構造モデル
3. データベースの基礎を成す関係モデル

等が基本的である．

1．と2．の理解には，グラフ理論的素養が欠かせないし，3．においては数学的な写像や関係に関する基礎知識が必要である．

グラフにおける種々の解法は，その頂点の機械的かつ網羅的な探索が基本的であり，その手法として，深さ優先探索と幅優先探索がある．また，グラフに対し，その (i,j) 要素が頂点 i から頂点 j への辺のラベル（本数，距離，遷移確率，遷移記号）を表す遷移行列を考えることができる．

様々な隣接行列：その要素と基本演算

要素	積演算	和演算	備考
本数	×	+	路の本数
辺の有無	∧	∨	路の有無
距離	Min	+	最短距離
遷移確率	×	+	マルコフ連鎖
遷移記号	∨	連接	路のラベル：正規集合

これらの隣接行列の利用については，後にグーグルランキングや代数的オートマトン理論のところで議論する．

関係データベースで基本となる関係演算は，集合演算（和，差，積）と直積，射影，選択であり，標準的な記述言語 SQL では

SELECT （射影） FROM （直積） WHERE （選択）

で問い合わせを行う．さらに表（関係）の正規化においては，その内部に関数関係を含まないことが本質的である．

例．以下の表で示された学生と成績という 2 つの関係が与えられたとする．

学生 $\subseteq$ 学生ID×氏名×学部×学年，成績 $\subseteq$ 科目×学生ID×教員ID×点数×年度

学生 ID	氏名	学部	学年
1101	田中	商	1
1102	中山	経	1
1103	山本	法	1
1104	本田	理	1

科目	学生 ID	教員 ID	点数	年度
線型	1101	0001	90	2010
微積	1102	0005	80	2010
線型	1102	0001	70	2010
線型	1103	0001	60	2010

例えば，線型で80点以上の学生の氏名を求める問い合わせは，以下のように書くことができる．
SELECT　氏名, 学部
FROM　学生，　成績
WHERE 学生.学生ID＝成績.学生ID ∧ 成績.科目＝線型 ∧ 成績.点数≧80
これは，学生×成績　から条件（学生.学生ID＝成績.学生ID ∧成績.科目＝線型∧成績.点数≧80）を満たすレコードを取り出し，その氏名と学部を表示することを意味し，

氏名	学部
田中	商

という関係を得る．

5　アルゴリズム

数学は歴史的には解法の学であるから，教育数学の様々な局面において，解法（アルゴリズム）の記述は基本的であろう．実際，代入記号（:=）の使用は比較的見られるようになってきたが，数学におけるアルゴリズム記述のために標準的な（実装を仮定しない抽象レベルの）以下のような人工言語が考えられるかもしれない．

基本構文

1. //以降行末までコメント
2. 〈__条件__〉ならば

   【＿＿＿＿】

   ことをし，そうでなければ

   【＿＿＿＿】

   ことをする
3. 〈__変数__〉の値が
   - 〈__値__〉のとき【＿＿＿＿】
   - 〈__値__〉のとき【＿＿＿＿】
   - …
   - 〈__値__〉のとき【＿＿＿＿】

   ことをする
4. まず【＿＿】として，〈__条件__〉まで

【＿＿＿＿】

ことを［【＿＿】しながら］繰り返す

5. 〈＿値の列＿〉中の（〈＿条件＿〉を満たす）各要素〈＿変数＿〉に対して順に

【＿＿＿＿】

ことを繰り返す

6. 以下の手順のうち

- 【＿＿＿＿】
- 【＿＿＿＿】
- …
- 【＿＿＿＿】

のどれかを実行する.

7. 〈＿条件＿〉を満たす〈＿変数＿〉を選ぶ

8. 手順（関数）定義

（入力列）手順名［結果の式表現 … ］

定義本体

9. 手順［関数］呼び出し

（入力列を） 手順名

（入力列を） 関数名 したもの

式表現

データ構造とその基本演算

- スカラー：加減乗除
  - 整数：剰余，大小関係，因数分解
  - 有理数：大小関係
  - 実数：大小関係
  - 複素数：極座標変換
  - 有理式：部分分数分解

- 実数区間　(a,b), (a,b], [a,b), [a,b]
- 整数区間　m..n または [m..n]
- 文字列：連接
- ベクトル（1次元配列）：定数倍，和，内積，外積
  - 行ベクトル
  - 列ベクトル
- 行列（2次元配列）：和，積，逆行列，基本操作，…
- 関係：集合和，共通部分，積，射影，選択，直積，…
- グラフ：強連結部分，深さ優先探索，幅優先探索，…　　ネットワーク　　木構造

これにしたがって，例えば，逆行列と微分の手順を記述してみると以下のようになるだろう．

例1. 行基本変形を用いて，n 次正方行列 A の逆行列 A^{-1} を求める手順

（$n \times m$ 行列 B の i 行と j 行を） 交換する [式表現 $B := P(i,j)B$]

$[1..m]$ の各要素 k に対して順に

$B_{i,k}$ と $aB_{j,k}$ を交換する

ことを繰り返す

（$n \times m$ 行列 B の i 行を a） 倍する [式表現 $B := Q(i,a)B$]

$[1..m]$ の各要素 k に対して順に

$B_{i,k}$ を a 倍する

ことを繰り返す

（$n \times m$ 行列 B の i 行に j 行の a） 倍を加える [式表現 $B := R(i,j,a)B$]

$[1..m]$ の各要素 k に対して順に

$B_{i,k}$ に $aB_{j,k}$ を加える

ことを繰り返す

手順本体　//n 次正方行列 A の逆行列 A^{-1} を求める

$B := n \times 2n$ 次行列 $A|E$
//行列の基本変形を用いて B を $E|A^{-1}$ に変形するために
$[1..n]$ の各要素 j 対して順に

〈$B_{j..n,j}$ が零ベクトル〉 ならば

終了　//A は逆行列を持たない

をし，そうでなければ

//B の j 列 $= \mathbf{e}_j$ にするために
〈$aB_{i,j} = 1$〉 を満たす数 a と行番号 $i \in [j..n]$ を選択
（B の i 行を a） 倍する
（B の j 行と i 行 を） 交換する　//$B_{j,j} = 1$ になった
列 $[1..n]$ の 〈j と異なる〉 各要素 i 対して順に
（B の i 行に j 行の $-B_{i,j}$） 倍を加える
ことを繰り返す　//B の j 列 $= \mathbf{e}_j$ になった

ことをする

ことを繰り返す　//行列の基本変形を用いて B を $E|A^{-1}$ になった
$A^{-1} := B$ の右半分

例２．$f(x)$ の微分を求める手順

まず $f(x)'$ として 〈' のつく部分式がなくなる〉 まで

（' のつく部分式）が

$(a\,g(x) + b\,h(x))'$ **のとき** それを $a\,g(x)' + b\,h(x)'$ で置き換える
$(g(x)h(x))'$ **のとき** それを $g(x)'h(x) + g(x)h(x)'$ で置き換える
$\left(\dfrac{g(x)}{h(x)}\right)'$ **のとき** それを $\dfrac{g(x)'h(x) - g(x)h(x)'}{h(x)^2}$ で置き換える
$(g(x)^n)'$ **のとき** それを $n\,g(x)^{n-1}g(x)'$ で置き換える
$(h(g(x)))'$ **のとき** それを $h'(g(x))g(x)'$ で置き換えて
（$h'(g(x))$）が
$\sin' g(x)$ **のとき** それを $\cos g(x)$ で置き換える
$\cos' g(x)$ **のとき** それを $-\sin g(x)$ で置き換える
$a'^{g(x)}$ **のとき** それを $\log a\, a^{g(x)}$ で置き換える
$\log'_a g(x)$ **のとき** それを $\dfrac{1}{\log a\, g(x)}$ で置き換える

…

$(h^{-1}(g(x)))'$ **のとき** それを $\dfrac{g(x)'}{h'(y)}$ で置き換えて　$//h(y)=g(x)$

（関数 $h(g(x))$）が

$\sin g(x)$ **のとき** $h'(y)$ を $\sqrt{1-g(x)^2}$ で置き換える

$//$主値 $[-\pi/2,\pi/2]$, $h'(y)=\cos(y)=\sqrt{1-h(y)^2}=\sqrt{1-g(x)^2}$

$\tan g(x)$ **のとき** $h'(y)$ を $1+g(x)^2$ で置き換える

$//$主値 $[-\pi/2,\pi/2]$,$h'(y)=1/\cos^2(y)=1+h(y)^2=1+g(x)^2$

…

x **のとき** それを 1 で置き換える

a **のとき** それを 0 で置き換える

ことを繰り返す

研究集会における兵頭先生のお話に引き移せば,「公式⇒問題演習」の繰返しの中で，解法（アルゴリズム）の構築は学習者任せにされてきたと言える．もちろん，数学や計算機科学の目的が，上記のようなアルゴリズムの習得のみにないことはいうまでもないが，アルゴリズムを明示的に意識させる（する）ことによって，学習者が，もっとも単純なアルゴリズム（＝解答パターンの丸暗記）から脱却して，より複雑な解法の（内的）構築を可能にし，本質的理解を深める助けになるものと思われる．

どちらにしろ，以下のようなアルゴリズムの基本的枠組みは教育数学の中で意識されてよい．

5.1 貪欲アルゴリズム

局所的な最適戦略が（最適）解の構成につながるケース

例．始点からの最短路問題：最短距離が確定済みの頂点に隣接する未確定頂点から，最短の頂点を選ぶことを繰り返す．

5.2 分割統治アルゴリズム

問題を部分問題に「分割」し，解いた部分解を結合する解法．

5.3 再帰アルゴリズム（数学的帰納法）

問題を「分割」した部分問題に，元の問題と同形な問題が含まれる場合．最小単位まで（再帰的に）分解された問題はほとんど自明なので，分割手順と結合手順を与えることによって，問題は再帰的に解ける．

例．微分の諸公式はこの再帰的な「分割」と「結合」の手順を示したもの．

5.4 ダイナミックアルゴリズム

問題の効率的な分割が（局所的には）不明な場合に，可能な部分問題の解のリストを保持する解法．

例．$d_i \times d_{i+1}$ 行列 A_i の積 $A_0 \cdot A_1 \cdots A_n$ の効率的な計算順序を求めるには，各 k, i に対し $A_i \cdot A_{i+1} \cdots A_{i+k}$ の計算の手間をボトムアップで求める．

6 教材としてのトピック

情報科学と数学との関連を述べる上での好教材と思われるものをあげてみたい．

6.1 Google のページランク [5]

Google のページランクは，Web ページのリンク構造が定める巨大なマルコフ連鎖を考え，その定常分布を求めることに相当する．各 Web ページを頂点とし，リンクを辺とする有向グラフの隣接行列 L は，$N(\doteqdot 81$ 億$)$ 次元正方 0,1 行列で，

$$L_{i,j} := \begin{cases} 1 & \text{ページ } i \text{ からページ } j \text{ へのリンクがある} \\ 0 & \text{o.w.} \end{cases}$$

で定義される．確率行列に対するペロン・フロベニウスの定理から，それが既約（⇔グラフが強連結）かつ非周期的（⇔グラフ中の閉路長の最大公約数が 1）ならばユニークな定常解を持つので，適当な $\alpha \in [0,1)$ に対して N 次元確率行列 $G(\alpha)$ を以下で定義する．

$$G(\alpha)_{i,j} := \begin{cases} \alpha L_{i,j}/c_i + (1-\alpha)/N & \text{出リンク数 } c_i := \sum_j H_{i,j} > 0 \text{ のとき} \\ 1/N & \text{出リンク数 } c_i = 0 \text{ のとき} \end{cases}$$

これは，リンクをたどる確率を α とし，すべてのリンクまたはページを等確率で選ぶことを意味する．また，出リンクを持たないページからは，任意のページに等確率で遷移するものとしている[2]．

このとき，ページランクのスコアは $\mathbf{p}G(\alpha) = \mathbf{p}$ かつ $\sum_j \mathbf{p}_j = 1$ を満たす確率ベクトル $\mathbf{p}$ で与えられる．しかし，G は巨大（で密）な正値行列であり，この連立方程式を直接解くのは，N^3 に比例する計算時間がかかり実行可能でない．

[2]実際使われる $G(\alpha)$ には，多くの工夫が施されている [5]

一方，既約かつ非周期的という条件から，任意の確率ベクトル$\mathbf{x}$に対し$\lim_{n\to\infty}\mathbf{x}G(\alpha)^n=\mathbf{p}$が成立つ．ここで，$c_i:=\sum_j L_{i,j}$とおくと，

$$(\mathbf{x}G(\alpha))_j=\sum_i \mathbf{x}_i G(\alpha)_{i,j}=\alpha\sum_{L_{i,j}=1}\mathbf{x}_i/c_i+(\alpha\sum_{c_i=0}\mathbf{x}_i+1-\alpha)/N$$

である．第2項$(\alpha\sum_{c_i=0}\mathbf{x}_i+1-\alpha)/N$は$j$と独立なので1回計算すればすむことと，出リンク数$c_i$が平均的には10程度であることを考慮すると，$\mathbf{x}G(\alpha)$の計算は$N$に比例する時間で実行可能である．

さらに，$G(\alpha)$の主固有値は1で他の固有値の絶対値はα以下であるから，$\mathbf{x}G(\alpha)^n\to\mathbf{p}$の収束の速さは，$\alpha^n\to 0$の収束の速さで抑えられる．グーグルは$\alpha=0.85$を採用していると言われており，50回程度の繰返しでおおよそ2〜3桁の精度が得られる．

このように，Googleのページランクにおいても線形代数の基本的定理が重要な役割を果たしている．

6.2 RSA公開鍵暗号系

インターネットで広く用いられるRSA暗号系は，素数に関わる美しい定理に理論的支柱があるとともに，コンピュータでの実行可能性に関する理解が必要になる．RSA暗号系の理解に必要とされる素材をピックアップすれば，以下のようになるだろう[3].

1. 理論的枠組は以下の定理から導かれる
   - フェルマーの小定理より，
     pが素数かつ$x\neq 0 \mod p$ならば$x^{p-1}=1 \mod p$
   - これに中国の剰余定理を適用して，
     p,qが素数ならば任意の自然数kに対し$x^{k(p-1)(q-1)+1}=x \mod pq$
2. アルゴリズムの実行可能性に関しては，
   - 素数の分布＋効率的な素数判定法により，
     十分大きな素数を見つけることは実行可能
   - modの世界での，
     - ユークリッドの互除法を用いた効率的な商演算により，
       p,qの各々を知っていれば公開鍵(e,pq)から$ed=1 \mod (p-1)(q-1)$で定まる秘密鍵(d,pq)の計算が実行可能
     - 効率的な巾乗計算により，
       鍵（(e,pq)または(d,pq)）を知っていれば暗号化（$x^e \mod pq$の計算）または複合（$y^d \mod pq$の計算）が実行可能

3. 暗号系の安全性に関しては，

   - 素因数分解の効率的解法が知られていないので，
     公開鍵 (e, pq) から p, q の各々を求めることは困難
   - 離散対数問題の効率的解法が知られていないので，
     公開鍵 (e, pq) から秘密鍵 (d, pq) を求めることは困難

6.3 代数的オートマトン理論

文字の有限集合を Σ すると，Σ 上の文字列の集合 $\Sigma^* := \{a_1 a_2 \dots a_n \mid a_1, a_2, \dots, a_n \in \Sigma, n \geq 0\}$ は，連接演算のもとで Σ から生成される自由モノイドになる．このとき，

（有限）オートマトンは，
　（有限個の）状態間の遷移を Σ^* の文字列でラベルづけしたネットワークであり，
　Σ^* の文字列（遷移）を作用素とする（有限）モノイドであると考えられる

正規集合は，
　有限オートマトンの開始状態から受理状態への路のラベルの集合として定義され
　自由モノイド Σ^* から有限モノイドへの準同型写像の逆像として特徴付けられる

正規集合のクラスは集合演算 $(\cup, \cap, \cdot^{\mathsf{C}})$ で閉じている

正規集合の等価問題は決定可能である

文字の有限集合 $\Sigma_1, \Sigma_2, \dots, \Sigma_n$ が与えられたとき，

n テープ文字列は，
　これらが生成する自由モノイドの直積 $\prod \Sigma_i^* := \Sigma_1^* \times \Sigma_2^* \times \cdots \times \Sigma_n^*$ の元

n テープ（有限）オートマトンは，
　（有限個の）状態間の遷移を $\prod \Sigma_i^*$ の文字列でラベルづけしたネットワークであり，
　$\prod \Sigma_i^*$ の文字列（遷移）を作用素とする（有限）モノイドであると考えられる

n テープ正規集合は，
　n テープ有限オートマトンの開始状態から受理状態への路のラベルの集合として定義され
　自由モノイド $\prod \Sigma_i^*$ から有限モノイドへの準同型写像の逆像として特徴付けられる

n テープ正規集合のクラスは集合演算で閉じているわけではなく，

2 テープ非決定性有限オートマトンの等価問題は決定不能であるが

nテープ決定性有限オートマトンの等価問題は
代数学における以下の定理を応用して決定可能であることが示された [6]

- free group は fully ordered
- fully ordered group の直積は fully ordered

7 おわりに

一橋大学では，社会科学系大学であることもあって，教養の情報・計算機関連科目は数理科目の一環として提供され，数学エリアが担当してきた．近年は情報基盤センターのスタッフが充実してきており，担当者が増えている．

情報教育においても，How To 的内容や倫理的内容以外に，数学的内容をきちんと教える必要があるだろう．それがそのまま『教育数学』に組み込まれるものであるとも考えないが，情報科学・計算機科学の中からも，幅広い題材が提供できるであろうし，逆に情報教育も、数学的な考え方をより重視して展開していくべきであろう．

参考文献

[1] 文部科学省, 情報教育の実践と学校の情報化～新「情報教育に関する手引」～，文部科学省 (2004)，http://www.mext.go.jp/a_menu/shotou/zyouhou/020706.htm

[2] 国立国会図書館調査および立法考査局文教科学技術課，「高等学校における情報科の現状と課題」，国立国会図書館 ISSUE BRIEF NUMBER 604(2008)，http://www.ndl.go.jp/jp/data/publication/issue/0604.pdf

[3] 山崎秀記，情報科学の基礎，サイエンス社 (2008)

[4] A. K. Dewdney, A Tinkertoy computer that plays tic-tac-toe, SCIENTIFIC AMERICAN October 1989
http://www.rci.rutgers.edu/ cfs/472_html/Intro/TinkertoyComputer/TinkerToy.html

[5] Langbille・Meyer 著，岩野・黒川・黒川訳，Google PageRank の数理，共立出版（2009）

[6] Tero Harju and Juhani Karhumäki, The Equivalence Problem of Multitape Finite Automata,
http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.107.1849&rep=rep1&type=pdf