

Combinatorial system and error-correcting code

広大 教養 岡本 雅典

§1 はしがき

多量のデータを計算機の記憶装置内に格納する場合、有限射影幾何にしたがう配置によって格納すれば検索に際し代数的 identifier を用いて引き出せば、通常の轉置ファイル構成より検索時間が短縮される場合のあることが Abraham, Ghosh and Ray-Chaudhuri (1968) により示された。その後有限射影幾何による配置をファイル構成に用いるこの種の研究は Abraham and Ghosh (1968), Ray-Chaudhuri (1968), Bose, Abraham and Ghosh (1969), Ghosh (1969) により進められて来ている。これらの研究を大別すると(1)記録の属性に有限射影幾何の点 E , 記憶装置内のバケットを flats に (多くは 1-flat) それぞれ対応させる方式と(2)記録の属性に flats (1以上の flat) を, バケットに点 E を対応さ

せる方式とがある。質問の形式もいくつかの属性に因してそれぞれ異なる個数の質問を課する方向の研究が進められていく (Ghosh, 1969)。本論はこの方向に沿ったもので、Ghosh (1969) とは異なり、Bose, Abraham and Ghosh (1969) の結果を拡張する一つの試みをした。

なお error-correcting code との関連としては Ray-Chaudhuri (1968) の行った file organization の研究では属性のすべての t -plets を覆う最小集合を見出すために BCH code の性質が利用された。また Bose, Abraham and Ghosh (1969) では error-correcting code 理論での P_t -性質を有する行列をファイル構成理論の出発点としている。またファイル構成に因して開弁すべき問題としては

(1) 属性集合 $A = \{A_1, A_2, \dots, A_m\}$ としたとき二つのカテゴリに属する $A_1 = \{A_{j_1}, A_{j_2}, \dots, A_{j_s}\}$, $A_2 = \{A_{k_1}, A_{k_2}, \dots, A_{k_t}\}$ に対する二つの質問 Q_1, Q_2 ,

$$Q_1 = \begin{pmatrix} A_{j_1} & A_{j_2} & \dots & A_{j_s} \\ v_{j_1} & v_{j_2} & \dots & v_{j_s} \end{pmatrix}, \quad Q_2 = \begin{pmatrix} A_{k_1} & A_{k_2} & \dots & A_{k_t} \\ v_{k_1} & v_{k_2} & \dots & v_{k_t} \end{pmatrix}$$

に同時に答を得るシステムの開弁。

(2) 検索時間と冗長度の概念の正確な定義。

(3) 各システムの評価検討・比較。

(4) PBI BD の応用。

(5) Dynamic filing system の開発。

等が考えられる。

§2 質問に関する属性が複数個の値を取り得る場合

いま i 番目の個人記録は n 個の属性 A_1, A_2, \dots, A_n に対してそれぞれ数値 $a_{i1}, a_{i2}, \dots, a_{in}$ をとるとし、 i 番目の項目に対して属性ベクトル $f(i) = (a_{i1}, a_{i2}, \dots, a_{in})$ を作る。 i と $f(i)$ が i 番目の記録を構成し、 $f(i)$ の集合はファイルである。 n 個の属性 A_1, A_2, \dots, A_m のいくつかについて質問を行う。 いま g 個の属性について質問が行われ、各属性は質問に答えて高々 t 個の相異なる値を取り得るものとする。 ただし $g \leq t < n$ とする。 n, t, g, t は正整数である。

(定義) j 番目の属性 A_j は次の行列 $L_{j \cdot}$ により表わされる。

$$L_{j \cdot} = \begin{bmatrix} L_{j1} \\ L_{j2} \\ \vdots \\ L_{jt} \end{bmatrix} = \begin{bmatrix} h_{j11}x_1 + h_{j12}x_2 + \dots + h_{j1r}x_r \\ h_{j21}x_1 + h_{j22}x_2 + \dots + h_{j2r}x_r \\ \dots \\ h_{jt1}x_1 + h_{jt2}x_2 + \dots + h_{jtr}x_r \end{bmatrix} \quad (1)$$

ここで $j = 1, 2, \dots, n$ とし, $h_{j \cdot 1} = [h_{j11}, h_{j21}, \dots, h_{jR1}]$,
 $h_{j \cdot 2} = [h_{j12}, h_{j22}, \dots, h_{jR2}]$, \dots , $h_{j \cdot r} = [h_{j1r},$
 $h_{j2r}, \dots, h_{jRr}]$ とおく。これらのベクトルの各成分は
 $GF(q)$ 上の要素である。(いま

$$H = \begin{bmatrix} h_{1 \cdot 1} & h_{1 \cdot 2} & \dots & h_{1 \cdot r} \\ h_{2 \cdot 1} & h_{2 \cdot 2} & \dots & h_{2 \cdot r} \\ \vdots & \vdots & & \vdots \\ h_{n \cdot 1} & h_{n \cdot 2} & & h_{n \cdot r} \end{bmatrix} \quad (2)$$

とみると H は n 行 r 列行列で, 各要素は $GF(q)$ 上から取られる。 $\text{rank } H = r < n$ かつどの r 個の行ベクトルも 1 次独立 (P_L -性質) とする。なお, q, n を与えてかかる性質をもつ最小の r を見出すことは coding の問題でもある。

次に雙向 Q を次のように定義する。

$$Q = \begin{bmatrix} A_{j_1} & A_{j_2} & \dots & A_{j_g} \\ V_{j_1} & V_{j_2} & \dots & V_{j_g} \end{bmatrix} \quad (3)$$

ただし $V_{j_1}' = [v_{j_1 1}, v_{j_1 2}, \dots, v_{j_1 R}]$, $V_{j_2}' = [v_{j_2 1}, v_{j_2 2}, \dots, v_{j_2 R}]$
 \dots , $V_{j_g}' = [v_{j_g 1}, v_{j_g 2}, \dots, v_{j_g R}]$ とする。この雙向 Q
 に対して次の各方程式を対応させる。

$$L_{j_1} = V_{j_1}, L_{j_2} = V_{j_2}, \dots, L_{j_g} = V_{j_g} \quad (4)$$

たとえば $L_{j_0} = V_{j_0}$ については

$$\begin{cases} h_{j_0,1}x_1 + h_{j_0,2}x_2 + \dots + h_{j_0,r}x_r = v_{j_0,1} \\ h_{j_0,2}x_1 + h_{j_0,2}x_2 + \dots + h_{j_0,2r}x_r = v_{j_0,2} \\ \dots \\ h_{j_0,k}x_1 + h_{j_0,k}x_2 + \dots + h_{j_0,k}x_r = v_{j_0,k} \end{cases}$$

となり (4) から g 個の一次方程式が得られる。このように n 個の linear forms の集合 $L = \{L_1, L_2, \dots, L_n\}$ の中から g 個の一次方程式の組をえらぶ。その作り方はまず L の中から g 個の $L_{j_0}, L_{j_2}, \dots, L_{j_g}$ をえらぶ。そのうちから各 L_{j_0} に対して一つずつ取られたものを $L_{j_0,1}, L_{j_2,1}, \dots, L_{j_g,1}$ とし、線形方程式系

$$L_{j_0,1} = v_{j_0,1}, L_{j_2,1} = v_{j_2,1}, \dots, L_{j_g,1} = v_{j_g,1}$$

を作る。この方程式系の係数行列 ($g \times r$) を G_1 とする。

次にすでにえらばれた $L_{j_0}, L_{j_2}, \dots, L_{j_g}$ の中から $L_{j_1,2}, L_{j_2,2}, \dots, L_{j_g,2}$ とはすべて異なる g 個の $L_{j_1,2}, L_{j_2,2}, \dots, L_{j_g,2}$ を取る。これから作った線形方程式系

$$L_{j_1,2} = v_{j_1,2}, L_{j_2,2} = v_{j_2,2}, \dots, L_{j_g,2} = v_{j_g,2}$$

の係数行列を G_2 とおく。以下同様にして G_3, G_4, \dots, G_R の係数行列を得る。これらはいずれも基本変形によりたとえ G_1 は

$$\left[\begin{array}{ccc|ccc} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \ddots & & \\ & & & & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{array} \right] = [G_{1g}, G_{1\lambda}]$$

$\xleftarrow{g} \quad \xrightarrow{\lambda=r-g}$

の形となる。したがって解 X'_{1g} は $X'_{1g} = G_{1g}^{-1} V_{j_1} - G_{1g}^{-1} G_{1\lambda} X'_{1\lambda}$ により $X'_{1g} = [x_1, x_2, \dots, x_g]$, $X'_{1\lambda} = [x_g, x_{g+1}, \dots, x_r]$ 。 $X'_{1\lambda} = 0$ とおけば解 $X'_{1\lambda} = [u_{11}, u_{12}, \dots, u_{1g}]$ を得る。したがって G_{1g} を構成する G_{11} の列番号を $\lambda_1, \lambda_2, \dots, \lambda_g$ とすると $u_{1\lambda} = 0$ ($\lambda \neq \lambda_1, \lambda_2, \dots, \lambda_g$) とし解

$$u_1' = (u_{11}, u_{12}, \dots, u_{1r})$$

を得る。同様にして

$$u_2' = (u_{21}, u_{22}, \dots, u_{2r})$$

$$\dots$$

$$u_g' = (u_{g1}, u_{g2}, \dots, u_{gr})$$

まとめて解行列

$$U = \begin{bmatrix} u_1' \\ u_2' \\ \vdots \\ u_g' \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1r} \\ u_{21} & u_{22} & \dots & u_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ u_{g1} & u_{g2} & \dots & u_{gr} \end{bmatrix} \quad (5)$$

を得る。このとき行列 Ω に対して属性 $L_{j.}$ は Ω において

$$L_{j.}(\Omega) = \begin{bmatrix} L_{j_1}(\Omega'_1) \\ L_{j_2}(\Omega'_2) \\ \vdots \\ L_{j_g}(\Omega'_g) \end{bmatrix} \quad (6)$$

なる値に attain するという。

(定理) いま $GF(q)$ 上の $(g \times r)$ 行列の集合を Ω_0 とする。ただし各行列のどの行ベクトルも r 個以下の要素は同時に零でないものとする。このとき質問 $Q = \begin{pmatrix} A_{j_1} & A_{j_2} & \dots & A_{j_g} \\ V_{j_1} & V_{j_2} & \dots & V_{j_g} \end{pmatrix}$ に対して方程式 (4) を満たす $GF(q)$ 上の長個の $(g \times r)$ 行列の組が少なくとも一つ Ω_0 内にある。

(したがって質問 Q に対して Ω_0 内の $(g \times r)$ 行列のある一つの組を対応させることができる。それは $L_{j.}$ が Ω において attain する仕方においてである。 Ω_0 の中の $(g \times r)$ 行列の数を b_0 とすると

$$b_0 = \{ 1 + (q-1)\binom{r}{1} + (q-1)^2\binom{r}{2} + \dots + (q-1)^r\binom{r}{r} \} \binom{m}{g}$$

であり、一方可能な質問の総数を b とすると

$$b = \{ q\binom{m}{1} + q^2\binom{m}{2} + \dots + q^r\binom{m}{r} \} \binom{N}{r}$$

ただし N はファイルの記録の総数である。 $\binom{N}{r} > \binom{m}{g}$ であるのは $m > r$ なることから $b > b_0$ である。(一般には $N \gg m$ ことが多いが、 r と g との相対的大きさについて

は何も云えない)。この条件の下で、雙向 Ω と \mathcal{L}_0 内の行列の一群との対応は many to one である。 $\Omega \supset \Omega$ の中の任意の $(q \times r)$ 行列の組に対して少なくとも一つの $V_{j_1}, V_{j_2}, \dots, V_{j_s}$ が対応しているような部分集合 Ω を考える。 Ω 中の行列の数を c とすれば $c \leq c_0$ である。

(定理) $r < nr \leq q^m - 1$ ならば次の条件を充す m 行 r 列の行列 H がある。(i) H の各要素は $GF(q)$ から取られる。(ii) $\text{Rank } H = r$, (iii) H は P_t -性質をもつ。ただし $r \leq m(t - \lfloor t/q \rfloor)$ 。証明の方針はまず $GF(q)$ の有限拡大体 $GF(q^m)$ をベクトル空間と見做す。 $GF(q^m)$ の要素で H を構成し、 $GF(q^m)$ の要素について $\alpha^{q^m} - 1 = 0$ の性質から α^{sq} (第 s 行, $s = \lfloor t/q \rfloor$ とす) なる列を取り除き H_1 を作る。更に他の列の一次結合となっているものがあればこれを除いて H を作る。この間 P_t -性質は保たれる。

(例) $q = 3, t = 4, nr = 8$ (前半は Bose, Abraham and Ghosh (1969) に于る)。 $r = nr \leq q^m - 1 = 8 \therefore m = 2$

$$GF(3^2) \longleftrightarrow V_m \text{ over } GF(3)$$

$$\begin{aligned} \alpha_i &= \alpha^{i-1} & (a_0, a_1) \\ \alpha_1 &= \alpha^0 = 1 & = (1, 0) \\ \alpha_2 &= \alpha^1 = \alpha & = (0, 1) \\ \alpha_3 &= \alpha^2 = 1 + 2\alpha & = (1, 2) \\ \alpha_4 &= \alpha^3 = 2 + 2\alpha & = (2, 2) \\ \alpha_5 &= \alpha^4 = 2 & = (2, 0) \\ \alpha_6 &= \alpha^5 = 2\alpha & = (0, 2) \\ \alpha_7 &= \alpha^6 = 2 + 2\alpha & = (2, 1) \\ \alpha_8 &= \alpha^7 = 1 + \alpha & = (1, 1) \end{aligned}$$

$$H_0 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ \alpha & \alpha^2 & \alpha^3 & \alpha^4 \\ \alpha^2 & \alpha^4 & \alpha^6 & \alpha^4 \\ \alpha^3 & \alpha^6 & \alpha^4 & \alpha^4 \\ \alpha^4 & 1 & \alpha^4 & \alpha^4 \\ \alpha^5 & \alpha^2 & \alpha^7 & \alpha^4 \\ \alpha^6 & \alpha^4 & \alpha^2 & 1 \\ \alpha^7 & \alpha^6 & \alpha^3 & \alpha^4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 \\ \alpha & \alpha^2 & \alpha^4 \\ \alpha^2 & \alpha^4 & \alpha^4 \\ \alpha^3 & \alpha^6 & \alpha^4 \\ \alpha^4 & 1 & \alpha^4 \\ \alpha^5 & \alpha^2 & \alpha^4 \\ \alpha^6 & \alpha^4 & 1 \\ \alpha^7 & \alpha^6 & \alpha^4 \end{bmatrix}$$

$$= \begin{bmatrix} 10 & 10 & 10 \\ 01 & 12 & 20 \\ 12 & 20 & 10 \\ 22 & 21 & 20 \\ 20 & 10 & 10 \\ 02 & 12 & 20 \\ 21 & 20 & 10 \\ 11 & 21 & 20 \end{bmatrix} \rightarrow H = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 2 & 2 \\ 1 & 2 & 2 & 0 & 1 \\ 2 & 2 & 2 & 1 & 2 \\ 2 & 0 & 1 & 0 & 1 \\ 0 & 2 & 1 & 2 & 2 \\ 2 & 1 & 2 & 0 & 1 \\ 1 & 1 & 2 & 1 & 2 \end{bmatrix}$$

$$\begin{aligned} L_1 & ; x_1 + x_3 + x_5 \\ L_2 & ; x_2 + x_3 + 2x_4 + 2x_5 \\ L_3 & ; x_1 + 2x_2 + 2x_3 + x_5 \\ L_4 & ; 2x_1 + 2x_2 + 2x_3 + x_4 + 2x_5 \\ L_5 & ; 2x_1 + x_3 + x_5 \\ L_6 & ; 2x_2 + x_3 + 2x_4 + 2x_5 \\ L_7 & ; 2x_1 + x_2 + 2x_3 + x_5 \\ L_8 & ; x_1 + x_2 + 2x_3 + x_4 + 2x_5 \end{aligned}$$

\mathcal{L} の逆も $\text{GF}(3)$ 上で 1 次方程式

$$\text{雙向 } Q = \begin{pmatrix} A_1 & A_2 & A_3 & A_4 \\ v_{11} & v_{21} & v_{31} & v_{41} \\ v_{12} & v_{22} & v_{32} & v_{42} \end{pmatrix} \text{ に対して}$$

$$L_{11} = L_1 = v_{11}, \quad L_{21} = L_2 = v_{21}, \quad L_{31} = L_3 = v_{31}, \quad L_{41} = L_4 = v_{41},$$

$$\omega'_1 = (2v_{21} + 2v_{41}, 2v_{11} + v_{21} + 2v_{31} + v_{41}, v_{11} + v_{21} + v_{41}, v_{21} + 2v_{31} + 2v_{41}, 0)$$

$$\text{同様} \quad L_{12} = L_5 = v_{12}, \quad L_{22} = L_6 = v_{22}, \quad L_{23} = L_7 = v_{32}, \quad L_{24} = L_8 = v_{42}$$

$$\text{よ} \quad \begin{cases} 2x_1 + x_3 + x_5 = v_{12} \\ 2x_1 + x_3 + 2x_4 + 2x_5 = v_{22} \\ 2x_1 + x_2 + 2x_3 + x_5 = v_{32} \\ x_1 + x_2 + 2x_3 + x_4 + 2x_5 = v_{42} \end{cases}$$

$$\text{より} \quad \begin{cases} x_1 & & & + x_5 = v_{22} + v_{42} \\ & x_2 & & + 2x_5 = v_{12} + 2v_{22} + v_{32} + 2v_{42} \\ & & x_3 & + 2x_5 = v_{12} + v_{22} + v_{42} \\ & & & x_4 + x_5 = 2v_{12} + v_{32} + v_{42} \end{cases}$$

において $x_5 = 0$ において

$$\omega'_2 = (v_{22} + v_{42}, v_{12} + 2v_{22} + v_{32} + 2v_{42}, v_{12} + v_{22} + v_{42}, 2v_{22} + v_{32} + v_{42}, 0)$$

したがって雙向に対して Ω の中で行列

$$\Omega = \begin{pmatrix} \omega'_1 \\ \omega'_2 \end{pmatrix}$$

が対応する。たとえば v_{ij} に数値を入れたときは、すなわち

$$Q = \begin{pmatrix} A_1 & A_2 & A_3 & A_4 \\ 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 2 \end{pmatrix} \rightarrow W = \begin{pmatrix} 2 & 1 & 2 & 0 & 0 \\ 2 & 1 & 0 & 2 & 0 \end{pmatrix}$$

$W \in \Omega$ に対して質問に関する属性 $L_j(W)$ が対応している。
そこで W に行列 $B_j = L_j(W)$ の組を対応させる

$$B(W) = (B_1, B_2, \dots, B_m) = (L_1(W), L_2(W), \dots, L_n(W))$$

この B を W に対応する Block という。

(定理) Ω の $W_1, W_2, W_1 \neq W_2 \rightarrow B(W_1) \neq B(W_2)$

かくて対応 (many to one); $Q \rightarrow W \rightarrow B(W)$

(定理) Block $B(W)$ が 質問 Q

$$Q = \begin{pmatrix} A_{j_1} & A_{j_2} & \dots & A_{j_g} \\ V_{j_1} & V_{j_2} & \dots & V_{j_g} \end{pmatrix}$$

に対応するならば

$$B_{j_1} = V_{j_1}, B_{j_2} = V_{j_2}, \dots, B_{j_g} = V_{j_g}.$$

Storage

M = 記録の accession number を格納する記憶部の番地
集合.

$[B(W)]_k$: $g \times n$ 行列の第 k 行 vector $\rightarrow M$ の subsets = $[M(W)]_k$
= buckets
行列 W の第 k 行 vector を buckets $[M(W)]_k$
の label という。 $k=1, 2, \dots, g$

$$A \subseteq [B(\mathbb{W})]_k \otimes f(i).$$

逆に $A' \subseteq Q(A)$ に対して $[M(\mathbb{W}, A')]_k$ を得れば $A' = [B(\mathbb{W})]_k \otimes f(i) \rightarrow (a_{ij_1}, a_{ij_2}, \dots, a_{ij_g}) \in V$ の中に定まり箇所を記す。箇所を記すすべての記録の accession number を見出すためには $M(Q) = \bigcup_{A' \subseteq Q(A)} M(\mathbb{W}, A')$

Retrieval rule

$$\text{箇所 } Q; \quad Q(A) = \{A_{j_1}, A_{j_2}, \dots, A_{j_g}\}$$

$Q \Rightarrow$ bucket $M(\mathbb{W})$ を決定。

$$M(Q) = \bigcup_{A' \subseteq Q(A)} M(\mathbb{W}, A')$$

参 考 文 献

- Abraham, C.T., S.P. Ghosh and D.K. Ray-Chaudhuri, 1968: File organization schemes based on finite geometries, Information and Control, 12, 143-163.
- Buchholz, W., 1963: File organization and addressing, IBM system J., 2, 86-111.
- Carmichael, R.D., 1937: Introduction to the theory of groups of finite order, Dover pub., (1956), New York.
- Ghosh, S.P., 1969: Organization of records with unequal multiple-valued attributes and combinatorial queries of order 2, Information sciences, 1, 363-380.
- Ghosh, S.P. and C.T. Abraham, 1968: Application of finite geometry in file organization for records with multiple-valued attributes, IBM J. Res. Develop., 12, 180-187.
- Lancaster, F.W., 1968: Information retrieval and systems. Characteristics, Testing and Evaluation, John Wiley & Sons, New York.
- Ray-Chaudhuri, D.K., 1968: Combinatorial information retrieval systems for files. SIAM J. Appl. Math., 16, 973-992.

Bose, R.C., C.T. Abraham and S.P. Ghosh, 1969: File organization of records with multiple-valued attributes for multi-attribute queries, Chapel Hill, 1967. in "Combinatorial Mathematics and its applications". Univ. North Carolina Press.

Vajda, S., 1967a: Patterns and configuration in finite spaces, Charles Griffin, London.

_____, 1967b: The mathematics of experimental design (Incomplete block designs and latin squares), Charles Griffin, London.

Yamamoto, S., T. Fukuda and N. Hamada, 1966: On finite geometries and cyclically generated incomplete block designs, J. Sci. Hiroshima Univ. Ser. A-I, 30, 137-149.