社団法人 情報処理学会 研究報告
IPSJ SIG Technical Report

2006－AL－107 (9)
2006／7／3

完全三部グラフを用いたRAIDのアクセス順序

足立智子

東邦大学 理学部 情報科学科
〒274-8510 千葉県船橋市三山 2-2-1
E-mail: adachi@is.sci.toho-u.ac.jp

概要： RAID とはハードディスクドライブの処理速度と安全性を高める技術である．RAID のアクセスコストを低減するため，Cohen 等(2001)によって cluttered ordering という概念が導入された．これは，RAID の information disk と check disk を完全グラフの辺と頂点に対応させて information disk の順序付けを考察する，というものである．また，Mueller 等(2005)は，二次元の RAID を完全二部グラフに対応させることで 数理モデル化を行った．本稿では，完全三部グラフを用いた場合の RAID のアクセス順序について報告する．

Accessed Ordering for RAID utilizing the Complete Tripartite Graph

Tomoko Adachi

Toho University, Department of Information Sciences
2-2-1 Miyama, Funabashi, Chiba 274-8510 Japan
E-mail: adachi@is.sci.toho-u.ac.jp

Abstract: The desire to speed up secondary storage systems has lead to the development of redundant arrays of independent disks (RAID) which incorporate redundancy utilizing erasure code. To minimize the access cost in RAID, Cohen, Colbourn and Froncek (2001) introduced (d, f)-cluttered orderings of various set system for positive integers d, f. For the complete graph, Cohen et al. gave some cyclic constructions of cluttered orderings. Mueller, Adachi and Jimbo (2005) gave cluttered orderings for the complete bipartite graph. In this paper, we will investigated ordering for the complete tripartite graph.

1. はじめに

RAID とはハードディスクドライブ（以下，ディスクと呼ぶ）の処理速度と安全性を高める技術である．この技術は，ネットワーク構築やサーバなど，高い信頼性と性能が要求されるコンピュータでは不可欠な存在となっている．RAID は基本的に，ディスクの読み込み・書き込みを複数のディスクで並列に行うことにより処理速度を高め，記憶すべきデータを格納した information disk の他に ディスクの破損箇所の発見・修復のための check disk と呼

ばれる冗長性を持たせたディスクを用いることによって安全性を高めている．しかし，安全性を高めるためといって check disk を多くすると，追加のコストが増えてしまう．そこで，安全性と追加コストのバランスを考えることが重要になってくる．

RAID のアクセスコストを低減するため，Cohen 等(2001, 文献[6])によって cluttered ordering という概念が導入された．これは，RAID の information disk と check disk を完全グラフの辺と頂点に対応させて information disk の順序付けを考察する，というものである．また，Mueller 等(2005, 文献[8])は，二次元の RAID を完全二部グラフに対応させることで 数理モデル化を行った．

本稿では，これらの研究をさらに発展させ，完全三部グラフを用いた場合の RAID のアクセス順序について報告する．

2. グラフを用いた RAID の数理モデル化

information disk には保存したいデータを分割して格納し，check disk には information disk 内のデータが破損した場合に復旧するための冗長データを格納する．そして今，n 個の information disk と c 個の check disk があるとする．例えば，図 1 では information disk は 9 個，対応する check disk は 6 個となる．この RAID をグラフで表現することで数理モデル化を行う．

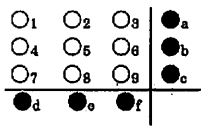

図 1. 二次元の RAID

RAID の check disk を頂点，information disk を辺とみなすことで，RAID をグラフで表現することができる．n(=m^2)個の information disk・c(=2m)個の check disk を持つ RAID は，上下に m 個ずつ計 c 個の頂点・n 本の辺を持つ完全二部グラフ $K_{m,m}$ に対応する．

先に図 1 で示した二次元の RAID は，図 2 のように完全二部グラフ $K_{3,3}$ で表現できる．

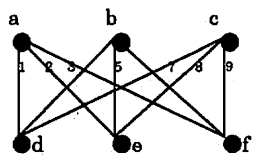

図 2. 図 1 に対応する完全二部グラフ

information disk 内のデータが破損した場合には，次のように check disk を用いて復旧する．n 個の information disk と c 個の check disk を持つ RAID に対し，これらの関係を 0, 1 を成分にもつ c×(n+c) 行列 H=[P | I] で表す．この行列 H はパリティ検査行列と呼ばれる（文献[2]）．ただし，I は単位行列であり，P は c 行 n 列の{0,1}-行列である．パリティ検査行列 H の最初の n 列は information disk に対応し，後半の c 列は chech disk に対応している．パリティ検査行列 H の 1 つの行に現れる information disk の内容の排他的論理和が計算され，その行に対応する check disk に書き込まれている．そして，1 つのディスクが壊れても復旧できるように，パリティ検査行列 H の列は mod 2 で線形独立になっている．パリティ検査行列の詳細については，文献[2]および[7]を紹介する．

3. Cluttered Ordering

あるグラフ G=(V, E)について，c=|V|，E={e_0, e_1,…, e_{n-1}} とし，nより小さい正の整数dを考える．また，{0, 1, …, n-1} 上の置換πに対して $V_i^{\pi,d}$ を「{$e_{\pi(i)}$, $e_{\pi(i+1)}$, …, $e_{\pi(i+d-1)}$}の各辺に含まれる点の集合」とする（インデックスは mod n で計算し，0≦i≦n-1 である）．

ここで，d本の辺を持つ部分グラフのアクセスコストを その部分グラフの頂点数で測る．するとアクセスコストの上限は $\max_i|V_i^{\pi,d}|$ で与えられる．このとき，$\max_i|V_i^{\pi,d}|=f$ となる辺の順序付けを(d,f)-cluttered ordering と呼ぶ．

完全二部グラフ $K_{3,3}$ の(3,4)-cluttered ordering を図3に示す．

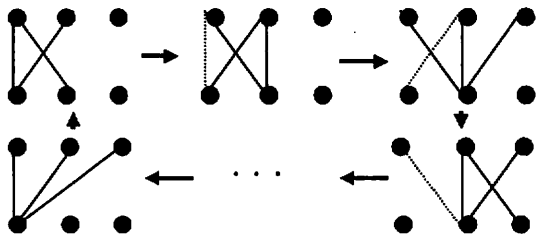

図3. $K_{3,3}$ の(3,4)-cluttered ordering

4. wrapped ⊿- labelling および (d, f)-movement

完全グラフにおける cluttered ordering の構成法は Cohen 等（文献[5]および[6]）によって与えられた．また，Steiner triple system での cluttered ordering の構成法は Cohen 等(文献[4])によって与えられた．二次元 RAID に自然に対応するような完全二部グラフにおける cluttered ordering の構成法は，Mueller 等（文献[8]）によって与えられた．Mueller 等は，このために，wrapped ⊿- labelling と (d,f)-movement という2つの概念を導入している．本稿では，これらの概念に基づき，完全三部グラフを用いた RAID のアクセス順序を調べている．そこで，本節では，wrapped ⊿- labelling および (d,f)-movement について説明する．

4.1. wrapped ⊿- labelling

二部グラフ H=(U, E)について U=V∪W，d=|E| とする．このとき，写像 δ：U →Z_d×Z_2 が 以下の二つの条件を満たすとき，この写像 δ のことを H の⊿- labelling と呼ぶ．但し，π_1：Z_d×Z_2→Z_d は第1成分の射影である．

1：δ(V)⊂ Z_d×{0}, δ(W)⊂ Z_d×{1} を満たす．

2：Z_d の各要素が { π_1(δ(v)) - π_1(δ(w)) | v∈V, w∈W, (v, w)∈E} に一つずつ存在する．

Δ-labelling δ は，二部グラフ H の頂点を Z_d の各要素でラベル付けしている．一般に，頂点のラベル付けは，グラフをその部分グラフに分解するツールとしてよく知られている．グラフの分解の詳細については，文献[1]を紹介する．

更に U の部分集合 X，Y に対し，Z_d×Z_2 において

$$\pi_1(\delta(Y)) = \pi_1(\delta(X)) + k \quad , \quad \gcd(k, d) = 1$$

を満たす整数 k が存在するとき，この⊿- labelling δ のことを H の wrapped ⊿- labelling と呼ぶ．

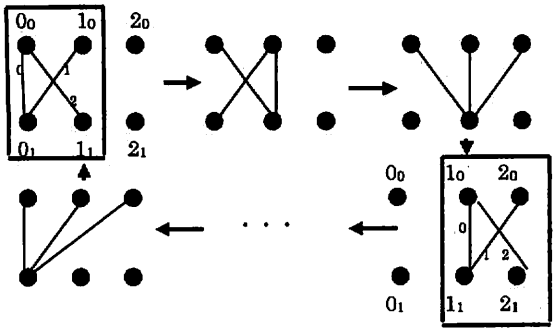

図 4. $K_{3,3}$ の wrapped ⊿· labelling

4.2. (d, f) - movement

次に (d,f̂) – movement について述べる.

同型な二つの二部グラフ H=(U, E), H'=(U', E') について

$U = V \cup W$, $U' = V' \cup W'$, $|V|=|V'|$, $|W|=|W'|$,

$E=\{e_0, e_1, \cdots, e_{d-1}\}$, $E'=\{e'_0, e'_1, \cdots, e'_{d-1}\}$

とする. また, {0, 1, ⋯, d-1}上の置換 π を用いて, 完全二部グラフ G を

$H_0:=H$, $H_i:=(U_i, E_i)$, $1 \le i \le d$

と, d+1 個の部分グラフに分割する. 但し

$E_i:=(E_{i-1} \setminus \{e_{\pi(i-1)}\}) \cup \{e'_{\pi(d+i-1)}\}$

U_i は E_i の各辺に含まれる頂点の集合

とする. このとき, $H_d = H'$ となり, $\max_{0 \le i \le d} |U_i| = f$ ならば, π を H から H' への(d,f̂)·movement と呼ぶ.

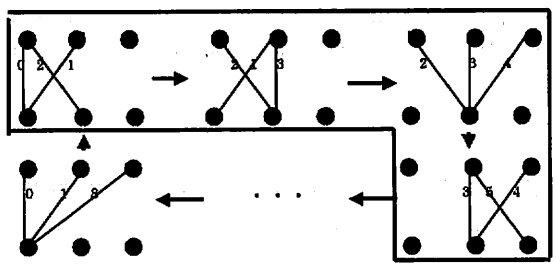

図 5. (3,4)·movement

wrapped ⊿· labelling と(d,f̂)·movement を用いることにより, 完全二部グラフにおける (d,f̂)·cluttered ordering の存在に関して, 次の定理が得られる.

定理 1 (文献[8]). 同型な二部グラフ H, H'に対し, wrapped ⊿· labelling と(d,f̂)·movement が存在するならば, 完全二部グラフ $K_{d,d}$ において(d,f̂)·cluttered ordering が存在する.

5. 完全二部グラフ $K_{3t,3t}$ の Ordering

Mueller 等 (文献[8])は, 特別な二部グラフ H(h;t)を定め, H(1;t)などの系列に対して, それぞれに同型な二部グラフへの wrapped ⊿· labelling の構成法を与えた. この考え方を基に, 次節では完全三部グラフの ordering を考える. そのため, 本節では H(1;t)の wrapped ⊿· labelling を構成し, 完全二部グラフ $K_{3t,3t}$ の ordering を構成する方法を紹介する.

5.1. 二部グラフ H(h;t)

本節では，自然数 h, t をパラメータとして，次で与えられる特別な二部グラフ H(h;t)=(U, E)について考察する．まず，頂点集合 U=V∪W を，次のように 各 h(t+1)個の頂点を持つ２つの部分集合 V, W に分ける．

$V:=\{v_i \mid 0\le i<h(t+1)\}$,
$W:=\{w_i \mid 0\le i<h(t+1)\}$,

よって頂点の個数は $|U|=2h(t+1)$ となる．

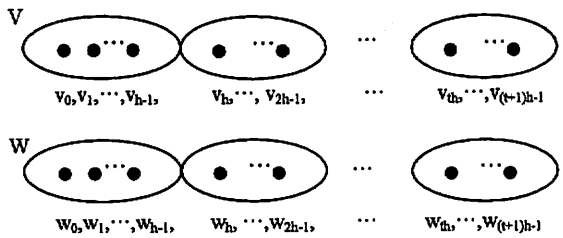

図 6.　二部グラフ H(h;t)の頂点集合

次に，辺集合を，次のように t 個の部分集合 $E_s(0\le s<t)$に分割する．更に，部分集合 E_s は，それぞれ，E_s', E_s'', E_s'''の３つの部分集合に分けられる．

$E_s' := \{\{v_i, w_j\} \mid s\times h\le i,j <s\times h+h\}$,
$E_s'' := \{\{v_i, w_{h+j}\} \mid s\times h\le j\le i<s\times h+h\}$,
$E_s''' := \{\{v_{h+i}, w_j\} \mid s\times h\le i\le j<s\times h+h\}$,
$E_s := E_s' \cup E_s'' \cup E_s'''$　　$0\le s<t$
$E := \cup_{0\le s\le t-1} E_s$

よって辺の本数は$|E| = t\times(h^2+h(h+1)/2+h(h+1)/2) = th(2h+1)$となる．下の図 7 は，h=3, t=1 の場合の二部グラフ H(3;1)の辺集合の分割を表している．

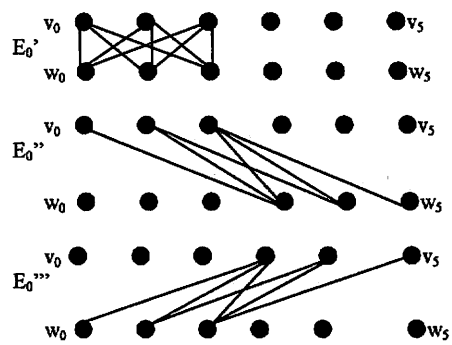

図 7.　二部グラフ H(3;1)の辺集合の分割

ここで，二部グラフ H(h;t)に関して次の定理により (d,f)-movement の存在が保証される．

定理 2 (文献[8]). 自然数 h, t（但し t≥2）に対し，d=h(2h+1) , f=4h とすれば，二部グラフ H(h;t)に関して E_0 から E_{t-1} への(d,f)-movement が存在する．

従って，H(h;t)に関して，同型な二部グラフへの wrapped ⊿-labelling の構成法を与えれば，定理 1 および定理 2 より，対応する完全二部グラフにおける cluttered ordering が与え

られる．このことより，次の定理が得られる．

定理 3（文献[8]）．自然数 h, t に対して，二部グラフ H(h;t)の任意の wrapped ⊿- labelling から，完全二部グラフ $K_{m,m}$の(d,f)-cluttered ordering が得られる．このときのパラメータの値は，m=th(2h+1), d=h(2h+1), f=4h となる．

5.2. H(1;t)の wrapped ⊿- labelling の構成

二部グラフ H(1;t)=(U, E)は，2(t+1)個の頂点と 3t 本の辺を持つ．自然数 t が与えられた時，頂点集合 U=V∪W 上の写像 δ : U→Z_{3t}×Z_2を次のように定める．

$$\delta(v_j) = \begin{cases} (jt, 0) & 0 \le j \le t-1 \text{ の時} \\ (t^2+1, 0) & j=t \text{ の時} \end{cases}$$

$$\delta(w_j) = \begin{cases} (j(t-1), 1) & 0 \le j \le t-1 \text{ の時} \\ (t^2+1, 1) & j=t \text{ の時} \end{cases}$$

但し，写像 δ による像の第 1 成分は，mod 3t で計算された整数である．

ここで，上で定めた δ の像の第 1 成分の差のリスト Δ(E)を計算すると以下のようになる．

$$\Delta(\cup_{0\le j\le t-1} E_j') = \{jt-j(t-1) \mid 0\le j\le t-1\} = (0, 1, 2, \cdots, t-1),$$
$$\Delta(\cup_{0\le j\le t-2} E_j'') = \{jt-(j+1)(t-1) \mid 0\le j\le t-2\} = (2t+1, 2t+2, \cdots, 3t-1),$$
$$\Delta(\cup_{0\le j\le t-2} E_j''') = \{(j+1)t-j(t-1) \mid 0\le j\le t-2\} = (t, t+1, \cdots, 2t-2),$$
$$\Delta(E_{t-1}'' \cup E_{t-1}''') = \{(t-1)t-(t^2+1), t^2+1-(t-1)^2\} = (2t-1, 2t),$$
$$\Delta(E) = \Delta(\cup_{0\le j\le t-1} E_j') \cup \Delta(\cup_{0\le j\le t-2} E_j'') \cup \Delta(\cup_{0\le j\le t-2} E_j''') \cup \Delta(E_{t-1}'' \cup E_{t-1}''') = (0, 1, 2, \cdots, 3t-1)$$

以上のように，Z_{3t}のすべての要素は，Δ(E)にちょうど 1 度ずつ現れることがわかる．

また，任意の t に関して，$k=t^2+1$ とおけば，k は 3t と互いに素である．よって明らかに上で定めた写像 δ は，wrapped ⊿-labelling の条件を満たしている．従って，この写像 δ を H(1;t)の wrapped ⊿- labelling と定めれば，定理 3 を適用することにより，次の結果が得られる．

定理 4（文献[8]）．任意の自然数 t に対し，パラメータの値が d=3, f=4 となるような完全二部グラフ $K_{3t,3t}$の(d,f)-cluttered ordering が存在する．

定理 5（文献[8]）．任意の自然数 t に対し，パラメータの値が d=3s+r, f=2(s+1)+r (s>0, r=0, 1, 2)となるような完全二部グラフ $K_{3t,3t}$の(d,f)-cluttered ordering が存在する．

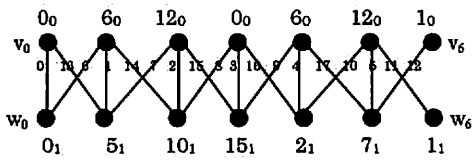

図 8. H(1;6), |E|=18, |V|=12, k=1 の wrapped ⊿-labelling

6. 完全三部グラフの Ordering

本節では，完全三部グラフを用いた RAID のアクセス順序を調べる．任意の自然数 m に対し，完全三部グラフ $K_{m,m,m}$ は 3m 個の頂点，m^2本の辺を持つ．頂点集合は $Z_m \times Z_3$ に対応し，一辺に接続する二頂点は $Z_m \times Z_3$ の第二成分が異なる．

6個の頂点と m本の辺を持つ三部グラフ H=(U, E)をとると，$U=V_0 \cup V_1 \cup V_2$, $V_0:=\{u_0, v_0\}$, $V_1:=\{u_1, v_1\}$, $V_2:=\{u_2, v_2\}$, $m=|E|$ と表せる．本稿では，$m=|E|=9$ の場合について調べる．頂点集合 $U= V_0 \cup V_1 \cup V_2$ 上の写像 $\delta : U \rightarrow Z_9 \times Z_3$ を次のように定める．

$$\delta(u_0) = (0, 0),\quad \delta(v_0) = (a, 0),$$
$$\delta(u_1) = (0, 1),\quad \delta(v_1) = (a, 1),$$
$$\delta(u_2) = (b, 2),\quad \delta(v_2) = (a+b, 2)$$

但し，a, b は Z_9 の元であり，写像 δ による像の第1成分は，mod 9 で計算された整数である．ここで，上で定めた δ の像の第1成分の差のリスト Δ(E)を計算すると

$$\Delta(E) = \{\pm a, \pm b, \pm(a+b), \pm(a-b)\}$$

となる．Z_9 のすべての要素が Δ(E)にちょうど1度ずつ現れる a,b の値は

$$(a, b)=(1, 3), (2, 3), (3, 1), (3, 2), (3, 4), (4, 3)$$

に限る．k=a が gcd(k, 9)=1 という条件を満たすように取ると，a=3 の場合は不適となり，a=1, 2, 4 となる．よって，

$$(a, b)=(1, 3), (2, 3), (4, 3)$$

のとき，上で定めた写像 δ は，wrapped ⊿-labelling の条件を満たしている．

次に，(d,f)-movement の存在性を調べる．(a, b) = (1, 3) の場合，図9のように，(9, 8)-movement が存在する．(a, b) = (2, 3) の場合，図10のように，(9, 8)- movement が存在する．(a, b) = (4, 3) の場合，図11のように，d=9 のときに(d, f)- movement が存在するとしても，f は9以上となる．

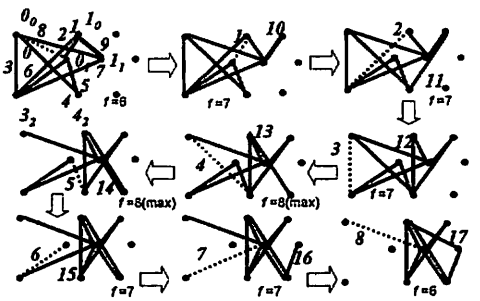

図9. (a, b)=(1,3)の場合の(d,f)-movement

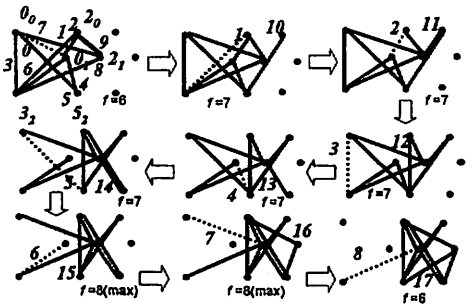

図10. (a, b)=(2,3)の場合の(d,f)-movement

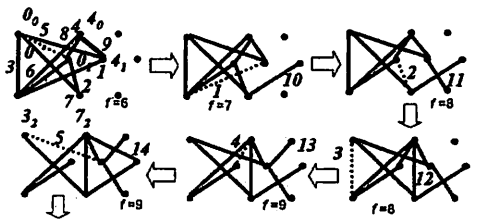

図11. (a, b)=(4,3)の場合の(d,f)-movement

したがって，(a, b) = (1, 3) , (2, 3)の場合に，wrapped ⊿-labelling と(9, 8)- movement が存在するので，次の結果が得られる．

定理 6. パラメータの値が d=9, f=8 となるような完全三部グラフ $K_{3,3,3}$ の(d,f)-cluttered ordering が存在する．

では，(a, b) = (1, 3) の場合と(a, b) = (2, 3) の場合では，どちらがより良いといえるであろうか．双方の(9, 8)- movement の図を比較してみよう．(a, b) = (1, 3) の場合は，図 9 のように，移動する 9 つのグラフのうち，2 つのグラフにおいて f が最大値 8 を取り，5 つのグラフにおいて f が 2 番目に大きな値 7 を取り，2 つのグラフにおいて f が最小値 6 を取る．他方，(a, b) = (2, 3) の場合は，図 10 のように，移動する 9 つのグラフのうち 2 つのグラフにおいて f が最大値 8 を取り，5 つのグラフにおいて f が 2 番目に大きな値 7 を取り，2 つのグラフにおいて f が最小値 6 を取る．どちらも f の値に対するグラフの数は同じなので，同程度の良さであるといえるであろう．

任意の自然数 t について，上のグラフを t 個つなげたグラフを考えると，同様のことがいえるであろう．したがって，任意の自然数 t に対し，パラメータの値が d=9, f=8 となるような完全三部グラフ $K_{3t,3t,3t}$ の(d,f)-cluttered ordering が存在する．

7. おわりに

ある完全三部グラフの系列について，wrapped ⊿-labelling を構成し，RAID のアクセス順序を探索した．

他の完全三部グラフや 一般に完全 n 部グラフを用いて場合について RAID の最適なアクセス順序を探索するのが今後の課題である．

文献

[1] J. Bosak, *Decompositions of Graphs*, Kluwer Academic Puclishers, Dordrecht, 1990.

[2] Y. Chee, C. Colbourn, and A. Ling, Asymptotically optimal erasure-resilient codes for large disk arrays, *Discrete Applied Mathematics*, vol.102, Issues 1－2, pp.3－36, 2000.

[3] P. Chen, E. Lee, G. Gibson, R. Katz and D. Patterson, RAID: High－performance, reliable secondary storage, *ACM Computing Surveys*, vol.26, pp.145－185, 1994.

[4] M. Cohen and C. Colbourn, Optimal and Pessimal Orderings of Steiner Triple Systems in Disk Arrays, *Theoretical Computer Science*, vol.297, Issues 1－3, pp.103－117, 2003.

[5] M. Cohen and C. Colbourn, Ladder orderings of pairs and RAID performance, *Discrete Applied Mathematics*, vol.138, no.29, pp.35－46, 2004.

[6] M. Cohen, C. Colbourn, and D. Froncek, Cluttered orderings for the complete graph, *COCOON 2001*: Lect. Notes Comp. Sci. 2108, pp.420－431, Springer Verlag, 2001.

[7] L. Hellerstein, G. Gibson, R. Karp, R. Katz and D. Patterson, Coding techniques for handling failures in large disk arrays, *Algorithmica*, vol.12, pp.182－208, 1994.

[8] M. Mueller, T. Adachi, and M.Jimbo, Cluttered orderings for the Complete Bipartite Graph, *Discrete Applied Mathematics*, in press. *Math*, vol.152, Issues 1－3, pp. 213-228, 2005.