Using graphs for some discrete tomography problems

M.-C. Costa* D. de Werra† C.Picouleau‡

Abstract

Given a rectangular array where entries represent the pixels of a digitalized image, we consider the problem of reconstructing an image from the number of occurrences of each color in every column and in every row. The complexity of this problem is still open when there are just three colors in the image. We study some special cases where the number of occurences of each color is limited to small values. Formulations in terms of edge coloring in graphs and as timetabling problems are used; complexity results are derived from the model.

Keywords : discrete tomography, complete bipartite graphs, perfect matchings, edge coloring, timetabling

1 Introduction

In discrete tomography (see [7] for an overview of theory and applications of this field), the image reconstruction problem is important since its solution is required for developing efficient procedures in image processing, data bases, crystallography, statistics, data compressing, ... It can be formulated as follows : an image of $(m \times n)$ pixels of p different colors has to be reconstructed. For convenience we consider that there is in addition a color $p + 1$ which is the ground color. We are given the number $a(i, s)$ of pixels of each color s in each row i and also the number $\alpha(j, s)$ of pixels of each color s in each column j. Can one reconstruct an image from the $a(i, s)$ and $\alpha(j, s)$? First of all, we are concerned with the consistency problem : given the $a(i, s)$ and $\alpha(j, s)$, does there exists an image corresponding to the values $a(i, s)$ and $\alpha(j, s)$? In the case of positive answer, one has to reconstruct

*CEDRIC CNAM, Paris (France)
†Ecole Polytechnique Fédérale de Lausanne (Switzerland)
‡CEDRIC CNAM, Paris (France)

efficiently such an image. The uniqueness of the solution is essential with respect to tomography, however we shall not deal on this specific issue.

The consistency and reconstruction problems have been extensively studied (see for instance [2],[5],[7],[8]). The main fact resulting from these studies is that the complexity of this reconstruction problem is still open for $p = 2$. Actually, determining the complexity of this problem is a challenge in discrete tomography.

The purpose of this note is to explore the boundary between easy and difficult problems. Since difficult problems already arise in very simple situations, the solvable cases which we shall describe here are very special.

Here we shall concentrate on the special case where each color s occurs at most once in each row and in each column (except of course the ground color). In section 2, we show that this problem is strongly NP-complete even with $n = 3$ columns by exploiting the analogy with some class-teacher timetabling problems. In the following sections we are deliberately using the language of graphs because many basic results of graph theory provide simple solutions to some of the problems. We will essentially be dealing with complete bipartite graphs. In section 3, we will in particular determine the maximum number of matchings which can be removed in such a way that the remaining graph still has a perfect matching. This will be exploited in the construction process of a solution of our problems. In sections 4 and 5, we will derive some special cases solvable in polynomial time from elementary graph theoretical properties.

For basic definitions about complexity, the reader is referred to [6] and for graph theoretical terms to [1].

2 Image reconstruction and timetabling

Let us first formulate the image reconstruction problem $RP(m, n, p)$ as follows : we are given an $(m \times n)$ array A together with a set of $p+1$ colors. In addition for each row i (resp. each column j of A) and each color s, $a(i, s)$ (resp. $\alpha(j, s)$) will be the number of occurrences of color s in row i (resp. column j). In $RP(m, n, p)$ we want to know whether there exists or not an assignment of one color to each entry of A which satisfies the requirements for each row i, for each column j and each color s.

For a solution to exist we must necessarily have

$$\begin{array}{rcll} \sum_{s=1}^{p+1} a(i,s) & = & n & (i = 1, \ldots, m) \\ \sum_{s=1}^{p+1} \alpha(j,s) & = & m & (j = 1, \ldots, n) \\ \sum_{i=1}^{m} a(i,s) & = & \sum_{j=1}^{n} \alpha(j,s) & (s = 1, \ldots, p+1) \end{array}$$

These conditions are necessary but not sufficient for the existence of a solution to $RP(m,n,p)$.

$RP(m,n,p)$ is polynomially solvable for $p+1=2$ (see [8]) and the associated consistency problem is strongly NP-complete for $p+1=4$ (see [2]). The complexity is unknown for $p+1=3$.

In this paper we shall derive some complexity results for special cases of $RP(m,n,p)$ and we will exhibit some polynomially solvable cases. For this purpose we will exploit the analogy of $RP(m,n,p)$ with some special types of class-teacher timetabling problems.

The problem $TIMETABLE(m,p,n)$ is defined as follows ([4]): we are given a collection of m classes $c_1,\ldots,c_m$ and a collection of p teachers $t_1,\ldots,t_p$ together with the $m\times p$ requirement matrix R whose entry r_{is} is the number of one-hour lectures which teacher t_s has to give to class c_i. In addition we have a set $H=\{1,\ldots,n\}$ of periods (hours) and for each teacher t_s a set $T_s\subseteq H$ which represents the periods where t_s is available.

Without loss of generality we also assume that the number $|T_s|$ of periods where teacher t_s is available is equal to the number $\sum_{i=1}^{m} r_{is}$ of lectures which t_s has to give.

In addition, for all classes c_i we assume that they are available at all periods in H.

$TIMETABLE(m,p,n)$ can be reformulated in the following way : we associate an $(m\times n)$ array A and define for each teacher t_s and each class c_i, $a(i,s)=r_{is}$. Similarly for each teacher t_s and each period j we define $\alpha(j,s)=1$ if $j\in T_s$ or 0 else. A feasible timetable is an assignment of each lecture (c_i,t_s) to some period $j\in T_s$ in such a way that no teacher (and no class) is involved in two lectures simultaneously; assigning a lecture (c_i,t_s) to some period $j\in T_s$ consists in setting $A_{ij}=s$; this amounts to giving color s to entry (i,j) of array A. This will be possible only if $\alpha(j,s)=1$ (t_s available at period j) and if $a(i,s)=r_{is}\geq 1$. Color s will occur exactly once in column j (t_s is involved in one lecture in period $j\in T_s$).

Furthermore since each class c_i is involved in at most one lecture at a time (with some teacher, say t_s) we will assign at most one color s to each entry of A. In addition the number of occurrences of t_s in row i will be $a(i,s)=r_{is}$.

The ground color of $RP(m,n,p)$ corresponds to idle periods for classes; more precisely, $a(i,p+1)=n-\sum_{s=1}^{p} a(i,s)$, $i=(1,..,m)$, is the number of periods where class i has no lecture and $\alpha(j,p+1)=m-\sum_{s=1}^{p}\alpha(j,s)$, $j=(1,..,n)$, is the number of classes having no lecture at period j.

Then we have a one-to-one correspondence between timetables and solutions of $RP(m,n,p)$ where the values $\alpha(j,s)$ are restricted to 0 or 1 ($s\leq p$).

Since timetabling problems have been rather extensively studied, we may exploit the analogy with $RP(m, n, p)$ to derive some complexity results.

A color s will be called *unary* if in $RP(m, n, p)$ we have $a(i, s) \in \{0, 1\}$and $\alpha(j, s) \in \{0, 1\}$ for each row i and for each column j.

$RPU(m, n, p)$ will be the problem $RP(m, n, p)$ where colors $1, \dots, p$ are unary (while color $p + 1$ may not be unary).

Proposition 1 *$RPU(m, n, p)$ is strongly NP-complete even if $n = 3$*

Proof : It is known that $TIMETABLE(m, p, n)$ is NP-complete even if $n = |H| = 3$ and $r_{is} \in \{0, 1\}$ for each class c_i and each teacher t_s (see [4]). From the correspondence between timetables and solutions of $RP(m, n, p)$ we deduce that $RP(m, n = 3, p)$ is strongly NP-complete. Since $a(i, s) = r_{is} \in \{0, 1\}$ the result follows for $RPU(m, n = 3, p)$. ∎

So $RPU(m, n = 3, p)$ is difficult even if array A has three columns. Observe also that each color s occurs at most three times in A. Now, we may examine the case where each one of the p unary colors occurs at most twice: the number $n(s)$ of occurrences of each color s satisfies $n(s) \leq 2$. In fact, we can consider a slightly more general case than the one where $n = 3$.

Proposition 2 *$RP(m, n, p)$ with*

1. $\sum_{i=1}^{m} a(i, s) \leq 2 \quad s = 1, .., p$
2. $0 \leq \alpha(j, s) \leq 1 \quad j = 1, .., m,\ s = 1, .., p$

can be solved in polynomial time.

Proof : This situation corresponds to a timetable problem where each teacher has to give at most two lectures. As before we assume that each teacher t_s is available during $|T_s| = \sum_i r_{is} = \sum_i a(i, s)$ periods. Here we allow $a(i, s) = r_{is}$ to be 2, which means that teacher t_s gives two lectures at the same class c_i. Such a teacher has thus a unique timetable. It is also the case for teachers t_s with $\sum_{i=1}^{m} a(i, s) = 1$. We start by constructing the timetable of all such teachers t_s with $a(i, s) = 2$ or with $\sum_{i=1}^{m} a(i, s) = 1$. Then, if no conflict has occurred, we are facing a special timetable problem with "binary teachers" which can be solved in polynomial time (see [4]) by reduction to a $2 - SAT$ problem. ∎

Remark

One may observe that in $TIMETABLE$ when the requirement matrix $R = (r_{is})$ satisfies $r_{is} \in \{0,1\}$, we may interchange the role of the set H of periods and the set C of classes.

It is easy to see that there is a feasible timetable for the first problem if and only if there is one for the second one.

Hence we could observe (see [4]) :

$TIMETABLE(m,p,n)$ is NP-complete even if $m = 3$ and $r_{is} \in \{0,1\}$ for each class c_i and each teacher t_s.

It is interesting to observe that this exchange is generally not possible when $r_{is} \notin \{0,1\}$.

We shall next discuss some solvable cases of RPU by exploiting the analogy with timetabling and by using an edge coloring formulation.

3 Graph theoretical formulation

From now on, we will use graph theoretical formulations in order to derive other properties. We will have a complete bipartite graph $G = K_{m,n}$ based on two sets of nodes R, S with sizes m and n. Each edge $[i,j]$ of $K_{m,n}$ corresponds to entry (i,j) in row i and column j of the $(m \times n)$ array A. We will assume $m \leq n$ unless otherwise stated.

We can interpret the image reconstruction problem in array A as follows: the entries of color s in A correspond to a subset B_s of edges (a partial subgraph of $K_{m,n}$) such that B_s has $a(i,s)$ edges adjacent to node i of R and $\alpha(j,s)$ edges adjacent to node j of S. We have to find a partition $B_1, B_2, .., B_{p+1}$ of the edge set of $K_{m,n}$ where each B_s satisfies the above degree requirements.

Clearly, having obtained such a partition of the edge set of $K_{m,n}$ we can trivially derive the solution of the image reconstruction problem.

It is known that there exists an edge coloring of $K_{m,n}$ (no two adjacent edges may have the same color) using n colors; $n = \Delta(G)$ is the maximum degree of G. This is a consequence of the König theorem (see [1]).

Each color class is a matching (a set of nonadjacent edges) saturating all m nodes in the first set of nodes. If $m = n$ each color class is a *perfect* matching (i.e. a matching saturating all nodes). So $K_{n,n}$ contains n disjoint perfect matchings; this property will be used later. One notices also that in such a case these matchings can be constructed in polynomial time (see [1]).

We also recall the König-Hall theorem which states that in a bipartite graph $G = (R, S, E)$ built on two subsets R, S of nodes, there exists a matching saturating all nodes of R if and only if, for any q, every subset of q nodes of R is linked to at least q nodes of S.

This result will be used for instance in designing a sequential edge-coloring procedure.

In the case where all colors (except possibly the last one) are unary, the entries of array A corresponding to a fixed color are represented by a matching M in $G = K_{m,n}$.

If we try to apply a sequential algorithm which constructs the color classes one after other, we have at each step the following situation :

we are given a complete bipartite graph K^s built on the sets R^s, S^s of nodes where color s must occur on exactly one edge at each node. In the graph $|R^s| = |S^s|$ (otherwise no solution exists) and there are $s-1$ partial matchings $M_1, M_2, \ldots, M_{s-1}$ corresponding to the edges already colored with colors $1, 2, \ldots, s-1$ respectively.

The following question arises naturally : when is it possible to find in K^s a perfect matching M_s such that $M_s \cap M_w = \emptyset$ for $w = 1, \ldots, s-1$?

In case such an M_s can be found, the matchings $M_1, \ldots, M_{s-1}$ can be kept as such.

Before answering this question, we shall give a more general statement which will suggest a sequential coloring procedure.

The number $n(s)$ of occurrences of any color s in $RPU(m, n, p)$ is $n(s) = \sum_{i=1}^{m} a(i, s) = \sum_{j=1}^{n} \alpha(j, s)$. Clearly in K^s we must have $|R^s| = |S^s| = n(s)$.

Proposition 3 *Consider a problem $RPU(m, n, p)$ and assume that a solution has been found for colors $1, 2, \ldots, s-1$ with $s-1 < p$. Then a solution can be found for colors $1, 2, \ldots, s$ (without modifying colors $1, 2, \ldots, s-1$) if $n(s) \geq 2s-2$.*

Proof : Let G^s be the bipartite graph built on node sets R^s, S^s associated with rows i (resp. columns j) for which $a(i, s) = 1$ (resp. $\alpha(j, s) = 1$).

G^s is obtained by removing the edges of matchings $M_1, M_2, \ldots, M_{s-1}$ corresponding to the first $s-1$ colors. We have the following facts :

In G^s every node has at most $s-1$ non neighbours in the other set; so it has at least $n(s) - (s-1) \geq s-1$ neighbours. Consider a subset B of nodes in R^s. If $|B| \leq s-1$, then the set $N(B)$ of neighbours of B satisfies $|N(B)| \geq s-1 \geq |B|$. If $|B| \geq s$, then every node in S^s has a neighbour in B so $N(B) = |S^s| = n(s) \geq |B|$.

Since for any subset B of R^s, we have $|N(B)| \geq |B|$, it follows from the theorem of König-Hall that G^s has a perfect matching M_s. It defines the assignment of color s. ∎

Proposition 4 *The largest number of arbitrary disjoint matchings M_i whose edges can be consecutively removed from a complete bipartite graph $K_{n,n}$ in such a way that the remaining graph G' still has a perfect matching is $q = \lceil \frac{n-1}{2} \rceil$.*

Proof : One verifies immediately that for $t = \lceil \frac{n-1}{2} \rceil + 1 = \lceil \frac{n+1}{2} \rceil$ there exists disjoint matchings $M_1, \ldots, M_t$ such that G' has no perfect matching. Take for $M_1 \cup \ldots \cup M_t$ the complete subgraph $K_{t,t}$ of $K_{n,n}$ constructed on the first t nodes of the left set and on the first t nodes of the right set.

Then in G' the first $t = \lceil \frac{n+1}{2} \rceil$ nodes of the left set are linked only to the last $\lfloor \frac{n-1}{2} \rfloor$ nodes of the right set. Since $t > \lfloor \frac{n-1}{2} \rfloor$ it follows from the theorem of König-Hall that there is no perfect matching in G'.

On the other hand, it follows from Proposition 3 that we may remove $s - 1 \leq \lceil \frac{n-1}{2} \rceil$ matchings : this is possible since in G' we have $n = n(s) \geq 2s - 1$ (n odd) or $n = n(s) \geq 2s - 2$ (n even). ∎

Remark

An 'inverse' problem consisting of finding in a bicolored complete bipartite graph a perfect matching M containing a given number of edges of each color has been handled in [9]; it can be solved in polynomial time. ∎

We can also mention the following consequence:

Corollary 1 *Assume that in $RPU(m, n, p)$ we have $n(1) \geq 1, n(2) \geq 2, \ldots, n(s) \geq 2s - 2, \ldots, n(p) \geq 2p - 2$. Then $RPU(m, n, p)$ has a solution. It can be obtained by constructing consecutively p perfect matchings of sizes $n(1), \ldots, n(p)$ in the corresponding graphs.*

We do not know whether for a fixed number of unary colors the problem $RPU(m, n, p = k)$ is solvable in polynomial time; as mentioned earlier for $p = 3$ arbitrary colors (non unary), $RP(m, n, p = 3)$ is however NP-complete.

We shall examine the case of $p \leq 3$ unary colors in section 4.

One should finally observe that Corollary 1 has an interpretation in terms of timetabling which is apparently original.

Assume we have a class-teacher timetabling problem where each teacher t_s is available during a set $T_s \subseteq H$ of periods and $|T_s| = \sum_i r_{is} = n(s)$ with $r_{is} \in \{0, 1\}$. If the number $n(s)$ of lectures that teacher t_s has to give satisfies $n(s) \geq 2s - 2$ for each teacher t_s, then there always exists a feasible timetable.

4 Case of 2 or 3 unary colors

In this section, we deal with the case where there are at most 3 unary colors ($a(i,s) \leq 1, \alpha(i,s) \leq 1$) in addition to the ground color. Notice that a unary color may occur up to $\min\{m,n\} = m$ times in the array.

Proposition 5 *$RPU(m,n,p=2)$ is solvable if and only if we do not have $n(1) = n(2) = 1$ and a pair (i,j) with $a(i,1) = a(i,2) = \alpha(j,1) = \alpha(j,2) = 1$.*

Proof : Clearly if we are in the bad case there is no solution since entry $[i,j]$ should receive at the same time color 1 and color 2.

Let us see that in all other cases a solution can be found : if $n(1) = n(2) = 1$ and $a(i,1) \neq a(i,2)$ or $\alpha(j,1) \neq \alpha(j,2)$ for some pair i,j, then obviously colors $1,2$ will have to be assigned to different entries and this is possible.

Assume now that $1 \leq n(1) \leq n(2)$ (with $n(2) \geq 2$). Then there is a solution from Corollary 1. ∎

We consider now the problem $RPU(m,n,p=3)$ and assume $n(1) \leq n(2) \leq n(3)$.

If $n(2) \geq 2$ and $n(3) \geq 4$, there is a solution from Corollary 1.

If $n(2) \leq 1$, then M_1 and M_2 are uniquely determined and it is immediate to verify whether a solution exists.

The only remaining case is $2 \leq n(2) \leq n(3) \leq 3$. When $n(3) = 3$, we consider the graph K^3 built on node sets R^3, S^3 (they are the nodes which must be adjacent to one edge of color 3). Consider $E^* = (M_1 \cup M_2) \cap E(K^3)$; if E^* contains a square, then in the remaining graph $K^3 - (M_1 \cup M_2)$ it will not be possible to construct a perfect matching M_3. But unless $n(1) = n(2) = 2$ a square in E^* can be removed by choosing a different matching M_2'. If there is no square in E^*, a perfect matching can always be constructed.

Finally when $n(3) = 2$, we also have $n(2) = 2$ and it is immediate to determine whether a solution exists.

Hence we have shown :

Proposition 6 *$RPU(m,n,p=3)$ can be solved in polynomial time.*

5 Some non unary cases

We shall first examine here the case where each color $s \leq p$ satisfies $a(i,s), \alpha(j,s) \in \{0,2\}$ for each row i and each column j. Such a color will be called binary.

As before the number of occurrences of color s is $n(s) = \sum_{i=1}^{m} a(i,s) = \sum_{j=1}^{n} \alpha(j,s)$. We shall assume in addition that $n(1) \leq n(2) \leq \cdots \leq n(p)$. The corresponding problem will be denoted by $RPB(m,n,p)$.

Let us examine the case when there are $p = 2$ such colors.

Before stating the results on this case, we need the following lemma. We recall the reader that a $k-$factor in a graph G is a subset F of edges such that each node of G is adjacent to k edges of F. So a $1-$factor is a perfect matching and a $2-$factor is a union of node disjoint cycles which meets all nodes.

Lemma 1 *Let C be a cycle in $K_{n,n}$ (with $n \geq 3$), then the smallest k for which there is a $k-$factor containing C is $k = 2$ if $|C| \neq 2(n-1)$ or $k = 3$ else.*

Proof : Here $|C|$ is the number of edges (or equivalently of nodes) in C; if $|C| = 2(n-1)$ it means that C visits all but two nodes of $K_{n,n}$. In such a case clearly one cannot find a $2-$factor containing C. Let us show how to find a $3-$factor F which contains all edges of C; let $R = 1, 2, .., n$ and $S = \bar{1}, \bar{2}, .., \bar{n}$ be the left set and the right set of nodes of $K_{n,n}$; if $n = 3$, we take all edges of $K_{3,3}$ to get F. If $n \geq 4$, we assume that C uses edges $[1, \bar{1}], [2, \bar{2}], \ldots, [n-1, \overline{n-1}], [1, \overline{n-1}], [2, \bar{1}], [3, \bar{2}], \ldots, [n-1, \overline{n-2}]$. There exists a perfect matching $M = \{[1, \bar{2}], [2, \bar{3}], \ldots, [n-1, \bar{n}], [n, \bar{1}]\}$ disjoint from C. We replace in M edge $[1, \bar{2}]$ by edges $[n, \bar{2}], [1, \bar{n}]$ and $[n, \bar{n}]$. These edges form a $3-$factor with the edges of C.

Assume now that $|C| \neq 2(n-1)$; clearly, if $|C| = 2n$, it is a $2-$factor, so $F = C$. If $|C| = 2(n-q)$ with $q \geq 2$, there are q nodes in R and q nodes in S which are not visited by C; we construct a cycle C' on these $2q$ nodes. Then, $F = C \cup C'$ is the required $2-$factor. ∎

We notice that, in order to have solutions of $RPB(m,n,p=2)$ we must have $n(1) \geq 4$. We will consider the complete bipartite graph K^s built on the sets R^s (resp. S^s) of nodes corresponding to rows i with $a(i,s) = 2$ (resp. to columns j with $\alpha(j,s) = 2$) for $s = 1, 2$. Clearly $|R^s| = |S^s|$. We shall assume w.l.o.g. $\left|R^1 \cap R^2\right| \geq \left|S^1 \cap S^2\right|$.

Proposition 7 *If we are in one of the following cases:*

1. $n(1) = n(2) = 4 \;:\; R^1 \cap R^2 \neq \emptyset,\; S^1 \cap S^2 \neq \emptyset$
2. $4 \leq n(1) \leq n(2) = 6 \;:\; R^1 \subseteq R^2,\; S^1 \cap S^2 \neq \emptyset$
3. $n(1) = 6, n(2) = 8 \;:\; R^1 \subseteq R^2,\; S^1 \subseteq S^2$

then $RPB(m,n,p=2)$ has no solution.

Proof : For cases 1 and 2 one takes in K^1 a $2-$factor M_1 and one verifies immediately that there are not enough edges in K^2 for constructing a $2-$factor corresponding to color 2.

In case 3, consider any $2-$factor built in K^1; it consists of a cycle C of length 6. From the assumption, C is a cycle of K^2 which meets all but two nodes of the graph. From lemma 1, there is no $2-$factor of K^2 containing C. Since the assignment of color 2 will correspond to a $2-$factor in K^2, a necessary and sufficient condition of existence of a solution for $RPB(m,n,p=2)$ in this case is that the edge set of K^2 be decomposable into two disjoint $2-$factors F, F' with $F \supseteq C$. Since no such pair F, F' can be found, the problem has no solution. ∎

Proposition 8 *$RPB(m,n,p=2)$ has a solution if and only if we are not in the cases of proposition 7. It can be found in polynomial time.*

Proof : It is obvious that if $R^1 \cap R^2 = \emptyset$ or $S^1 \cap S^2 = \emptyset$ a solution exists since the edge sets of K^1 and K^2 are disjoint. From now we assume that $R^1 \cap R^2 \neq \emptyset$ and $S^1 \cap S^2 \neq \emptyset$; we also suppose that $n(2) \geq 6$ (if not we are in the case 1 of Proposition 7).

We show that if it exists a solution for an instance I of $RPB(m,n,p=2)$ for which $|R^1 \cap R^2| = k$, then there is a solution for every instance $\bar{I}$ of $RPB(m,n,p=2)$ defined as follows : $R^2 = \bar{R}^2, S^2 = \bar{S}^2, S^1 = \bar{S}^1$ and $|R^1| = |\bar{R}^1|, |\bar{R}^1 \cap R^2| = k-1$ and $|\bar{R}^1 \triangle R^1| = 2$. Let r be the node such that $r \in \bar{R}^1, r \notin R^1$ and let x be the node such that $x \notin \bar{R}^1, x \in R^1$. Let $[x,a]$ and $[x,b]$ the two edges colored with the first color in the solution of I : replacing $[x,a]$ and $[x,b]$ with $[r,a]$ and $[r,b]$ we obtain a $2-$factor which gives a solution for $\bar{I}$.

So it remains only to consider the cases where $R^1 \subseteq R^2$ and $S^1 \subseteq S^2$. From Lemma 1 we have a solution when $n(1)+4 \leq n(2)$. When $n(1)+2 = n(2), n(1) \geq 8$, the $2-$factor of K^1 consists in a collection of C_4 (cycle of length four) and at most one C_6 (cycle of length six). We have to distinguish 2 cases :

a) $n(1) \equiv 0 (\mod 4)$, then the $2-$factor of K^1 is a collection of C_4; the construction of the $2-$factor of K^2 is given in Figure 1.

b) $n(1) \equiv 2 (\mod 4)$, then the $2-$factor of K^1 is a collection of C_4 with an additional C_6; the construction of the $2-$factor of K^2 is given in Figure 2. ∎

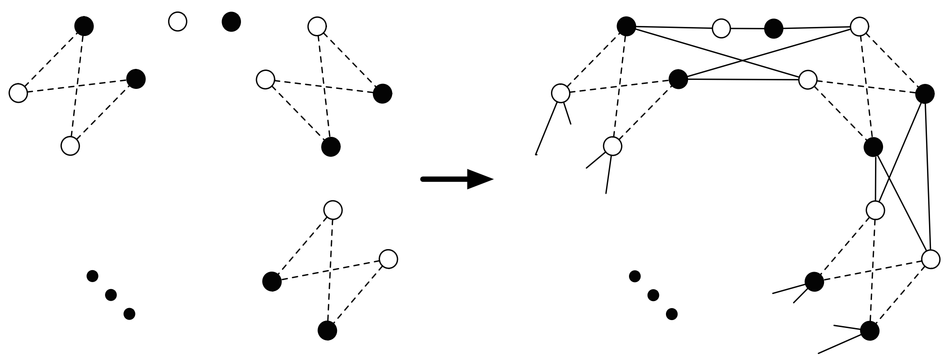

Figure 1 : *the* $2-$*factor of* K^1 *is a collection of* C_4

Corollary 2 *$RPB(m, n, p = 2)$ has a solution if $n(2) \geq 10$.*

For $RPB(m, n, p)$ we can derive sufficient condition for solutions to exist from proposition 4.

Proposition 9 *$RPB(m, n, p)$ has a solution if the number of occurrences of each color s satisfies $n(s) \geq 8s - 4,\ s = 1, \ldots, p$. Moreover, such a solution can be constructed by sequentially assigning colors $1, \ldots, p$.*

Proof : This follows from Proposition 3 by observing that $RPB(m, n, p)$ corresponds to a problem $RPU(m, n, 2p)$ defined as follows; for every color s we do the following : for every row i of RPB with $a(i, s) = 2$ we set $a'(i, s) = 1 = a'(i, p + s)$ in RPU and for every column j of RPB with $\alpha(j, s) = 2$ we set $\alpha'(j, s) = 1 = \alpha'(j, p + s)$. Clearly every solution of RPU will produce a solution of RPB by identifying colors s and $s + p$. Since the condition of RPB implies that the RPB satisfies the condition of Proposition 3, the result follows. ∎

Remark

The above result is in some sense best possible. It is indeed worth observing that such a construction would not work for $3p$ colors, i.e., when $a(i, s), \alpha(i, s) \in \{0, 3\}$ for all i, j, s. The reason is that in $K_{2,2}$ if M_1 is an arbitrary perfect matching, any edge $e \notin M_1$ can be extended to a perfect matching M_2 with $M_1 \cap M_2 = \emptyset$, while in $K_{3,3}$ if we have a perfect matching

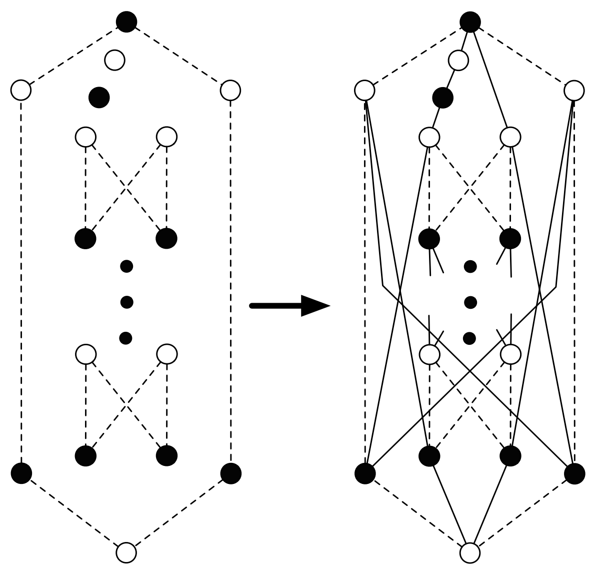

Figure 2 : *the* $2-$*factor of* K^1 *is a collection of* C_4 *and one* C_6

M_1, a pair e_1, e_2 of edges not in M_1 may not always be extended to a perfect M_2 (with $M_1 \cap M_2 = \emptyset$). ∎

Finally we shall deal with a case related to timetabling problems : $a(i,s)$ being the number of lectures that teacher t_s has to give to class c_i, it is generally a nonnegative integer r_{is}; $\alpha(j,s)$ represents however the availability of teacher t_s at period j. It is therefore either 1 or 0 for each teacher.

We shall say that a color s is *semi-unary* if $\alpha(j,s)$ is 0 or 1 for any column j or $a(i,s)$ is 0 or 1 for any row i.

$RPSU(m,n,p)$ will be the problem $RP(m,n,p)$ where all p colors (except the ground color $p+1$) are semi-unary. Then we can state :

Proposition 10 $RPSU(m,n,p=2)$ *is solvable in polynomial time.*

Proof : Consider the complete bipartite graph $\hat{K}$ constructed on nodes sets $\hat{R} = \{\text{rows } r_i \text{ with } a(i,1) + a(i,2) > 0\}$ and $\hat{S} = \{\text{columns } c_j \text{ with } \alpha(j,1) + \alpha(j,2) > 0\}$; first we construct a subgraph $\hat{G}$ with degrees $d(r_i) = a(i,1) + a(i,2)\ \forall r_i \in \hat{R}$ and $d(c_j) = \alpha(j,1) + \alpha(j,2)\ \forall c_j \in \hat{S}$.

If no such $\hat{G}$ can be found, then clearly $RPSU(m,n,p=2)$ has no solution; such a construction is a maximum flow problem.

If such a $\hat{G}$ can be found, then we show that in all cases a solution can be derived. We have to assign a color 1 or 2 to each one of its edges. We first construct an assignment of colors which satisfies the requirements in $\hat{S}$, i.e., we have $\alpha(j,s)$ edges of color s at node c_j $(s = 1, 2)$. This can be done by coloring sequentially the edges of $\hat{G}$.

If this assignment satisfies also the requirements in $\hat{R}$, i.e., we have $a(i,s)$ edges of color s at node r_i $(s = 1, 2)$, then we are done.

Otherwise there is a node $r_i \in \hat{R}$ with $d^1(r_i) > a(i,1)$ edges of color 1 and a node $r_h \in \hat{R}$ with $d^2(r_h) > a(h,2)$ edges of color 2. Notice that necessarily $h \neq i$.

If there exists, in $\hat{K}$, an alternating chain C starting at r_i with an edge of color 1 and ending at r_h with an edge of color 2, we can exchange the colors in C and we have improved the coloring.

Assume there is no such chain. Then there exists an edge $[r_i, c_j]$ of color 1 and an edge $[r_h, c_k]$ of color 2; $c_j \neq c_k$ (otherwise $[r_i, c_j], [r_h, c_k]$ would be an alternating chain from r_i to r_h).

Edge $[r_h, c_j]$ of $\hat{K}$ cannot have color 2 (it would form an alternating chain from r_i to r_h); it cannot have color 1 : since c_j already has an edge of color 1 we would have $\alpha(j,1) \geq 2$. Since color 1 is semi-unary, we should have $a(h,1) \leq 1$, but we have $d^2(r_h) > a(h,2)$; since $d^1(r_h) + d^2(r_h) = a(h,1) + a(h,2)$, we have $d^1(r_h) < a(h,1) \leq 1$, i.e. $d^1(r_h) = 0$, which is a contradiction with the assumption that $[r_h, c_j]$ has color 1. Hence $[r_h, c_j]$ has no color.

For similar reasons $[r_i, c_k]$ can have neither color 1 nor color 2; so it is uncolored.

Now removing the colors of $[r_i, c_j], [r_h, c_k]$ and giving colors 2 and 1 to $[r_i, c_k]$ and $[r_h, c_j]$ respectively, we get an improved coloring at r_i and at r_h without deteriorating it at the other nodes.

Repeating this procedure as long as the requirements are not satisfied in $\hat{R}$ will give us the required coloring. ∎

6 Final remarks

Our purpose was to explore the neighborhood of some open problems in discrete tomography. After having observed that the problem with unary colors and n=3 is NP-complete, we have exhibited several special cases which are solvable in polynomial time.

To obtain these results, we have exploited analogies with some timetabling and graph theoretical problems. As a byproduct we have derived some propoerties of complete bipartite graphs.

We have examined some solvable cases of the problems $RPU(m,n,p)$, $RPB(m,n,p)$ and $RPSU(m,n,p)$. With similar tools one could examine also $RPU(m,n,p=4)$ and $RPB(m,n,p=3)$, but it would lead to tedious case analyses. The question arises to know whether these problems can be solved in polynomial time for a fixed number $p=k$ of colors. Further research is needed in this area; its results will be useful at the same time for timetabling purposes.

Acknowledgments:

The authors express their deep gratitude to the referees whose comments have substantially contribued to improve the presentation of the paper.

Part of this work was carried out during a visit of the second author to CNAM and visits of the first and the third authors to EPFL ; the support of these institutions is gratefully acknowledged.

References

[1] C. Berge, *Graphes*, Gauthier-Villars, (Paris, 1983).

[2] M. Chrobak and C. Dürr, *Reconstructing polyatomic structures from discrete X-rays: NP-completeness proof for three atoms* Theoretical Computer Science, 259 (1-2) (2001) pp. 81-98

[3] D. de Werra, A. S. Asratian and S. Durand, *Complexity of some special types of timetabling problems* Journal of Scheduling, 5 (2002) 171-183

[4] S. Even, A. Itai and R. Shamir, *On the complexity of timetable and multicommodity flows problems* SIAM J. on Computing, 5 (1976) pp. 691-703

[5] R. J. Gardner, P. Gritzmann and D. Prangenberg, *On the computational complexity of determining polyatomic structures by X-rays* Theoretical Computer Science, 233 (2000) pp. 91-106

[6] Garey M. R. and Johnson D. S., *Computers and Intractability, a Guide to the Theory of NP-Completeness*, ed. Freeman, (New York 1979).

[7] Herman G. and Kuba A., *Discrete Tomography : Foundations, Algorithms and Applications*, ed. Birkhauser, 1999.

[8] H. Ryser, *Combinatorial Mathematics*, Mathematical Association of America and Quinn & Boden, Rahway, New Jersey, 1963.

[9] T. Yi, K. G. Murty and C. Spera, *Matchings in colored bipartite networks* Discrete Applied Math., 121 (2002) pp. 261-277