Discrete Applied Mathematics 8 (1984) 25-30
North-Holland

THE COMPLEXITY OF COMPLETING PARTIAL LATIN SQUARES

Charles J. COLBOURN*
Department of Computational Science, University of Saskatchewan, Saskatoon, Saskatchewan, S7N OWO, Canada

Received 22 April 1982

Completing partial Latin squares is shown to be NP-complete. Classical embedding techniques of Hall and Ryser underly a reduction from partitioning tripartite graphs into triangles. This in turn is shown to be NP-complete using a recent result of Holyer.

1. Introduction

A *Latin square* of *order n* is an $n \times n$ array; each entry is an element from the set $\{1, \ldots, n\}$. Each row contains each element exactly once, and each column contains each element exactly once. Latin squares are fundamental tools in experimental design (see [11], for example). Algebraically, a Latin square is the multiplication table of a *quasigroup*; studies in this context have led to applications in algebra, combinatorics, graph theory, and other areas [3].

One of the major research drives on Latin squares concerns partial Latin squares. A *partial Latin square* of order n is an $n \times n$ array; each entry is either empty, or else it contains an element from $\{1, \ldots, n\}$. Each row (column) contains each element at most once. One important investigation of the structure of Latin squares aims to characterize partial Latin squares which can be *completed* to Latin squares, without the addition of rows, columns, or elements. This study was initiated by Hall [5] and Ryser [13], and has been pursued by many researchers; an excellent survey is given in [10].

The computational complexity of completion has been considered by Rosa [12]; in fact, he asked whether the problem is NP-complete. Giles, Oyama, and Trotter [4] also asked whether completion is NP-complete; in addition, they showed that a somewhat more general problem is NP-complete. The expectation of obtaining a good algorithm is further lessened by a recent result showing that deciding whether a symmetric partial Latin square can be completed to a symmetric square is NP-complete [2]. One might expect an easy modification of this construction to remove the 'symmetric' restriction. However, this does not appear to be so - all of the symmetric partial Latin squares constructed in [2] can be completed, to (possibly) asym-

* Research partially supported by NSERC Canada Grant A5047.

metric Latin squares! The result in the symmetric case may nonetheless lead us to expect a similar result in the asymmetric case.

Our purpose here is to show that this is indeed so: completing partial Latin squares is NP-complete. We show this in two steps. In the first, we develop a correspondence between completion and a graph-theoretical problem, partitioning tripartite graphs into triangles. We show that this latter problem is NP-complete, using a proof technique of Holyer [8]. In the second step, we use classical Latin square embedding techniques of Hall [5] and Ryser [13] to reduce triangle-partition of tripartite graphs to completion.

This result demonstrates that a good characterization of completable partial Latin squares is unlikely. Moreover, it is an important addition to the class of NP-complete problems in the areas of algebra and combinatorial design theory.

2. Triangle-partition of tripartite graphs

We assume familiarity with standard graph-theoretic terminology [1]. A graph $G=(V,E)$ is *tripartite* if its vertex set V can be partitioned into three sets V_1, V_2, V_3 each inducing an independent set in G; equivalently, G is 3-colourable. A *triangle-partition* of a graph $G=(V,E)$ is a partition of E into sets, each containing three edges which form a triangle (or K_3) in G. A tripartite graph is *uniform* if every vertex exhibits the following regularity: a vertex in V_1 (or V_2, or V_3) has the same number of neighbours in V_2 and V_3 (or V_1 and V_3, or V_1 and V_2, respectively). A tripartite graph with a triangle-partition must be uniform.

The relation between completion of partial Latin squares and triangle-partition of tripartite graphs is a close one. Given a partial Latin square P of order n, we form a graph $G(P)$, called the *defect* of P. $G(P)$ has vertex set $\{r_i \mid \text{row } i \text{ contains an empty square}\} \cup \{c_j \mid \text{column } j \text{ contains an empty square}\} \cup \{e_k \mid \text{element } k \text{ appears in fewer than } n \text{ squares}\}$. The edges of $G(P)$ are as follows:

(1) If the (i,j) square of P is empty, (r_i, c_j) is an edge.
(2) if row i does not contain element k, (r_i, e_k) is an edge.
(3) If column j does not contain element k, (c_j, e_k) is an edge.

$G(P)$ has a triangle-partition if and only if P can be completed. Thus completion reduces to triangle-partition of tripartite graphs. In fact, we will show that the converse also holds, i.e. that these two problems are polynomial time equivalent. We first show that triangle-partition of tripartite graphs is NP-complete.

Holyer [8] has shown that triangle-completion of graphs is NP-complete. Essentially the same proof shows that

Theorem 2.1. *Deciding whether a tripartite graph has a triangle-partition is NP-complete.*

Proof. Membership in NP is immediate. To show completeness, we use Holyer's construction [8], with minor modifications to ensure that the graph constructed is tripartite. We do not include the details of Hoyler's method here, but instead describe only the necessary modifications. The basic ingredient is a graph $H_{3,p}$ which can be partitioned into triangles in two distinct ways. $H_{3,p}$ is defined as follows: its vertex set is

$$\{(x_1, x_2, x_3) \in Z_p^3 \mid x_1 + x_2 + x_3 \equiv 0 \pmod{p}\}.$$

Two vertices (x_1, x_2, x_3) and (y_1, y_2, y_3) are adjacent if there exist distinct i, j, and k so that

(1) $x_i \equiv y_i \pmod{p}$,
(2) $x_j \equiv y_j + 1 \pmod{p}$, and
(3) $x_k \equiv y_k - 1 \pmod{p}$.

We first observe that $H_{3,p}$ is 3-colourable (tripartite) if and only if $p \equiv 0 \pmod{3}$. In fact, any 3-colouring of the 'standard' K_3 (the triangle $(0,0,0), (1,-1,0), (1,0,-1)$) forces a unique 3-colouring of $H_{3,p}$.

We modify Hoyler's construction in two ways. First, we always choose the building blocks $H_{3,p}$ so that $p \equiv 0 \pmod{3}$; having done so, we arbitrarily 3-colour the $H_{3,p}$. Second, when Hoyler identifies 'patches' in these graphs, we ensure that the identification is done so as to identify vertices of the same colour. This does not alter Hoyler's argument, but ensures that the graph remains 3-coloured, or tripartite. □

3. Completion of partial Latin squares

In this section, we show that completion of partial Latin squares is NP-complete. We do this by establishing that every uniform tripartite graph is the defect of some partial Latin square. We require some preliminary definitions.

Given a tripartite graph $G = (V, E)$ with tripartition $V_1 \cup V_2 \cup V_3$, we label the vertices in V_1 arbitrarily with distinct labels from $\{r_1, r_2, \ldots, r_x\}$, $x = |V_1|$. We similarly label V_2 from $\{c_1 c_2, \ldots, c_y\}$ and V_3 from $\{e_1, e_2, \ldots, e_z\}$. We assume this arbitrary labelling henceforth.

A *Latin framework* for such a tripartite graph G, denoted LF$(G; r, s, t)$, is an $r \times s$ array. Each entry is either empty, or else is from the set $\{1, \ldots, t\}$. Each row (column) contains each element at most once. Three additional constraints are imposed, based on the structure of G:

(1) If G contains the edge (r_i, c_j), the (i, j) entry of the LF$(G; r, s, t)$ is empty; otherwise, it contains an element from $\{1, \ldots, t\}$.

(2) If G contains the edge (r_i, e_k), row i of the LF$(G; r, s, t)$ does not contain element k.

(3) If G contains the edge (c_j, e_k), column j of the LF$(G; r, s, t)$ does not contain the element k.

When $r=s=t$, G is precisely the defect of the LF$(G; r, r, r)$. In fact, LF$(G; r, r, r)$ is a partial Latin square which can be completed if and only if G has a triangle-partition. Thus, our next task is to produce Latin frameworks.

Lemma 3.1. *For an n-vertex uniform tripartite graph G, there is a Latin framework* LF$(G; n, n, 2n)$.

Proof. Define the $n \times n$ array L as follows. If (r_i, c_j) is an edge of G, leave the (i,j) entry of L empty. Otherwise, let the (i,j) entry contain the element $1+n+((i+j) \bmod n)$. L is an LF$(G; n, n, 2n)$. □

We adapt a theorem of Ryser [13] to recursively expand these Latin frameworks:

Lemma 3.2. *Let L be an* LF$(G; r, s, t)$ *for a uniform tripartite graph G. Denote by $R(k)$ the number of times element k appears in L plus half the degree of e_k in G. Then whenever $R(k) \ge r+s-t$ for all $1 \le k \le t$, L can be extended to an* LF$(G; r, s+1, t)$ *L' in which $R'(k) \ge r+(s+1)-t$ for all $1 \le k \le t$.*

Proof. Denote by S_i the set of all elements $\{k \mid k$ does not appear in row i of L *and* (r_i, e_k) is not an edge of $G\}$. Let M be the set of all elements $\{k \mid R(k)=r+s-t\}$. We show that the sets $S_1, \dots, S_r$ have a system of distinct representatives containing all elements in M. This system of distinct representatives is added as the $(s+1)$st column, to create L' from L. The remaining details parallel Ryser's proof closely, but we include them here for completeness.

Using the theorem of Hoffman and Kuhn [7], we must only show that the sets $S_1, \dots, S_r$(1) have an s.d.r., and (2) for every $M' \subseteq M$, at least $|M'|$ of the sets $S_1, \dots, S_r$ have nonempty intersection with M'. First we address (1). Let $1 \le m \le r$ and choose any m of the sets. Each set contains $t-s$ elements, since G is uniform. Thus the sum of cardinalities of these m sets is $m(t-s)$. On the other hand, each of the elements $1, 2, \dots, t$ appears at least $r+s-t$ times in L and G; from this, note that each element k appears in at most $t-s$ of the $\{S_i\}$. Now examine the union of the m chosen sets. This union contains some number p of elements. Since each element appears in at most $t-s$ of the sets, we have $p(t-s) \ge m(t-s)$, so $p \ge m$. Thus any m sets have at least m elements in their union, and P. Hall's theorem [6] ensures the existence of an s.d.r.

Next we address (2). Let M' be any subset of M, and suppose there are p of the $\{S_i\}$ having nonempty intersection with M'. Each of these p sets contains $t-s$ elements; hence the sum of cardinalities of these sets is $p(t-s)$. However, each element of M appears in exactly $r-(r+s-t)=t-s$ of the $\{S_i\}$, and thus $|M'|(t-s) \le p(t-s)$, whence $|M'| \le p$. This establishes that at least $|M'|$ of the sets have nonempty intersection with M' and thus the Hoffman-Kuhn theorem guarantees the existence of an s.d.r. containing all elements of M. □

We use Lemma 3.2 to repeatedly add columns. To add rows, one approach is to simply transpose the array and use Lemma 3.2 again. In fact, the addition of rows does not need this amount of machinery. M. Hall [5] showed that an $r \times s$ Latin rectangle based on $1, 2, \ldots, s$ can always be completed to a Latin square. This is proved by establishing that an s.d.r. exists for the sets $C_1, \ldots, C_s$, where C_i consists of the elements not appearing in column i. Precisely the same proof establishes that

Lemma 3.3. *A Latin framework* $\mathrm{LF}(G; r, s, s)$ *for a uniform tripartite graph G can be extended to a Latin framework* $\mathrm{LF}(G; s, s, s)$. □

Lemmas 3.1, 3.2 and 3.3 in conjunction lead to

Theorem 3.4. *Given an n-vertex uniform tripartite graph G, a Latin framework* $\mathrm{LF}(G; 2n, 2n, 2n)$ *can be produced in polynomial time.*

Proof. Initially, we produce a $\mathrm{LF}(G; n, n, 2n)$ using the technique of Lemma 3.1. To this, columns are added by repeatedly solving bipartite matching problems to find systems of distinct representatives; the Hopcroft–Karp algorithm [9] solves bipartite matching in polynomial time. We repeat this until we have an $\mathrm{LF}(G; 2n, 2n)$. Finally rows are added, again by using bipartite matching to find systems of distinct representatives, to produce the needed $\mathrm{LF}(G; 2n, 2n, 2n)$. This involves $O(n)$ bipartite matching problems, and hence requires polynomial time. □

This theorem provides the necessary reduction to show

Theorem 3.5. *Deciding whether a Latin square can be completed is NP-complete.*

Proof. Membership in NP is immediate. To show completeness, we reduce triangle-partition of tripartite graphs, which is NP-complete by Theorem 2.1. Given an n-vertex tripartite graph (with specified tripartition), we first determine whether it is uniform. If it is not, there is no triangle-partition. If it is, we apply Theorem 3.4 to produce a Latin framework $\mathrm{LF}(G; 2n, 2n, 2n)$ in polynomial time. This Latin framework is a partial Latin square. We need only show that this partial Latin square can be completed if and only if G has a triangle-partition. This follows directly from the observation that, by construction, G is the defect of the partial Latin square. □

4. Conclusions

Theorem 3.5 severely limits any hopes of obtaining a good characterization of completable partial Latin squares. In addition, the construction and use of Latin frameworks may prove useful in investigating the complexity of other algebraic and combinatorial problems, particularly using Latin frameworks in combinatorial constructions in place of Latin squares. The potential of this approach awaits further research.

Acknowledgements

Thanks are due to Marlene Colbourn, Pavol Hell, and Kevin Phelps for many helpful discussions on this matter.

References

[1] J.A. Bondy and U.S.R. Murty, Graph Theory with Applications (Macmillan, New York, 1976).
[2] C.J. Colbourn, Embedding partial Steiner triple systems is NP-complete, J. Combin. Theory (A) 35 (1983) 100–105.
[3] J. Denes and A.D. Keedwell, Latin Squares and their Applications (Academic Press, New York, 1976).
[4] F.R. Giles, T. Oyama, and L.E. Trotter, On completing partial Latin squares, Proc. Eighth Southeastern Conf. Combinatorics, Graph Theory, and Computing (1977) 523–543.
[5] M. Hall Jr., Distinct representatives of subsets, Bull. Amer. Math. Soc. 54 (1948) 922–926.
[6] P. Hall, On representatives of subsets, Proc. London Math. Soc. 10 (1935) 26–32.
[7] A.J. Hoffman and H.W. Kuhn, Systems of distinct representatives and linear programming, Amer. Math. Monthly 63 (1956) 455–460.
[8] I. Holyer, The NP-completeness of some edge-partition problems, SIAM J. Computing 10 (1981) 713–717.
[9] J.E. Hopcroft and R.M. Karp, An $n^{5/2}$ algorithm for maximum matching in bipartite graphs, SIAM J. Computing 2 (1973) 225–231.
[10] C.C. Lindner and T. Evans, Finite Embedding Theorems for Partial Designs and Algebras (Les Presses de L'Universite de Montreal, Montreal, Canada, 1977).
[11] D. Raghavarao, Constructions and Combinatorial Problems in Design of Experiments (Wiley, New York, 1971).
[12] A. Rosa, Problem session, Proc. Second Caribbean Conf. Combinatorics and Computing (1977).
[13] H.J. Ryser, A combinatorial theorem with application to Latin rectangles, Proc. Amer. Math. Soc. 2 (1951) 550–552.