Discussiones Mathematicae
Graph Theory 33 (2013) 57–69
doi:10.7151/dmgt.1648

Dedicated to Mieczysław Borowiecki on his 70th birthday

WHEN IS AN INCOMPLETE $3 \times n$ LATIN RECTANGLE COMPLETABLE?

Reinhardt Euler

Université Européenne de Bretagne and Lab-STICC, CNRS, UMR 6285,
Université de Brest, B.P.809,
20 Avenue Le Gorgeu, 29285 Brest Cedex, France

e-mail: Reinhardt.Euler@univ-brest.fr

and

Paweł Oleksik[1]

AGH University of Science and Technology,
Faculty of Geology, Geophysics and Environment Protection,
Department of Geoinformatics and Applied Computer Science,
30-059 Cracow, Poland

e-mail: oleksik@agh.edu.pl

Abstract

We use the concept of an availability matrix, introduced in Euler [7], to describe the family of all minimal incomplete $3 \times n$ latin rectangles that are not completable. We also present a complete description of minimal incomplete such latin squares of order 4.

Keywords: incomplete latin rectangle, completability, solution space enumeration, branch-and-bound.

2010 Mathematics Subject Classification: 05B15, 05C65.

1. Introduction and Basic Results

An $n \times n$ array L each cell of which contains exactly one symbol $i \in N = \{1, \ldots, n\}$ such that each symbol occurs in each row and in each column exactly once is a *latin square* (of order n). If we replace "exactly" by "at most" and if not all cells

[1]Work supported by the AGH-UST Fac. of GG&EP Research Plan, Task No. 11.11.140.032.

are filled, we obtain an *incomplete* latin square, and if for $r, s \in N$ the non-empty cells of an incomplete latin square form a rectangle of r rows and s columns, we will speak of an $r \times s$ *latin rectangle*. Given $n \in \mathbb{N}$ the problem of characterizing those incomplete latin squares that are *completable* (to a latin square of the same order), is an open question. There are however partial results: Evans' conjecture [8] (proved by Smetaniuk [13], and independently by Andersen and Hilton [2]) states that an incomplete latin square containing at most $n-1$ filled cells is always completable. A similar result is due to Hall [10], who showed that no condition is required to complete an $r \times n$ latin rectangle. A third result due to Ryser [12] states that an $r \times s$ latin rectangle can be completed if and only if each symbol $n \in N$ appears at least $r + s - n$ times. By using the concept of an availability matrix the first author has shown in [7] how further such results can be obtained for the completability of (one or more) incomplete rows of specific structure.

The result presented in this paper is of different a nature. Let $E_n = \{e_{ijk} : 1 \leq i, j, k \leq n\}$ be an arbitrary set of n^3 elements. Call an $n \times n$ array L *feasible*, if each cell of L contains a symbol k at most once. Obviously, we can identify the selection of the element e_{ijk} with the appearance of symbol k in cell ij, and hereby obtain a $1-1$-relation between subsets of E_n and feasible arrays over N. For convenience, we will make no real distinction between a feasible array and its corresponding subset. In particular, any latin square corresponds to a specific n^2-element subset of E_n, and the system $\mathcal{B}_n$ of all these sets constitutes a clutter, say, of *bases*, a notion well known from matroid theory. Any such clutter induces a (unique) clutter $\mathcal{C}_n$ of *circuits*, i.e., subsets of E_n that are not contained in any member of $\mathcal{B}_n$ and that are minimal with respect to this property. As a consequence, the complete knowledge of $\mathcal{C}_n$ would answer the completability question in the following sense: an incomplete latin square can be completed if and only if it does not contain any circuit.

In 1985 (see Euler *et al.* [6]), we have initiated the study of $\mathcal{C}_n$ by considering circuits arising from two distinct symbols in one cell or two identical symbols in one row or in one column, that we call *elementary*, and others arising from particular latin rectangles. Our motivation was the application of linear programming techniques to solve the *planar* 3*-dimensional assignment problem* (P), the solutions of which correspond to the latin squares of the given order. Observe that (P) also contains our completability question, shown by Colbourn [5] to be NP-complete, as a special case. In this context, circuits are useful for providing facet-defining inequalities for associated polyhedra. For surveys on 3-dimensional assignment problems we refer the reader to Burkard *et al.* [4] and Spieksma [14].

The main objective of this paper is to study the clutter of circuits associated with the collection of all $r \times n$ latin rectangles for given r. We will limit ourselves to *non-elementary* circuits, i.e., the collection $\mathcal{C}_n^r$ of all those incomplete $r \times n$ latin rectangles, that are not completable and minimal with respect to this property.

A complete answer for all $r \in \{1, \ldots, n\}$ would provide necessary and sufficient conditions for the completability of *any* incomplete latin square. In the following, we will fully answer this question for $r = 3$. Just observe that by Hall's theorem [10], $\mathcal{C}_n^r$ is a subfamily of $\mathcal{C}_n$, and our result therefore contributes to a better knowledge of $\mathcal{C}_n$. We also point to the work of Brankovic *et al.* [3], who studied circuits under the name of *premature partial latin squares*, and to the recent work of Adams *et al.* [1] for which the knowledge of circuits could open a different approach. We also give a complete description of $\mathcal{C}_n$ for $n = 4$ that we have obtained by computer calculations. We just mention that generating the family of circuits associated with a clutter of bases is a special case of *transversal hypergraph generation* (as for instance studied by Khachiyan *et al.* [11]), which has many applications in combinatorics and computer science and whose exact complexity status is still open. We refer to Hagen [9] for recent results on this topic.

The basis of our analysis is the following theorem:

Theorem 1 (Frobenius-König)**.** *A* $(0,1)$*-matrix* A *of size* $n \times n$ *contains* n $1's$ *no two of which lie in the same row or column if and only if* A *does not contain a* 0*-submatrix of size* $u \times v$ *such that* $u + v = n + 1$*.*

It is the application of this theorem to a very special matrix that will lead us to a complete description of the family $\mathcal{C}_n^3$.

Definition 2. Let L be an incomplete latin square the m-th row of which contains $0 < t < n$ empty cells. Moreover, let $S(m)$ denote the set of symbols not appearing in that row and $J(m)$ the set of column indices of its empty cells. The *availability matrix* $A(L,m)$ is the $t \times t$ matrix obtained from the $n \times n$ matrix A by deleting rows A_i for $i \in N \setminus S(m)$ and columns A^j for $j \in N \setminus J(m)$, and with an element $A_i^j(L,m)$, $i \in S(m)$, $j \in J(m)$ marked with an asterisk as "non-available" if and only if symbol i appears in column j of L.

$$A = \begin{bmatrix} 1 & \ldots & 1 \\ 2 & \ldots & 2 \\ \vdots & & \vdots \\ n & \ldots & n \end{bmatrix}$$

What is the use of $A(L,m)$?

- If it is possible to select within this matrix t available elements, one per row and one per column, then row m is completable;
- if, however, this is not possible, by Frobenius-König's Theorem 1, $A(L,m)$ has to contain a $p \times q$ submatrix of non-available elements such that $p + q = t + 1$ (see Figure 1 for an illustration of that case).

In case of an $r \times n$-latin rectangle L we hereby obtain necessary and sufficient conditions for the completability of a new type of incomplete latin square (that

L :

4	1	5	6	2	3
2	3	1	5	6	4
3	4	6	1	5	2
1	2				

$$A(L,4): \begin{bmatrix} 3 & 3 & 3 & 3^* \\ 4 & 4 & 4 & 4^* \\ 5^* & 5^* & 5^* & 5 \\ 6^* & 6^* & 6^* & 6 \end{bmatrix}$$

Figure 1. An incomplete, non-completable latin square and its availability matrix with non-available elements marked by a "*".

can be checked in polynomial time by solving an assignment or bipartite matching problem over $A(L, m)$):

Theorem 3. *Let L be an $r \times n$ latin rectangle with $0 < t < n$ empty cells in row $r + 1$. Then L is completable if and only if any subset I of $S(r + 1)$ is contained in at most $t - |I|$ of the columns $L^j, j \in J(r + 1)$.*

2. A Complete Description of $\mathcal{C}_n^3$

Before turning to three rows we just mention that the case $r = 1$ is obvious: an incomplete latin row is always completable, i.e., $\mathcal{C}_n^1$ is empty, and for the case $r = 2$ the family $\mathcal{C}_n^r$ for $n \geq 3$ is fully represented by the two types of circuits illustrated in Figure 2 (up to row- and column interchangements, and for any symbol $i \in N$).

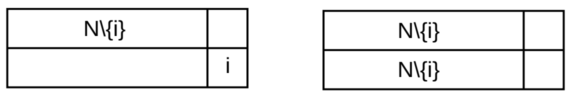

Figure 2. The 2 types of circuits for $r = 2$.

As to $r = 3$ we start with those circuits that arise from the non-completability of a single row, so-called 1-*row-circuits*. Applying Theorem 3 we come up with 4 different types as illustrated in Figure 3 (again for $n \geq 3$ and throughout the paper, up to row- and column interchangements, and for any distinct $i, j, k \in N$). We just remark that this family is well understood for any $r \in \{1, \ldots, n\}$; a description (in its conjugate form rows $\leftrightarrow$ symbols) has already been given in Euler *et al.* [6].

Now let us turn to 2-*row-circuits*, i.e., those incomplete $3 \times n$ latin rectangles, which are not completable and minimal with respect to this property, which do not contain any 1-row-circuit but which contain two rows, say row 1 and row 2,

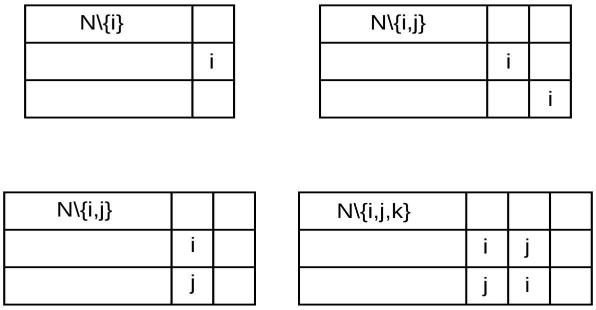

Figure 3. The 4 types of 1-row-circuits for $r = 3$.

that are not completable. Clearly, the second circuit depicted in Figure 2 is a first such 2-row-circuit.

To describe the others, we consider the availability matrices A_1 and A_2 of rows 1 and 2, and observe the following: first, both matrices must have a line (i.e., row or column) in common, since otherwise the two rows would be completable; second, a common line can contain at most *one* non-available symbol, since an asterisk can only arise from a symbol in row 3. Therefore, a forbidden submatrix within A_1, engendered by the completion of row 2 (or vice versa) can only be of size 2×2, 2×1 or 1×2. In the first case, A_2 cannot be of size 1×1 only: a symbol k, already appearing in row 2, is marked as non-available for row 1, and thus also for row 2 and its empty cell. Therefore, we can delete symbol k from row 2, a contradiction to minimality. We are thus lead to a second 2-row-circuit, depicted as type 2 in Figure 4, for which we also indicate the 2 availability matrices and the way to complete row 1 after deletion of an arbitrary symbol l. For $n \geq 6$, row 3 is then always completable due to Theorem 3. The proof for row 2 is similar, and for row 3 it is straightforward.

As to forbidden submatrices of size 2×1 or 1×2 we come up with 3 possibilities, types 3 to 5 in Figure 4, illustrated in a similar way as type 2. We just mention that types 1, 3 and 4 exist for $n \geq 3$, type 2 for $n \geq 3$ and $n \neq 4$, and type 5 for $n \geq 4$. We also ask for the following properties:

- in type 3, i and j have to be different from k but i may be equal to j;
- in type 4, the empty cells in rows 1 and 2, not in the last column, may appear in a same column;
- in type 5, i, j and k must all be different ($i = k$ would contradict minimality).

Finally, if one of the availability matrices is of size 1×1 and the other of size 2×2, we obtain a contradiction to minimality. Our 5 types of 2-row-circuits are therefore exhaustive.

We now come to 3-*row-circuits*, i.e., those incomplete $3 \times n$ latin rectangles, which are not completable, minimal with respect to this property and which do

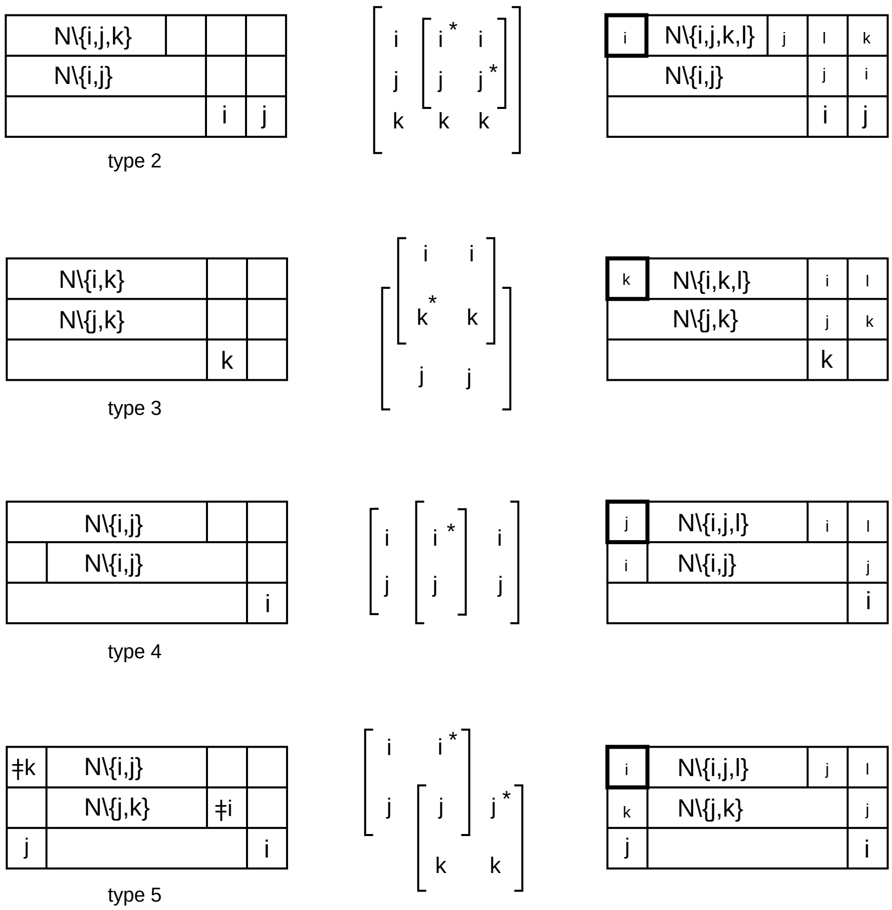

Figure 4. The 4 remaining types of 2-row-circuits, their availability matrices and completability for one element deleted in row 1.

not contain any 1- or 2-row-circuit. Figure 5 exhibits 5 different types. The first of them exists for $n \geq 5$ and the remaining for $n \geq 4$.

As before, the following conditions are required:

- all symbols $i, j, k, l \in N$ have to be different with the exception of type 5, in which we may have $l = j$ (if $l = j$ or $l = k$ in type 4 for instance, then C would properly contain a 2-row-circuit);
- in type 2, the first 3 empty cells in rows $1, 2, 3$ have to be in different columns;
- in type 4, the first 2 empty cells in rows $1, 2$ may appear in a same column;
- in type 5, the first empty cell in row 3 must not appear in the column containing k or i.

To show that this list is exhaustive, let us consider a 3-row-circuit C together with

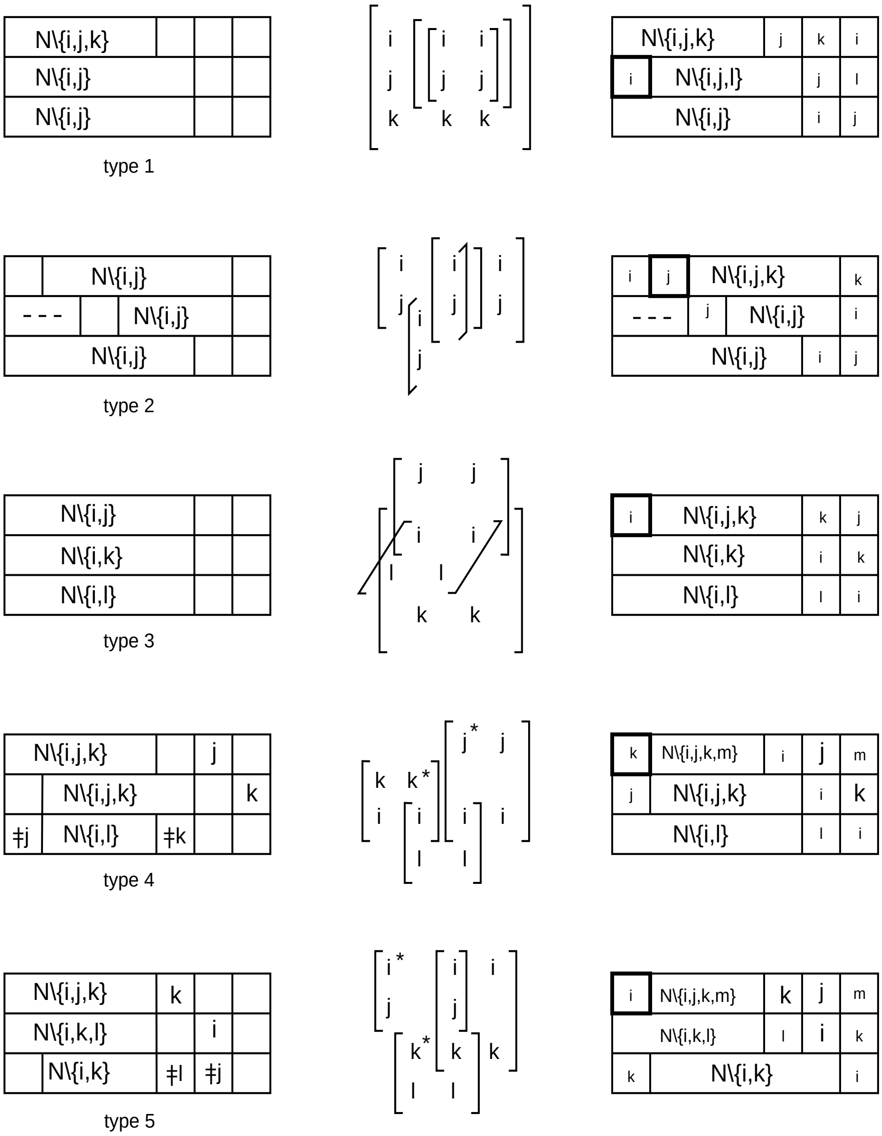

Figure 5. All types of 3-row-circuits.

the associated availability matrices A_1, A_2 and A_3. We claim that A_3 is of size at most 3×3, i.e., there are at most 3 empty cells in row 3 of C. By assumption the first two rows of C are completable. By Theorem 3 and for any such completion, there must be a set $I \subseteq S(3)$ which is contained in $t - |I| + 1$ columns C^j, for $j \in J(3)$. Such an I can only appear in rows 1 and 2, and therefore, $t - |I| + 1 \leq 2$ and $|I| \in \{1, 2\}$, which implies $t \leq 3$. By symmetry, any other of the 3 rows of C cannot have more than 3 empty cells.

Can there be more than one row with *exactly* 3 empty cells? For an answer to this question let row 3 be empty in columns C^{n-2}, C^{n-1} and C^n, and let $I = \{i, j\}$. The only ways to make I appear twice, say in columns C^{n-1} and C^n, are the following:

1. symbol i is in cells $1, n-1$ and $2, n$, and symbol j is in cells $1, n$ and $2, n-1$, but then C would contain a 1-row-circuit (type 4 in Figure 3);

2. symbol i is in cell $1, n-1$, and symbol j is in cell $1, n$, the cells $2, n-1$ and $2, n$ being empty, and row 2 containing all symbols from $N \setminus \{i, j\}$, but then C would contain a 2-row-circuit (type 2 in Figure 4);

3. rows 1 and 2 both contain $N \setminus \{i, j\}$ in their first $n-2$ columns the remaining cells being empty.

Only case 3 applies and, therefore, row 3 is the only row with 3 empty cells. Obviously, C induces a unique 3-row-circuit (type 1 in Figure 5).

We are left with the situation that A_1, A_2 and A_3 are all of size at most 2×2. Since any of the 3 matrices has to share at least one line with one of the others, the possibility of size 1×1 for all 3 contradicts non-completability. The only possible case in which two of the matrices are of size 1×1 is illustrated in Figure 6.

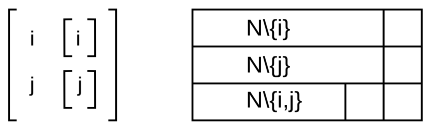

Figure 6

If symbols i and j are in cells $2, n-1$ and $1, n-1$, respectively, C properly contains a 1-row-circuit of type 3, and if both symbols appear in a same column C^j, $j \in \{1, \ldots, n-2\}$, C properly contains a 3-row-circuit of type 1. If only symbol i (or symbol j, respectively) appears in column $n-1$, then C contains a 2-row-circuit of type 3. And finally, if symbols i and j are in different columns C^j, $j \in \{1, \ldots, n-2\}$, C properly contains a 3-row-circuit of type 2. Altogether, C cannot represent a circuit.

All other cases, in which one of the 3 matrices is of size 1×1, are depicted in Figure 7 (up to symmetries). They can be treated along the same line leading to the same conclusion that C cannot be a circuit.

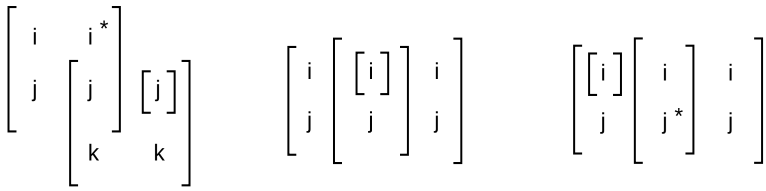

Figure 7

We still have to treat the case that all 3 matrices are of size 2×2. It turns out that *any* induced 3-row-circuit is among those presented in Figure 5.

If two of the matrices, say A_1 and A_2, coincide, i.e., for row 1 and row 2 we have $S(1) = S(2)$ and $J(1) = J(2)$, C either contains a 3-row-circuit of type 1 or it is completable.

If A_1 and A_2 just coincide in one column, the assumption on the completability of any pair of rows gives us a 3-row-circuit of type 2 or 3, or we obtain completability of C.

If finally, A_1 and A_2 coincide in just one cell, let α denote the number of cells that A_3 can have in common with the superposition of A_1 and A_2. In case that $\alpha = 1$ the only way to make C non-completable leads to a 3-row-circuit of type 5 or, by symmetry, of type 4, for $\alpha = 2$ it is such a circuit of type 4 (with the first two empty cells in rows $1, 2$ appearing in a same column) or, by symmetry, of type 5 with $l = j$, and for $\alpha = 3$ the assumption on the completability of any pair of rows always implies completability of C.

Altogether, we have shown:

Theorem 4. *An incomplete* $3 \times n$ *latin rectangle is completable if and only if it does not contain any* 1*-row-circuit of type* 1 *to* 4*,* 2*-row-circuit of type* 1 *to* 5*, or* 3*-row-circuit of type* 1 *to* 5*.*

We just remark that by interchanging the role of column indices and symbols (so-called *conjugacy*), the number of types in each class can be reduced further to 3, 4 and 3 for 1-row, 2-row and 3-row-circuits, respectively.

To conclude this section we would like to show how most of the circuits arise as special cases of more general descriptions.

i) Let L be an $r \times s$ latin rectangle of order n such that $r + s = n + 1, r \geq 2$, symbol k not occurring in L and the other members all appearing at least two times, in rows $1, \ldots, r$ and columns $1, \ldots, s$. By Ryser's theorem [12] L cannot be completed, but the removal of any symbol leads to completability.

L therefore represents a member of $\mathcal{C}_n$ covering in particular type 3 of the 3-row-circuits.

ii) Let L be as above with the difference, that all symbols are from $\{1, \dots, s\}$ and that $n - s - 1$ arbitrary symbols have been removed. Again, by [12] L cannot be completed but the removal of any symbol leads to completability, i.e., L represents a member of $\mathcal{C}_n$ (already described in Euler *et al.* [6]) and covering type 1 of the 2-row-circuits as well as type 1 of the 3-row-circuits.

iii) Let L be an $r \times s$ latin rectangle over N and let symbol k appear l times in rows $r + 1, \dots, n$ and columns $s + 1, \dots, n$, at most once in a same row or column. In Euler *et al.* [6] we have shown that this incomplete latin square can be completed if and only if each symbol from $N \setminus \{k\}$ appears at least $r + s - n$ times in L and symbol k appears at least $r + s + l - n$ times in L. If $r + s = n - l + 1$ and symbol k does not appear in L we obtain a member of $\mathcal{C}_n$ in general form covering in particular any 2-row-circuit of type 3. By conjugacy rows $\leftrightarrow$ symbols all types of 1-row-circuits are members of $\mathcal{C}_n$, too.

iv) The last case to be treated will be type 2 of the 2-row-circuits. Again we start with an $r \times s$ latin rectangle L as given in ii). Since $n - s - 1 = r - 2$, we may choose one of the (at least) 2 rows, say row r, containing s symbols. We remove these symbols from row r and place symbols $s+1, \dots, n$ (arbitrarily) in columns $s + 1, \dots, n$ of that same row. We obtain an incomplete latin square L' which is not completable, since any completion would have to contain symbols $1, \dots, s$ in the first s columns of row r. Removing any symbol from the first $r - 1$ rows leads to completability, and moving any symbol in row r to a cell within columns $1, \dots, s$ also leads to completability, as in ii). Again, L' is shown to be a member of $\mathcal{C}_n$ covering type 2 of the 2-row-circuits.

v) Finally, type 4 of the 3-row-circuits and type 5 of the 2-row-circuits have already been studied in Euler [7] with respect to a general form. This was, however, only possible in the context of *circulant* latin rectangles.

3. Complete Descriptions of $\mathcal{C}_n$ for Small n

This section is on the generation of a complete description of $\mathcal{C}_n$ for small n on a computer. For $n = 3$ such a description consisting only of 1- and 2-row-circuits has already been given in Euler *et al.* [6]. Up to row- column- and symbol interchangements (so-called *isotopy*) and the exchange of row- column- or symbol indices (i.e., conjugacy) we are left with two types which are represented in Figure 8.

1	2	
		3

1	2	
2	1	

Figure 8. All types of circuits for $n = 3$.

As to $n = 4$ we have generated the associated family of circuits by means of a computer program, whose basic idea is traversing the solution space with the application of heuristics (like branch-and-bound technique and prediction of properties).

In view of the previous results it is sufficient to exhibit those 4×4 arrays that contain all 4 symbols and no empty row or column. Up to isotopy and conjugacy we come up with 2 types, both 2-row-circuits with respect to rows 1 and 2. Together with their availability matrices they are represented in Figure 9.

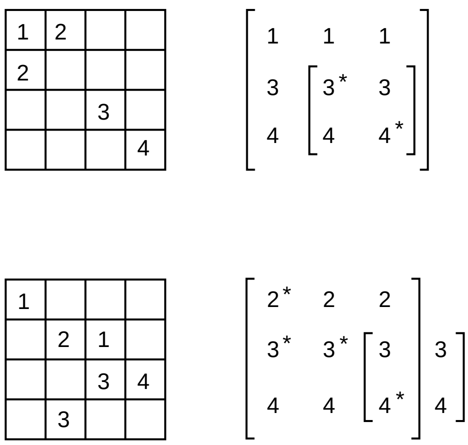

Figure 9. The 2 remaining types of circuits for $n = 4$.

These results allow us to present a complete description of $\mathcal{C}_4$:

Theorem 5. *An incomplete latin square of order* 4 *is completable if and only if, up to isotopy and conjugacy, it does not contain any of the following arrays as a subarray:*

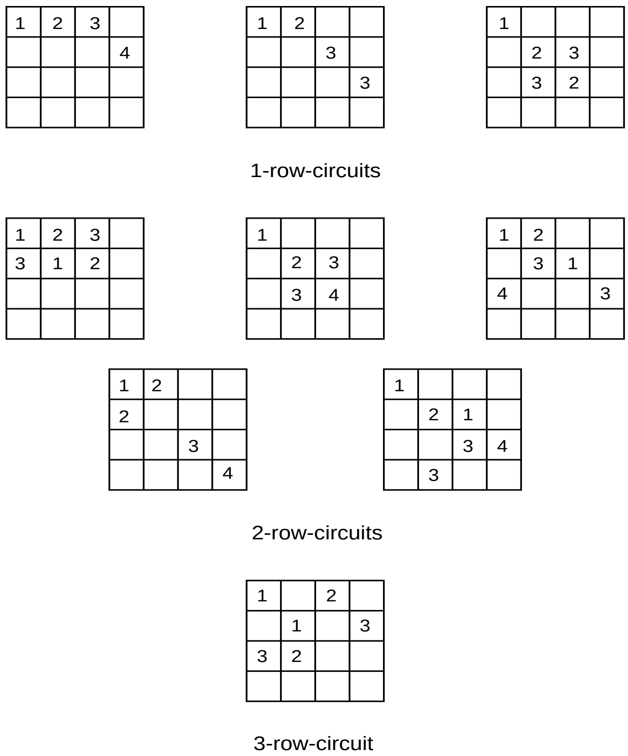

Figure 10. All types of circuits for $n = 4$.

4. Conclusion and Future Work

A first direction for future research could be a complete characterization of 2-row-circuits. Also, to obtain a complete description of $\mathcal{C}_n$ via an extension of this work to $r \times n$ latin rectangles for any r, the idea arises of generalizing 1-row, 2-row and 3-row-circuits to m-row-circuits for $m > 3$. For this a detailed study of our computational results should be helpful.

In this respect our results for a complete description of $\mathcal{C}_5$ are of much a wider scope than those for $\mathcal{C}_4$. They are currently being analysed.

References

[1] P. Adams, D. Bryant and M. Buchanan, *Completing partial latin squares with two filled rows and two filled columns*, Electron. J. Combin. **15** (2008) #R56.

[2] L.D. Andersen and A.J.W. Hilton, *Thank Evans!*, Proc. Lond. Math. Soc. (3) **47** (1983) 507–522.
doi:10.1112/plms/s3-47.3.507

[3] L. Brankovic, P. Horák, M. Miller and A. Rosa, *Premature partial latin squares*, Ars Combin. **63** (2002), 175–184.

[4] R. Burkard, M. Dell'Amico and S. Martello, Assignment Problems (SIAM, 2009).

[5] C. Colbourn, *The complexity of completing partial latin squares*, Discrete Appl. Math. **8** (1984) 25–30.
doi:10.1016/0166-218X(84)90075-1

[6] R. Euler, R.E. Burkard and R. Grommes, *On latin squares and the facial structure of related polytopes*, Discrete Math. **62** (1986) 155–181.
doi:10.1016/0012-365X(86)90116-0

[7] R. Euler, *On the completability of incomplete latin squares*, European J. Combin. **31** (2010) 535–552.
doi:10.1016/j.ejc.2009.03.036

[8] T. Evans, *Embedding incomplete latin squares*, Amer. Math. Monthly **67** (1960) 958–961.
doi:10.2307/2309221

[9] M. Hagen, *Lower bounds for three algorithms for transversal hypergraph generation*, Discrete Appl. Math. **157** (2009) 1460–1469.
doi:10.1016/j.dam.2008.10.004

[10] M. Hall, *An existence theorem for latin squares*, Bull. Amer. Math. Soc. **51** (1945) 387–388.
doi:10.1090/S0002-9904-1945-08361-X

[11] L. Khachiyan, E. Boros, K. Elbassioni and V. Gurvich, *A global parallel algorithm for the hypergraph transversal problem*, Inform. Process. Lett. **101** (2007) 148–155.
doi:10.1016/j.ipl.2006.09.006

[12] H.J. Ryser, *A combinatorial theorem with an application to latin rectangles*, Proc. Amer. Math. Soc. **2** (1951) 550–552.
doi:10.1090/S0002-9939-1951-0042361-0

[13] B. Smetaniuk, *A new construction on latin squares* - I: *A proof of the Evans Conjecture*, Ars Combin. **11** (1981) 155–172.

[14] F.C.R. Spieksma, *Multi-index assignment problems: complexity, approximation, applications*, in: Nonlinear Assignment Problems, Algorithms and Applications, L. Pitsoulis and P. Pardalos, (Eds.), Kluwer (2000) 1–12.

Received 24 March 2012
Revised 26 July 2012
Accepted 27 July 2012