S-72.2420/T-79.5203 Graph Theory (5 cr) Spring 2006 Programming Project

A *Latin square* of order n is an $n \times n$ array over n symbols such that each symbol appears exactly once in every row and in every column. See Figure 1 for an illustration. Let us assume in what follows that the symbols are $1, 2, \ldots, n$.

1	2
2	1

1	2	3
2	3	1
3	1	2

1	2	3	4
2	1	4	3
3	4	1	2
4	3	2	1

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

Figure 1: Latin squares

A *partial* Latin square has some of its entries undefined. A *completion* of a partial Latin square is an assignment of symbols to the undefined entries that completes the partial square to a valid Latin square. Not all partial Latin squares can be completed. For example, the left-hand square in Figure 2 cannot be completed, whereas the right-hand square has exactly one completion.

1	2	3	4	5
2	3	.	.	.
3	.	4	.	.
4	.	.	5	.
5	.	.	.	1

1	2	3	4	5
2	3	.	.	.
3	.	4	.	.
4	.	.	5	.
5	.	.	.	2

Figure 2: Partial Latin squares

In the language of graphs, the completion of partial Latin squares can be studied using the following two alternative frameworks.

Clique. Consider a graph with the n^3 ordered triples of the form (i, j, k) as vertices, $i, j, k \in \{1, 2, \ldots, n\}$. Any two distinct triples are connected by an edge if and only if the triples agree in at most one coordinate. The resulting graph for $n = 2$ is illustrated in Figure 3. The intuition is that each triple (i, j, k) indicates that we assign the value k to the entry at row i, column j in an $n \times n$ array. Observe now that Latin squares of order n correspond to cliques of size n^2 in the graph. Furthermore, a partial Latin square can be completed if and only if the corresponding triples occur in a clique of size n^2.

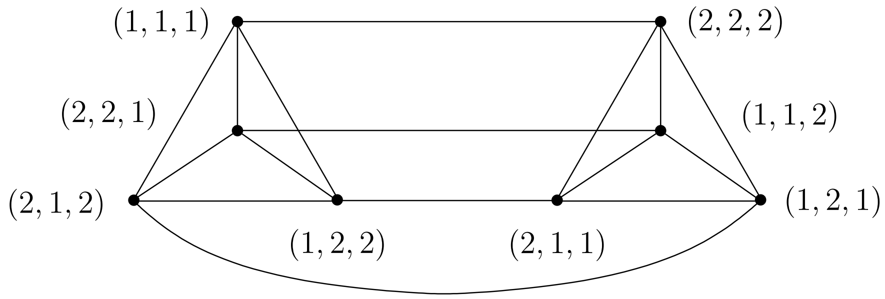

Figure 3: The graph for $n = 2$

Covering. An alternative modeling strategy is to view the task as a hypergraph problem, in which we are to cover all three types of vertices r_ic_j, r_iv_k, c_jv_k with $i, j, k \in \{1, 2, \ldots, n\}$ exactly once using hyperedges of the form $\{r_ic_j, r_iv_k, c_jv_k\}$. For example, for $n = 2$ the vertices are

$$r_1c_1,\ r_1c_2,\ r_2c_1,\ r_2c_2,\ r_1v_1,\ r_1v_2,\ r_2v_1,\ r_2v_2,\ c_1v_1,\ c_1v_2,\ c_2v_1,\ c_2v_2$$

and the hyperedges are

$$\{r_1c_1, r_1v_1, c_1v_1\},\ \{r_1c_1, r_1v_2, c_1v_2\},\ \{r_1c_2, r_1v_1, c_2v_1\},\ \{r_1c_2, r_1v_2, c_2v_2\},$$
$$\{r_2c_1, r_2v_1, c_1v_1\},\ \{r_2c_1, r_2v_2, c_1v_2\},\ \{r_2c_2, r_2v_1, c_2v_1\},\ \{r_2c_2, r_2v_2, c_2v_2\}.$$

The Project. The aim of this project is to develop algorithms for completing partial Latin squares. The groups, consisting of 1–3 students, should try *both* of the aforementioned approaches. The groups may choose to develop their own software from scratch, but the use of existing software tools for clique search and cover search is permitted. Such tools include Cliquer (`http://users.tkk.fi/pat/cliquer.html`) and Algorithm DLX (`http://www-cs-faculty.stanford.edu/~knuth/` – see paper P159 and the program DANCE).

The time scale for the project is as follows:

15.3. Presentation of problem, forming groups. Please register your group via e-mail to ⟨`jdubrovi@tcs.hut.fi`⟩. Groups of less than 3 students that would like more members should indicate this upon registration.
17.3. Deadline for group registration.
28.4. Project review (oral presentation & written report).

The project objectives are as follows:

- Using the developed programs, count the number of distinct Latin squares up to order 5.
- In general, the problem of deciding whether a given partial Latin square admits a completion is known to be NP-complete.[1] Thus, it is to be expected that some partial Latin squares are very difficult to complete. Find and document such squares with as small as possible order that are difficult to complete with the algorithms employed. Do you think that the squares will be difficult for other algorithms as well?
- *Sudoku* has recently become a popular recreational exercise. Model the task of completing a sudoku square using the language of graphs. Modify the developed programs to handle sudoku completion. Are the squares published in daily newspapers difficult for the algorithms employed? Can you construct more difficult squares? What about 16×16 squares?

The techniques and the findings should be documented in a written report of length approximately 5 pages, excluding possible tables of computational results. The reports are subjected to peer review. The marking of your fellow students is in the formula for the mark of the course (see the general information of the course). The reports should be handed out during project review: one copy for each group and each of the two teachers. Each group is also to give a short oral presentation during the review describing the findings and the techniques employed.

[1] C.J. Colbourn, The complexity of completing partial Latin squares, *Discrete Applied Mathematics* **8** (1984) 25–30.

Learning Objectives. Upon completion of the project, the students are able to

- model combinatorial problems using the language of graphs;
- classify and compare published algorithms for graph-theoretic problems;
- apply results from the scientific literature in developing practical algorithms;
- collaborate with fellow students in a science project;
- report project results concisely orally and in writing;
- assess algorithms and results produced by others.