Permutation Covers

Charles J. Colbourn

School of Computing, Informatics, and Decision Systems Engineering
Arizona State University

Workshop on Graphs and Algorithms, Fields Institute, May 2014

Permutation t-Coverings

- A t-subpermutation of $\{0, \ldots, v-1\}$ is a t-tuple $(x_1, \ldots, x_t)$ with $x_i \in \{0, \ldots, v-1\}$ for $1 \leq i \leq t$, and $x_i \neq x_j$ when $i \neq j$.
- A permutation π of $\{0, \ldots, v-1\}$ *covers* the t-subpermutation $(x_1, \ldots, x_t)$ if $\pi^{-1}(x_i) < \pi^{-1}(x_j)$ whenever $i < j$.
- (In other words, the permutation is a linear extension of the subpermutation.)
- For example, $(4, 0, 3)$ is a 3-subpermutation that is covered by the permutation 4 2 0 3 1.

Permutation t-Coverings

- A *permutation covering* of *order* v and *strength* t is a set $\Pi = \{\pi_1, \ldots, \pi_N\}$ where π_i is a permutation of $\{0, \ldots, v-1\}$, and every t-subpermutation of $\{0, \ldots, v-1\}$ is covered by at least one of the permutations $\{\pi_1, \ldots, \pi_N\}$.
- Call one a PermC$(N; t, v)$.
- When written as an array, often called a *sequence covering array* SeqCA$(N; t, v)$.

Permutation t-Covering

Example

$t = 3$, $v = 5$, $N = 8$

SeqCA	CSSP
4 2 0 3 1	2 4 1 3 0
1 4 3 0 2	3 0 4 2 1
3 1 2 0 4	3 1 2 0 4
0 2 4 1 3	0 3 1 4 2
2 1 3 4 0	4 1 0 2 3
0 3 4 1 2	0 3 4 1 2
3 0 2 1 4	1 3 2 0 4
4 1 2 0 3	3 1 2 4 0

Scrambling Sets

- A *completely t-scrambling set of permutations*, CSSP($N; t, v$) is an $N \times v$ array $A = (a_{ij})$ for which
  - every row forms a permutation of the v symbols, and
  - in every set of t columns $c_1, \ldots, c_t$, and for every permutation ψ of $\{1, \ldots, t\}$, there is a row ρ such that $a_{\rho c_{\psi(i)}} < a_{\rho c_{\psi(i+1)}}$ for $1 \leq i < t$.
  - (in other words, in every set of t columns, every 'pattern' appears on these t columns in at least one row)
- This is *equivalent* to a SeqCA($N; t, v$) – just interchange the roles of columns and symbols.

Sequence Covering Arrays

The Existence Question

- Given t and v, what is the smallest N for which a SeqCA($N; t, v$) exists?
- Call this number SeqCAN(t, v).
  - SeqCAN(t, v) $\geq t!$
- SeqCAN($2, v$) $= 2$ for all $v \geq 2$ – Just take any permutation and its reversal!
- SeqCAN(t, v) $= t!$ when $v \leq t + 1$ (Levenshtein), and SeqCAN(4,6) = 4! (Mathon and Tran Van Trung).
- But SeqCAN(t, v) $> t!$ when $v \geq 2t$ and $t \geq 3$.

Sequence Covering Arrays
The Existence Question when $t \geq 3$

- A connection with "covering arrays" demonstrates that SeqCAN(t, v) is $\Omega(\log v)$.
- Choosing N permutations uniformly and independently at random, the expected number of uncovered t-subpermutations is $\frac{v!}{(v-t)!}\left(\frac{t!-1}{t!}\right)^N$.
- When t is fixed, this shows that SeqCAN(t, v) is $O(\log v)$.
- And indeed, an efficient greedy algorithm produces solutions!

Sequence Covering Arrays

The Existence Question when $t \geq 3$

- There is also one direct and one recursive construction when $t = 3$.
- But for $t \geq 4$, we are currently reliant on algorithmic methods.
- In addition to greedy methods, answer set programming, constraint programming, and cooperative search methods have been applied.

A Post-Optimization Method

- Choose an arbitrary order on the permutations.
- Determine all t-permutations covered by each permutation that is not covered by an earlier one.

Example

SeqCA	First Covered
4 2 0 3 1	031 201 203 231 401 403 420 421 423 431
1 4 3 0 2	102 130 132 140 142 143 302 402 430 432
3 1 2 0 4	104 120 124 204 304 310 312 314 320 324
0 2 4 1 3	013 021 023 024 041 043 213 241 243 413
2 1 3 4 0	134 210 214 230 234 240 340
0 3 4 1 2	012 032 042 034 341 342 412
3 0 1 4 2	014 301
1 4 2 0 3	103 123
3 2 4 1 0	321 410

A Post-Optimization Method

- Choose an arbitrary order on the permutations.
- Determine all t-permutations covered by each permutation that is not covered by an earlier one.
- For each permutation, form a poset on the v elements in which $x \prec y$ when there is some subpermutation in which x precedes y and that is covered for the first time by this permutation.
- Choose an arbitrary linear extension of each poset, and replace the permutation using this linear extension.
- Example: From permutation 1 4 2 0 3, $\{103, 123\}$ has the poset $1 \prec 0$, $1 \prec 2$, $0 \prec 3$, $2 \prec 3$; one linear extension is 4 1 2 0 3.

Example

SeqCA	First Covered
4 2 0 3 1	031 201 203 231 401 403 420 421 423 431
1 4 3 0 2	102 130 132 140 142 143 302 402 430 432
3 1 2 0 4	104 120 124 204 304 310 312 314 320 324
0 2 4 1 3	013 021 023 024 041 043 213 241 243 413
2 1 3 4 0	134 210 214 230 234 240 340
0 3 4 1 2	012 032 042 034 341 342 412
3 0 2 1 4	014 301 321
4 1 2 0 3	103 123 410
3 2 4 1 0	–

A Post-Optimization Method

- Choose an arbitrary order on the permutations.
- Determine all t-permutations covered by each permutation that is not covered by an earlier one.
- For each permutation, form a poset on the v elements in which $x \prec y$ when there is some subpermutation in which x precedes y and that is covered for the first time by this permutation.
- Choose an arbitrary linear extension of each poset, and replace the permutation using this linear extension.
- If there is a permutation that covers no subpermutation for the first time, remove it.
- Repeat the steps above until some stopping criterion is met.

Using the Post-Optimization Method

$t = 4$		
v	Initial	Final
5	26	24
6	34	24
7	41	36
8	44	41
9	52	46
10	57	51
13	71	62
15	78	67
25	104	91
90	180	162

$t = 5$		
v	Initial	Final
6	148	122
7	198	175
8	242	218
9	284	261
10	318	300
11	354	335
12	386	360
13	419	390
15	475	451
20	590	574
30	748	725

Conclusion

- Randomly choosing different linear extensions to alter the structure of the permutation covering appears to provide useful improvements in solutions that were the best known.
- But perhaps this suggests that the other constructions are themselves not particularly good?