Online Supplement

for

Creating clear designs: a graph-based algorithm and a catalog of clear compromise plans

by

Ulrike Grömping, Beuth University of Applied Sciences Berlin, Germany

1. Run times for finding differently structured class 3 compromise plans

This section of the online supplement shows timings for finding 64 run and 128 run clear class 3 compromise plans. The tables compare search times for structures with G1 consisting of the first or the last *m* factors. For 64 runs (Table 1), these are always very similar, whereas for 128 runs, there are sometimes huge differences.

64 runs	*m*=2		*m*=3	
factors	first *m*	last *m*	first *m*	last *m*
7	0.34	0.33	0.33	0.33
8	0.36	0.36	0.36	0.36
9	0.45	0.43	0.43	0.45
10	0.50	0.50	0.5	0.49
11	0.56	0.54	0.57	0.55
12	0.61	0.62	0.14*	0.16*
13	0.70	0.69		
14	0.84	0.85		
15	1.03	1.05		
16	1.19	1.22		
17	1.53	1.53		
18	0.16*	0.16*		

Table 1: Time in seconds for finding the best clear class 3 compromise design in 64 runs (on a dual core 2.4GHz processor Windows PC)
* design does not exist

For 128 runs (Table 2), the left-hand side part of the table shows timings from the small catalog of designs included in the software package **FrF2** (Grömping 2007-2011), the right-hand side part from the much larger complete catalog. The times are subject to random sampling error, which apparently can be large for longer times, as evidenced by comparing times for 128 run 15 factor plans for *m*=3 with the last *m* factors where searching the complete catalog took less time than searching a smaller catalog. This is not an issue, as the timings reported in this section are meant to illustrate substantial differences only.

128 runs	only designs in **R** package **FrF2**				complete 128 run design catalog			
	m=2		*m*=3		*m*=2		*m*=3	
factors	first *m*	last *m*	first *m*	last *m*	first *m*	last *m*	first *m*	last *m*
8	0.36	0.37	0.58	0.34	0.54	0.53	0.76	0.53
9	0.43	0.43	0.42	0.42	0.58	0.59	0.61	0.58
10	0.48	0.48	0.47	0.48	0.64	0.66	0.66	0.66
11	0.55	0.56	0.56	0.55	0.72	0.70	0.72	0.72
12	0.75	0.69	0.71	0.69	0.88	0.87	0.92	0.90
13	0.80	0.81	0.83	0.78	0.98	0.98	0.98	0.99
14	0.92	5.78	0.94	0.92	1.15	5.95	1.28	1.18
15	1.04	5.48	1.07	5157.14	1.43	5.93	1.64	4933.95
16	1.33	34.92	1.36	32273.86	1.98	35.77	3.47	32553.68
17	1.48	14.92	1.53	>174000.00	2.96	16.14	5.12	not run
18	1.86	1.81	1.92	23.58	4.77	4.53	6.31	26.99
19	2.32	2.34	2.34	2.36	12.66	11.94	9.10	7.08
20	2.99	2.98	0.19*	0.15*	16.29	12.19	8.14*	7.61*
21	3.84	3.82			15.83	13.87		
22	2.80	2.79			20.03	16.36		
23	3.31	3.23			29.25	27.63		
24	3.95	3.88			42.64	40.05		

Table 2: Time in seconds for finding the best clear class 3 compromise design in 128 runs (on a dual core 2.4GHz processor Windows PC))
* design does not exist

2. Extending the algorithm to larger designs

Following a referee's suggestion, the 508177 best (MA) 256 run designs with 18 factors from Ryan and Bulutoglu (2010; generators provided through personal communication by K. Ryan; contain all designs with up to 9 words of length 4) have been used for investigating the behavior of the algorithm for situations even larger than 128 runs. The resulting clear compromise plans from this investigation and any clear compromise plans from the MA resolution IV 256 run designs with up to 24 factors have been included into Tables 1 to 4 in the paper's appendix. As complete catalogs have not been available for 256 runs, it is not possible to state non-existence within 256 runs based on these searches. Effort and search times for the 256 run situations are discussed below.

After obtaining and loading the generator file by Ryan and Bulutoglu (2010), word length patterns (up to A_6) were calculated, and the designs were sorted from best to worst, which took about a week in a 32-bit Windows version of **R** software on a 2.4GHz Dual core computer under Windows XP (the code for obtaining the word length patterns was not optimized for efficiency and could presumably be sped up substantially). Note that there is a substantial amount of ties – only 128 of the designs are uniquely defined through their word length pattern with words up to length 6; for an interesting very large tied group of more than 2000 designs, the number of words of length 7 was subsequently used for partial resolving of ties (a sub sample showed that going out to much longer words did not resolve more ties). From a practical point of view, for most purposes the differences in numbers of longer words will not be crucial. Once the sorted designs were available, the CIGs and a few more additional pieces of information were added to them. This task took about three and a half hours – here, the critical aspect was not run time but storage space; the catalog had to be split into two portions (for subsequent usage, the larger portion had to be split even further). The actual search times for clear compromise plans differed strongly between situations: search times for clear class 3 compromise plans were 4.54, 5.31, 38.79, 35.19, 45.53, 34.69, 33.06 seconds, respectively (m_1=5 to 11). Search times for clear class 1 compromise plans were about 37.4, 3100 and 3125 seconds for m_1=9 to 11; for m_1=12, existence of a clear class 1 compromise plan can be inferred from the class 3 plan for m_1=11; however, a search was not successful and was estimated to take years rather than weeks. Thus, the *MA* clear class 1 compromise plan with m_1=12 for 18 factors in 256 runs could not be identified with the algorithm; for practical purposes, the MA clear class 3 plan for 11 factors can be used. A search for MA class 1 compromise plans with m_1 higher than 12 (Ke et al. upper bound: m_1 = 15) was not attempted; MA clear class 4 compromise plans were found within about 30, 7, and 9 seconds (m_1=7 to 9).

Searches of the overall MA 256 run designs in more than 18 factors for clear compromise plans (i.e. the only available design searched for each case) yielded further MA clear compromise plans in 256 runs. The search for clear class 1 compromise plans in 19 factors took about 4, 5 and 10 seconds for finding designs for m_1=5 to 7 and 81 seconds for declaring non-existence of a clear class 1 compromise plan based on the MA design for m_1=8. Further timings were below 10 seconds for finding most designs and at most about 30 seconds for finding or rejecting further 256 run clear class 1 compromise plans for Table 2. Further clear class 3 compromise plans from the 256 run MA designs do not exist, which is immediately obvious from the lack of factors with all 2fis clear and was also confirmed by the algorithm fast (less than a second in every case). Class 4 compromise plans were most demanding, also because – contrary to the other classes – non-existence for a given m_1 does not imply non-existence for larger values of m_1. For some class 4 compromise plans in 256 runs, the fastest strategy to find out whether the MA design can accommodate a clear class 4 compromise plan was to have the algorithm eliminate impossible cases and to manually inspect the graph – as proposed by Wu and Chen – for cases for which the algorithm did not indicate impossibility within a few minutes. Most clear class 4 compromise plans in 256 runs for 19 to 22 factors were found in this way; the interactive graph representation available in **R** package **igraph** (Csardi and Nepusz 2006) proved very helpful for this purpose, particularly when choosing the layout according to Reingold and Tilford (1981). This is illustrated in Figure 1 for the 20 factor MA design: Five groups of four factors are identified which do not interact with each other within group. Consequently, a clear class 3 compromise plan based on this design is impossible. For a clear class 4 compromise plan, factors of the same group must stay together within either G1 or G2; hence, sizes of G1 and G2 can only be multiples of 4. This consideration rules out, for example, the 20 factor clear class 4 compromise plan with 5 factors in G1 that the algorithm neither found nor rejected within about 10 hours. Search for clear class 4 compromise plans with 19 factors and m_1=7 or 8 took very long and was interrupted; the search for m_1=9 was successful. The G1 / G2 subdivision for m_1=9 together with plotting the design's graph allowed to manually find plans for m_1=7 or 8.

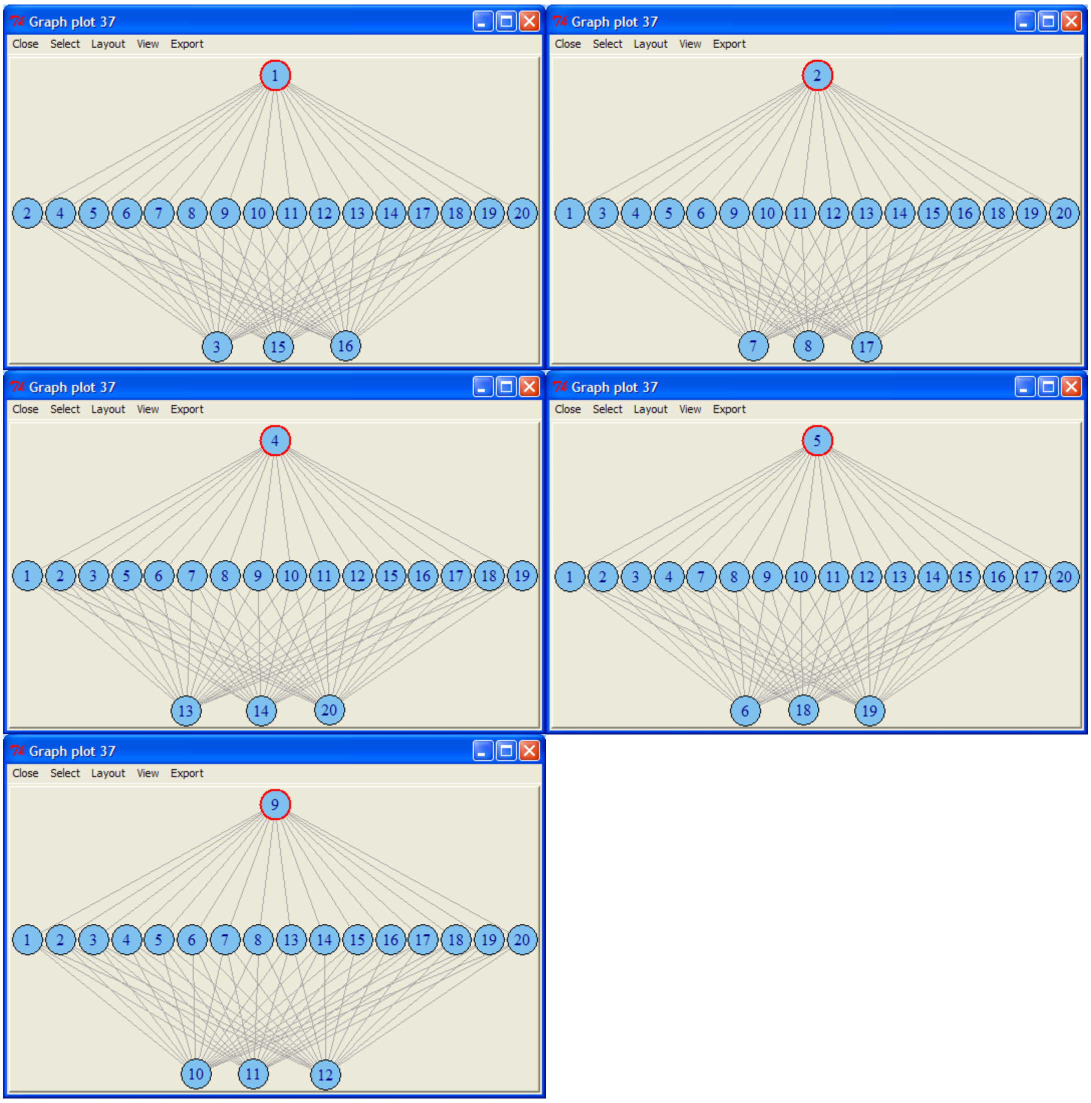

Figure 1: CIG for the MA Design 20-12.1 in 256 runs with 20 factors:
Five representations in Reingold-Tilford layout
(interactive graphs created with **R** package **igraph**; in case of doubt which edges exist, moving vertices usually helps)