

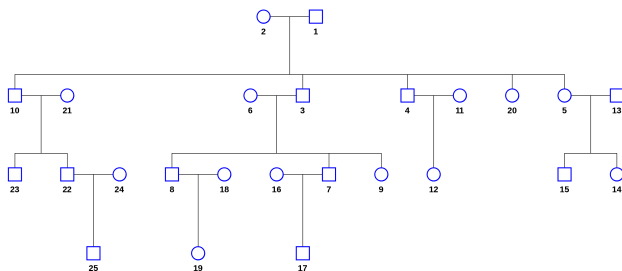
SUPPLEMENTARY MATERIAL

Pedigrees for Haplotyping Evaluation

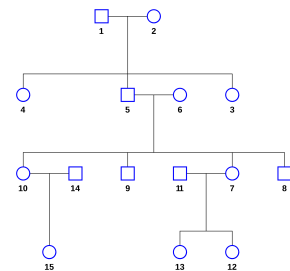
The pedigrees on which the haplotyping performance has been evaluated are shown in Figure S1 below. Also included are structural parameters of the corresponding Bayesian networks as processed by DAOPT (number of variables n , max. domain size k , induced width w).

Notice that although “pedigree19” (Figure S1b) has relatively the smallest parameter values, it is actually the hardest to solve for the haplotyping task. This demonstrates the intricacy of predicting problem complexity for an optimization task using search algorithms like branch-and-bound, and illustrates well the load balancing challenge we had to address. The structural parameters of the problem can predict its difficulty only to a low degree. Namely, optimization problems having a large number of variables, and even high induced width (also known as treewidth) can be very easy to solve if they possess significant levels of determinism (zeros in the probability tables) that can be exploited in search schemes. Even more significantly, the probability values themselves and their interactions can facilitate significant pruning of the search space. Known as *context-sensitivity* in the Bayesian network literature, neither issue is captured by structural parameters.

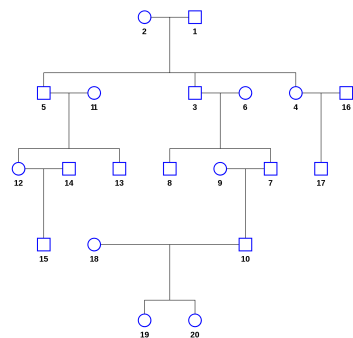
We note that likelihood problems on the other hand (which correspond to LOD score queries) exhibit far less context-sensitivity and therefore their performance is often better captured by the structural parameters.



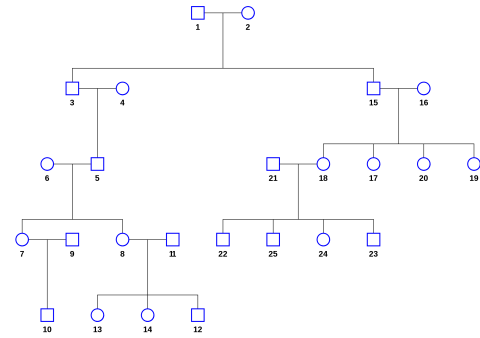
(a) *pedigree7*, 25 individuals, 20 markers ($n = 1068, k = 4, w = 32$).



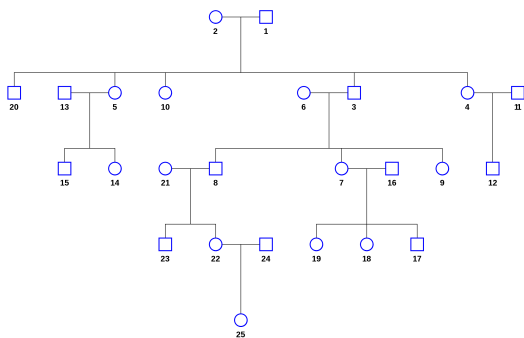
(b) *pedigree19*, 15 individuals, 20 markers ($n = 793, k = 5, w = 25$).



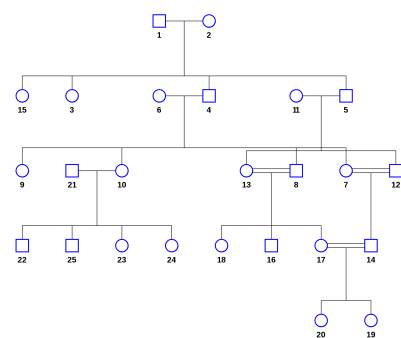
(c) *pedigree13*, 20 individuals, 20 markers ($n = 1077, k = 3, w = 32$).



(d) *pedigree31*, 25 individuals, 20 markers ($n = 1183, k = 5, w = 30$).



(e) *pedigree41*, 25 individuals, 20 markers ($n = 1062, k = 5, w = 33$).



(f) *pedigree51*, 25 individuals, 20 markers ($n = 1152, k = 5, w = 39$).

Fig. S1: Pedigrees used for haplotyping performance evaluation. Structural parameters of each pedigree’s corresponding Bayesian network are reported as number of variables n , max. domain size k , and induced width w .