Unsolved Problems in Graph Theory Arising from the Study of Codes*

N. J. A. Sloane

Mathematical Sciences Research Center
AT&T Bell Laboratories
Murray Hill, NJ 07974

1. Two fundamental questions in coding theory

Two of the most basic questions in coding theory are: (i) what is $A(n,d)$, the maximal number of binary vectors of length n with Hamming distance d apart?, and (ii) what is $A(n,d,w)$, the maximal number of binary vectors of length n, Hamming distance d apart, and where each vector contains precisely w ones?

The *Hamming distance* $\mathrm{dist}(u,v)$ between two binary vectors $u=(u_1, \ldots, u_n)$ and $v=(v_1, \ldots, v_n)$ is the number of i such that $u_i \neq v_i$; the *weight* $\mathrm{wt}(u)$ of u is the number of nonzero u_i (so that $\mathrm{dist}(u,v)=\mathrm{wt}(u-v)$); and a code in which $\mathrm{dist}(u,v) \geq d$ for every pair of distinct vectors u,v can correct

$$\left\lfloor \frac{d-1}{2} \right\rfloor$$

errors. We may always assume (without loss of generality) that d is even. See [25] for further information about codes.

A. E. Brouwer, J. B. Shearer, W. D. Smith and the author have recently made an extensive study [4] of the functions $A(n,d)$ and $A(n,d,w)$, and in particular have computed a table of lower bounds on these functions for $n \leq 28$. Portions of these tables are shown in Tables 1 and 2 below. In the course of this work a number of unsolved graph theory problems were encountered.

---

* This paper appeared in *Graph Theory Notes of New York*, Vol. **18**, 1989, pp. 11-20.

2. Finding maximal cliques

The *Hamming graph* $H(n,d)$ has 2^n vertices labeled by the binary vectors of length n, two vertices being joined by an edge if and only if the Hamming distance between the corresponding vectors is at least d. Then $A(n,d)$ is simply the size of a maximal clique in $H(n,d)$.

Table 1, which is taken from [4], shows the best lower bounds presently known on $A(n,d)$ for $n \le 28$. A period after an entry indicates that this is the exact value of $A(n,d)$. In particular it can be seen that the exact value is known for all $n \le 10$.

For example, it is known that $A(16,8) = 32$. In this case the optimal code is a first-order Reed-Muller code of length 16, formed from the rows of a Hadamard matrix of order 16 and its negative [25]. On the other hand for $A(17,8)$ it is known only that

$$36 \le A(17,8) \le 37 .$$

(Upper bounds on $A(n,d)$ can be found for example in [25] or [5], p. 248.)

The first undetermined value is $A(11,4)$, for which it is known only that

$$72 \le A(11,4) \le 79 .$$

$H(11,4)$ is a graph with only 2048 vertices: would some graph theorist please determine the size of a largest clique in it? (A number of algorithms for clique-finding have been published in recent years [1], [2], [14], [15], [26].) Every entry in Table 1 not followed by a period is an unsolved problem of this type.

If a reader should find a larger clique than is presently known in any of these problems, please write down the vertices of the clique (which form a record-breaking code), and notify the author!

TABLE 1
Lower bounds on $A(n,d)$

n,d	4	6	8	10	12	14	16	18	20
5	2.	1.	1.	1.	1.	1.	1.	1.	1.
6	4.	2.	1.	1.	1.	1.	1.	1.	1.
7	8.	2.	1.	1.	1.	1.	1.	1.	1.
8	16.[1]	2.	2.	1.	1.	1.	1.	1.	1.
9	20.[2]	4.	2.	1.	1.	1.	1.	1.	1.
10	40.[3]	6.	2.	2.	1.	1.	1.	1.	1.
11	72	12.	2.	2.	1.	1.	1.	1.	1.
12	144[4]	24.[8]	4.	2.	2.	1.	1.	1.	1.
13	256.	32.[8a]	4.	2.	2.	1.	1.	1.	1.
14	512.	64.	8.	2.	2.	2.	1.	1.	1.
15	1024.	128.	16.	4.	2.	2.	1.	1.	1.
16	2048.[1]	256.[9]	32.[14]	4.	2.	2.	2.	1.	1.
17	2720[5]	256	36[2]	6.[17]	2.	2.	2.	1.	1.
18	5248	512	64	10.	4.	2.	2.	2.	1.
19	10496[6]	1024	128	20.	4.	2.	2.	2.	1.
20	20480[7]	2048[10]	256	40.[8]	6.[17]	2.	2.	2.	2.
21	36864	2560[11]	512.	42[18]	8.[17]	4.	2.	2.	2.
22	73728	4096	1024.	48[17]	12.	4.	2.	2.	2.
23	147456	8192	2048.	68[19]	24.	4.	2.	2.	2.
24	294912[7]	16384[12]	4096.[15]	128[20]	48.[8]	6.[17]	4.	2.	2.
25	524288	16384	4096	151[21]	52[2]	8.[17]	4.	2.	2.
26	1048576	32768	4096	256	64	14.	4.	2.	2.
27	2097152	65536	8192	512	128[23]	28.	6.[17]	4.	2.
28	4194304[1]	131072[13]	16384[16]	1024[22]	128	56.[8]	8.[17]	4.	2.

KEY TO TABLE 1

Unmarked entries are either trivial or are obtained by shortening the code below.
An entry followed by a period is known to be exact.

1 = Extended Hamming code ([25], p. 23).
2 = Conference matrix code ([25], p.585).
3 = Found by M. Best ([5], p. 140).
4 = From the Steiner system $S(5,6,12)$ ([5], p. 139, [25], p. 585).
5 = Romanov – see Sect. VI of [4].
6 = From Hamming code over GF(5) [16].
7 = From the $u \mid u+v$ construction ([25], p. 76).
8 = Hadamard matrix code ([25], p. 49).
8*a* = ‘‘Nadler’’ code ([25], pp. 75, 79).
9 = Nordstrom-Robinson code ([25], p. 73).
10 = Nonlinear code from Construction X ([25], p. 583).
11 = From Construction X4 ([25], p. 585, Example 7).
12 = Wagner [34].
13 = Shortened non-primitive BCH code of length 32 ([25], p. 586).
14 = Reed-Muller code ([25], Chap. 13).
15 = Golay code ([5], Chaps. 3,11, [25], Chap. 20).
16 = Self-dual double circulant code ([5], p. 189, [25], p. 509).
17 = From Hadamard matrices using Levenshtein’s construction ([25], p. 50).
18 = Extended quasi-cyclic code [18].
19 = Extended cyclic code [19] (see Table 11 of [4]).
20 = Hashim-Pozdniakov linear code [17].
21 = Cyclic code (see Table 11 of [4]).
22 = Piret [29].
23 = Linear code (Eq. (51) of [4], [25], p. 593).

Similar problems arise in studying constant weight codes. The *Johnson graph* $J(n,d,w)$ has $\binom{n}{w}$ vertices labeled by binary vectors of length n and weight w, two vertices being joined by an edge if and only if the Hamming distance between the corresponding vectors is at least d. Equivalently, the vertices represent w-subsets of an n-set, two vertices being joined by an edge if and only if the corresponding subsets intersect in at most $w - \frac{1}{2} d$ points. Then $A(n,d,w)$ is the size of a maximal clique in $J(n,d,w)$.

Table 2, also taken from [4], shows the best lower bounds on $A(n,10,w)$ for $n \le 28$. Again every entry not followed by a period is an open problem. The first open case is $A(20,10,9)$, where we know only that

$$20 \le A(20,10,9) \le 24 \ .$$

The lower bound is a cyclic code: take all cyclic shifts of the vector

00010001001010011111

(see Table 11 of [4]), while the upper bound comes from [3].

It should be said at this point that there is a tiny bit of doubt about some of the upper bounds on $A(n,d,w)$ in [3] in the case $d = 10$; we are planning to recompute them. While it is usually easy to verify a lower bound on $A(n,d)$ or $A(n,d,w)$ (by checking the Hamming distance between the codewords) upper bounds are much harder to verify.

TABLE 2
Lower bounds on $A(n,10,w)$

n,w	6	7	8	9	10	11	12	13	14
12	2.	2	1	1	1	1	1	0	0
13	2.	2	2	1	1	1	1	1	0
14	2.	2.	2	2	1	1	1	1	1
15	3.[j]	3.[j]	3	3	3	1	1	1	1
16	3.	4.[j]	4.[j]	4	3	3	1	1	1
17	3.	5.[j]	6.[j]	6	5	3	3	1	1
18	4.[j]	6.[j]	9.[q2]	10.[s]	9	6	4	3	1
19	4.	8.[x]	12.[sb]	19.[c]	19	12	8	4	3
20	5.[s]	10.[q2]	17.[m]	20[c]	38.[hm]	20	17	10	5
21	7.[a]	13.[xh]	21.[c]	27[pc]	38	38	27	21	13
22	7.	16.[pc]	24[sd]	35[pc]	42[ec]	46[c]	42	35	24
23	8.[x2]	20[y]	33[pc]	45[pc]	54[pc]	63[pc]	63	54	45
24	9.[x2]	24[c]	38[pc]	56[c]	72[c]	90[pc]	96[c]	90	72
25	10.[s]	28[ec]	48[ec]	72[ec]	100[c]	125[c]	130[ec]	130	125
26	13.[q2]	28	54[pc]	84[pc]	130[c]	168[pc]	185[y]	191[y]	185
27	14.[q9]	36[q3]	66[pc]	111[c]	159[pc]	213[ya]	257[y]	283[ya]	283
28	16.[m]	37[q4]	78[pc]	132[pc]	195[yd]	280[ya]	356[ya]	414[ya]	435[yd]

KEY TO TABLE 2
An entry followed by a period is known to be exact.
Section and table references are to [4].

a = From a trivial design or its dual (Sect. III).
c = Cyclic code (Table 11).
ec = Extended cyclic code (Table 12).
hm = Hadamard matrix code (Theorem 10).
j = Juxtaposing (Eq. (1) of Sect. III).
m = Miscellaneous construction (Sect. XI).
pc = Orbits under a single permutation (Table 14).
qi = Quasi-cyclic code, for $2 \le i \le 9$ – fixed by a permutation containing *i* cycles of length *n/i* (Table 13).
s = Section of code below or diagonally down to right, obtained from (5) of Sect. III.
sb = Section of code below, obtained by direct examination of the code (Sect. III).
sd = Section of code diagonally down to right, obtained by direct examination of the code (Sect. III).
x = Lexicographic code (Sect. VIII).
xh = Lexicode with seed (Table 8).
*x*2 = Complement of lexicode with sum constraint (Table 7).
y = No known structure (Table 16).
ya = Obtained by extending the code above it in the table; no other structure.
yd = Obtained by extending the code diagonally above it to left; no other structure.

3. Finding maximal weighted cliques

Weighted clique problems arise in the following way. In many cases a code containing the maximal number ($A(n,d,w)$) of vectors has a nontrivial symmetry group. For example $A(18,8,6)=21$ is realized by the code

(110100)(100000)(110000),
(000010)(110100)(100001),
(000011)(100001)(010100),
(010101)(010101)(000000),
(000000)(000000)(111111).

The parentheses indicate that the permutation

$$(1,2,3,4,5,6)(7,8,9,10,11,12)(13,14,15,16,17,18)$$

of order six is to be applied to the indicated vectors. The first three vectors each give rise to six codewords, the fourth vector to two codewords and the last vector to a single codeword (since it is fixed by the permutation).

To generalize this construction, we choose some permutation group G on n letters, and divide the vectors of length n into *orbits* under G (two vectors are in the same orbit if and only if some permutation in G sends one to the other). We form a weighted graph as follows. The vertices represent good orbits (orbits in which the Hamming distance between any two distinct vectors is at least d), and two vertices are joined by an edge if and only if the Hamming distance between every vector in one orbit and every vector in the other orbit is at least d. Each vertex is weighted by the size of the corresponding orbit. Then the size of a maximal weighted clique in this graph is the largest constant weight code invariant under G.

Tables 11-15 of [4] contain many examples of codes found in this way. Better weighted-clique finding algorithms should produce many more.

4. Graph coloring problems

Besides looking for group-invariant codes we used many other constructions in [1], as can be seen in Table 2 above. A particularly powerful construction, applicable to codes with $d=4$, is the *partitioning construction* (see [4], [13]). For this one needs good colorings of the graphs $H(n,4,w)$.

Let $\Pi(n,w)=(X_1, \ldots, X_m)$ be a collection of disjoint sets or *color classes* $X_1, \ldots, X_m$, each of which is a code of length n, distance 4 and constant weight w, and whose union contains all $\binom{n}{w}$ vectors of weight w. In other words $\Pi(n,w)$ is a coloring of the graph $H(n,4,w)$. We assume $|X_1| \geq \cdots \geq |X_m|$. The vector $\pi(n,w)=(|X_1|, \ldots, |X_m|)$ is the *index vector* of $\Pi(n,w)$, and

$$\pi(n,w) \cdot \pi(n,w) = \sum_{i=1}^{m} |X_i|^2$$

is its *norm*. When several different colorings are known for a given n and w we denote them by $\Pi_1(n,w)$, $\Pi_2(n,w)$, ..., and their index vectors by $\pi_1(n,w)$, $\pi_2(n,w)$, ...

The reader is referred to [4] or [13] for details of the partitioning construction. The important point here is that the best colorings $\Pi(n,w)$ to use in the construction are those that are maximal in the following sense. We say that one coloring $\Pi(n,w)$ with index vector $\pi(n_1, w_1) = (a_1, ..., a_m)$ *dominates* another $\Pi(n', w')$ with index vector $\pi'(n_1, w_1) = (b_1, ..., b_{m'})$ if and only if

$$\sum_{i=1}^{j} a_i \geq \sum_{i=1}^{j} b_i$$

holds for all $j = 1, ..., \max\{m, m'\}$. A coloring is *maximal* if it is not dominated by any other.

In [4] we made an extensive investigation of colorings with small values of n, and found over a thousand $\Pi(n,w)$ with $n \leq 16$, no one of which is dominated by any other. A portion of this list is shown in the following table.

TABLE 3
Good colorings $\Pi(n,w)$

n	w	i	m	Norm	Notes	Index vector of $\Pi_i(n, w)$
6	3	1	6	72	*	4,4,4,4,2,2
7	3	1	6	211	*	7,7,6,6,5,4
8	3	1	7	448	*	8,8,8,8,8,8,8
8	4	1	6	844	*	14,14,12,12,10,8
9	3	1	7	1008	*	12,12,12,12,12,12,12
9	4	1	8	2066		18,18,18,18,16,15,15,8
9	4	2	10	2036		18,18,18,18,18,14,13,7,1,1
10	3	1	10	1530	*	13,13,13,13,13,13,13,13,13,3
10	4	1	10	5620	[13]	30,30,30,30,30,22,22,12,2,2
10	4	2	9	5614		30,30,30,30,26,25,22,15,2
10	5	1	8	8044		36,36,34,34,29,29,27,27

An asterisk indicates that the coloring is known to be maximal. In general it does not seem that there is unique maximal coloring for given values of n and w. In some rare cases it is possible to color $H(n,4,w)$ so that each color class is a t-design: such colorings are said to form a *large set* of designs. See [6]-[13], [20]-[24], [27], [28], [30]-[33], and other references cited in [4].

The problem of finding a good coloring is a generalization of usual problem of finding a good code. *Each* color class is code with $d = 4$, so now the goal is to find a small number of large disjoint codes.

We would like to know whether any of the colorings in Table 3 not marked with an asterisk (or those in Table 6 of [4]) can be improved.

A more important problem, however, is to bring some order into this subject: at present almost all the best colorings known have no mathematical structure, and can be described only by listing the vectors in each color class.

BIBLIOGRAPHY

[1] E. Balas, V. Chvátal and J. Nesetril, ‘‘On the maximum weighted clique problem’’, Math. of Operations Research, vol. 12 (1987), pp. 522-535.

[2] E. Balas and C. S. Yu, ‘‘Finding a maximal clique in an arbitrary graph’’, SIAM J. Comput., vol. 15 (1986), ppl 1054-1068.

[3] M. R. Best, A. E. Brouwer, F. J. MacWilliams, A. M. Odlyzko and N. J. A. Sloane, ‘‘Bounds for binary codes of length less than 25’’, IEEE Trans. Info. Theory, vol. 24 (1978), pp. 81-93.

[4] A. E. Brouwer, J. B. Shearer, N. J. A. Sloane and W. D. Smith, ‘‘A new table of constant weight codes’’, IEEE Trans. Info. Theory, to appear.

[5] J. H. Conway and N. J. A. Sloane, ‘‘Sphere Packings, Lattices and Groups’’, Springer-Verlag, N.Y. 1988.

[6] R. H. F. Denniston, ‘‘Some maximal arcs in finite projective planes’’, J. Comb. Theory, vol. 6 (1969), pp. 317-319.

[7] R. H. F. Denniston, ‘‘Some packings with Steiner triple systems’’, Discrete Math., vol. 9 (1974), pp. 213-227.

[8] R. H. F. Denniston, ‘‘Sylvester’s problem of the 15 schoolgirls’’, Discrete Math., vol. 9 (1974), pp. 229-233.

[9] R. H. F. Denniston, ‘‘Some new 5-designs’’, Bull. Lond. Math. Soc., vol. 8 (1976), pp. 263-267.

[10] R. H. F. Denniston, ‘‘Four doubly resolvable complete 3-designs’’, Ars Comb., vol. 7 (1979), pp. 265-272.

[11] R. H. F. Denniston, ‘‘Enumeration of symmetric designs (25, 9, 3)’’, Annals of Discrete Math., vol. 15 (1982), pp. 111-127.

[12] T. Etzion and A. Hartman, ‘‘Towards a large set of Steiner quadruple systems’’, preprint.

[13] T. Etzion and C. L. M. van Pul, ‘‘New lower bounds for constant weight codes’’, preprint.

[14] L. Gerhards and W. Lindenberg, ‘‘Clique detection for nondirected graphs: two new algorithms’’, Computing, vol. 21 (1979), pp. 295-322.

[15] M. Goldberg and T. Spencer, ‘‘A new parallel algorithm for the maximal independent set problem’’, SIAM J. Comput., vol. 18 (1989), pp. 419-427.

[16] H. Hämäläinen, ‘‘Two new binary codes with minimum distance three’’, IEEE Trans. Info. Theory, vol. 34 (1988), p. 875.

[17] A. A. Hashim and V. S. Pozdniakov, ‘‘Computerized search for binary linear codes’’, Electronic Lett., vol. 12 (1976), pp. 350-351.

[18] M. K. Kaikkonen, ‘‘A New 4-Error Correcting Code of Length 20’’, preprint.

[19] M. K. Kaikkonen, personal communication.

[20] E. S. Kramer, ‘‘Some triple system partitions for prime powers’’, Utilitas Math., vol. 12 (1977), pp. 113-116.

[21] E. S. Kramer and S. S. Magliveras, ‘‘Some mutually disjoint Steiner systems’’, J. Comb. Theory, vol. 17A (1974), pp. 39-43.

[22] E. S. Kramer and D. M. Mesner, ‘‘Intersections among Steiner systems’’, J. Comb. Th., vol. 16A (1974), pp. 273-285.

[23] E. S. Kramer and D. M. Mesner, ‘‘The possible (impossible) systems of 11 disjoint $S(2,3,13)$’s ($S(3,4,14)$’s)’’, Utilitas Math., vol. 7 (1975), pp. 55-58.

[24] J. X. Lu, ‘‘On large sets of disjoint Steiner triple systems I-VI’’, J. Comb. Theory, vol. 34A (1983), pp. 140-146, 147-155, 156-183; vol. 37A (1984), pp. 136-163, 164-188, 189-192.

[25] F. J. MacWilliams and N. J. A. Sloane, ‘‘The Theory of Error-Correcting Codes’’, North-Holland, Amsterdam, 1979.

[26] P. M. Pardalos and G. P. Rodgers, ‘‘A branch and bound algorithm for the maximum clique problem’’, preprint.

[27] K. T. Phelps, ‘‘On cyclic Steiner quadruple systems $S(3,4,20)$’’, Annals of Discrete Math., vol. 7 (1980), pp. 277-300.

[28] K. T. Phelps, ‘‘A class of large sets of Steiner triple systems of order 15’’, preprint.

[29] P. Piret, ‘‘Good linear codes of length 27 and 28’’, IEEE Trans. Info. Theory, vol. 26 (1980), p. 227.

[30] L. Teirlinck, ‘‘On the maximum number of disjoint Steiner triple systems’’, Discrete Math., vol. 6 (1973), pp. 299-300.

[31] L. Teirlinck, ‘‘On making two Steiner triple systems disjoint’’, J. Comb. Theory, vol. 23A (1977), pp. 349-350.

[32] L. Teirlinck, ‘‘On large sets of disjoint ordered designs’’, Ars Comb., vol. 25 (1988), pp. 31-37.

[33] L. Teirlinck, ‘‘A completion of Lu’s determination of the spectrum for large sets of disjoint Steiner triple systems’’, preprint.

[34] T. J. Wagner, ‘‘A search technique for quasi-perfect codes’’, Info. Control, vol. 9 (1966), pp. 94-99.

Unsolved Problems in Graph Theory Arising from the Study of Codes*

N. J. A. Sloane

Mathematical Sciences Research Center
AT&T Bell Laboratories
Murray Hill, NJ 07974

ABSTRACT

Recent work on binary codes has revealed a number of unsolved problems in graph theory. Three types of problems arise: finding maximal cliques in certain graphs, finding maximal weighted cliques, and finding good colorings.

* This paper appeared in *Graph Theory Notes of New York*, Vol. **18**, 1989, pp. 11-20.