KNIGHT-GRAPHABLE WORDS

TRISTAN MILLER
Darmstadt, Germany
http://www.nothingisreal.com/

MIKE KEITH
New Jersey, USA
http://www.cadaeic.net/

Imagine an arbitrarily large chessboard where some of the squares are labelled with distinct letters of the alphabet. A word is said to be *knight-graphable* (NG) if there exists some labelling such that a knight can spell out the word by starting on the square labelled with its first letter, then making a valid move to the square labelled with its second letter, and so on until the square labelled with the last letter is reached. Repeated runs of the same letter are permitted, for which we assume the knight simply stays put for the "move". (See Figure 1 for an example.) This article considers how we can characterize mathematically the set of knight-graphable words and what the longest knight-graphable words in English are.

Let's begin by considering the easiest cases. It is obvious that all words with only one or two distinct letters are NG. However, not all words with exactly three distinct letters are NG—for example, consider the word DEAD. Figure 2 illustrates the only distinct chessboard labelling where D is one knight move away from E (all the others being translations, rotations, and reflections). If the knight starts on the D and moves to the E, then the only possible places to put the A are marked by the black circles—these are the squares one knight move away from E. However, none of these squares are one knight move away from D, which is the next letter which must be visited after A.

In general, we can observe that it is not possible for the third distinct letter in a word to be adjacent to both the first and second distinct letters, because a knight that has made two moves away from a particular starting square requires at least two more moves to return to it. To convince yourself of this, consider that every time a knight moves, it changes square colour. A knight starting on a black square and making two moves away must now be on another black square. From there it can move only to a white square, which excludes the starting black square as an immediate destination.

Considering words with more than three distinct letters gets somewhat complicated unless we formalize the problem mathematically. We therefore adopt and extend the approach employed by Mike Keith to study the related problems of king- and queen-graphable words [2]. Specifically, we define an alphabet A as a set of letters, and a *word graph* for a word with $1 \leq n \leq |A|$ distinct letters as the connected graph of order n that has an edge (a_1, a_2) if the letters $a_1, a_2 \in A$ are adjacent to each other in the word, irrespective of order. We define the *infinite knight's graph* to be the graph representing all knight moves on an infinite chessboard, where each vertex represents a square and

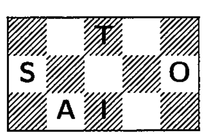

Figure 1: A chessboard labelling for TATTOOISTS which corresponds to a valid knight subgraph

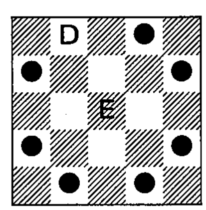

Figure 2: An attempted chessboard labelling for DEAD

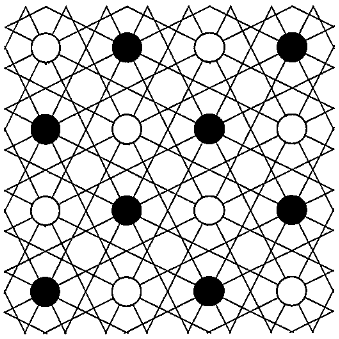

Figure 3: Detail of the infinite knight's graph

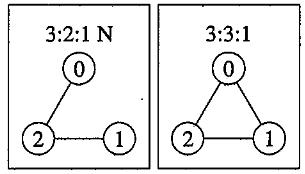

Figure 4: All connected graphs with 3 vertices. In this and subsequent figures, each graph is marked with a unique identifier of the format n:e:i, where n is the number of vertices, e is the number of edges, and i is the graph's index in nauty's [4] enumeration of all connected graphs with n vertices and e edges. The letter N follows the identifier of graphs which are knight subgraphs.

each edge is a legal move (see Figure 3). A graph is called a *knight subgraph* if its vertices can be mapped to different squares of a chessboard such that the connected pairs of vertices in the graph are a knight's move apart. More formally, a graph is a knight subgraph if it is an isomorphism of some connected subgraph of the infinite knight's graph. Finally, a word is knight-graphable (NG) if its word graph is a knight subgraph.

The number of distinct connected graphs of each order n, starting with $n = 2$, is 1, 2, 6, 21, 112, 853, 11 117, 261 080, ... [1], and so any word graph must be a labelling of one of these graphs of the same order. Similarly, every finite subgraph of the infinite knight's graph must also correspond to one of the connected graphs. There are a couple properties of the infinite knight's graph which are worth noting: first, every vertex has exactly eight edges. Second, if, as in Figure 3, we colour each vertex as black or white according to the chessboard square it represents, we observe that every edge joins a white vertex to a black vertex; there are no edges which directly link two white vertices or two black vertices. Any graph whose vertices can be coloured in this manner is said to be *bipartite*. We can therefore make the following statement about NG words:

> *The word graph of a knight-graphable word is bipartite, and the degree of every vertex is 8 or less.*

This reduces the number of candidate graphs, starting at $n = 2$, to 1, 1, 3, 5, 17, 44, 182, 730, 4 031, 25 588, 212 646, 2 239 096, ...

Figure 4 shows the two connected graphs for $n = 3$; of these, only the first is a knight subgraph. The second graph corresponds to (among others) the word DEAD, which we showed earlier to be not knight-graphable. Another way of concluding that the second graph is not a knight subgraph is to note that its triangle forms a cycle of length 3. It can be shown that any graph that contains a cycle with an odd number of edges is not bipartite, so this graph is not bipartite and therefore not a knight subgraph.

Given the two order-3 graphs, we can work out how many words with three distinct letters are NG. The second edition of the North American Scrabble tournament's Official Tournament and Club Word List (TWL06) used by Keith [2] contains 2 147 words with three distinct letters. Of these 1 874 (about 87%) have a word graph corresponding to the first graph, and are therefore knight-

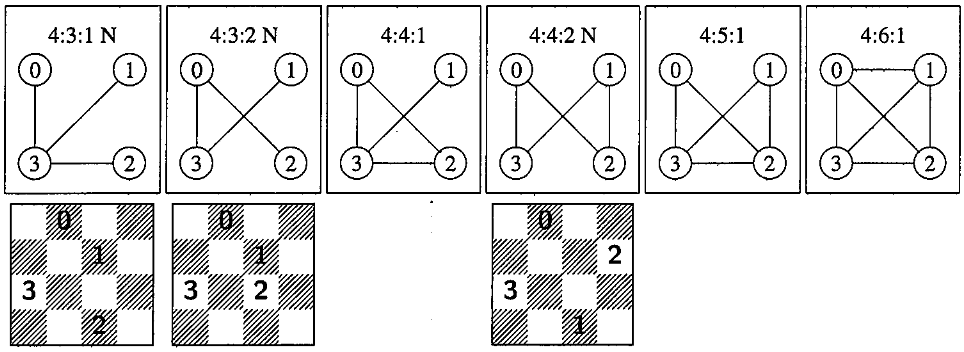

Figure 5: All connected graphs with 4 vertices, and corresponding chessboard labellings.

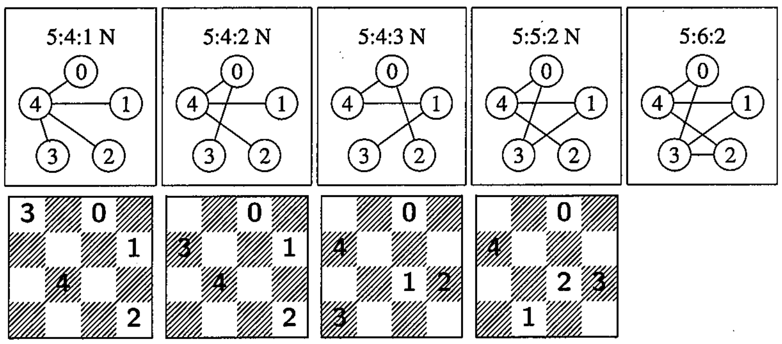

Figure 6: All connected bipartite graphs with 5 vertices, and corresponding chessboard labellings.

graphable. The longest of these words is the 9-letter SEERESSES. This happens to be longer than the longest non-NG word in the set—the 8-letter GREEGREE. If we expand our search beyond TWL06, we find the 11-letter NG word KINNIKINNIK.

We now proceed to the six order-4 graphs, which are shown at the top of Figure 5. Graphs 4:4:1, 4:5:1, and 4:6:1 contain triangles and therefore cannot be knight subgraphs; for each of the other three graphs the figure shows a sample chessboard labelling as proof that they are knight subgraphs. TWL06 lists 8 463 words with four distinct letters. The number of words corresponding to the six graphs, in order, are 812, 5 287, 1 403, 758, 202, and 1, so a total of 6 857 words (81%) are knight-graphable. The longest NG word in this group is SERENENESSES (Graph 4:3:1, 12 letters) and the longest non-NG word is SENSELESSNESSES (Graph 4:4:1, 15 letters). We can find nothing longer by consulting sources beyond TWL06.

For the case $n = 5$ there are 21 connected graphs; Keith [2] illustrates the complete set but here in Figure 6 we reproduce only the five which are bipartite. It is easy to find chessboard labellings for the first four of these (also in Figure 6), but Graph 5:6:2 presents us with a problem. We can see that this graph contains two cycles of length 4, one containing the edges $(0,3)$, $(2,3)$, $(2,4)$, and $(0,4)$, and the other $(1,4)$, $(2,4)$, $(2,3)$, and $(1,3)$. Note that the two cycles have two edges in common. But we have the following theorem:

> **Theorem:** *In a knight subgraph, two distinct 4-cycles cannot share more than one edge.*

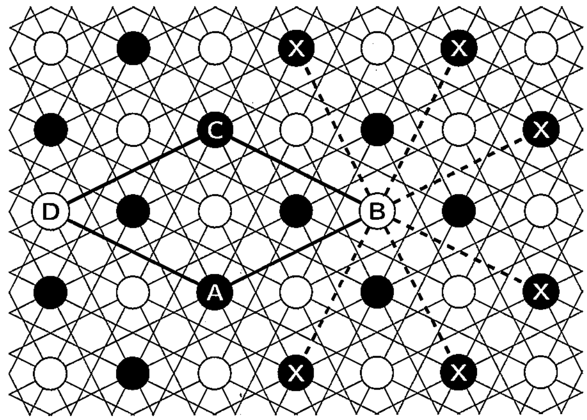

Figure 7: Proof that in a knight subgraph, two distinct 4-cycles cannot share more than one edge.

Proof: Without loss of generality, take one of the 4-cycles to be the one marked with thick lines (*ABCDA*) in Figure 7. Let's assume the theorem is false, and try to construct a second distinct 4-cycle sharing more than one edge with *ABCDA*. Without loss of generality we can start this 4-cycle on vertex *A* and use edge *AB* for the first edge (an edge that's shared with the first 4-cycle). We can't take *BC* as the second edge because then the second 4-cycle has to be *ABCDA*, which violates the "distinct" part of the theorem. So the second edge has to be one of the dashed ones. The third edge can't be *BC*, *CD*, or *DA*, so the fourth edge of the second 4-cycle has to be the one that's shared with the first 4-cycle; that means that the fourth edge of the second 4-cycle has to be *DA* and the cycle has to go through *D*. This means there has to be an edge between one of the vertices marked *X* and vertex *D*. But no such edge exists. We have arrived at a contradiction, thus completing the proof.

Given the above, we can conclude that Graph 5:6:2 is not a knight subgraph, and we can refine our characterization of what constitutes a knight-graphable word:

NG theorem: *The word graph of a knight-graphable word is bipartite, has a maximum degree of 8, and has no two 4-cycles which share more than one edge.*

This theorem reduces the number of possible knight subgraphs, starting at $n = 2$, to 1, 1, 3, 4, 12, 26, 88, 257, ...

TWL06 lists 19 965 words with five distinct letters, of which 13 440 (67%) are knight-graphable. The longest NG words are of length 12, and there are seven of them: EFFETENESSES (Graph 5:4:1), JEJUNENESSES (Graph 5:5:2), SEVERENESSES (Graph 5:4:1), SODDENNESSES (Graph 5:5:2), SUDDENNESSES (Graph 5:5:2), SULLENNESSES (Graph 5:5:2), and UNEVENNESSES (Graph 5:4:1). However, there are thirteen words which are even longer but not knight-graphable, topped by the 15-letter SLEEPLESSNESSES (Graph 5:7:2). Going beyond TWL06, we find the 13-letter NG word LOLLAPALOOSAS (Graph 5:5:2).

Figure 8 shows all 17 connected bipartite graphs with six vertices. Five of these (Graphs 6:7:2, 6:7:12, 6:8:2, 6:8:16, and 6:9:13) contain two 4-cycles with more than one shared edge, so by

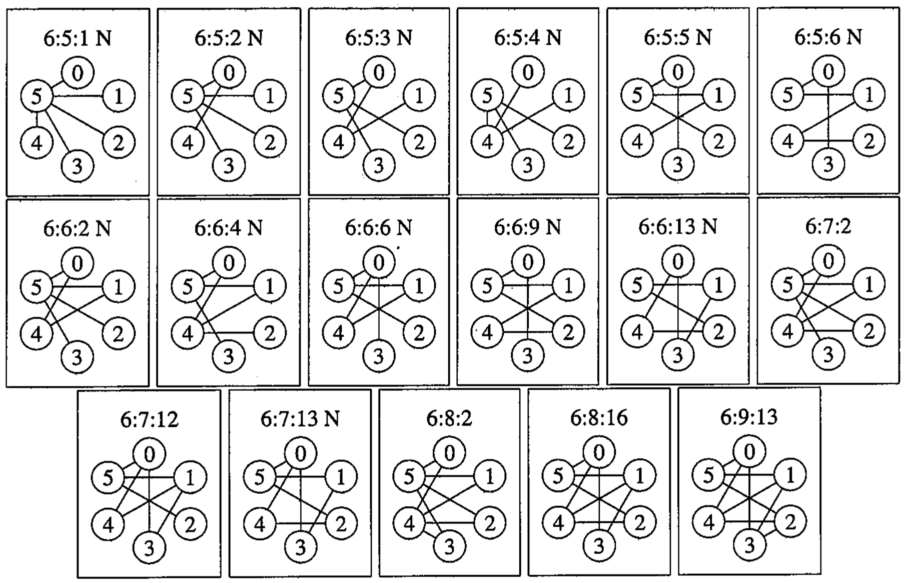

Figure 8: All connected bipartite graphs with 6 vertices.

the NG theorem they are not knight subgraphs. We have verified empirically that the remaining twelve are all NG, so for $n = 6$ the necessary conditions of the NG theorem are also sufficient and provide a complete characterization of NG words. There are 33 562 words with six distinct letters in TWL06; of these 20 981 (63%) are NG. The longest NG words in this list are the 14-letter CHINCHERINCHEE (Graph 6:7:13), DIVISIVENESSES (Graph 6:6:4), and FEMININENESSES (Graph 6:6:2). There are 19 non-NG words which are even longer, all with 15 letters, such as INSTANTANEITIES (Graph 6:10:2). Outside of TWL06 there are a couple more length-14 NG words: MISSISSIPPIANS (Graph 6:6:2) and IMMINENTNESSES (Graph 6:5:5). For non-NG words we find the 17-letter INSENSITIVENESSES (Graph 6:8:10) and DEFENSELESSNESSES (Graph 6:6:1).

We verified by computer analysis that, somewhat surprisingly, the necessary conditions in the NG theorem are also sufficient for $n = 7$, 8, and 9. We did not exhaustively investigate $n = 10$ or higher, but we strongly suspect that the conditions in the NG theorem will need to be expanded to handle these larger values of n.

There are too many knight subgraphs of order 7 and above to illustrate here, though Table 1 gives the full count for orders up to 9, and Table 2 shows the number and percentage of NG words in TWL06 for orders up to 15. Table 3 gives the longest NG words for each order, both for TWL06 and for a broader search. (Which words are non-NG ceases to be so interesting as the number of distinct letters n increases. NG words get rarer as n increases, and we have empirically observed that the non-NG ones are just the longest words for a given n.)

There is one missing entry from Table 3 which we reserve as a puzzle for the reader: Figure 9 shows a chessboard labelling for the single longest NG word of 14 distinct letters. It is 22 letters long and begins with an H, and is presently the longest known knight-graphable word in English. What is it? The solution can be found in this issue's *Answers and Solutions*.

n	$\lvert G_C\rvert$	$\lvert G_B\rvert$	$\lvert G_N\rvert$
2	1	1	1
3	2	1	1
4	6	3	3
5	21	5	4
6	112	17	12
7	853	44	26
8	11 117	182	88
9	261 080	730	257

Table 1: Number of connected graphs of various types: all (G_C), bipartite (G_B) and knight subgraph (G_N)

n	total	# NG	% NG
2	259	259	100
3	2 147	1 874	87
4	8 463	6 857	81
5	19 965	13 440	67
6	33 562	20 981	63
7	40 859	22 948	56
8	34 530	18 259	53
9	22 544	10 414	46
10	11 101	4 553	41
11	4 117	1 637	40
12	999	404	40
13	129	66	51
14	11	8	73
15	2	2	100
total	178 688	101 702	57

Table 2: Number words in TWL06 and the number and percentage of which are NG, broken down by number of distinct letters (n)

n	TWL06		any source	
2	WEEWEE	(6)	DEEDEED	(7)
3	SEERESSES	(9)	KINNIKINNIK	(11)
4	SERENENESSES	(12)	SERENENESSES	(12)
5	UNEVENNESSES	(12)	LOLLAPALOOSAS	(13)
6	FEMININENESSES	(14)	FEMININENESSES	(14)
7	OBSESSIVENESSES	(15)	POSSESSIVENESSES	(16)
8	PARALLELEPIPEDS	(15)	SUGGESTIBILITIES	(16)
9	PRESENTABLENESS	(15)	OVEREXPRESSIVENESSES	(20)
10	VERISIMILITUDES	(15)	NONPRESENTABLENESSES	(20)
11	ALPHABETIZATION	(15)	INDESTRUCTIBILITIES	(19)
12	ACCOMPLISHMENTS	(15)	UNACCOMODATINGNESSES	(21)
13	BLAMEWORTHINESS	(15)	HYPEREXCITABILITIES	(19)
14	OVERSPECULATING	(15)	*(see article text)*	(22)
15	UNCOPYRIGHTABLE	(15)	PSEUDOHALLUCINATORY	(19)

Table 3: Longest NG words by number of distinct letters (n) from TWL06 and from any source. The length of each word is given in parentheses.

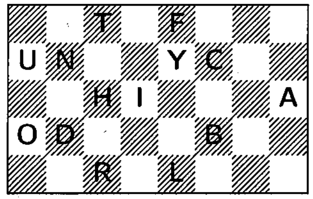

Figure 9: A chessboard labelling for the longest known knight-graphable English word (22 letters, 14 distinct letters).

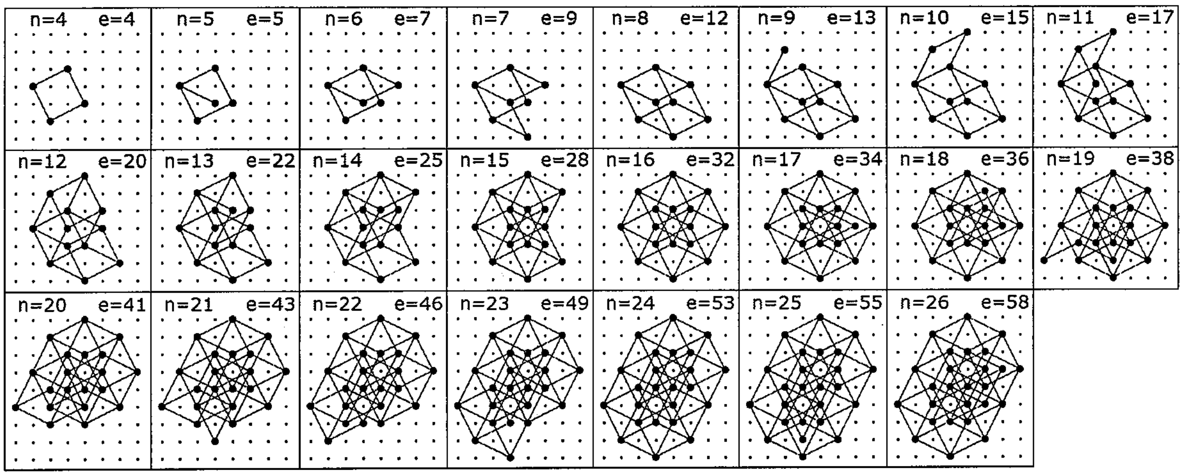

Figure 10: Knight subgraphs of orders 4 to 26 with the (conjectured) maximum number of edges

In closing, we consider the knight's-move version of a topic discussed in [2]: for each n, the number of different letters in a knight-graphable word or phrase, what is the largest possible number of distinct unordered bigrams (where, e.g., TR is considered the same bigram as RT)? In graph theory terms: what is the largest possible number of edges in a knight subgraph with n vertices? This appears to be a non-trivial problem.

Our best solutions for $n = 4$ to 26 are shown in Figure 10. We have proved that these have the most possible edges only up to $n = 9$, though we conjecture they are optimal up to $n = 16$. After $n = 16$ we are less sure; the graphs shown here are the best we have been able to construct using a heuristic computer search. Note that the graph with most edges for $n = 4$ is a square (all vertices have degree 2), for $n = 8$ is a cube (all vertices have degree 3), and for $n = 16$ is a hypercube (all vertices have degree 4). One might be tempted to think that after $n = 16$ the strategy would be to work towards constructing the graph where all vertices have degree 5. Unfortunately, this doesn't work:

Theorem: *The k-regular graph (graph where all vertices have degree k) is a knight subgraph only for $k \leq 4$.*

Proof: Assume the opposite, that one exists with $k > 4$, and consider its representation with knights on squares of an infinite chessboard. In this arrangement there must be a column that has no knights anywhere to the left of it. Pick any knight in that column.

Since there are no knights in the two columns to the left, this knight can attack at most 4 knights. So as a graph this vertex has degree at most 4. This contradicts the assumption that all vertices have degree > 4 and completes the proof. [3]

Now that we have the (conjectured) knight subgraphs with maximum number of edges, the next challenge is to turn them into knight-graphable words (or phrases, if a single word is not possible) by labelling their vertices with distinct letters. Those words with the fewest overall number of letters are the rarest so we concentrate on these. By exhaustive computer search we found minimal-length Scrabble words, or pairs of words, for each maximal graph from $n = 4$ to $n = 11$; these are shown in Figure 11. For $n = 4$, 5, and 6 there are many other words that work, but the $n = 7$, $e = 9$, length-11 word (UNDISGUISED) is unique.

Can you find labellings for all the other graphs in Figure 10 that spell out a sequence of words? Especially interesting is the last one, $n = 26$, whose labelling must use the whole alphabet.

References

[1] Frank Harary. The number of linear, directed, rooted and connected graphs. *Transactions of the American Mathematical Society*, 78:445–463, March 1955.

[2] Mike Keith. Some new results on king- and queen-graphable words. *Word Ways: The Journal of Recreational Linguistics*, 45(2):151–159, May 2012.

[3] S. C. Locke, Florida Atlantic University, private communication, January 2014.

[4] Brendan D. McKay and Adolfo Piperno. *nauty and Traces User's Guide (Version 2.5)*, January 2013.

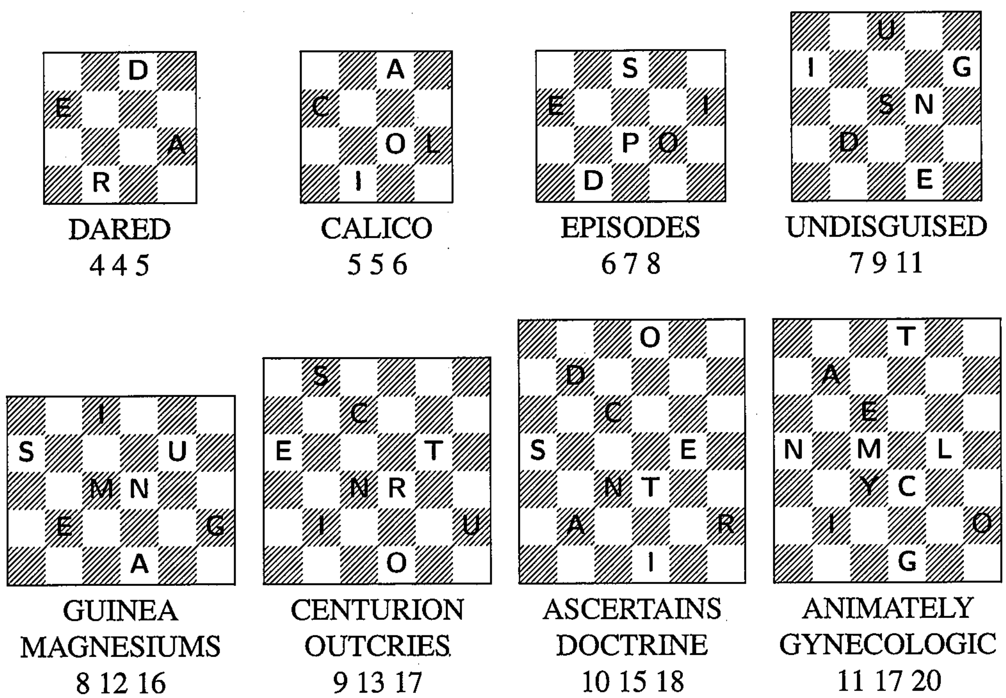

Figure 11: Knight-graphable words or phrases for $n = 4$ to 11 whose graph has the largest number of edges for that n (see corresponding graphs in Figure 10). The three numbers listed are, respectively: number of vertices, number of edges, and total number of letters.