SOME NEW RESULTS ON KING- AND QUEEN-GRAPHABLE WORDS

Mike Keith
Furlong, Pennsylvania

A word (or phrase, sentence, etc.) is said to be *king-graphable* (KG) if each of its distinct letters can be placed on a square of an infinite chessboard such that the entire word or phrase can be spelled out by starting on the first letter and making only the moves of a chess king to get from letter to letter. As shown below, the word ANTIINSTITUTIONALISTS can be king-graphed. Note that doubled letters, like the "I I" in this word, are allowed; the king just sits on the same square to spell the two I's.

ANTIINSTITUTIONALISTS

U	T	S
	I	N
L	A	O

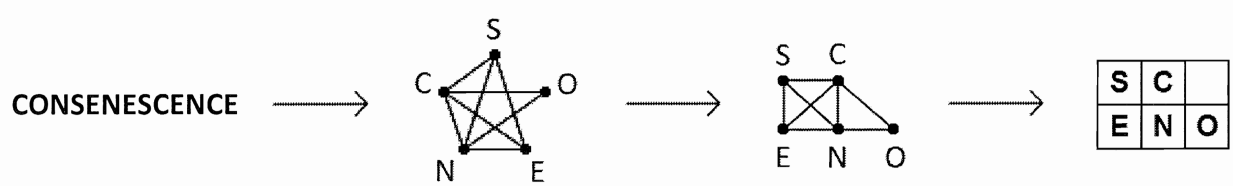
REPOSITORIES

	O			P
T	S	R		
	I	E		

A word or phrase is *queen-graphable* (QG) if the same thing can be done using the moves of a chess queen. (In this case, when moving from one letter to another all intervening squares must be empty.) Every king-graphable word is queen-graphable but the converse is not the case. The diagram on the right shows a queen-graphed word for which it can be proved that king-graphing is impossible. In the Feb. 2012 "Look Back!" Ross Eckler asked several questions related to QG words. In this article we answer some of his questions and present some additional results on king- and queen-graphability.

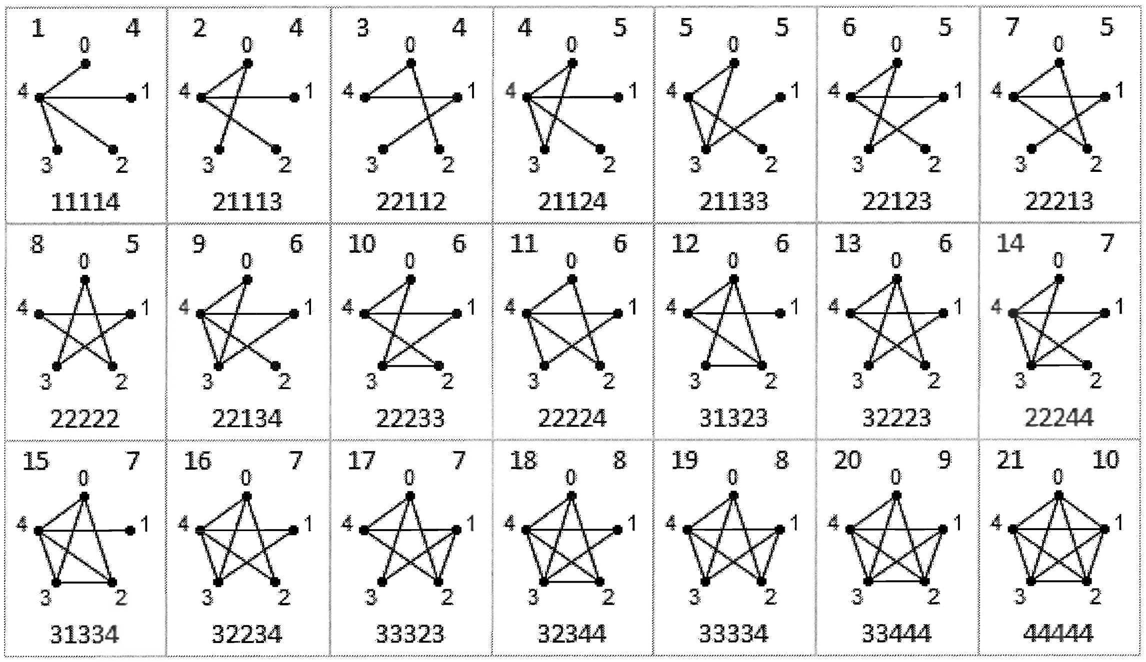
To better understand the graphability of words we turn to the undirected graphs studied in graph theory. The *word graph* for a word W is the graph with n nodes (which we might think of as labelled with the n different letters in W) that has an edge between two nodes labelled X and Y if the letters X and Y are adjacent to each other in the word (in either order, XY or YX). A graph is called *connected* if there is a path, using one or more edges, between any two nodes. A word graph is always connected because by definition it has a single path that touches all the nodes, and therefore has a path between *any* two nodes. In summary, a word W corresponds to some connected graph with n nodes and e edges, where e is the number of unique unordered bigrams in the word. A word is KG or QG if the nodes of its graph can be placed on different squares of a chessboard such that connected nodes are the required king's moves or queen's moves apart.

The diagram below shows the equivalence between a word, its word graph, a KG representation of the graph and letters on the chessboard, for a particular king-graphable word with n=5 and e=8.



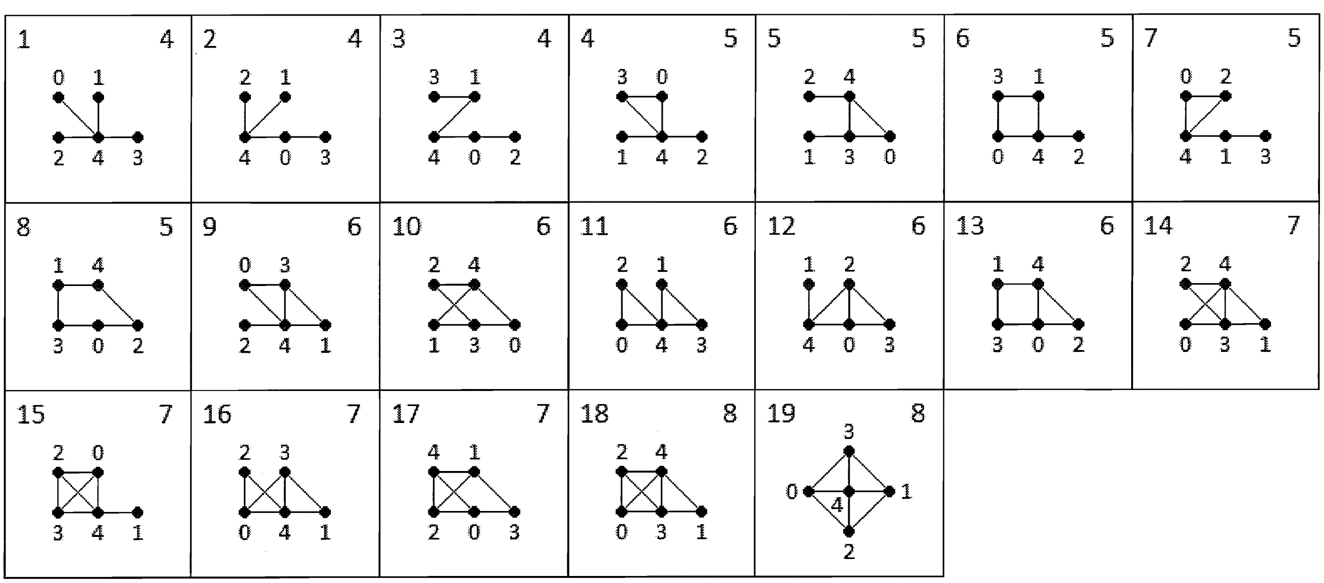
Which English words can be king-graphed or queen-graphed? It is easy to see that all words with $n \leq 4$ are KG (and so also QG), so the cases of interest start with $n = 5$. In the TWL06 North American

Scrabble word list I used for this study there are 19,965 words with n=5, each of which maps to some connected graph on five nodes, but there are only 21 connected graphs with n=5, as shown in the diagram below. To determine the graphability of these 19,965 words with n=5 only these 21 graphs need to be examined.



The connected graphs on 5 nodes. In each box, the upper right number is e (number of edges), the string of numbers at the bottom gives the degrees of vertices 0,1,2,3,4, and the upper left number is an "ID number" (1 to 21) identifying each graph. The graphs are listed in order of increasing number of edges.

The first 19 of these are king-graphable, as the figure below demonstrates. Nodes are numbered the same as in the above figure to clearly show the equivalence between the two drawings of each graph.



It is relatively easy to show that graphs #20 and #21 cannot be king- or queen-graphed. There is only one Scrabble word, INTENSITIES, that maps to graph #20, and none for graph #21, so for $n = 5$ only one word out of 19,965 cannot be king- or queen-graphed. Going beyond the Scrabble words, INSCIENCES also works for graph #20 and is the shortest possible non-KG word.

Since #20 and #21 are the only graphs in this list with more than eight edges, we can summarize the conditions for word graphability when n=5 in a simple statement:

A word with 5 different letters can be king-graphed (and queen-graphed) if and only if $e \leq 8$.

The previous figure demonstrates the following interesting result, which is a consequence of the fact that each graph from 1 to 17 is a subgraph of graph #18:

Any word with 5 different letters that can be king-graphed can have its lettered squares arranged in the shape of either the P or the X pentomino.

Here is the longest word we found for each graph (allowing sources beyond TWL06):

1: EFFETENESSES (12)
2: UNEVENNESSES (12)
3: LOOSENESSES (11)
4: TENSELESSNESS (13)
5: USELESSNESSES (13)
6: SUDDENNESSES (12)
7: SENSUOUSNESSES (14)
8: SKINNINESSES (12)
9: SLEEVELESSNESS (14)
10: NONNICOTINIC (12)
11: NEEDLESSNESSES (14)
12: UNSENSUOUSNESS (14)
13: NONSENUOUSNESS (15)
14: ACYCLICALLY (11)
15: UNSONOROUS (10)
16: SLEEPLESSNESSES (15)
17: SUPERPRESSURE (13)
18: CONSENESENCE (13)
19: MILLIOSMOLS (11)

To proceed past n =5 it is helpful to employ computer assistance due to the large number of cases that need to be considered: 112 different connected n=6 graphs, 853 with n=7, 11117 with n=8, etc. We used a computer program that takes each connected graph with n nodes and determines if it can be king-graphed or queen-graphed via exhaustive search. The results for n=6, 7, and 8 are shown in the tables below.

Graphs with 6 nodes 91 KG, 94 QG out of 112 graphs

e	5	6	7	8	9	10	11	12	13	14	15	
KG	6	13	19	22	19	10	2	0	0	0	0	← the number of graphs that can be king-graphed
QG	6	13	19	22	**20**	**11**	**3**	0	0	0	0	← the number that can be queen-graphed
x	0	0	0	0	0	3	6	5	2	1	1	← the number that can be neither KG nor QG.

Graphs with 7 nodes 535 KG, 585 QG out of 853 graphs

e	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
KG	11	33	67	107	129	113	60	13	2	0	0	0	0	0	0	0
QG	11	33	67	107	132	126	86	21	2	0	0	0	0	0	0	0
x	0	0	0	0	0	12	40	74	62	40	21	10	5	2	1	1

Graphs with 8 nodes 4308 KG, 5422 QG out of 11117 graphs

e	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
KG	23	89	236	486	804	1053	954	500	140	21	2	0	0	0	0	0	0	0	0	0	0	0
QG	23	89	236	486	814	1132	1261	995	345	39	2	0	0	0	0	0	0	0	0	0	0	0
x	0	0	0	0	0	37	193	584	1170	1251	968	658	400	220	114	56	24	11	5	2	1	1

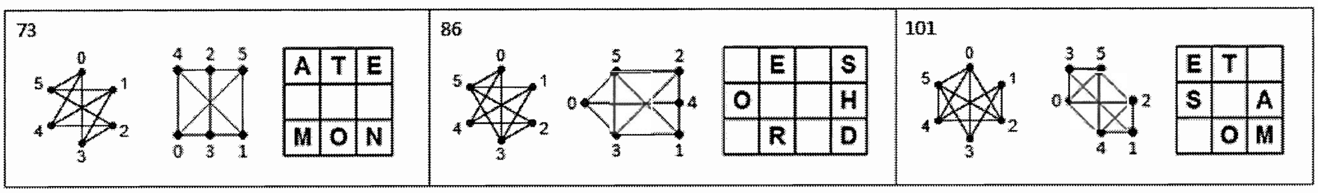
$L(n)$ . . . . $U(n)$

The tables above have three sections. In the left section every graph is both KG and QG, in the middle section a graph *may* be KG and/or QG but it depends on the graph (not just the value of e), and in the right section every graph is neither KG nor QG. As shown under the $n = 8$ table, the lower and upper limits of e in the middle section are denoted $L(n)$ and $U(n)$. We see that the simple necessary and sufficient condition for $n > 5$ does not have a direct analogue for larger n, being replaced by the following three statements:

- Every graph with $e < L(n)$ is KG.
- Every graph with $e > U(n)$ is neither KG nor QG.
- For $L(n) \leq e \leq U(n)$, whether a graph is KG or QG depends on the graph, not just e.

The number of KG graphs on n nodes starting with n=2 is 1, 3, 6, 19, 91, 535, 4308..., while the number of QG graphs is 1, 3, 6, 19, 94, 585, 5422, etc.

The three bold entries in the $n = 6$ table contain the first instances of graphs that are QG but not KG. Each of these three graphs is shown below with a QG representation and an example letter grid.



The letter grid for graph #101 is for the Web3 word METASOMATOSES. I did not find a single word corresponding to the other two graphs (can you?); their letter grids are for the phrases MONOTONE MANATEE (#73) and REORDERED HORSESHOES (#86). The number of QG-but-not-KG graphs starting with n=6 is 3, 50, 1114, ...

For a given value of n, two conflicting phenomena affect the number of non-graphable words. On the one hand every graph with more than $U(n)$ edges is non-graphable, but on the other hand words with $U(n)$ or more edges are rare. To see this in greater detail, the three tables below show how all the words in TWL06 with n = 5, 6, and 7 map to the 21, 112, and 853 graphs with n = 5, 6, and 7. (Drawings of all the graphs up to n = 7, numbered with the same ID numbers, can be viewed online by going to cadaeic.net/graphpics.htm.) Each cell in a table contains a number or a symbol specifying how many words correspond to the graph with a given ID number; if there are no words for that ID, the cell is left blank. The first row of each table has ID numbers 1 to 50, the next row 51 to 100, and so on, as indicated by the column-head and row-head numbers.

n = 5: 21 graphs, 19965 words

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
▪	■	■	■	▪	■	■	■	▪	▪	▪	▪	▪	32	6	85	14	9	1	◸ 1	◸

Legend

▪	100 to 999		KG and QG
■	1000 to 9999	□	QG but not KG
█	≥ 10000	◸	not KG or QG

n = 6: 112 graphs, 33562 words

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
0	12	▪	32	■	▪	█	▪	▪	■	▪	■	▪	▪	■	■	■	14	▪	■	▪	▪	▪	▪	▪	▪	▪	17	97	■	▪	▪	▪	4	90	▪	50	65	50	43	12	40	1	18	2	12	97	▪	18	▪	56
50	19	▪	7	1		2	12	3	12	17	3	4	19	8	3	10	23		6		7	2	□		2	3	1	5	4	2		3	◸	1	◸	□	1		◸	1	2		1	1	◸	◸		◸	◸	
100	□	◸	◸	◸	◸	◸	◸	◸	◸	◸	◸	◸																																						

$n = 7$: 853 graphs, 40859 words

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
0		24	2	■	37	67	62	█	█	21	█	22	16	■	13	75	■	82	■	█	█	67	■	11	█	4	45	■	■	■	27	■	█	35	■	█	■	█	█	13	■	■	■	█	11	35	25	73	90	65
50	51	■	■	■	■	■	3	9	■	2	22	■	■	11	19	66	44	12	■	7	■	■	39	■	2	66	26	■	44	27	■	14	3	3	65	■	16	■	16	■	38	■	■	26	3	■	74	31	2	25
100	26	50	33	■	75	57	■	13	4	66	86	12	2	9	21	19	18	5	1	5		16	1	58	19	15	15	1	33	45	6	9	15	39	38	14	18	20	10	4	24	79		8	5	7	60	34	7	4
150	1	70	15	38	3		12	1	1	10	9		4	11	7	40	27	24	1	1	31	5	5	13		56	23	90	54	■	4	8		8	10	7	9	2	3	2	3		4	2	3			4	2	9
200	13	6	3	33	25	67		8	5	6	12	58	4	7	20	8	2	3	2		2		11			1	5	1	6	3	4	4	2	6	3	2	7		4	6		5				2	1		1	
250	1	10	17	9			1	2	2			4		8	2	1		2		1	10	2	3	1	2		6				2	1	6					1	4	7	1	9	1	6	10	6	1	11	4	
300		1	2	4		2			1		6		4						1				2	2		1	1	1	3	3	4	1	1		12	2	13		3	1	4									
350	1		1							5		1								3		1			1											1						3			2	1				1
400																							2	1	1						1	1	1		1			2			1	1			2	1	1			1
450	2	2	1	1	3											1								2	1	1		2		1			2			1									1					
500							2																																								1			
550				1																	1																					1								
600		1										2																																						
650																																																		
700																																																		
750																																																		
800																																																		
850																																																		

We always sort the graphs for a given n by increasing number of edges, so we can see from these tables that the words corresponding to each graph rapidly thin out as the number of edges increases.

In these tables the shaded cells highlight those graphs which are either not KG or not QG *and* for which at least one word maps to that graph. There are only a few such words in these tables, and this remains true as we examine the words with larger values of n. In fact:

- Only 29 of the 178,694 Scrabble words are not king-graphable
- Only 5 of the 178,694 Scrabble words are not queen-graphable

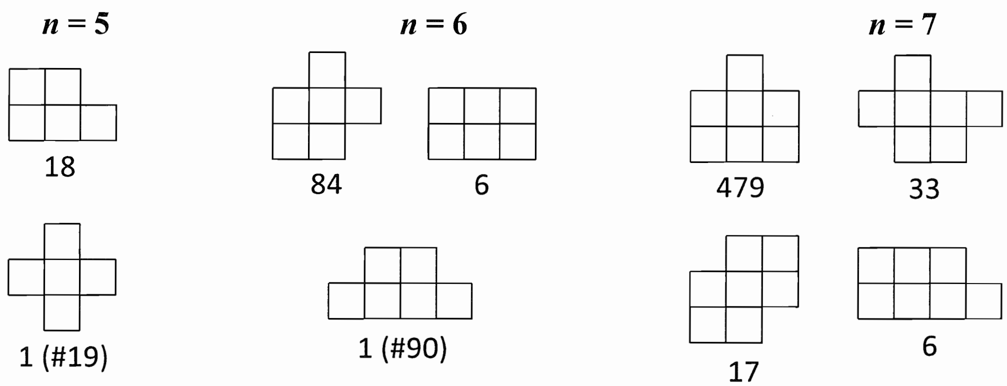
The 29 non-KG words are shown below; the five in bold face are the ones which are not QG either.

$n = 5$:	**INTENSITIES**			
$n = 7$:	**CONTORTIONISTIC**	**EXCRESCENCIES**	**OMNISCIENCES**	ENREGISTERING
	EPISCOPACIES	REPOSITORIES	SUPERSATURATES	
$n = 8$:	**OVERINTENSITIES**	ANIMALIZATION	ANTICAPITALIST	ANTIRATIONALIST
	INTERDISTRICT	SAPONACEOUSNESS	SILVICULTURIST(S)	TACHYARRHYTHMIA
$n = 9$:	ANIMALIZATIONS	CANNIBALIZATION	CHORDAMESODERMS	ETHNOHISTORIANS
	OLIGOPSONISTIC	SIGMOIDOSCOPIES	SIGNALIZATIONS	
$n = 10$:	ANTIRADICALISM(S)	MARGINALIZATION	RADICALIZATION	
$n = 11$:	RADICALIZATIONS			

Of course these results are for a specific word list, and one that does not include any words longer than 15 letters. Undoubtedly the statistics will be similar for other words sources, however, so even though additional non-KG and non-QG words can be found, they are few and far between.

Every empty cell in the tables above can be taken as a puzzle: can you find a word that maps to that graph? Notice that the first cell in the n=7 table is empty! This corresponds to what might be called the “asterisk graph”, with one node connected to six surrounding nodes. We haven’t found a word for this graph in a dictionary source but there is at least one reasonable candidate, a term for having made telemetry measurements again: RETELEMETERED.

The P/X-pentomino theorem for $n = 5$ can be extended to any n by seeking the minimum number of distinct n-ominoes which can be used to represent any KG word with n different letters. The figure below shows a minimal set of two pentominoes, three hexominoes, and four septominoes that suffice for all KG words with $n = 5$, 6, and 7. (The number below each polyomino specifies the number of graphs corresponding to it.) We have found a set of eight octominoes that work for $n = 8$ but do not know for sure if that set is minimal.



In closing, we consider the problem of whether it is possible to derive the values of $L(n)$ and $U(n)$ that define the three regions of graphability (always possible, maybe possible, and always impossible) for each n. From the foregoing results we see that for $n = 5, 6, 7, 8$ the values of $U(n)$ are 8, 11, 14, 17, which are just $3n$-7, and there are always two different KG graphs with this value of e. Do these trends continue? The answer is no, as shown by the actual numbers up to n=26 given in the table below.

n	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
U(n)	8	11	14	17	20	23	26	30	33	36	39	43	46	49	53	56	60	63	66	70	73	77
#KGs	2	2	2	2	2	4	5	1	1	3	7	1	3	11	1	5	1	4	17	2	12	2

The values of $U(n)$ seem quite irregular, so it is a nice surprise to find that there is an exact formula:

The most distinct bigrams possible in a KG or QG word having n distinct letters ($n \geq 5$) is

$$U(n) = \max_{w,h} \{ 4wh - 3(w+h) - E(wh - n) + 2 \}$$

where

$$E(k) = 4k - 4R(k) - \left[\frac{k - 4T(R(k))}{R(k)+1} \right], \quad R(k) = \left[\frac{\sqrt{2k+1} - 1}{2} \right], \quad T(k) = \frac{k(k+1)}{2}$$

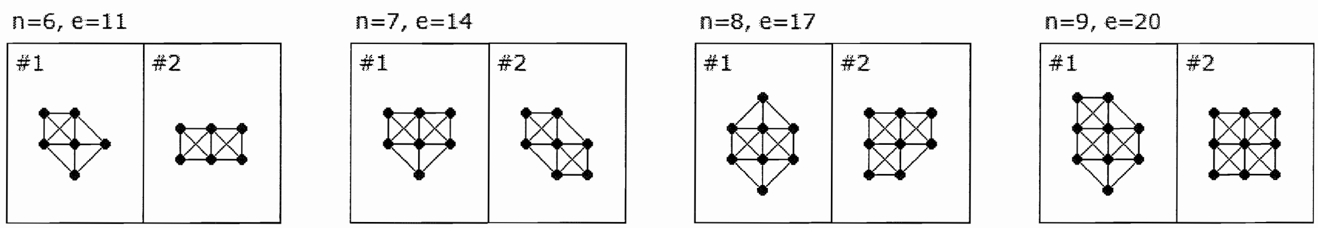
and h and w are all integers such that $3 \leq w \leq 2\sqrt{n}$, $n/w \leq h \leq w$ and $wh \geq n$.

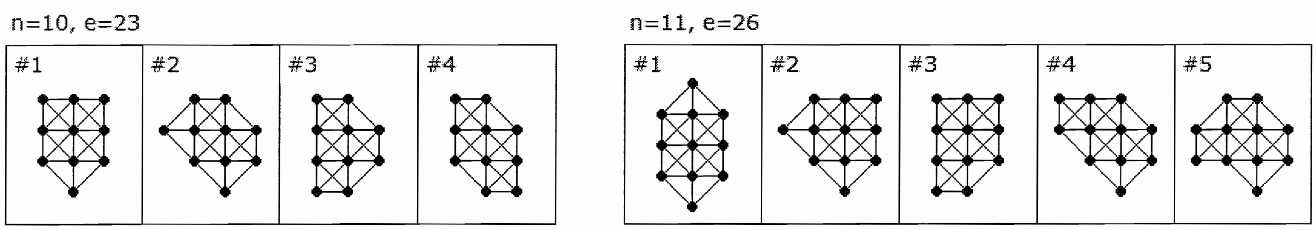
(*as usual,* [ ] *is the greatest integer function*)

A full proof of this theorem would take us too far afield but here is a brief summary. To obtain the largest number of edges for a KG graph with n nodes the best strategy is to have a full rectangular (w x h)-sized array of n nodes, perhaps with parts of one to four corners lopped off, with all nodes connected to king-adjacent nodes in all possible ways. In the $U(n)$ formula, $4wh - 3(w+h) + 2$ is the number of edges in a full w x h rectangle and the $E(wh - n)$ term represents the number of edges lost by removing a total of $wh - n$ nodes from the four corners in the most optimal way. The *max* in the $M(n)$ formula selects the maximum value of this computation over all relevant-sized rectangles.

The number of queen's graphs having $M(n)$ edges is the same as the number of king's graphs, because every king's graph is also a queen's graph and the extra flexibility provided by the queen's move does not produce a denser packing. The only exception is $n = 6$, $e = 11$ (having 2 KG's but 3 QG's), which is caused by a special QG-only graph (graph #101 depicted earlier) that can only occur when $n = 6$.

Here are pictures, for $6 \leq n \leq 11$, of all the different KG graphs with the maximum number of edges:





We did not find any single words corresponding to these graphs, so instead we looked for two-word phrases to represent them. The table below contains one shortest and longest word or phrase for each maximal-edge king graph starting from $n = 5$:

$n = 5$, $e = 8$, #18: **REPURSUES** (9), CONSENESCENCE (13)
$n = 5$, $e = 8$, #19: MILLIOSMOLS (11)
$n = 6$, $e = 11$, #1: PENITENT SPIES (13), INCANTATION CONNOTATION (22)
$n = 6$, $e = 11$, #2: BEARABLE LASERS (14), CLASSICAL CLINICIANS (19)
$n = 7$, $e = 14$, #1: GOING UNCONSCIOUS (16), INCONSISTENT CONNECTIONS (23)
$n = 7$, $e = 14$, #2: **CENTRIST CRETINS** (15),
$n = 8$, $e = 17$, #1: **FLAGRANT AIRLIFTING** (18), CORNINESS RESTRICTIONISTS (24)
$n = 8$, $e = 17$, #2: (*)
$n = 9$, $e = 20$, #1: INSUBSTANTIAL OBNUBILATION (25)

Note: the graph ID numbers in this table (#1, #2, etc.) correspond to the ID numbers in the figure above, except for n=5, where the ID numbers correspond to the full list of 21 n=5 graphs.

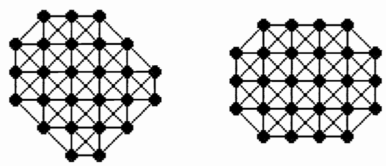
The entries shown in bold are very special, as they use e+1 letters, the fewest letters that a word or phrase corresponding to a graph with e edges can possibly have. A word with e+1 letters can only exist if the graph has an *Euler path* - a walk through the graph that uses each edge exactly once. A well-known theorem states that a graph has an Euler path if and only if the number of nodes with odd degree is two or fewer. The three graphs in the list above having an Euler path are *the only maximal-*

edge KG graphs with n ≥ 5 having two or less odd-degree nodes, and therefore the only ones for which a minimal-length word can exist.

We could not find a two-word phrase for the graph marked (*) even with an exhaustive search of all pairs of words in our largest word list, so for that graph, and those beyond the end of this table, it will be necessary to use three-word or longer phrases. Can you construct phrases for the n = 8 (#2) and n = 9 (#2) graphs and those with n = 10 and 11? How short can these phrases be?

The most extreme version of this question is the *maximal-edge king-graphable pangram challenge*: construct a text, preferably as short as possible, with n = 26 and the largest possible number of bigrams, 77. This text must result in one of these two graphs, with the nodes arranged as shown:

n=26, e=77



It is surely not possible to make a text like this (of any length) that makes grammatical sense, so we merely require the text to be a series of words, all different. How few total letters can be used? Here is our best attempt:

	Q	Z	S	K	
B	U	I	L	Y	V
D	P	C	R	E	G
X	O	W	A	T	N
	M	F	H	J	

Word string:

AW CUB QUIZ QI PUD BUZZ SLIP
DUB POX MOD OX DOM FOX POW PI
COW MOW FOP CLIP CRY LEA TAJ NTH
FAT RAJ HAW HAT JAW RE YEN GET
GYVE ELK VEG VERY SKY SIR CARL ZIP

This letter grid and word sequence were created entirely by hand in collaboration with Ross Eckler. All words are in TWL06 and there are 119 letters. Our strategy was to use short words exclusively, since this allows us to capture many bigrams (which would otherwise be uncapturable, because there are no English words that contain them) using the last letter of a word in combination with the first letter of the following word. Longer words can be employed (this letter layout allows one to spell JACUZZI, VENTRICLE or WATERLEAF or, for example) but doing so usually comes at the expense of using more overall letters. Readers are encouraged to consider these related puzzles:

- Find a solution for this graph using fewer letters (perhaps with a different letter grid).
- Find a solution, with the fewest letters possible, for the other maximal n = 26 graph.
- What is the fewest number of *words* possible in a solution (allowing any number of letters)?
- What is the longest word that can be incorporated in a solution?

When working on these problems, be warned: it is easy to think one has a solution only to find, upon careful checking, that one or two of the 77 bigrams are missing.

Having presented some results on the value of $U(n)$ we now consider the lower limit of the middle region, $L(n)$. The situation here is even simpler:

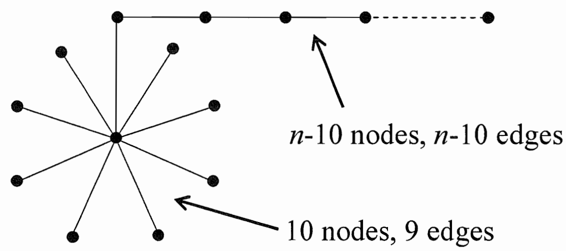
King-graphability:

$$L(n) = \begin{cases} n+3 & \text{for } 6 \le n \le 9 \\ n-1 & \text{for } n \ge 10 \end{cases}$$

Queen-graphability

$$L(n) = \begin{cases} n+4 & \text{for } 6 \le n \le 9 \\ n-1 & \text{for } n \ge 10 \end{cases}$$

The $L(n)$ values for $n = 6$ to 9 were obtained by computer, though their simple form suggests that there may be a theoretical explanation. For $n = 10$ and beyond, a different factor comes into play: the following graph with $n \ge 10$ nodes and n-1 edges has a node connected to nine other nodes:



Since the king and queen can only move in eight different directions, a node with degree 9 cannot be accommodated. Therefore the possibly-ungraphable region starts at $e = n - 1$. Since $n - 1$ edges is the smallest number of edges a connected graph can have, this means that for $n \ge 10$, there is no value of e for which king- or queen-graphing is always possible.