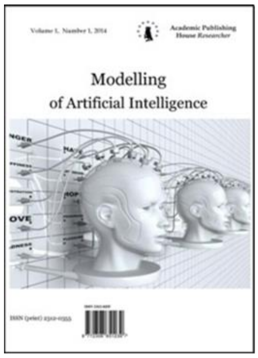

Published in the Slovak Republic
Modeling of Artificial Intelligence
Has been issued since 2014.
E-ISSN: 2413-7200
2018, 5(1): 3-21

DOI: 10.13187/mai.2018.1.3
www.ejournal11.com

Articles and statements

Investigation of the Regularities in the Formation of Solutions n-Queens Problem

Eric Grigoryan [a, *]

[a] Independent researcher, France

This publication is devoted to all Google employees...
You help to find necessary information to go ahead...

Abstract

The n-Queens problem is considered. A description of the regularities in a sequential list of all solutions, both complete and short, is given. Determined that:

1. The fraction of total solutions in the general list of all solutions decreases, with increasing value of n.

2. Complete solutions are distributed in a sequential list of all solutions in such a way that the most likely solutions are complete solutions located in the list close to each other.

3. There is a symmetry in the order of the location of the complete solutions in the general list of all solutions. If the solution is complete in the i-th position from the beginning of the list, then the symmetric solution from the end of the list, located in the position n-i + 1, is also complete (rule of symmetry of solutions).

4. Any pair of solutions, both short and full, arranged symmetrically in the list of all solutions, are complementary – the Queen position indices sum of the corresponding rows is a constant and is equal to n + 1 (the rule of complementarity of solutions). This suggests that only the first half of the list of all complete solutions is "unique", any complete solution from the second half of the list can be obtained on the basis of the complementarity rule. The consequence of this rule is the fact that for any value of n, the number of complete solutions will always be an even number.

For an arbitrary matrix of a solution of size n x n, it is established that:

5. The probability of completion to a full solution an arbitrary composition of k queens, gradually decreases with increasing value of k to a certain minimum, and then increases, with a further increase in the value of k.

6. There is some minimum value of the size of the composition k0, such that any composition whose size is less than or equal to k0 can always be completed to a complete solution. As the value of n increases, the value of k0 also increases.

7. The activity of row cells in solution matrix is symmetric with respect to the horizontal axis passing through the middle of this matrix. This means that the cells activity in the i-th row always coincides with the cells activity in the row n- i + 1. By activity is meant the frequency with which

[*] Corresponding author
E-mail addresses: ericgrig@gmail.com (E. Grigoryan)

the cell index occurs in the corresponding row of the list of complete solutions. Similarly, the activity of the cells of the columns of the solution matrix is symmetrical about the vertical axis dividing the matrix into two equal parts

8. For any n, in the sequential search for all solutions, the first complete solution appears only after some sequence of short solutions. The size of the initial sequence of short solutions increases with increasing n. The length of the list of short solutions until the first complete solution for even values of n appears is much larger than for the nearest odd values.

9. The row in the solution matrix, on which difficulties begin to move forward, and the first short solution is formed, divides the matrix according to the rule of the golden section. For small values of n, such a conclusion is approximate, but with an increase in the value of n, the accuracy of such an output asymptotically increases to the level of the standard rule.

Keywords: n-Queens problem, constraint satisfaction problems, non-deterministic problem, state space, search for regularities.

1. Introduction

I was always sure that if the data is not random, then there must be a certain regularity in them, if even this regularity we can't find. It was this conviction that was the reason for the search for regularities in the n-Queens problem solutions.

The formulation of the problem is quite simple: it is necessary to distribute n queens on a chessboard of size n x n so that there is not more than one queen in each row, each column, and on the left and right diagonals passing through the cell where the queen is located. This task is easy to understand or explain to anyone, but it is difficult to solve it. The fact is that there is no rule (or set of rules) based on which we can arrange the queens in each row so get a solution. The solution can be obtained only on the basis of a search of various variants of the arrangement of queens in certain rows. However, the complexity of this method of solution is that the number of all variants of the arrangement of queens grows exponentially with increasing n. In addition, the execution of any step forward, for placing the queen in the free position of some row, changes the list of free positions in the remaining rows, and when we go back one step, in order to form a search branch, we must clear the traces of previously performed actions.

The problem of the distribution of n queens on a chessboard of size n x n has a long history. Originally it was formulated in 1848 by M. Bezzel (Bezzel, 1848), as an intellectual task for a conventional chessboard. Over time, the statement of the problem was extended by F. Nauck (Nauck, 1850), and the size of the chessboard could take on an arbitrary value.

There are a large number of publications related to various aspects of solving the n-Queens problem. Some of these publications are publicly available, and the other part requires payment to view the publication. I did not watch paid publications and can't refer to them. Among open publications there are many interesting and informative works that relate to several areas of research: the search for all complete solutions for a given chessboard size n, the development of a fast algorithm for finding one solution for different values of n, the computational complexity of algorithmic calculations, and also various modifications of the initial statement of the problem. To familiarize with these areas, I would recommend the remarkable work of Bell & Stevens (Bell, Stevens, 2009) and also I.P. Gent, C. Jefferson, P. Nightingale (*Gent et al., 2017*), which provides a fairly detailed overview of the various areas of the study. Some directions are considered in more detail in the works of H.A. Priestley, M.P. Ward (Priestley, Ward, 1994), R. Sosic and J. Gu (Sosic, Gu, 1994), J. Mańdziuk (Mańdziuk, 2002), A.S. Farhan, W.h. Z. Tareq, F. H. Awad (Farhan et al., 2015). Especially noteworthy is an online publication (Kosters, 2017), supported by W. Costers, which was prepared by a group from Universiteit Leiden and contains a link to 339 publications (as of 2018) related to the problem of n-queens.

Although the problem of n-Queens has remained active for more than 150 years, and during this time research has been conducted, and a huge number of publications have appeared, I haven't been able to find a publications that would have to do with the search for regularities in the results of solving this problem.

Definitions

Here and below, the size of the chessboard will be denoted by the symbol n. The solution will be called complete if all n queens are consistently arranged on a chessboard. All other solutions, when the number of correctly arranged queens is less than n, will be called short. By the length of the solution we mean the number (k) of correctly arranged queens. The number of all solutions (short and complete), for a given value of n, will be denoted by the symbol m. As an analog of the "chessboard" of size n x n.

2. Results

Beginning

In order to conduct the research, an algorithm for finding all solutions for an arbitrary value of n was developed. We did not use recursion or a nested loop system. For large values of n, such an approach would be rather problematic. The algorithm was built on the basis of a set of interacting events, in each of which, a certain system of actions, interconnected, takes place. This makes it possible to simply implement the mechanism of changing the state space when selecting the next node in the Forward Tracking branch, or clearing the traces of previously performed actions, when returning back to one or more steps (Back Tracking). In the algorithm there are no special requirements to the amount of memory or the speed of the processor, so the calculations can be made on any home computer (laptop). Based on this algorithm, all successive solutions (both short and complete) were found for different values of the solution matrix (n = 7, ..., 16). Since the size of the list of complete solutions is a named sequence (sequence A000170 [10]) and is indicated in many publications, it seems to me interesting to bring the size of the list of all solutions, for the values n: 7 (194), 8 (736), 9 (2936), 10 (12774), 11 (61076), 12 (314730), 13 (1716652), 14 (10030692), 15 (62518772), 16 (415515376).

Further, using the solutions found, we give formulations of some problems, methods for their solution, and a description of the results obtained. Also, we present the results of the computational experiments carried out to evaluate the probability of completion to full solution an arbitrary composition of k queens.

1. About the state space in which solutions are being sought.

The search of various variants of the arrangement of queens in certain rows leads to the formation of a state space. If there were no prohibitions on the location of queens in any cell, then the size of the state space would be equal to n^n. If we only consider a rule that forbids the placement of more than one queen in each column, then we get a subset of the state space whose size will be equal to n! This subset of the state space corresponds to the problem of the distribution of n rooks. If, at the same time, we also take into account the rule that prohibits the arrangement of more than one queen on the left and right diagonals passing through the cell where the queen is located, we get a search space whose size is less than n!. We call such a subset of the state space a productive search space, starting from the fact that only in this subspace are all possible solutions. Any completed branches in the productive search space are solutions with a certain number of correctly arranged queens. Basically - these are short solutions, and only a small part of the rest are complete solutions.

Figure 1 shows the graphs of the dependence of the natural logarithm of three indicators: a) the factorial values (n!) of the size of the solution matrix; b) the number of all solutions (both short and complete); c) the number of complete solutions, depending on the size of the solution matrix (n). As expected, all curves have exponential growth, and, evidently, the logarithm of the factorial grows much faster than the corresponding values of all solutions and complete solutions. Also, the growth rates of the number of all solutions and the number of complete solutions vary, although this is not so noticeable on the graph, due to the small size of the sample of values of n. For example, for n = 13, the difference between the logarithms of these indicators is 3.148, and for n = 16 this difference increases by 0.190 and is 3.338. Obviously, with a further increase in the value of n, this discrepancy will be more significant.

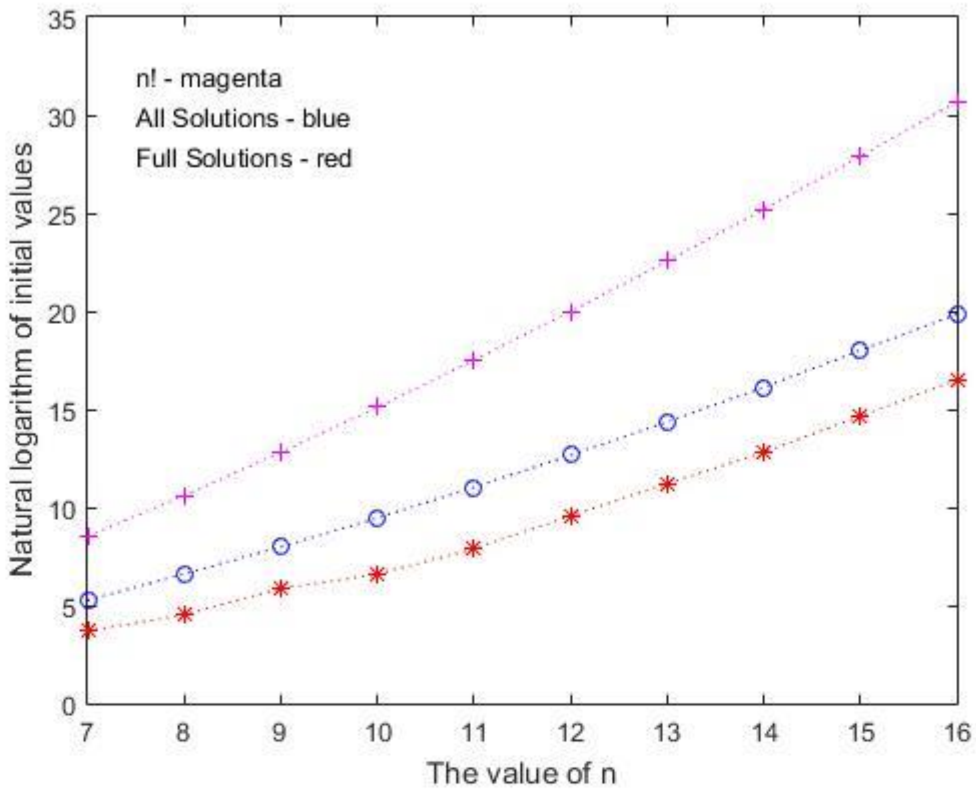

Fig. 1. Dependence the logarithm of the size of difference subsets of State Space from **n**

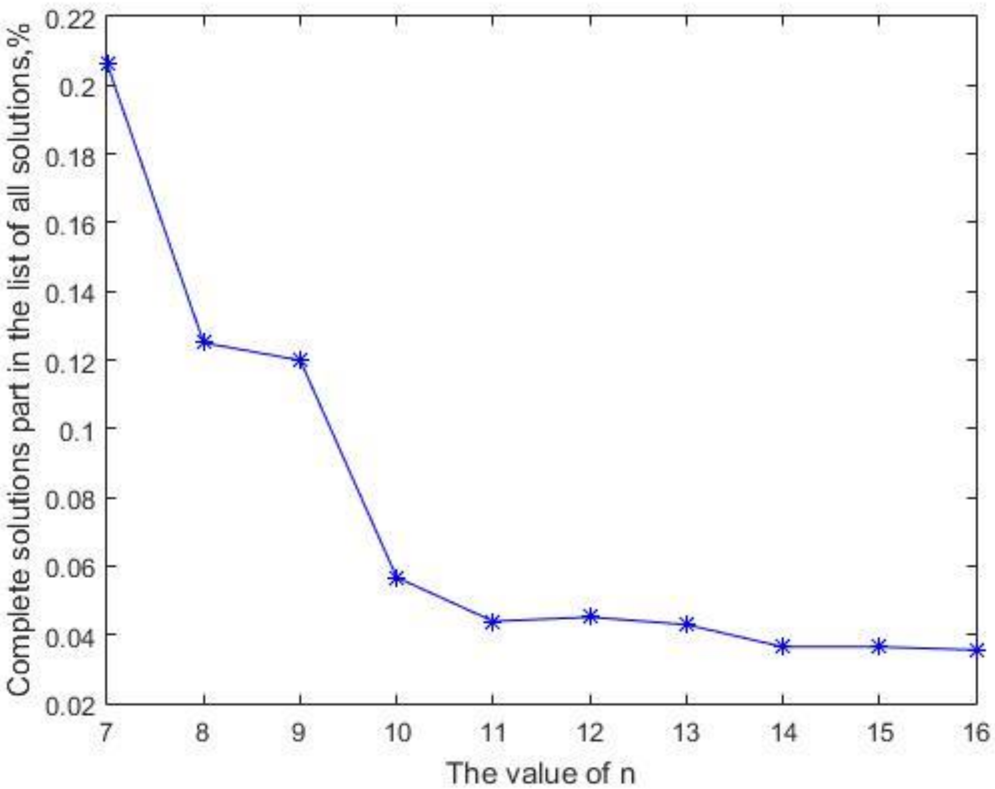

Fig. 2. Decreasing complete solutions share in the list of all solutions with increasing **n**

2. How does the share of complete solutions change in the general list of all solutions?

Figure 2 shows the plot of the fraction of complete solutions in the general list of all solutions from the value of n. It is seen that as the size of the solution matrix increases, the share of all complete solutions in the general list decreases. For initial values n = 7, ..., 14, this value decreases rapidly from the value 0.2062 to 0.0364, and then a gradual, asymptotic decrease of this value continues. Here there is a formal contradiction between the two statements: on the one hand, the number of complete solutions exponentially increases with increasing value of n, on the other hand, the probability of obtaining a complete solution in a sequential list of all solutions is constantly decreasing. This seeming paradox is explained very simply, the size of the productive space and the associated size of the list of all solutions grows faster with increasing n than the number of complete solutions. It's like trying to find a needle in a haystack – the amount of hay "with increasing n" grows faster than the number of needles that are lost there.

3. What is the frequency of solutions of different lengths in the list of all solutions?

Table 1. Relative frequency (**%**) of solutions of different length (**k**) for a matrix of size **n**x**n** (**n** = 7, ..., 16).

n\k	4	5	6	7	8	9	10	11	12	13	14	15	16
7	10.31	**31.23**	27.84	20.62									
8	2.45	20.38	**34.78**	29.89	12.50								
9	0.34	5.79	21.73	**35.83**	24.32	11.99							
10	0.05	1.35	8.41	25.62	**32.94**	25.96	5.67						
11		0.15	2.12	11.80	26.71	**34.47**	20.36	4.39					
12		0.01	0.29	3.28	13.56	29.88	**31.29**	17.18	4.51				
13		**e**	0.03	0.68	4.72	16.57	28.76	**28.84**	16.11	4.29			
14			**e**	0.90	1.14	6.47	17.49	**28.01**	27.45	15.70	3.64		
15				0.01	0.18	1.80	7.53	18.46	**27.07**	26.63	14.64	3.68	
16				**e**	0.02	0.35	2.23	8.23	18.89	**27.56**	25.39	13.77	3.56

e – denotes values that are less than 0.01

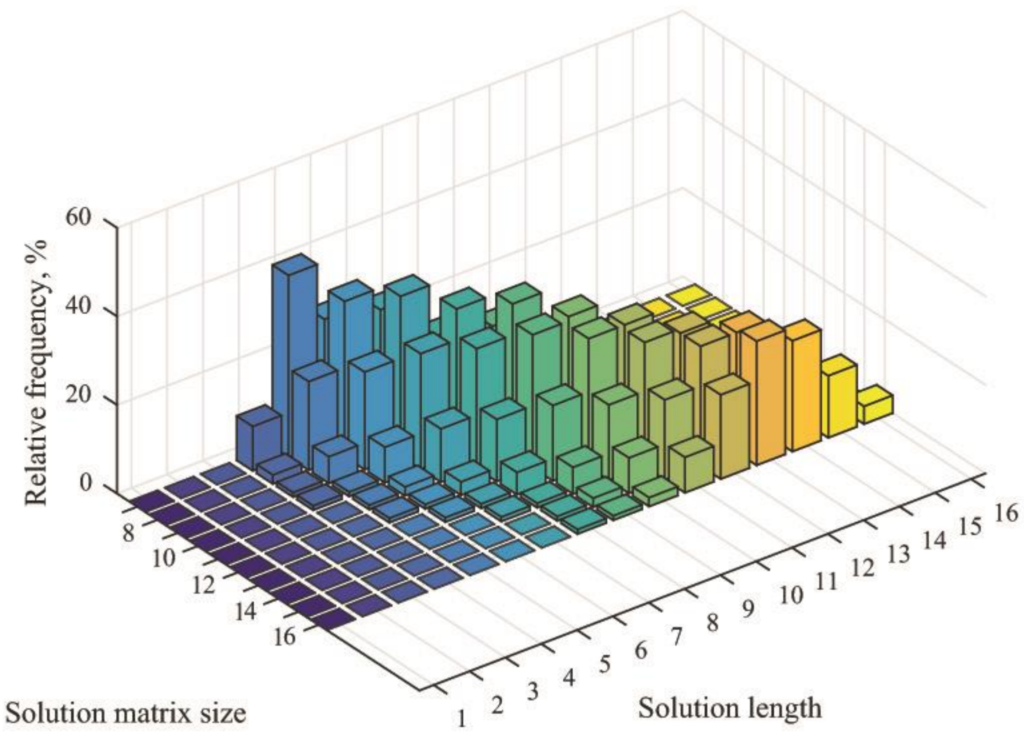

Fig. 3. Frequency of solutions of different lengths depending on the size of the solution matrix, **n** = 7, ..., 16

As mentioned earlier, all completed branches in the productive search space are solutions with a different number of correctly arranged queens. It is of interest to us how often solutions of different lengths are found in the general list of all solutions. Table 1 shows the corresponding values of the relative frequencies for solutions having different lengths that were obtained from the list of all solutions for n = 7, ..., 16. The corresponding visual representation of these data is given in Figure 3. The analysis of the table allows us to draw the following conclusions:

a) for each solution matrix of size n, there is some length of the solution that has the maximum frequency (these values are shown in bold).

b) as the value of n increases, the number of solutions with different lengths increases. Accordingly, the relative frequency of all solutions decreases.

c) for each solution matrix of size n, there is a certain minimum size of the solution length, below which short solutions do not occur. With increasing value of n, the value of this threshold increases. For example, for n = 8, the threshold value is 4, respectively, for n = 16, the threshold value is 7.

4. How complete solutions are located in a sequential list of all solutions?

In the formulation of the n-Queens problem there are no reasons that would give grounds for making any assumption about the order of following complete solutions in the general list of all solutions. We were interested in whether the complete solutions are distributed uniformly in the general list, randomly, or it has some peculiarities. To find this out, we analyzed the distances between the nearest complete solutions in a sequential list of all solutions. As can be seen from Fig. 4, where for n = 12, a histogram of the distribution of the corresponding frequencies is presented, with the greatest frequency there are complete solutions that directly follow one another. These are cases of the formation of the search branch, when the relationships of free positions in the last rows allow one to form two or more consecutive full solutions. Further, the maximum frequency has those complete solutions, between which are located: one short solution, two short solution, etc.

In order to find the regularities in the location of the complete solutions in the general list of all solutions, we analyzed the lists of all solutions for n = 7, ..., 16. To graphically demonstrate the results, in Figure 5, for the value n = 8, the length of each solution in the list of all 736 solutions is indicated. Here, only 92 solutions are complete, and they are highlighted in red, the remaining 644 solutions are short, and are highlighted in blue. It can be seen that the complete solutions are not evenly distributed in the list of all solutions. There are zones where full solutions are found more often, but there are places where complete solutions are rare or do not occur at all. However, another thing is important here. If we look closely at the blue-red barcode, we can notice one very important feature, all the red lines are symmetrical with respect to some conditional vertical line passing through the middle of the list of solutions. In fact, as the check shows, if at the i-th step from the beginning of the general list there is a complete solution, then, respectively, the complete solution will necessarily be found at step m - i + 1, where m is the size of the list of all solutions. For example, for n = 8, if the first complete solution in the sequential search of all solutions appears at the step 43, then, respectively, the last complete solution in the list will be found in step 736 - 43 + 1 = 694. If the 17th solution for a 10x10 matrix appears in the list at step 368, the symmetric complete solution will appear in the list of all solutions in step 12774 - 368 + 1 = 12407. This rule is valid for a matrix of a solution of any size.

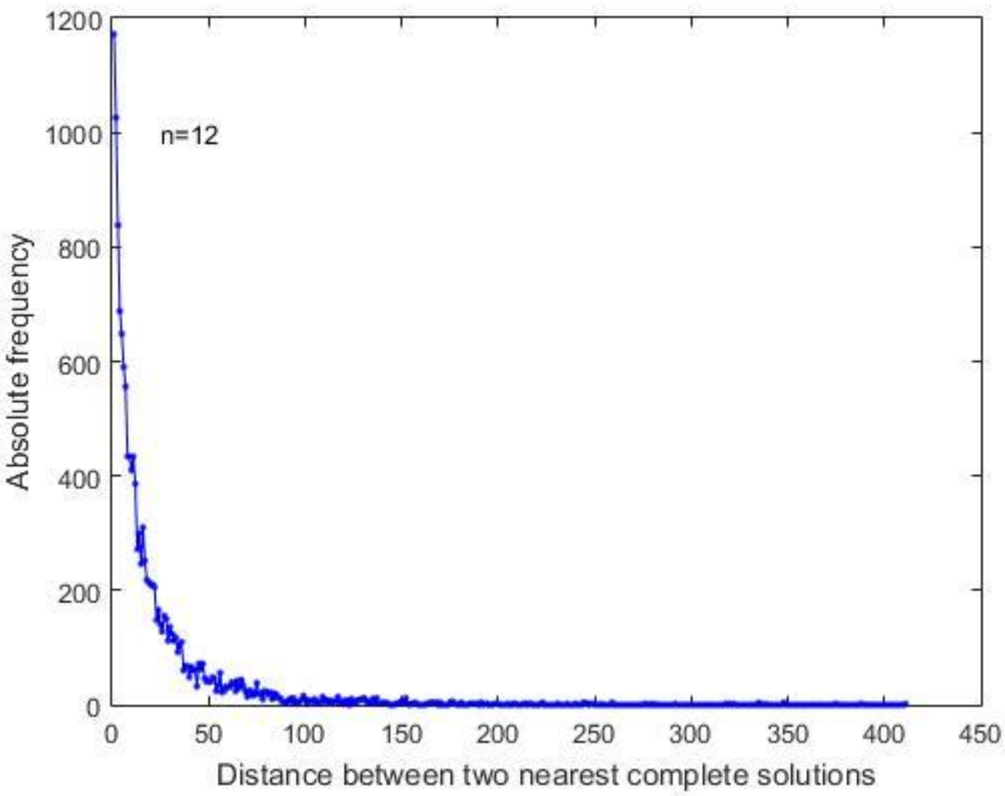

Fig. 4. Dependence of the frequency on the distance between two nearest complete solutions

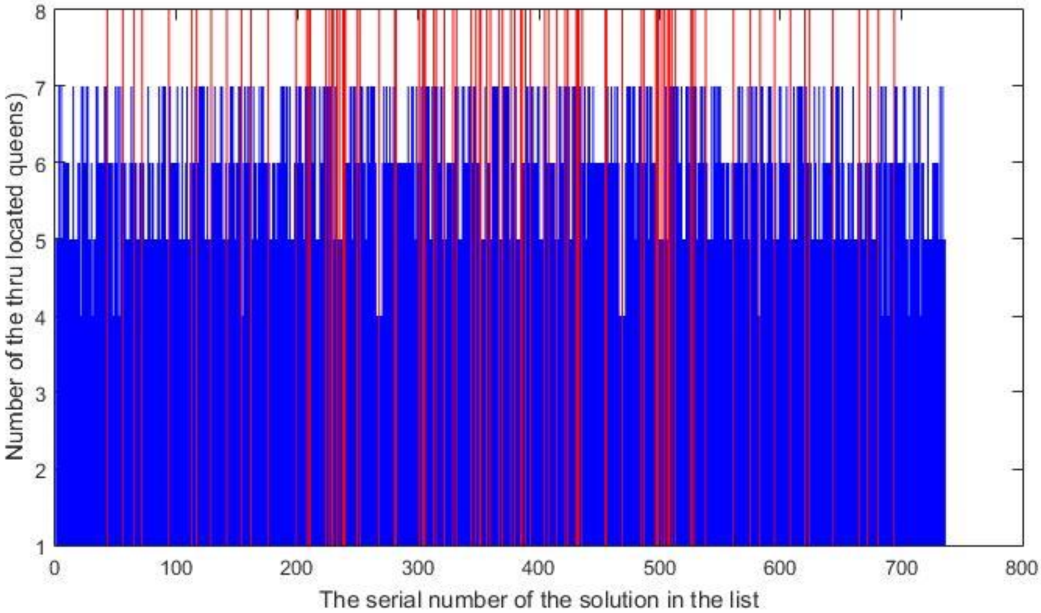

Fig. 5. The length of each solution in a sequential list of all solutions, for a matrix of size **8**x**8** (**red** – full solutions, **blue** – short solutions). The total number of all solutions is **736**

Therefore, we can formulate a rule. For any value of n, if the solution is complete in the sequential list of all solutions in the i-th position from the beginning of the list, then the symmetric solution from the end of the list in the position m - i + 1 will also be complete (rule of symmetry of solutions). The consequence of this rule is the fact that for any value of n, the number of complete solutions will always be an even number. (All the lists of complete solutions found so far are even numbers).

If we compare the queen position indices of any two symmetric solutions, then we can find a critically important feature: symmetric pairs of solutions are complementary. The sum of the corresponding values of the indices of the queens of symmetric solutions is n + 1. For example, the 17th solution for n = 10 in the list of all solutions is in the 368th step from the beginning of the list and the indexes of the queen positions at this step are (1, 5, 7, 10, 4, 2, 9, 3, 6, 8). The corresponding symmetric solution is at step 12407 and has queen positions indexes (10, 6, 4, 1, 7, 9, 2, 8, 5, 3). If we add the corresponding values of the indices of each pair, we obtain (11, 11, ..., 11). This rule is valid for any value of n, moreover, both for complete symmetric solutions and for short symmetric solutions. This allows us to formulate the second rule. For a matrix of solutions of any size n, any pairs of solutions (both short and complete) arranged symmetrically in a sequential list of all solutions are complementary - the sum of the indices of the positions of the corresponding rows is a constant and is equal to n + 1 (the complementarity rule for solutions). If we denote by Q (I) and Q1 (I) the arrays of indices of complementary solutions, then the rule

$Q (i) + Q1 (i) = n + 1, \quad i = (1, n)$

This rule means that if a complete solution is obtained at the i-th step, then the symmetric complete solution at step m - i + 1 becomes known. Therefore, when searching for all complete solutions, it is sufficient to find only the first half of all complete solutions. The second half of the list of complete solutions can be determined from the solutions already obtained, on the basis of the complementarity rule. The criterion that half of the list of complete solutions is reached is the fulfillment of the complementarity rule between the previous complete solution Q (i - 1) and the subsequent Q (i). that is, it is necessary that the sum of each pair of corresponding values of the indices of two consecutive solutions be equal to n + 1. Since any complete solution from the list of all complete solutions is unique, only those consecutive full solutions will be complementary, which are on both sides of the border separating the list in half.

These two rules will allow in the future, in the search for all complete solutions for any next value of n, to reduce the amount of calculations and, correspondingly, the counting time in half.

5. Visualization the sequence of steps to find the first complete solution

How is the process of performing forward steps (Forward Tracking) and backward (Back Tracking) when forming the solution search branch. In order to answer this question, we, for a matrix of 10x10, determined a sequence of 194 transitions between the rows until the first complete solution appeared. The corresponding graph is shown in Figure 5. Blue lines mean forward movement, and red lines - return back. During these 194 steps, 35 short solutions were created, there were transitions between different lines without creating any solutions and, in the end, a complete solution was obtained. The figure shows that most of the transitions (84.5 %) occur between the lines (5, 6, 7, 8). This is, in a way, a "bottleneck" on the way to the "goal". As follows from the figure, there are only 7 cases of transition to the 4th row and one case of transition to the third row. There are also 13 cases of transition to the 9th row. Three attempts to go to the 10th row were unsuccessful, since there was no free position in these search branches on the 10th row. This example clearly demonstrates all the branches of the formation of short solutions, until the first complete solution is obtained.

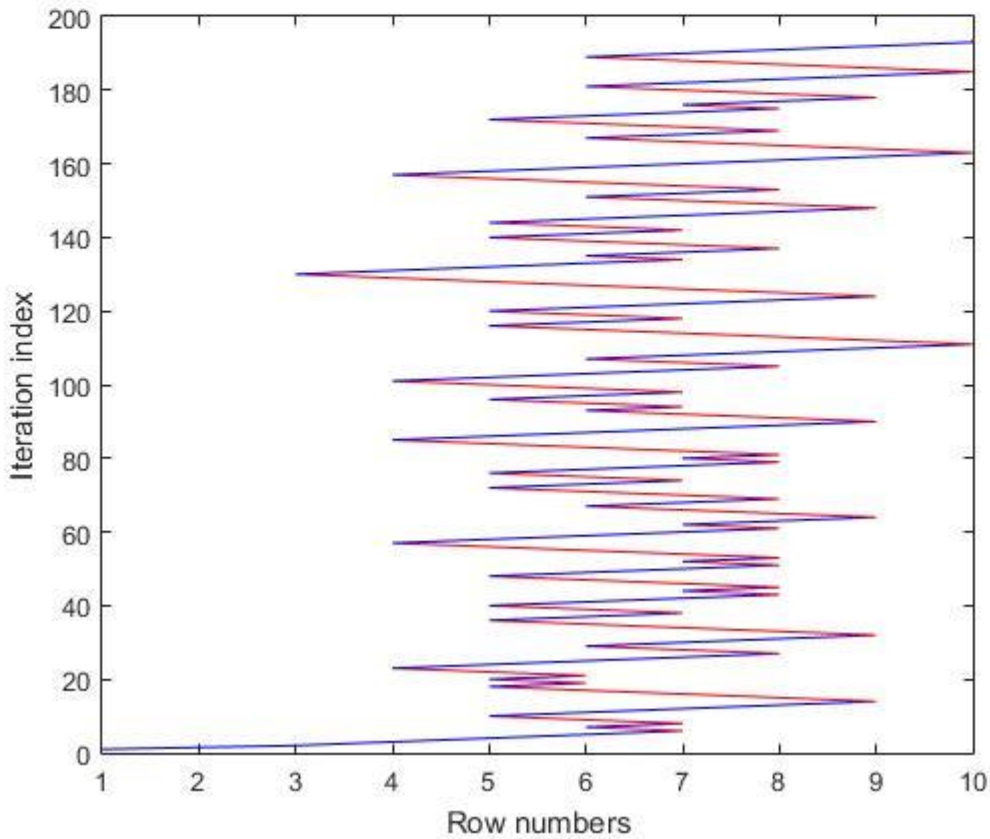

Fig. 6. Visualisation of BackTracking (**red**) and ForwardTracking (**blue**) for first **194** steps of search of solution (**n**=10)

Any algorithm for solving such a problem will be effective if it contains a mechanism that will exclude all (or part) branches leading to short solutions.

6. After what number of short solutions does the first complete solution appear?

Considering that complete solutions appear unequally at different sections of the list of all solutions, it is important to find out through what number of short solutions the first complete solution appears. To this end, for the values n = 7, ..., 35, all short solutions were sequentially calculated before the first complete solution appeared. As can be seen from Figure 6, where the graph of the dependence on n, the natural logarithm of the step number is presented, when the first complete solution appears, for even values of n the first complete solution appears much later than for the nearest odd values of n. For example, for an odd value of n = 21, the first complete solution appears at step 3138, and for the next, even value of n = 22, the first complete solution appears at

628169 step. Accordingly, for the next, odd value of n = 23, the first complete solution appears at step 9155. The number of iteration steps for even n = 22, respectively, is 200.2 and 68.6 times greater than for the nearest odd values. Especially evident in the counting time, this is manifested for n = 34. Here, the first complete solution appears on the step 826 337 184, and for the nearest odd numbers (33, 35), respectively at step 50 704 900 and 84 888 759. It should also be said about the violation of the monotonicity of the growth of the number of short solutions until the appearance of the first complete solution, with increasing n. For even values of n, this occurs for n = 19, for odd ones, for n = 24 and n = 26.

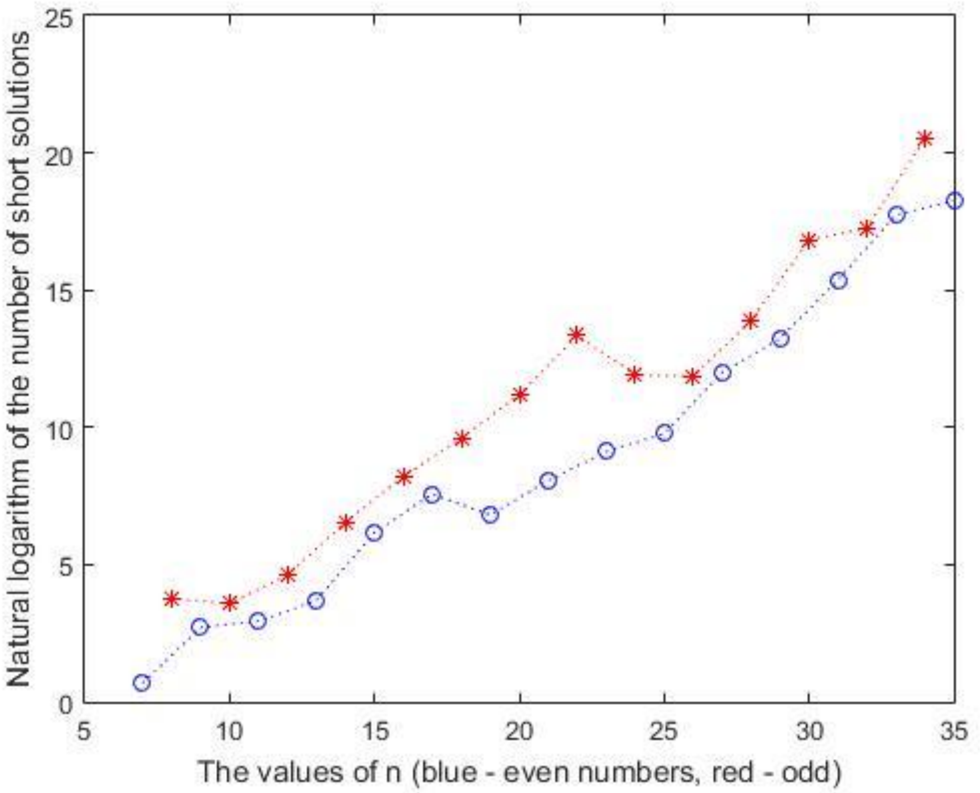

Fig 7. Number of short solutions until first complete solution appears for different **n**

7. Is the frequency of occurrence of cells of each line in the list of all complete solutions dentical?

The n x n size solution matrix, which is an analog of the chessboard, is like the scene where all events occur. Any complete solution formed on this scene consists of cell indices of different rows, since each such cell is a node in the solution search branch into the depth. The question that will interest us – is the activity of cells in each row the same, when forming a list of all complete solutions? Obviously, the higher the value of the frequency, the higher the activity of this cell will be in the formation of the list of complete solutions. To establish this, we define for each row on the basis of a list of all complete solutions, the relative frequency of occurrence of different indices. First, we perform an analysis for a solution matrix of size n = 8. Let's consider sequentially each row of the storage array of complete solutions and determine the frequency of different index values. In Table 2, the corresponding values of the absolute frequencies of the activity of the different cells in each of the eight rows are presented, and in Figure 6 shows a group of 4 graphs, where each graph characterizes the change in relative frequencies within a single row. One of the fundamentally important conclusions that can be drawn from an analysis of all the data obtained is as follows:

- for a matrix of a solution of arbitrary size n x n, the activity of the cells of the i-th row coincides with the activity of the cell n - i + 1, i.e. the activity of the cells of the first row always coincides with the activity of the cells of the last row, respectively, the activity of the cells of the second row coincides with the activity of the cells of the penultimate row, and so on. In Table 2, for clarity, the frequencies of the first and eighth row and column cells are highlighted in bold.

Table 2. Absolute frequency of cell activity in each of the eight rows of the solution matrix 8x8, obtained on the basis of an analysis of the list of all complete solutions

row\col	1	2	3	4	5	6	7	8
1	**4**	**8**	**16**	**18**	**18**	**16**	**8**	**4**
2	**8**	16	14	8	8	14	16	**8**
3	**16**	14	4	12	12	4	14	**16**
4	**18**	8	12	8	8	12	8	**18**
5	**18**	8	12	8	8	12	8	**18**
6	**16**	14	4	12	12	4	14	**16**
7	**8**	16	14	8	8	14	16	**8**
8	**4**	**8**	**16**	**18**	**18**	**16**	**8**	**4**

- in the case when n is odd, only the median row of the solution matrix does not have a symmetric pair, for all other cells the above rule is valid.

We call this, "the property of horizontal symmetry of the activity of cells of different rows of the solution matrix". For this reason, we gave only 4 graphs for a matrix of a solution of size n = 8, since the cell activity graphs for rows (1, 8), (2, 7), (3, 6) and (4, 5) are completely identical.

It should also be noted that all graphs are symmetric with respect to the vertical axis dividing the matrix into two equal parts (in the case of an even value of n), or passing through a median column (in the case of an odd value of n). We call this, "the property of vertical symmetry of the activity of cells of different rows of the solution matrix". From Table 2, it is seen that the 1st column completely coincides with the last column and they are completely identical to the values of the first row. Similarly, the second column completely coincides with the seventh column, and their values completely coincide with the second row, etc. This means that the frequencies in the solution matrix are symmetric with respect to the left and right main diagonals.

I think that the presence of limiting rules in the formulation of the problem, and the associated property of nondeterminism, "create" some kind of harmonious relationship between nodes in different lines. Those branches of search that fit into these rules - lead to the formation of a complete solution. The remaining branches of the search, at some point violate these rules, and in the end, "complete their way" in the form of short solutions. Here it should be noted that the cells of the solution matrix have only a local relationship within the projection impact group. There are no prescribed rules for concerted action between them. Collective activity of cells is only a consequence of the result of the impact of restrictive rules. Therefore, an interesting question remains open, how the restrictive rules, as factors of nondeterminism, influence the cells of the solution matrix, which ultimately leads to the formation of a "harmonious" matrix of cell activity – symmetric with respect to the horizontal and vertical axes, as well as relative to the left and right principal diagonals. Is this a characteristic property of only this task, or does it occur for other nondeterministic tasks?

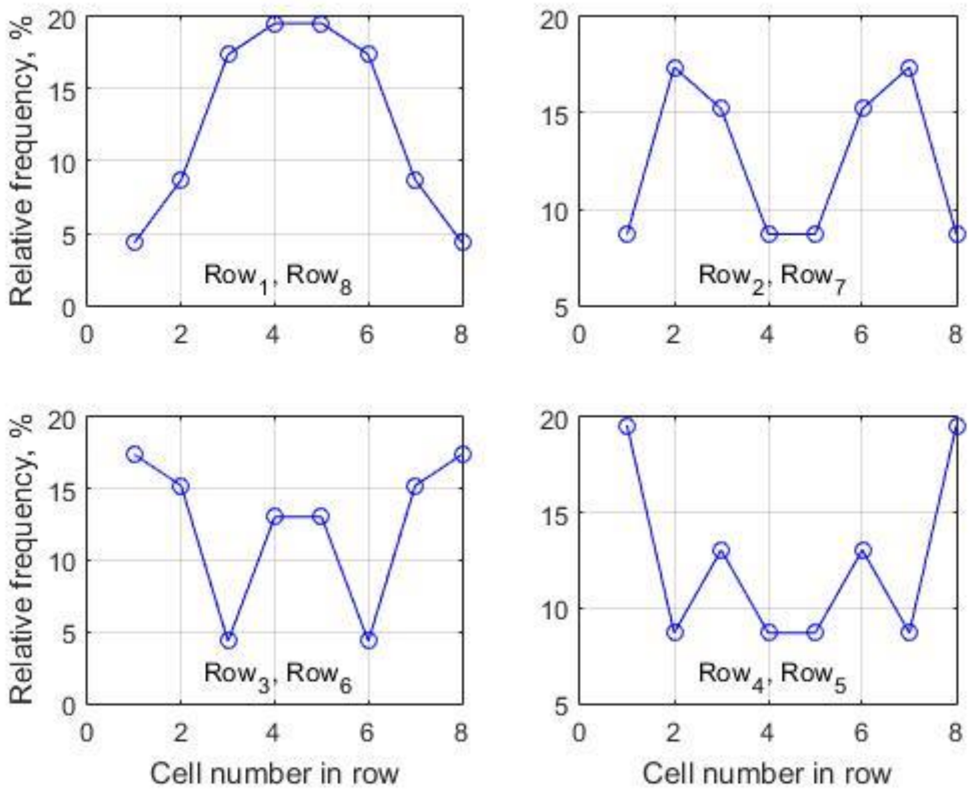

Fig. 8. The activity of the cells of each row when forming a list of complete solutions, **n**=**8**

8. From which row number is the Forward Tracking – Back Tracking algorithm included?

If we follow the sequence of the algorithm's actions, when a row is selected in the solution matrix for the location of the queen, we can see that starting with some row, which we will call "StopRow", there is a "slowdown" of the process of moving forward. In the search branch, this row is the first, where there are problems with the presence of a free position for the location of the queen. It is from this line that the Forward Tracking algorithm is used to move forward or Back Tracking – to clear traces of previously performed actions, and to return back. This is the row on which the first short solution appears.

The index of "StopRow", with which difficulties begin to move forward, depends on the size of the solution matrix. If we consider the ratio of this index, which we denote by StopInd to the size of the solution matrix n, then, as can be seen from the Figure 9-1, where the calculation results for the initial values n = 7, ..., 99 are presented, this ratio varies more or less and tends to decrease. As the value of n = (100, ..., 300) increases, this ratio ranges from 0.619 to 0.642 (Figure 9-2), and with a further increase in n, we get the following results (successively: n (StopInd, StopInd / n): 1000 (619, 0.6190), 2000 (1239, 0.6195), 3000 (1856, 0.6187), 4000 (2473, 0.6182), 5000 (3091, 0.6182). It is surprising, but it can be argued that the stop line divides the matrix according to the rule of the golden section: namely, the StopInd / n relation differs from (n-StopInd) / StopInd by a small value, which tends to zero with increasing n. For example, for n = 5000, the difference between the ratios 3091/5000 and 1909/3091 is 0.006, which means less than 0.1% of the average of these two ratios.

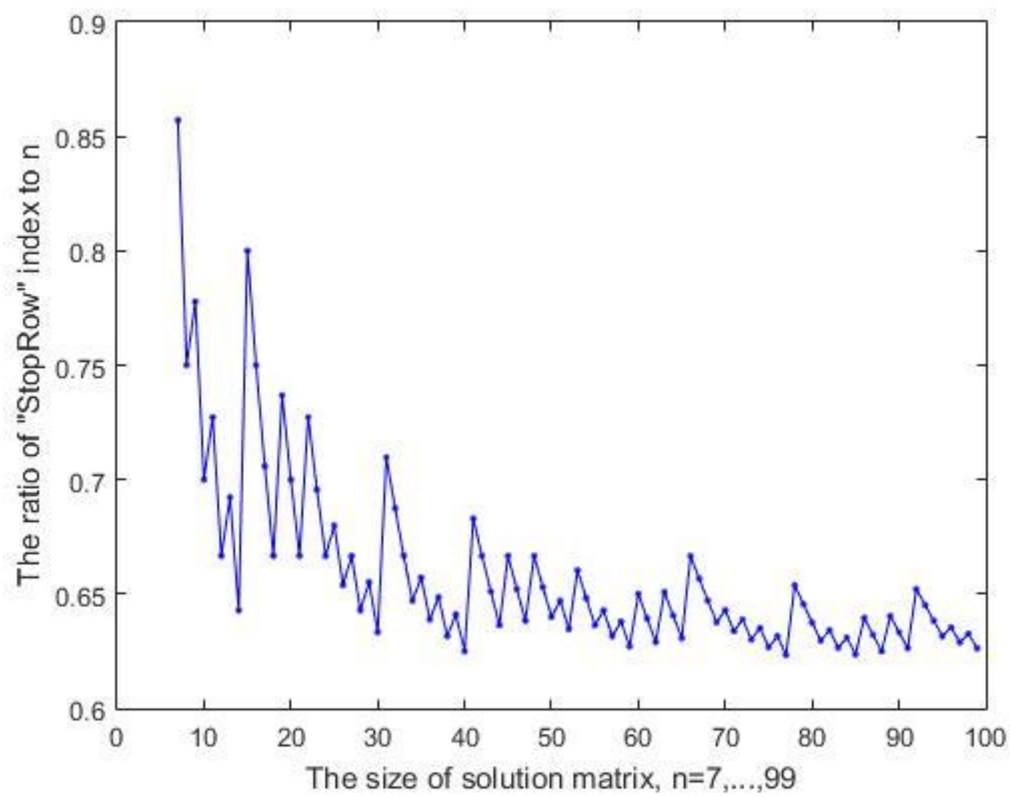

Fig. 9-1. Dependence the ratio of **StopRow** index to **n** on the solution matrix size (part-1)

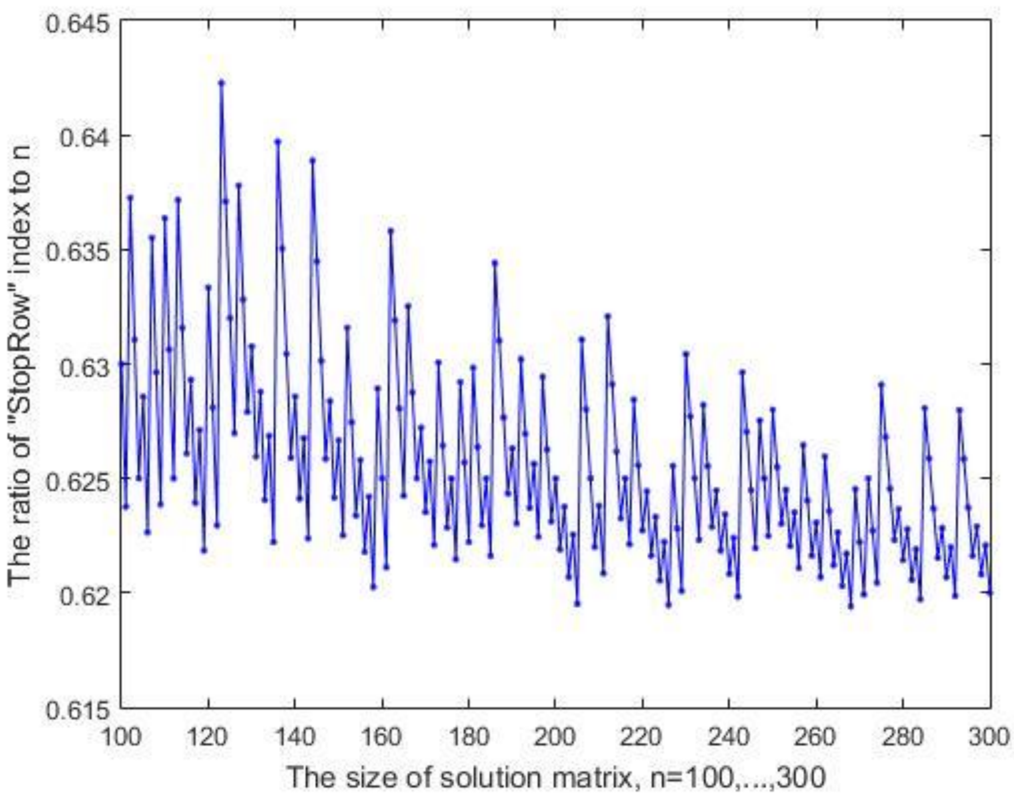

Fig. 9-2. Dependence the ratio of **StopRow** index to **n** on the solution matrix size (part-2)

The graph presented in two figures Figure 9-1, 2 has not a random form of variability, which resembles a record on the "music camp". One can see repeated jumps upward and a stepwise fall down with some irregular periodicity. Obviously, there is some reason for this behavior of the curve, and perhaps this will be of interest for the study. For this reason, for a more detailed visualization, the graph was presented in two figures.

9. What is the probability of completion an arbitrary composition of k queens to a complete solution?

To answer this question, we need to determine all possible admissible combinations of k queens (k = 2, ..., n-1) for each of the considered values of the size n of the solution matrix. Here, by permissible compositions, we mean such combinations that do not contradict the conditions of the task. After this, we can compare the resulting list of compositions with a list of complete solutions and determine the number of those compositions that are at least once found in any of the complete solutions. Obviously, in the process of generating these compositions, we must keep the queen's position to the row index in order to make a correct comparison. Such a format of data representation will allow us, for example, to confirm that for n = 10, the composition of two queens (0,0,7,0,0,0,0,3,0,0) can be supplemented to a complete solution, since the active positions in this composition, coincide with the corresponding positions in the solution (1, 5, 7, 10, 4, 2, 9, 3, 6, 8), which is the first in the list of complete solutions for n = 10.

We made a selection of all possible combinations of k queens based on a pseudo-random number generator, taking into account the limiting rules: "in each line, each column, and on the left and right diagonals passing through the cell where the queen is located, there should not be more than one queen". At the same time, we observed two conditions: a) the generation was carried out for a sufficiently large number of samples in order to cover all possible combinations with a high probability; b) before the analysis, classification of the obtained data was carried out, and those samples that coincided with the already selected compositions were excluded from the sample. It should be noted that this is a fairly time-consuming computational task, since the sample sizes are very large. Further analysis was carried out only on the basis of unique compositions. This way of generating combinations of k queens, quite accurately characterizes the variety of compositions in a real situation.

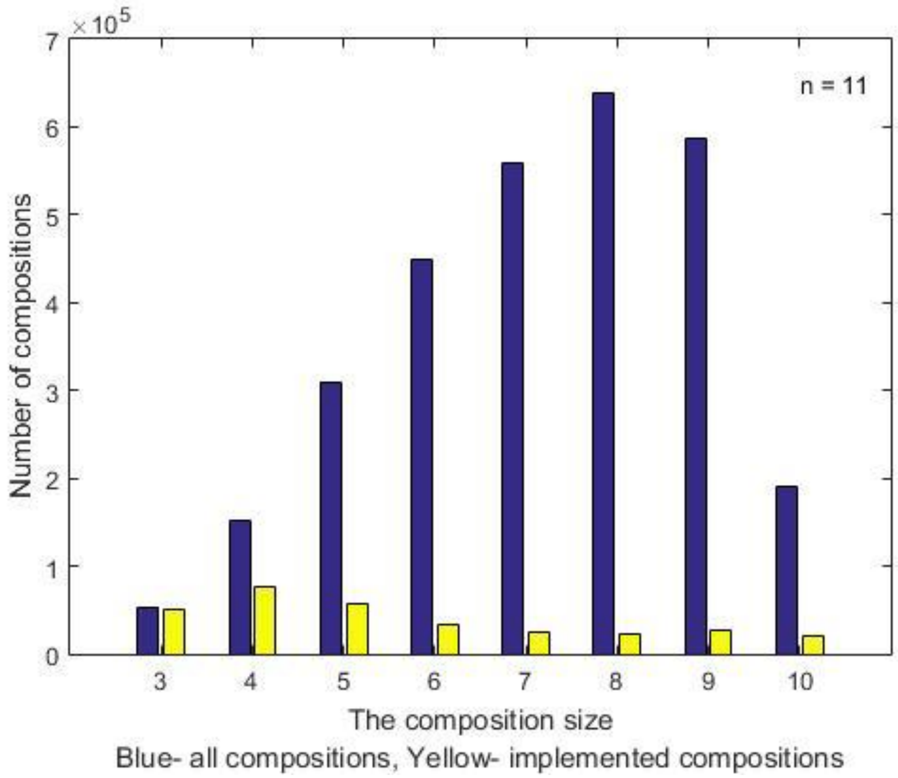

Fig. 10. The number of initial and implemented compositions of different size for the solution matrix 11x11

It should be noted that the calculations that we carried out for the same value of k yielded very similar results. This is due to the observance of the generation condition for a sufficiently large sample of random compositions. For example, for n = 11, the following samples of unique compositions were obtained (the value of k is given, and the sample size of unique compositions is indicated in parentheses): 3 (53190), 4 (151786), 5 (309764), 6 (449629), 7 (558196), 8 (637556), 9 (586628), 10 (191414). In order to obtain such samples, we each time generated one million compositions, some of which, then, as a result of classification, were filtered out. Calculations were carried out by us for ten values of the solution matrix, n = 7, ..., 16.

Table 3. The probability of completion to obtain the complete solution of an arbitrary composition of k queens located on a n x n-size solution matrix (n = 8, 9, 10)

k \ n	8	9	10
2	0.8168	0.9817	1.0000
3	0.3907	0.6569	0.7326
4	0.1763	0.2942	0.2944
5	**0.1136**	0.1463	0.1170
6	0.1145	**0.1061**	**0.0615**
7	0.2306	0.1185	0.0471
8		0.2290	0.0589
9			0.1284

Some of these results are presented in Table 3, (n = 8, 9, 10). In Figure 10, as an example, for a 11x11 size solution matrix, a set of bar charts is presented that correspond to the sample size of the generated compositions and the sample size of those compositions that have been confirmed, at least in one solution, from the list of complete solutions. Accordingly, and in Figure 11, for n = 8, 10, we presented graphs of the change in the probability of completion to obtain a complete solution. The most important conclusions that can be drawn from the analysis of the results obtained are as follows:

a) the probability of completing to obtain the complete solution of an arbitrary composition of k queens is not a constant value. The value of this probability gradually decreases to a certain minimum value, and then increases with increasing value of k. This is true for a matrix of a solution of any size. For example (Table 3, Figure 11), for n = 10, a minimal probability is 0.0471. This corresponds to compositions consisting of 7 queens. If we increase the size of the composition to 9, then the probability of completion increases to 0.1284, which is 2.73 times greater than at the minimum point.

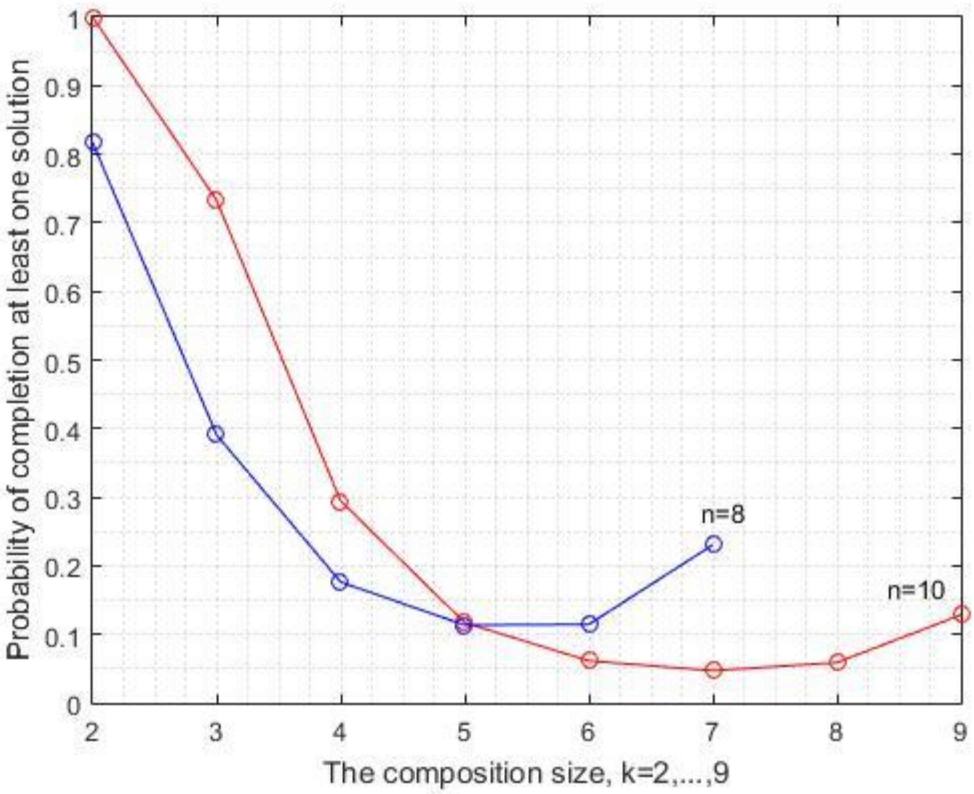

Fig. 11. Probability of completion to full solution depending on the number of queens in the composition. **Blue**- solution matrix **8**x**8**, **red - 10**x**10**

b) for each value of the solution matrix n, there is a certain minimum value of the composition size k0, such that any composition whose size is less than or equal to k0 can be completed to a complete solution. For values of n = 7, 8, 9, the value of k0 = 1. This means that not every composition of the two queens, from the list of acceptable compositions, can be completed to a full solution. For example, for n = 8, the probability of completing to obtain the complete solution of an arbitrary composition of two queens is equal to 0.8168. As the value of n increases, the corresponding value of k0 increases. For example, for n = 16, the value k0 = 5. This means that for a n=16 solution matrix, any composition formed in an allowable search space, whose size is less than or equal to 5, can be completed to a full solution.

In Figure 12, depending on the value of n (7, ..., 16), two graphs are presented: the first is the size of the composition k, at which the probability of completing to the full solution is minimal, on the second – the size of the composition, which can always be completed to obtain full solution. It is seen that the values of both indicators increase with increasing n.

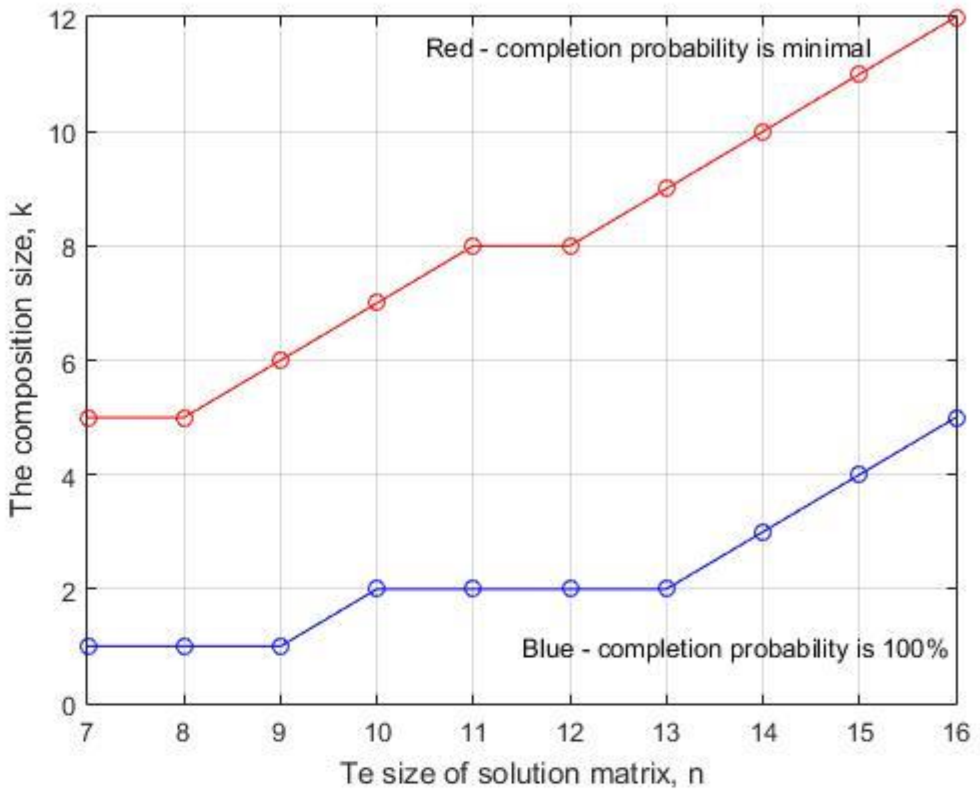

Fig. 12. The composition size at which the probability of completing is minimum (red) or equal 100 % (blue)

3. Conclusion

An analysis was carried out of the sequence of all solutions (both short and complete) for different values of the solution matrix (n = 7, ..., 16). As a result, it was established that for an arbitrary solution matrix of size n x n, the following statements hold:

1. Although the number of complete solutions increases exponentially with increasing value of n, however, their share in the general list of all solutions decreases.

2. Complete solutions are distributed in a sequential list of all solutions in such a way that they are found in the list with the greatest frequency, located close to each other.

3. There is a symmetry in the order of the location of the complete solutions in the general list of all solutions, with respect to the axis passing through the middle of the general list. For any value of n, if the solution is complete in the sequential list of all solutions in the i-th position from the beginning of the list, then the symmetric solution from the end of the list in the position m - i + 1 will also be complete (rule of symmetry of solutions). Here m is the size of the general list of all solutions.

The consequence of this rule is the fact that for any value of n, the number of complete solutions will always be an even number. (All the lists of complete solutions found so far are even numbers).

4. Any pairs of solutions (both short and full) arranged symmetrically in a successive list of all solutions, with respect to the axis passing through the middle of the list, are complementary – the sum of the indices of the positions of the corresponding rows is a constant value and is equal to n + 1 (the complementarity rule for the solutions).

This rule means that if a complete solution is obtained at the i-th step, then the symmetric complete solution at step m - i + 1 becomes known. Therefore, when searching for all complete solutions, it is sufficient to find only the first half of all complete solutions. The second half of the list of complete solutions can be determined from the solutions already obtained, on the basis of the complementarity rule.

The criterion that half of the list of complete solutions is reached is the fulfillment of the complementarity rule between the previous complete solution and the subsequent one, that is, it is necessary that the sum of each pair of corresponding values of the indices of two successive solutions be equal to n + 1.

5. In a sequential list of all solutions, the first complete solutions appear after a certain number of short solutions. For even values of n, the first complete solution appears much later than for the nearest odd values of n. For example, for n = 34, the first complete solution appears at 826 888 759th step, and for the nearest odd numbers (33, 35), respectively at 50 704 900th and 84 888 759th steps.

6. The activity of the cells of the solution matrix is symmetrical about the axis passing through the middle of this matrix. This means that the activity of the cells of the i-th row always coincides with the activity of the cell n - i + 1, i.e. the activity of the first line always coincides with the activity of the cells of the last row, respectively, the activity of the second line - coincides with the activity of the penultimate line, etc. By activity is meant the frequency with which the cell appears in the corresponding row of the list of complete solutions.

7. The row in the solution matrix, where difficulties begin to move forward, and the first short solution is formed, divides the matrix according to the rule of the golden section. For small values of n, such a conclusion is approximate, but with an increase in the value of n, the accuracy of such an output asymptotically increases to the level of the standard rule.

8. The probability of completion to a full solution an arbitrary composition of k queens, gradually decreases with increasing value of k to a certain minimum, and then increases, with a further increase in the value of k.

9. There is some minimum value of the size of the composition k0, such that any composition whose size is less than or equal to k0 can be completed to a complete solution.

Another rule that I would like to add to the list is the following:-in any problem connected with the formation of the branch of the search for solutions in the state space, in the presence of constraints, there must exist some rules for harmonious relations between all nodes of the branch of the search for solutions. The presence of bounding rules in the formulation of the problem, and the related property of nondeterminism "form" some kind of harmonious relation between the nodes of the search branch. This means that the relationship between the nodes of the branch of search for a solution is not accidental. Only those branches of the search, which fit into the harmonious rules of relations inherent in the given problem, lead to the formation of the correct solution. The remaining branches of the search, which at some point "violate" these rules, are ultimately excluded from consideration, as they lead to incorrect solutions. I think that such a fact can also occur in some other nondeterministic problems of forming a search branch in the state space under constraints.

I have considered only some questions that can be formulated on the basis of the results of solving the n-Queens Problem. I hope that the obtained results will make the mechanisms of the formation of nondeterministic processes and changes in the state space more transparent for understanding. Perhaps this will serve as a fulcrum for formulating new tasks and moving ahead.

References

Bezzel, 1848 – *Bezzel M.* (1848). Proposal of 8-queens problem, Berliner Schachzeitung. 3, 363.

Nauck, 1850 – *Nauck F.* (1850). Briefwechseln mit allen fur alle, Illustrirte Zeitung. 15 (377): 182. September 21 ed.

Bell, Stevens, 2009 – *Bell J., Stevens B.* (2009). A survey of known results and research areas for n-queens. *Discrete Mathematics*. 309: 1–31.

Gent et al., 2017 – Gent I. P., Jefferson C., Nightingale P. (2017). Complexity of n-Queens complmpletion. Journal of Artificial Intelligence Research, 59: 815-848.

Priestley, Ward, 1994 – *Priestley H.A., Ward M.P. Ward* (1994). A Multipurpose Backtracking Algorithm. *Journal Of Symbolic Computation,* 18.1: 1-40.

Sosic, Gu, 1994 – *Sosic. R., Gu J.* (1994). Efficient local search with conflict minimization: A case study of the n-queens problem. *IEEE Transactions on Knowledge and Data Engineering*, 6(5): 661-668.

Mańdziuk, 2002 – *Mańdziuk J.* (2002). Neural networks for the N-Queens Problem: a review. *Control and cybernetics*. v.31, № 2.

Farhan et al., 2015 – *Farhan A.S., Tareq W. Z., Awad F. H.* (2015). Solving N Queen Problem using Genetic Algorithm. *International Journal of Computer Applications*. Volume 122 – No.12.

Kosters, 2017 – Kosters W. (2017). n-Queens – 339 references. [Electronic resource]. URL: www.liacs.leidenuniv.nl/~kosterswa/nqueens/

Sloane, 2016 – *Sloane N.J.A.* (2016). The on-line encyclopedia of integer sequences. [Electronic resource]. URL: https://oeis.org/search?q=A000170&language=english&go=Search