American Journal of Engineering Research (AJER)
e-ISSN: 2320-0847 p-ISSN: 2320-0936
Volume-15, Issue-6, pp-58-64
www.ajer.org
Research Paper Open Access

Optimal Knowledge Entropy

Kucherov O. P.[1], and Mudryk A.M.[2]

1 Institute for Information Recording of the National Academy of Sciences of Ukraine.
ID: 0009-0009-9881-2691
2 Company "GIGAHARD" Ltd. EUROPE, Plovdiv, Bulgaria.
Correspondence: O.Kucherov@i.ua

ABSTRACT: *In this paper, an algebraic solution of the Kucherov equation* ***N=Ω^α*** *for the optimal knowledge entropy is found. A vector knowledge graph is presented, which consists of a set of information objects and is described by a route sheet.*
The volume of knowledge in the graph is determined by entropy (the degree of ordering into a hierarchical structure), which is measured by the average length of routes from the root to the objects. The route consists of zit vectors {visit →, transit ↓}. The zit (short for visit → and transit ↓) is a unit of measurement of knowledge in a structure. Similarly, the unit of measurement of information is the bit. Routes are recorded by structural numbers in the route sheet. Operations are performed by a structural operating system. Operations include visualization of the knowledge graph, management of a hierarchical database, data integration, complex comprehensive analytics, performing quantum calculations (e.g., quantiles), and the construction of diagrams and cartograms.

KEYWORDS: *State management, hierarchical database, route, route sheet, vector knowledge graph, knowledge entropy, Kucherov equation , zit – the unit of measurement of knowledge in a structure.*

Date of Submission: 05-06-2026 Date of acceptance: 16-06-2026

I. INTRODUTION

As P. Drucker stated in the book "Tasks of Management for the 21st Century" [1], knowledge is used to analyze a variety of facts and formulate strong arguments. To lead management teams, an experienced manager must ask: "Where is what I need?" Knowledge must answer this question that is, according to Drucker, knowledge serves as a means of accessing target information within in a structure.

However, the question remains: how can we build an optimal hierarchical knowledge structure that provides the shortest access to necessary information - a structure with minimal entropy?

II. TASK

The purpose of this work is to establish a formula for calculating the most efficient knowledge structure for a specific set of information objects. This formula will allow leaders to address challenges in managing ecological, social, and economic systems.

III. OPTIMIZATION OF NATUREAL LANGUAGE ALPHABET

Creating an optimal structure for quick access to a target object is similar to synthesizing an optimal alphabet for fast message transmission. The theory of constructing an optimal alphabet for transmitting information over communication channels was pioneered by Claude Shannon [2]. Optimization is based on reducing message entropy by shortening the code length for frequency of use symbols. Research into the frequency of letters in English revealed that the language allows for nearly double text compression. The development of an algorithm for constructing an optimal alphabet for natural languages was furthered by O. P. Kucherov [3]. Specifically, an optimal Ukrainian alphabet using Latin characters was developed [4], allowing for 10% compression. In subsequent research, an algorithm for reducing entropy in knowledge structure graphs was applied. Thus, the optimal path to access target information should be shorter for information that is accessed more frequently.

IV. ROUTING A SET OF OBJECTS IN A KNOWLEDGE GRAPH

The process of cognition is preceded by the accumulation of information about a set of objects and their subsequent grouping by properties. The groups thus obtained become new objects of the set, which are also grouped, but already at a higher level of the hierarchy.

Thus, a vector graph of knowledge appears, covering all objects of the set.

Each object in the set is assigned a route, which accurately and uniquely determines the position of each object in the vector graph of knowledge and establishes its subordination. According to the theory of knowledge of O.P. Kucherov [5], the path to the object is given by a route. The route is denoted by the sign ^ (read: route). Thus, ^ (route) accurately and uniquely determines the position of each object in the vector graph of knowledge.

The route sheet contains a list of the objects in the set.

These graphs cover a wide range of knowledge sets, which include classifiers, codifiers, organizational structures of enterprises, states, and unions. The most common and understandable route sheet is the table of contents of a book. We will demonstrate the routing of the UN structure.

Table 1.

Route sheet of the UN structure.

Knowledge structure	Object name
^	**The United Nations structure**
^A	**The General Assembly**
^AA	***Main Committees***
^AAA	Deals with disarmament, global security, and related international issues.
^AAB	Focuses on economic and financial matters.
^AAC	Addresses social, humanitarian, and cultural issues.
^AAD	Handles special political questions and decolonization.
^AAE	Manages the UN's administrative and budgetary affairs.
^AAF	The primary forum for international legal questions.
^AB	***Other Key Committees***
^ABA	General Committee: Composed of the Assembly President, Vice-Presidents, and Chairmen of the Main Committees.
^ABB	The Assembly can establish other necessary committees.
^B	**Security Council**
^C	**Economic and Social Council**
^D	**The Trusteeship Council**
^E	**The International Court of Justice**
^F	**The Secretariat**

According to the route sheet, each object has a route ^ and a name. For example, one of the objects of the set given in Table 1 is ^AB, Other Key Committees. Therefore, the route sheet describes the organizational structure.

Thus, the route to the object is a set of numbers. To denote the object`s number within a class, a Latin symbol is often chosen. The route and the name of the object are recorded in the route sheet. The number of route hierarchies is determined by the number of symbols. Since the number of hierarchical levels and the number of objects in the knowledge graph are unlimited, the route sheet for the hierarchical structure of the knowledge graph is also unlimited. Any infrastructure can be inserted or removed in any part of the graph, and the route sheet itself can be made part of another, larger knowledge graph.

As will be shown below, the presented structural arithmetic allows us to calculate the amount of knowledge in any knowledge graph, and also allows us to find the location of objects in a graph with both maximum chaos and maximum order using optimal knowledge entropy.

V. KNOWLEDGE VECTOR

A knowledge graph is usually understood as a directed acyclic graph with a defined vertex (the root of the graph), from which there is a unique, route-defined path to each object in the structure. Knowledge graphs are used in optimization theory, mathematical statistics, operations research, pattern recognition, reliability theory, etc. A vector knowledge graph describes a hierarchical structure in which objects of a set are organized. The structure consists of child and parent classes.

The vector knowledge graph is built on the basis of the vector knowledge theory of O. P. Kucherov [5]. We will provide the main definitions of this theory.

A vector knowledge graph is defined by routes to each object of the set. Each route consists of a certain number of vectors.

Let us consider the vector form of knowledge representation. A graphic description of a structure capable of covering a large set of objects and the connections between them, which allows for a simple and understandable representation of their structure, is carried out using vectors.

A knowledge base over a set of objects is a directed vector knowledge graph where each object has one vector at the input and two vectors at the output. At the output, the vectors create a bifurcation. If all information can be represented by a sequence of two digits {0,1}, then knowledge can be represented by an access route, which is a uniquely defined combination of two zit vectors {visit →, transit ↓}. Vector Visit → means transition to a child class. Vector Transit ↓ means transition to the next object in the current class.

Fig. 1. Information recording digits.

Fig. 2. Knowledge recording vectors.

Structures consist of objects connected by zit vectors {visit →, transit ↓}. All objects of the structure are numbered ^ by routes leading from the root to the object.
According to the route sheet, Tabl1e, the vector knowledge graph for the UN structure takes the following form.

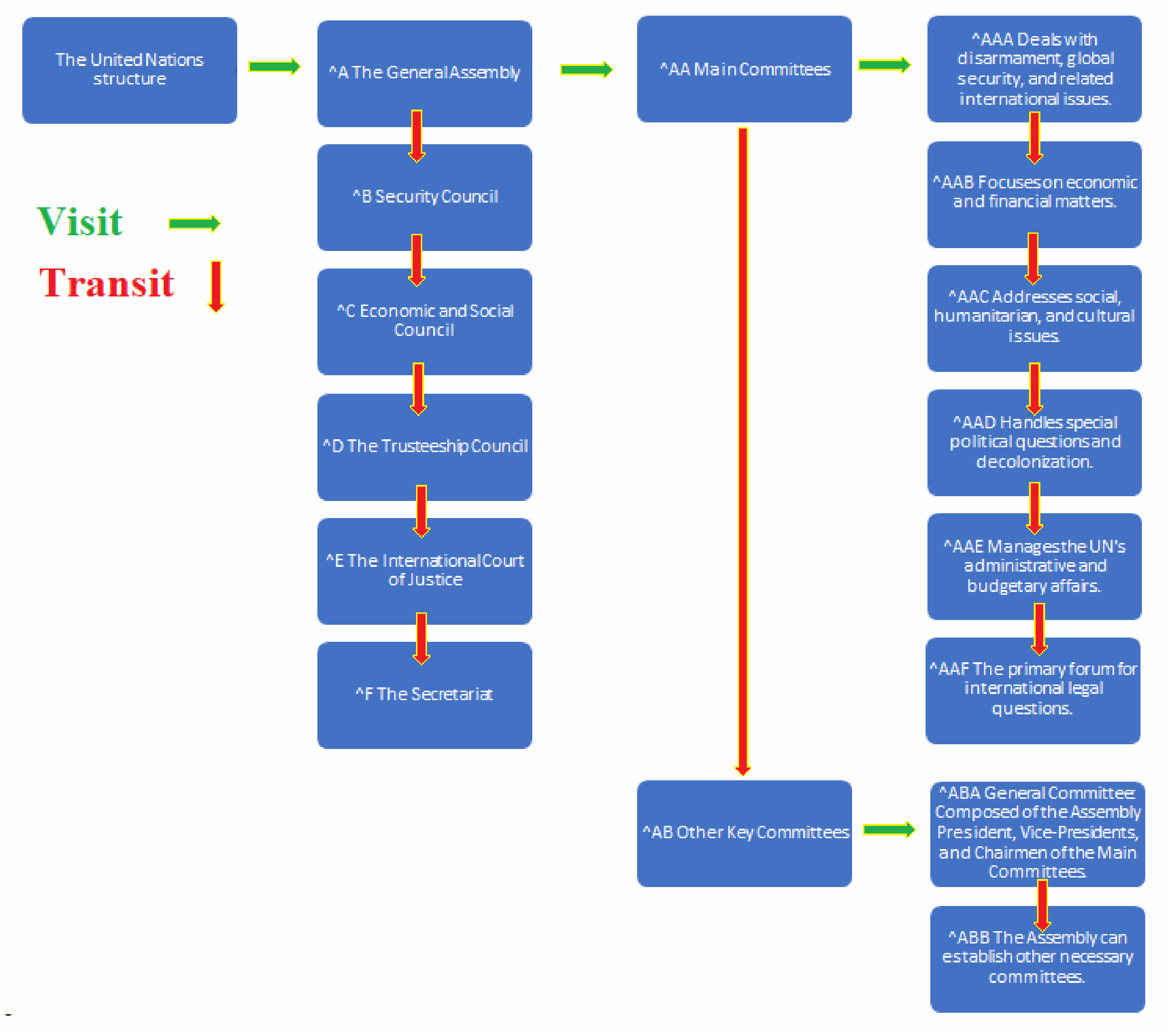

Fig. 3. Vector knowledge graph of the UN structure.

The vector knowledge graph describes a hierarchical structure in such a way that from each object one can move in the direction of one of two zit vectors: visit → or transit ↓. Figure 3 shows that there is a unique path to each object. The length of the path to an object is measured by the number of zit vectors. It is possible to find the average path length for all objects in the graph. As we will show below, the average path length across a vector knowledge graph is of fundamental importance to knowledge theory.

VI. UNIT OF MEASUREMENT FOR THE AMOUNT OF KNOWLEDGE IN STRUCTURE

The number of zit vectors determines the route length S_n to object n.

For example, consider the route to the object n "^ABB The Assembly can establish other necessary committees" with the route ^ABB. The route is movement along zit vectors {visits → and transits ↓}. In vector form, it has the following route: →→↓→↓, consisting of five zit vectors: three visits → vectors and two transits ↓ vectors. That is, the route length $S_{\wedge ABB} = 5$ zits.

In accordance, the binary zit vector {visit → or transit ↓} is a unit of measurement of knowledge in structure. This definition was first proposed by O.P. Kucherov in the monograph [5].

However, there is an alternative way to calculate the object access length using structural arithmetic. The length of the route to the same element ^ABB, calculated by means of structural arithmetic by adding structural numbers, will be $S_{^{\wedge}ABB}$ = $A+B+B$ = $1+2+2$ = 5 zits. That is, the two methods of calculating the length of the route give the same result, because they are equivalent.

This allows you to build route sheets, which are algorithms for a structural operating system. Operations include visualization of the knowledge graph, management of a hierarchical database, data integration, complex comprehensive analytics, performing quantum calculations (e.g., quantiles), construction of diagrams, and cartograms.

VII. AMOUNT OF KNOWLEDGE

Knowledge reduces chaos, increases order, and builds a working structure. Knowledge is accumulated to create the most ordered structure with the shortest access routes to target objects, which significantly simplifies the architecture of large databases and speeds up economic calculations for making management decisions.

The definition of basic operations on the knowledge graph, such as the route length in zits, allowed laying the foundations of the mathematical theory of knowledge. The volume of knowledge in a hierarchical structure is determined by the shortness of the average route in zits. Let us demonstrate this.

Let the route to the object n have a length S_n, in terms of the number of zit vectors: visits → and transits ↓. Then the amount of knowledge is determined by entropy, which is equal to the length of access according to the formula of O.P. Kucherov [6]:

$$S=\frac{1}{N}\sum_{n=1}^{N} S_n. \tag{1}$$

Thus, the average length of access S to the target information in the vector graph of knowledge is the entropy of knowledge.

Longest path to the target information.

For any set of N objects, there is the most inconvenient structure with zero knowledge, that is, with the longest average length S. This is an unstructured set in which all objects are arranged in a single row without hierarchy. The length of the route S_n of access to object n in this case is equal to the serial number of this object, $S_n = n$. Accordingly, formula (1) takes the form:

$$S_{max}=\frac{1}{N}\sum_{n=1}^{N} n=\frac{(N+1)}{2}. \tag{2}$$

This formula describes the case of complete chaos, when the entropy takes its maximum value $S = S_{max}$. Please note that in many traditional databases, including relational databases, knowledge is absent; records are arranged randomly without thematic hierarchy. Access to each object is performed by a complete search for the computer or by sequential scrolling for the user.

Here, the entropy of knowledge is large but finite, and its value is calculated according to the simplified formula (2).

Shortest average path to the target information.

The theorem on the existence of a knowledge structure with a minimum access path (minimum entropy) was proved by O. P. Kucherov [6]. According to this theorem, for any set of objects, there exists the most convenient structure with the maximum amount of knowledge and the smallest entropy, for which the number of objects N and the shortest access length Ω are related by the equation:

$$N=\Omega^{\Omega}. \tag{3}$$

This is the Kucherov equation, which establishes the relationship between the number of objects N in the set and the minimum access length Ω for the optimal knowledge structure. That is, it determines the structure with minimal entropy or maximal ordering.

VIII. FORMULA FOR THE OPTIMAL KNOWLEDGE ENTROPY

We can solve equation (3) to find a function $\Omega(N)$ relating the optimal knowledge entropy Ω to the number of information objects N in the knowledge graph. First, we will plot the dependence of Ω on $Ln(N)$, as shown in Figure 4.

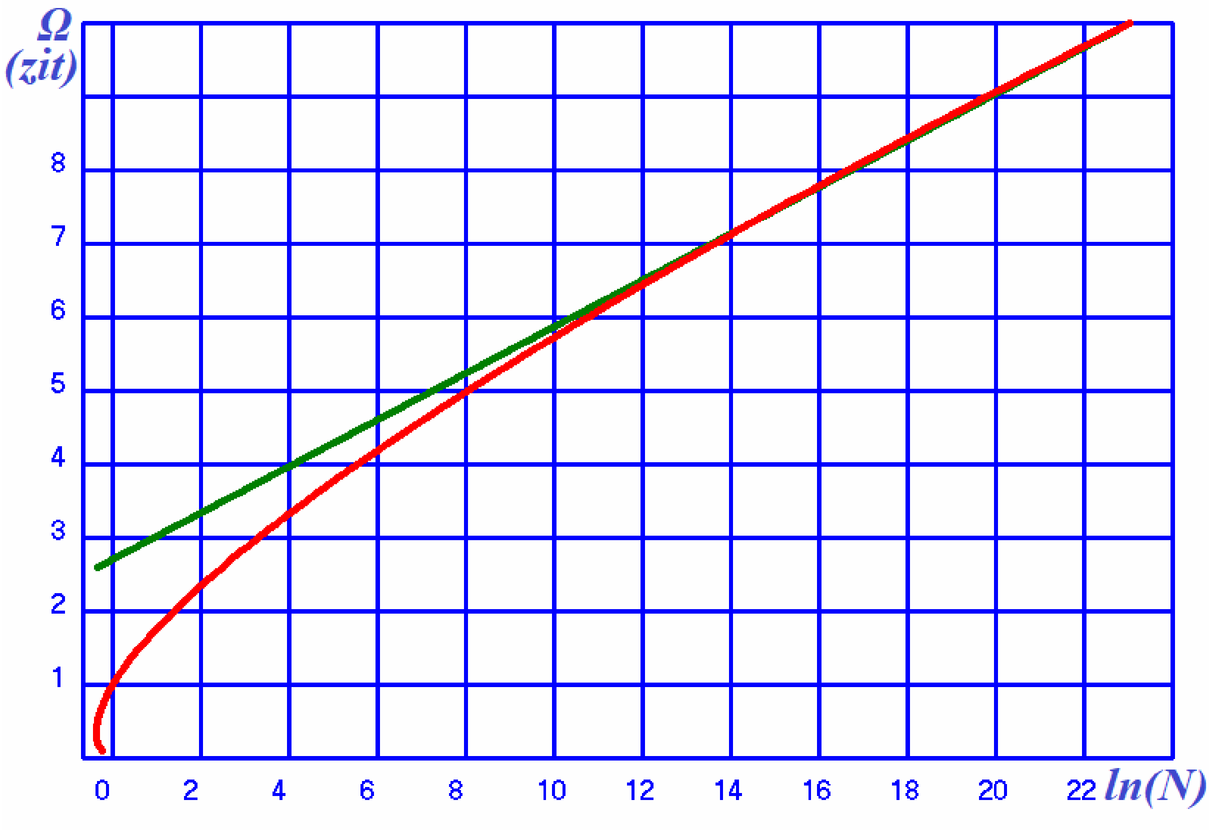

Fig. 4. Plot of Ω versus $Ln(N)$.

Figure 4 shows that the dependence of Ω on $Ln(N)$ asymptotically approaches a straight line. Therefore, we will look for a solution in the form of a linear function, the green line, from which a small component is subtracted, which goes to zero at infinity:

$$\Omega=a+b\ln N-c^{d\ln N}. \quad (4)$$

Since the exponentiation operation is absent in calculators and programming languages, it is replaced by the exponent of the logarithm. We will rewrite expression (4) in a more convenient form for calculation:

$$\Omega=a+b\ln N-c/\exp(d\ln N). \quad (5)$$

The coefficients a, b, c, d are found by the gradient descent method, we obtain: a=2.84; b=0.312; c=1.68; d=0.192. Thus solving equation (3) in elementary algebraic functions has the form:

$$\Omega=M(N)=2.84+0.312\ln N-1.68/\exp(0.192\ln N). \quad (6)$$

The function $\Omega(N)$ is proposed to be called the Mudryk Ω function.

IX. DERIVATION OF MUDRYK Ω FUNCTION VIA LAMBERT W FUNCTION

The Lambert ***W*** function is important in various fields, such as combinatorics, number theory, finance, and diffusion-convection equations [11]. We will use it to find the Mudryk ***Ω*** function.
Take the natural logarithm of the right and left sides of the Kucherov equation $N=\Omega^{\Omega}$ (3):

$$ln(N)=ln(\Omega^{\Omega})=\Omega\, ln(\Omega). \quad (7)$$

Since by definition $\Omega=e^{ln(\Omega)}$, we have:

$$ln(N)= ln(\Omega)e^{ln(\Omega)}. \quad (8)$$

By definition, the Lambert ***W*** function is:

$$N =W(N)e^{W(N)}. \quad (9)$$

As a result, we find:

$$ln(\Omega)= W(ln(N)). \quad (10)$$

Or simplified:

$$\Omega= ln(N)/W(ln(N)). \quad (11)$$

Thus, to calculate the optimal knowledge entropy ***Ω***, we can use the Lambert ***W*** function, which is implemented in many programming languages.

X. CONCLUSIONS

The volume of knowledge in a graph is determined by entropy (the degree of ordering into a hierarchical structure), which is measured by the average length of all routes from the root to objects. The route consists of zit vectors {visit →, transit ↓}, so the unit of measurement of knowledge is naturally called a zit. Similarly, the unit of measurement of information is the bit. Routes are recorded by structural numbers in the route sheet. Operations are performed by a structural operating system. Operations include visualization of the knowledge graph, management of a hierarchical database, data integration, complex comprehensive analytics, and the construction of diagrams and cartograms. In the work, the Kucherov equation $N=\Omega^{\Omega}$ for the optimal knowledge entropy is solved algebraically. The found function ***Ω(N)*** is proposed to be called the Mudryk ***Ω*** function.

REFERENCES

[1]. Друкер П. Ф. Задачи менеджмента в XXI веке. – К.: Вильямс, 2000.-С.270. https://pqm-online.com/assets/files/lib/books/druker1.pdf

[2]. Shannon, C. E. (1948). A Mathematical Theory of Communication. Bell System Technical Journal. 27: 379—423. https://ia803209.us.archive.org/27/items/bstj27-3-379/bstj27-3-379_text.pdf

[3]. Кучеров О.П., Українська латинка оптимальна. Інформаційні технології та спеціальна безпека – 2018. – №1. – С. 63–67. https://www.researchgate.net/publication/405205660_Optimal_Ukrainian_Latin_Alphabet

[4]. Кучеров О.П., Вимірювання рівня стислості алфавітів природних мов. Інформаційні технології та спеціальна безпека – 2018. – №1. – С. 63–67. https://www.researchgate.net/publication/405148250_Measuring_the_level_of_brevity_of_alphabets_of_natural_languages

[5]. Кучеров О.П. Теорія управління економікою знань. – К.: НДІ «Украгропромпродуктивність», 2008. – 252 с. https://www.researchgate.net/publication/400345727_Teoria_upravlinna_ekonomikou_znan

[6]. Кучеров О.П., Особливості створення знань у системі «Вектор К», // Актуальні проблеми економіки. – 2002, №9. – С. 37–39. https://www.researchgate.net/publication/400151683_Osoblivosti_stvorenna_znan_u_sistemi_Vektor_K

[7]. Кучеров, О.П.; Лавровський, С.Е. Образний комп'ютер для уряду. Інформаційні технології та спеціальна безпека, 2021, 1(007), с. 2–105. https://assets.zyrosite.com/A3Qw1aNbyDtxk61E/itsb-6-mk3zBaO2PEi5bow5.pdf

[8]. О.П. Кучеров, Є.В. Сарахан. Оптимізація управління процесами планування багаторічних насаджень на основі економіко-математичних моделей. Комп'ютерні засоби, мережі та системи. 2013, № 12, с. 107-116. https://nasplib.isofts.kiev.ua/server/api/core/bitstreams/1cab5611-7385-4f49-9671-fb1571325b48/content

[9]. О.П. Кучеров, О.Є. Стрижак. Формування операційного середовища інформаційно-аналітичних систем на основі онтологій / // Математичне моделювання в економіці: Зб. наук. пр. — 2013. — Вип. 3. — С. 40-47. https://nasplib.isofts.kiev.ua/handle/123456789/86363

[10]. І.М. Демчак, О.П. Кучеров, О.П. Савицька, Економіко-математичне моделювання сільських населених пунктів як основа для формування аграрної політики. / Інформаційні технології та спеціальна безпека, 2016, 1(002), с. 65–72. https://assets.zyrosite.com/A3Qw1aNbyDtxk61E/itsb-2-A1awnGbo4JToJnon.pdf

[11]. Stephen M. Disney, Roger Warburton. On the Lambert W Function: Economic Order Quantity applications and pedagogical considerations. (2012) International Journal of Production Economics. 140:756-764. DOI: 10.1016/j.ijpe.2011.02.027