

Scaling of Knowledge in Random Conceptual Networks

Lora J. Durak and Alfred W. Hübler

Center for Complex Systems Research
Department of Physics, University of Illinois, Urbana, USA

Abstract. We use a weighted count of the number of nodes and relations in a conceptual network as a measure for knowledge. We study how a limited knowledge of the prerequisite concepts affects the knowledge of a discipline. We find that the practical knowledge and expert knowledge scale with the knowledge of prerequisite concepts, and increase hyper-exponentially with the knowledge of the discipline specific concepts. We investigate the maximum achievable level of abstraction as a function of the material covered in a text. We discuss possible applications for student assessment.

1 Introduction

Defining quantitative measures for knowledge and understanding has long been a challenging problem [1]-[4]. Assessment of understanding and knowledge occurs so frequently in every day life, and technology, and is of such considerable importance that a comprehensive mathematical theory of knowledge and understanding would seem to have been required long ago. Historically, there have been multiple attempts to define understanding and knowledge. One of the earlier written theoretical descriptions of knowledge can be found in Plato's dialogues. Much of Theaetetus is devoted to the rejection of Pythagora's view that all knowledge is perception[5]. Plato inherited from Socrates the view that there can be knowledge in the sense of objective universally valid knowledge, such as the properties of numbers and geometrical objects [6]. In the *Republic*, Plato introduces the "levels of knowledge" of the development of the human mind on it's way from ignorance to true knowledge[7].

Still true knowledge exists only if a foundation of axioms or a priori knowledge is assumed to be true. Defining a minimum set of a priori knowledge is still a subject of active research in metaphysics. Aristotle values the most abstract knowledge as wisdom[8], but he encourages to test knowledge with our senses[6]. St. Augustine emphasizes abstract knowledge too, in *De Beata Vita* [9] "only

the wise man can be happy, and wisdom postulates knowledge of truth". For him 'a priori knowledge' has to be experientable either by our own senses or others[9]. Leibnitz disagrees with the empiricists such as Locke that all our concepts are ultimately derived from experience [6]. Kant sides with Leibnitz and states: "That all of our knowledge begins with experience, there is no doubt... But though all our knowledge begins with experience, it does not follow that all arises out of experience"[10] and details in *On the Form and Principles*: "Since then, in metaphysics, we do not find empirical principles, but in the very nature of pure intellect, not as experientable concepts, but as abstracted from the intrinsic laws of the mind (attending to its actions on the occasion of experience), and so as acquired" [6][11].

A recent and largely successful attempt of a quantitative theory of knowledge are formal languages in computer science, ranging from epsilon machines[12] and Bayesian networks to object oriented computer languages such as JAVA[13]. While expert systems have become a standard ingredient of manufacturing tools and consumer electronics, very little attempt has been made to assess and compare the amount of knowledge of these systems. The measure that comes closest to a measure of knowledge is Shannon's entropy[14], or average information content of a message. In information theory, the actual meaning of the message is unimportant. Instead the moment of surprise of the message, and the quality of transmission are measured.

Assessment of knowledge is a standard issue in education. National and International standardized test attempt to assess the knowledge of the students. Concept maps were developed by Ausubel [15], Novak and Gowin[16] to measure the change of the structure and organization of an individual's knowledge during the learning process (Anderson[17], Bruer [18], Stoddart [19]) in many areas [20] but little is known about the statistical and topological properties, in particular individual reproducibility and the predictive power of the test results. At this point, the philosophy of knowledge, the computer science approach to knowledge, and the assessment of knowledge in education appear to be quite disconnected.

In this paper, we try to bridge the gaps. We introduce several quantities to measure the amount of knowledge of an agent. One of these measures emphasis the value of the most abstract knowledge in line with Plato and St. Augustine, another measures emphasis the value of less abstract and a priori knowledge in line with Locke. From computer science, we adapt the definition of an agent, as an information processing system, including students, AI computer programs, and electronic courseware[21]. We study the scaling of these measures as a function of the prerequisite knowledge.

We classify knowledge in levels of abstraction, similar to Plato's classification, except that we use such levels only for classifying knowledge within a given discipline. A discipline is a given field of interest or a set of tasks with common properties, such as Algebra or Physics. We label knowledge as common knowledge if all of the agents know it. We don't enter into the discussion about

a priori knowledge and derived common knowledge, therefore adapting Kant's view. However, we introduce a special name for knowledge that is required to describe the concepts of a given discipline, but is neither common knowledge, nor part of that discipline and call it prerequisite knowledge. We distinguish between common knowledge concepts and prerequisite concepts. Discipline specific concepts are for solving typical tasks within a discipline, whereas, prerequisite concepts solve other tasks, often more basic tasks. The collection of discipline specific concepts and relations, is our definition of knowledge. More specific, the collection of least abstract discipline specific knowledge we call know-how or practical knowledge[23], whereas the collection of most abstract discipline specific knowledge is labeled as wisdom or expert knowledge.

We assume that in each level of abstraction, knowledge is structured in terms of concepts, where each concept is referenced by an identifier, such as a name or a symbol and contains (i) an objective (ii) a definition complemented by (iii) a collection of "like-this"-examples [22] and "hands-on examples" [1], (iv) a collection of applications, and (v) a list of related concepts. We assume that each component, is given as one or several "trains of reasoning" [1]. A "train of reasoning" is a sequence of sentences in a spoken or a formal language that include cartoon-style illustrations or animations, which are comprehensible by the agent.

The objective specifies a task such as "this concept establishes a relation between force, mass and acceleration". This task is solved in the definition. The definition of an abstract concept, is typically a short sequence of sentences and illustrations that define a quantity, such as "density" or describe a relationship between concepts. For example, Newton's law describes a relation between mass, force, and acceleration.

The examples in a concept are problem-solution pairs, where the solution is a derived from the definition of the concept. The applications in a concept are problem-solution pairs which require repeated use of the concept or illustrate relations to other disciplines. Some of the applications are typically hands-on [1]. Applications relate the concept to common knowledge [24]. In contrast to the methods in the applications, the definition leaves the sequence of sentences and illustrations ambiguous wherever permissible and employs discipline specific concepts that are as abstract as possible. This briefness and high level of abstraction can make the definition appear ambiguous. We feel that it is important to employ the use of examples because they are less abstract and resolve some of that ambiguity.

An agent knows a concept, if it can tell the identifiers, reproduce the definition, give examples, applications, solve problems which are very similar to the given examples and applications, and can list related concepts. This implies that the agent knows a concept, if it knows all sub-concepts of the concept and can name, but does not necessarily know related concepts. Sub-concepts are those concepts, which are used in the definition, examples, and applications. In this

paper, we study how the unknown prerequisite concepts limit knowledge. Knowing a concept is a prerequisite for understanding.[25] If the agent is able to give it's own interpretations and translations of the concept, the agent has an understanding of the concept.

If two concepts have a relation, i.e. similarities in their facts, methods, explanation, or typical application, they are called related. The concepts and their relations form a hierarchical conceptual network. If two concepts have a strong relation, this relation can be a concept by itself, an abstract concept[26].

2 A Simple Model of Conceptual Networks

We study conceptual networks for a given field of interest or discipline. A conceptual network is a cross-referenced library of concepts $C_i = 1 \dots N$ [27]. We extracted kinematics concepts from the highlighted equations in Serway and Faughn, College Physics[28]. Each concept C_i solves a typical task. The agent is described with a set of indicators, $c_i = 1, 2, 3, \dots, N$, and $p_j, j = 1, 2, 3, \dots, N_p$. c_i is set equal to 1 if the agent knows all the prerequisite concepts which are used in the description of the concept, otherwise it is set equal to zero. p_j is set equal to 1 if the agent knows the prerequisite concept P_j perfectly, otherwise it is set equal to zero $p_j = 0$, where $j = 1, 2, 3, \dots$. $p_{i,D}$ is set equal to 1 if the agent knows the description of the concept, otherwise it is set equal to zero, $p_{i,D} = 0$. Then c_i , the knowledge of the agent, is defined as:

$$K = \sum_{i=1}^N w_i c_i, \quad (1)$$

where w_i is a measure for the importance or weight of each concept. If $\langle c_i \rangle$ is the probability that an agent knows the concept C_i then the expectation value for the knowledge is:

$$\langle K \rangle = \sum_{i=1}^N w_i \langle c_i \rangle. \quad (2)$$

We assume that the agent knows common knowledge, whereas it knows prerequisite concepts $P_j, j = 1, 2, \dots, N_p$ only with probability p .

A concept is called a base concept if it contains only prerequisite concepts and common knowledge. The number of base concepts is N_1 . Abstract concepts of level $L = 2$ contain at least one base concept in the same discipline. If the base concept is replaced by it's own description in a level 2 concept, then this is called a substitution. For example, in a physics mechanics class, position $x(t)$ is a prerequisite concept, displacement $dx = x(t + dt) - x(t)$ is a base concept and $v(t) = dx/dt$ is a level 1 concept. If dx is replaced by $x(t + dt) - x(t)$ in

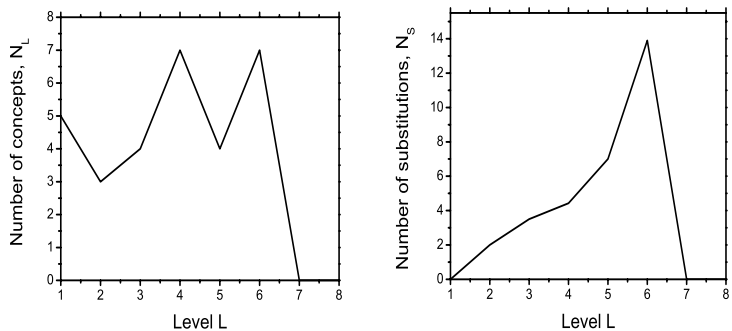


Fig. 1. The number of concepts N_L as a function of the level of abstraction L in kinematics (left). The number of substitutions required N_S for a full substitution of a level L concept in kinematics.

$v(t) = dx/dt = (x(t + dt) - x(t))/dt$, the displacement is substituted by its description of the velocity. If a concept contains an abstract concept of level L_i and possibly some lower levels, it is called an abstract concept of level $L_i + 1$. The number of concepts of abstraction level L within a given discipline is called N_L . We assume the N_L is constant for $L \leq \hat{L}$ and $N_L = 0$ else . Fig. 1 shows the N_L versus the level L for kinematics[28]. The mean value is $N_L=5$ and $\hat{L} = 6$.

If the lower level concepts are recursively substituted by their descriptions, then the concept is called "fully-substituted". N_s is the average number of sub-concepts required to describe a level L concept in terms of prerequisite concepts and common knowledge concepts. We assume that N_S increases exponentially with L :

$$N_S = \alpha * \lambda^L. \tag{3}$$

This assumption is supported by the data shown in Fig. 1. For kinematics we find $\alpha = 1$ and $\lambda = 1.6$.

When a concept is fully substituted, it's description contains a certain number of prerequisite concepts N_P and a certain number of common knowledge concepts N_C . Fig. 2 suggests that N_P increases linearly for small L , $N_P = 0.5 + 2 * L$ and is constant $N_P = 6.5$ for $L \leq 3$. In this paper, we assume that N_P is constant. The number of common knowledge concepts increases linearly with the level of abstraction:

$$N_C = N_{C,0} + n_C * L. \tag{4}$$

For kinematics we find $N_{C,0} = 1$, and $n_C = 0.6$, hence $N_C = 1 + 0.6 * L$.

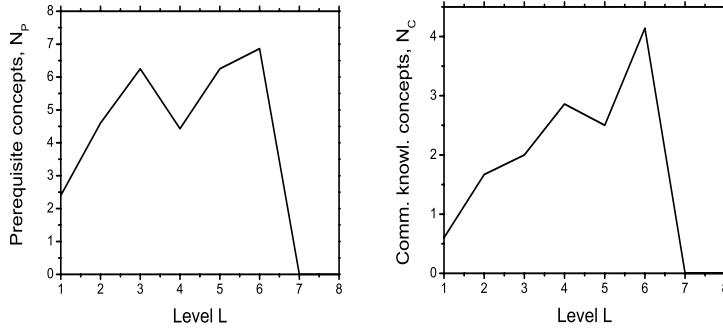


Fig. 2. The number of prerequisite concepts in a fully-substituted concept N_P as a function of the level of abstraction L for kinematics (left). The number of common knowledge concepts in a fully-substituted concept N_C as a function of the level of abstraction L for kinematics.

The probability of knowing a concept with n sub-concepts is $p_{i_1}, p_{i_2}, p_{i_3}, \dots, p_{i_n}$. For each agent the sequence of indicators $\{p_{i_1}, p_{i_2}, p_{i_3}, \dots, p_{i_n}\}$ is a sequence of zeros and ones $\{0, 1, 1, 0, \dots, 0\}$ of length n . If we suppose that the indicators are statistically independent, then the probability that the agent knows a concept of level L is:

$$P_L = p^{N_P} * p_D^{N_s} = p^{N_P} * p_D^{\alpha * \lambda^L}, \quad (5)$$

where p is the probability of knowing a prerequisite concept, and p_D is the probability of knowing the description of a discipline-specific concept. In Eqn. 5, the first factor reflects the requirement that the agent knows all prerequisite concepts in the fully-substituted concept. The second term guarantees that the agent knows the descriptions of sub-concepts required for a full substitution. We define the maximum level of abstraction, which the agent can reach by the condition $N_L * P_L = 1$, where the agent has to know at least one level L concept. Solving this equation for L yields the maximum level of abstraction L_{max} .

$$L_{max} = \frac{\ln \frac{-\ln N_L * p^{N_P}}{a * \ln p_D}}{\ln \lambda}. \quad (6)$$

To determine the knowledge of an agent, one needs to model the importance of the concepts w_i . There are two simple models. In the first model we assume, that all concepts are equally important since each concept matches one task. Then $w_i = 1$, for all i . In this case, knowledge (practical knowledge) is :

$$\langle K \rangle_P = \sum_{L=1}^{\hat{L}} N_L * P_L = p^{N_P} * \sum_{L=1}^{\hat{L}} p_D^{\alpha * \lambda^L}. \quad (7)$$

Hence, the practical knowledge scales with the probability of knowing a prerequisite concept p and as a sum (or integral) of hyper-exponential functions of p_D in the level of abstraction L .

In a second model, we assume that the weight w_i is proportional to information content of the concept. The probability of picking a concept with $N_P + N_C$ sub-concepts out of a set of M common knowledge and prerequisite concepts is $\rho = M^{N_P + N_C}$ if $N_P + N_C \ll M$. Therefore, the information content is $I = \ln(1/\rho) = (N_P + N_C) * \ln(M)$ and $w_i = I$. In this case, knowledge (expert knowledge) is:

$$\langle K \rangle_E = \ln(M) * \sum_{L=1}^{\hat{L}} (N_P + N_C) * N_L * P_L \quad (8)$$

$$= \ln(M) * p^{N_P} \sum_{L=1}^{\hat{L}} (N_P + N_{C,0} + n_c * L) * N_L * p_D^{\alpha * \lambda^L}. \quad (9)$$

Hence, the expert knowledge scales with the probability of knowing a prerequisite concept p like the practical knowledge, but increases even faster than the practical knowledge as a function of p_D in the level of abstraction L .

3 Related Models

In the previous section, we assumed that the number of substitutions increases exponentially with the level of abstraction. In a linear graph, the number of substitutions would increase linearly, such as $N_S = N_{S,0} * L$. In this case, the practical knowledge is:

$$\langle K \rangle_P = \sum_{L=1}^{\hat{L}} N_L * P_L = p^{N_P} * \sum_{L=1}^{\hat{L}} p_D^{N_{S,0} * L} \approx p^{N_P} * p_D^{N_{S,0} * \hat{L} + 1}. \quad (10)$$

Consequently, the practical knowledge scales both with the prerequisite knowledge p and the knowledge of the field-specific concepts p_D .

In the previous section, we assumed that the number of prerequisite concepts N_P does not depend on the level of abstraction L . However, the experimental data indicate that at least for small L , N_P increases linearly in L , such as $N_P = N_{P,0} + n_P * L$. In this case, the practical knowledge increases rapidly as a function of p :

$$\langle K \rangle_P = \sum_{L=1}^{\hat{L}} p^{N_{P,0} + n_P * L} * N_L * p_D^{\alpha * \lambda^L}. \quad (11)$$

Even if the experimental data for kinematics do not suggest this, one might assume that the number of abstract concepts N_L decreases with the level of

abstraction L . We use $N_L = N_{L,0} * \delta^L$, an exponential model, where $0 < \lambda < 1$. In this case, the practical knowledge is:

$$\langle K \rangle_P = \sum_{L=1}^{\hat{L}} N_{L,0} * \delta^L * p^{N_P} * p_D^{N_P}. \quad (12)$$

4 Discussion

If we assume the agent is a perfect learner, the probability of knowing a discipline-specific concept p_D is determined by the amount of material covered by a textbook or a course. The maximum level of abstraction has a logarithmic singularity at $p_D = 100\%$, see Eqn. 6. The means, that if the textbook or course is omitting just a very small fraction of the discipline concepts, this has a huge negative impact on the maximum level of abstraction that the student or computer agent can achieve.

Knowledge scales only as a power-law of p , the probability of knowing knowing prerequisite concepts, see Eqn. 7 and Eqn. 8. This means that the knowledge does not decrease very rapidly, when p decreases. Perhaps, this accommodates students who have learned a comparatively low level of prerequisite knowledge.

Currently, many printed textbooks are following a linear sequence of concepts through a conceptual network, where each new idea is deduced from the previous concept with very little reference to the branched topological structure of conceptual networks. In contrast, most electronic courseware makes the student aware of all the relations when a new concept is introduced; such as on-line *C*-computer language manuals, Windows NT Help, Yahoo Search Engine, Kaplan software and SAT English test preparation software. Courseware with intense cross-referencing exposes the student to concepts, which she or he may learn much later or never learn. In the past, there has been extensive discussion by Hamlyn[29] about the advantages and disadvantages of referring to unknown concepts for teaching. Cross-referencing is much easier in electronic courseware. In kinematics, 20 out of 25 concepts describe relations (80%). Considering that the largest part of knowledge is attributable to abstract knowledge, electronic courseware may help to improve learning environments since it supports relations.

The kinematics data suggest that substitutions required to substitute a concept of level L by prerequisite and common knowledge concepts increases exponential in L . If the number of substitution would increase linearly, each concepts would be at a higher level of abstraction. Representing concepts at a comparatively low level of abstraction, might be beneficial for some learners. In kinematics the highest level of abstraction is $\hat{L} = 6$ (see Fig. 1). Humans can memorize 7 ± 2 chunks of information in short term memory[30] depending on their level of intelligence. Possibly, there is a connection between the level of abstraction in kinematics $\hat{L} = 6$ and human short term memory.

It is surprising, that both the number of concepts N_L at a given level of abstraction (see. Fig. 1) and the number of prerequisite concepts N_P required for a full-substitution of a concept (see Fig. 2) do not seem to depend on the level of abstraction L .

We give thanks to M. Osborne, E.A. Jackson, P. Melby, and D. Smyth for thoughtful discussion. This research was supported by the Office of Naval Research Grant No. N00014-96-1-0335.

References

- [1] Hendley B.P. (ed.): Dewey, Russell, Whitehead, Philosophers as Educators, South. Illinois Univ. Press, Carbondale (1986) 16, 83.
- [2] Peters, R.S. (ed.): The Philosophy of Education, Oxford Univ. Press, (1973).
- [3] Grant, R.W., Tarcov N. (eds.): John Locke, Some Thoughts Concerning Education and Of the Conduct of the Understanding, Hackett Publishing Company, Inc., Indianapolis/Cambridge (1996).
- [4] Osborne M.D.: Constructing Knowledge in the Elementary School Classroom: Teachers and Students, Falmer, New York (1999); Osborne M.D.: Teacher as Knower and Learner, Reflections on Situated Knowledge in Science Teaching, Journal of Research on Science Teaching **35/4** (1998) 427.
- [5] Jowett B.: The Works of Plato (3rd edition), Oxford University Press, Oxford (1892) Theaetetus 152 c 5-7.
- [6] Copleston F.: A History of Philosophy, **I**, Doubleday, NY (1993) 143, 200, 288.
- [7] Jowett B.: The Works of Plato (3rd edition), Oxford University Press, Oxford (1892) Republic, 509 d 6 - 511 e 5.
- [8] Ross Sir W.D.: Aristotle's Metaphysics, 2 vol, Oxford University Press, Oxford (1924) Metaphysics A, 980 a I.
- [9] Quasten J. , Plumpe J.C. (eds.): Ancient Christian Writers: the Works of the Fathers in Translation, Westminster, Maryland (1946) De Beata Vita 2,10, and 14; 4; 27ff; De Trinitate, 15, 12, 21.
- [10] Kant I.: Critique of Pure Reason, 1, translated by Meiklejohn J.M.D., 2.ed., Everyman's Library, London (1933).
- [11] Kant I.: Gesammelte Schriften, Critical Edition, 22 vols., Prussian Academy of Sciences, Berlin (1902-42) On the Form and Principles 2, 8; W., II, p. 395.
- [12] Crutchfield J.P., Young, K.: Inferring statistical Complexity, Physical Review Letters **63** (1989) 105-108.
- [13] Harold E.R.: JAVA Network Programming, O'Reilly, Cambridge (1997).
- [14] Sloane N.J.A., Wyner A.D. (eds.): Claude Elwood Shannon: Collected Papers, IEEE Press, New York (1993).
- [15] Ausubel D.: Ed. Psych.: A Cognitive View, Werbel and Pack, NY (1968).
- [16] Novak J.D., Gowin D.B.: Learning how to Learn, Cambridge U. Press, (1984), Novak J.D.: Learning, Creating, and Using Knowledge, Lawrence Erlbaum Assoc., London (1998).
- [17] Anderson, O.R.: Some Interrelationships between Constructivist Models of Learning and Current Neurobiological Theory, with Implications for Science Education, Journal of Research and Science Teaching **29/10** (1992) 1037-1058.

- [18] Bruer, J. T.: *Schools for Thought*, MIT Press, Massachusetts (1993).
- [19] Stoddart, T., Abrams R., Gasper, E., Canaday, D., *Concept Maps as Assessment in Science Inquiry Learning - a Report of Methodology*, *International Journal of Science Education* **22**, (2000) 1221-1246.
- [20] Roth W.-M., Roychoudhury, A.: *The Concept Map as a Tool for the Collaborative Construction of Knowledge: A Microanalysis of High School Physics Students*, *J Res. Sci. Teaching* **30** (1993) 503-534; Stensvold, M.S., Wilson, J.T.: *The Interaction of Verbal Ability with Concept Mapping in Learning from a Chemistry Laboratory Activity*, *Sci. Ed.* **74** (1990) 473-480; Markow, P. G., Lonning, R.A.: *Usefulness of Concept Maps in College Chemistry Laboratories*, *J. Res. Sci. Teaching* **35** (1998) 1015-1029.
- [21] Schwarz, E., Brusilovsky P., Weber G.: *World-Wide Intelligent Textbooks in Ed. Telec.* 1996, Patricia Carlson and Fillia Makedon eds., AACE, Charlottesville (1996) 302; Weber, G.: *Individual Selection of Examples in an Intelligent Programming Environment*. *J Art. Int. Ed.* **7(1)** (1996) 3-33; Weber, G.: *Episodic Learner Modeling*, *Cog. Sci.* **20**, (1996) 195-236.
- [22] Wittgenstein, L: *Philosophical Investigations*, translated by Anscombe, G. E. M., Blackwell Publishers Ltd., Malden (1997) 490.
- [23] Ryle, G.: *The Concept of the Mind*, Hutch. Univ. Lib., London (1949) 26-51.
- [24] Whitehead, A. N.: *An Introduction to Mathematics*, Home University Library, London, (1911); rev. ed., Oxford University Press, New York (1958); Dewey, J.: *The Relation of Theory to Practice in Education in The middle Works of John Dewey*, Boydston, J. A. (ed.), Southern Illinois University Press, Carbondale (1977) **3** 249-272.
- [25] Jackson E.A.: *The Unbounded Vistas of Science: Evolutionary Limitations*, *Complexity* **5** (2000) 35-44.
- [26] Whitehead, A.N.: *The Principles of Mathematics in Relation to Elementary Teaching*, in Whitehead, A.N.: *The Organization of Thought: Educational and Scientific*, Williams and Norhgate, London (1917), reprint, Greenwood Press, Westport (1974), 101-2.
- [27] Beland, A., Mislevy, R. J.: *Probability-Based Inference in a Domain of Proportional Reasoning Tasks*, *Journal Educational Measurement* **33** (1996) 3-27.
- [28] Serway, R. A., Faughn, J. S.: *College Physics*, fourth ed., Saunders College Publishers, Fort Worth (1995).
- [29] Hamlyn, D.Y.: *Human Learning in The Philosophy of Education*, by R.S. Peters ed. , Oxford University Press, Oxford (1973).
- [30] Miller, G., 1956. *The magic number seven, plus or minus two: some limits of our capacity for processing information*. *Psychological Review* **63** (1956) 81-97.