



Institute of Information and  
Computational Technologies

ISSN : 2788-7677 (Online)  
ISSN : 2788-7987 (Print)

# ADVANCED TECHNOLOGIES AND **COMPUTER SCIENCE**

**2023**  
**No3**

[www.atcs.iict.kz](http://www.atcs.iict.kz)

---

Institute of Information and Computational Technologies

# **Advanced Technologies and computer science**

**№3**

Almaty 2023

ISSN: 2788-7677 (Online)  
ISSN : 2788-7987 (Print)

Institute of Information and Computational Technologies,

Advanced Technologies and computer science

This journal is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The publisher, the authors, and the editors are safe to assume that the advice and information in this journal are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published works and institutional affiliations.

28 Shevchenko str., Almaty, Republic of Kazakhstan  
7 (727) 272-37-11  
atcs@iict.kz

## About the Journal

Advance technologies and computer science is a bilingual scientific peer-reviewed, interdisciplinary, electronic journal of open access, including thematic areas:

- Section "**Applied mathematics, computer science and control theory**" includes papers describing modern problems in these areas.
- Section "**Information and telecommunication technologies**" also includes the following topics:
  - Data transmission systems and networks.
  - Internet technologies.
  - Cloud technologies.
  - Parallel computing.
  - Distributed computing.
  - Supercomputer and cluster systems.
  - Big data processing (Big-data).
  - Geographic Information Systems and Technologies.
- In the section "**Artificial intelligence technologies**" in addition to technology, there are works on topics:
  - Intelligent Management Systems.
  - Speech technology and computer linguistics.
  - Pattern Recognition and Image Processing.
  - Bioinformatics and biometric systems.
  - Human-machine interaction.
  - Machine learning.
  - Intelligent Robotic Systems.
- The section "**Information Security and Data Protection**" also covers topics:
  - Software and hardware information protection.
  - Mathematical methods for ensuring information security of complex systems.
- The section "**Modeling and optimization of complex systems and business processes**" may include:
  - Computational mathematics, numerical analysis and programming, mathematical logic.
  - Theory of Statistics.
  - Statistical Methods.

## **Editorial Team**

### **Chief Editor**

Kalimoldayev M.N., Academician of NAS RK, Doctor of Physics and Mathematics, Professor, DG of RSE "Institute of Information and Computational Technologies" SC MES RK (Kazakhstan)

**Deputy chief editor:** PhD, Mamyrbayev O.Zh (Kazakhstan)

### **Editorial team**

- Amirgaliev Ye.N., Doctor of Technical Sciences, Professor, Kazakhstan
- Arslanov M.Z., Doctor of Physics and Mathematics, Professor, Kazakhstan
- Berdyshev A.S., Uzbekistan
- Biyashev R.G., Doctor of Technical Sciences, Professor, Kazakhstan
- Ischukova Ye.A., Candidate of Technical Sciences, Docent, Russia
- Krak, Ukraine
- Posypkin M.A., Doctor of Physics and Mathematics, Russia
- Khairova N.F., Doctor of Technical Sciences, Ukraine
- Keylan Alimhan, Japan (Tokyo Denki University)
- Marat Ahmet, Turkey
- Mohamed Othman, Малайзия (Universiti Putra Malaysia)
- Naohisa Otsuka, Japan (Tokyo Denki University)
- Ravil Muhamediev, Latvia
- Waldemar Wójcik, Poland

## Contents

<b>Operational calculus of modified subset construction</b> Syzdykov M.	<b>4</b>
<b>Марковские идеи в байесовских сетях</b> Ахметова А.М., Абдилдаева А.А., Литвиненко Н. Г., Литвиненко А. Г.	<b>9</b>
<b>Применение технологий машинного обучения при диагностике глаукомы</b> Павлов С.В., Мамырбаев О.Ж., Оралбекова Д.О., Момынжанова К.	<b>17</b>
<b>Artificial intelligence for complexity theory</b> Syzdykov M.	<b>27</b>
<b>Method of constructing models of information attacks</b> Barakova A.Sh., Ussatova O.A.	<b>32</b>

UDC 004.02  
IRSTI 20.53.15

## OPERATIONAL CALCULUS OF MODIFIED SUBSET CONSTRUCTION

Mirzakhmet Syzdykov

Satbayev University, Almaty, Kazakhstan  
[mspmail598@gmail.com](mailto:mspmail598@gmail.com)

ORCID ID: <https://orcid.org/0000-0002-8086-775X>

**Abstract** We present the continuation of studying Extended Regular Expression (ERE) on the view of modified subset construction within the overridden operators like intersection, subtraction, and re-written complement. As before we have stated that in this case the complexity has a decreasing nature and tendency. We will give the strict definition of the operational part of this modified subset construction which is due to Rabin and Scott. The complexity of algorithm remains a magnitude less than NP-hard problems for which we have given the strict proof of equivalence in the prior work, so this work continues the studying of the comparable proof for a variety of problems to be computationally complex, however, explainable in terms of unified approach like operational calculus. In this calculus the general points of research are given to the representation of modified subset construction with at least two operands which are to be computed by subset construction and in terms of complexity of the effective algorithm they are computed using modified subset construction.

**Keywords:** subset construction, extended regular expressions, modification, operations, calculus.

### Introduction

The subset or Rabin-Scott construction which was full described in [1] represents conservative system of choosing between determinism and non-determinism in both aspects, however, lacks the efficiency of complexity in case of deterministic machine operating on the finite set of states, thus, it's obvious that it will lead the number of states as well as number of operations to grow exponentially in time of  $O(2^n)$ .

The latter case isn't limited to the usage of the classical Thompson algorithm [2], which is less complex and requires asymptotic explosion of complexity in  $O(m \cdot n)$ , where  $m$  is the number of symbols in sought or input string and  $n$  is the number of elements in matched regular expression. To the present time Thompson's constructions weren't used for extended regular expression matching.

Samuel C. Hsieh showed a more effective algorithm for intersection operator [3], however, it's still NP-hard as its complexity can be denoted by  $O(n^t)$ , where  $n$  is the average size of length of operands in ERE and  $t$  is the number of &-operators.

We have shown that ERE for intersection problem can be computed on both deterministic and non-deterministic finite automata, NFA and DFA respectively [4]. We have also introduced the sliced model of computation for our algorithm which tends to be in magnitude faster by applying operational calculus [5].

Our algorithm for NFA or DFA supersedes previous results [6, 7] which operate on the cross-product construction of the DFA by applying the operational calculus in the form of the operational logic for the set of operands to be performed in-time and in-memory. The non-emptiness intersection problem was shown to be NP-hard for sparse set of automata [7], however, we give another argument towards our conjecture of equivalence of complexity classes.

Aho-Corasick trees [8] and Lempel-Ziv-Welch [9] streams for regular expression matching are also discussed in this article as well as the "P versus NP" conjecture [10] for common case of the problem of deciding whether the intersection of the given languages is empty or not.

### Modified Subset Construction

This construction differs from the usual approach by Rabin-Scott in implementation of additional operators for the closure function which is defined as follows:

$$\varepsilon - \text{closure}(S) = \{t: f^i(s, \varepsilon) = t, \forall s \in S, \forall i = 1..n, n \in N\}. \quad (1)$$

Where in (1),  $f(x, y)$  is the reaching state function in the NFA and  $S$  is the set of states. We extend this construction with the additional operator in our defined calculus as it's given in prior work [4]:

$$\varepsilon - \text{modified} - \text{closure}(S) = \{t: t \in \varepsilon - \text{closure}(S) \cup t: g(t) = 0\}. \quad (2)$$

Where in (2) the  $g(t)$  is the base function which is computed during the approximation of the algorithm to the given point. By this point we define the artificial states which are implied for the subset construction with modifications the complexity of which is  $2^{o(n)}$ . The summary for this function can be found in [4].

Thus, for intersection operator  $g(t)$  is defined as follows:

$$g(t) = \text{deg}^+(t) - |\text{visited}(t)|. \quad (3)$$

Whereas in (3)  $\text{deg}^+(t)$  denotes the number of incoming edges for the given state  $t$  which is artificial by the definition as it wasn't implemented or introduced in prior works [1, 2, 3],  $\text{visited}(t)$  is the function denoting the number of visited edges during the closure computation process – we conclude that at each step this function is evaluated to its default value of zero.

The function  $g(t)$  for subtraction operator is defined as the logical gate consisting of binary input:

$$g(t) = \begin{cases} 0, & \text{visited}(t) = \{L\} \\ -1 & \end{cases}. \quad (4)$$

Where in (4)  $L$  is the left operand to be visited and right operand is omitted if it wasn't visited before, otherwise, the logical expression defined by function  $g(t)$  in (4) is evaluated to false and no further calculation is permitted.

For the complement the function  $g(t)$  is defined within the modified construction and re-writing of this operator within the same expression (4).

## On P versus NP

Since it was shown that non-emptiness intersection problem can be decided in non-polynomial time for specific cases, our algorithm decides it in time  $O(\text{PQI})$  for any case [5].

Thus, we conclude that there's another argument towards the proof of  $P = NP$ . As the function  $g(t)$  in (2) is invariant and is to be computed for the implied states to model the existence of primarily intersection operator, it's still well-known that it was extended for the case of subtraction operator and re-writing of this operator for complement.

According to functional hypothesis there could be a set of positive transformations leading the algorithmic logic to be reduced to polynomial memory and time complexity, however, this question remains open as per our prior works we based our proof on the observation of the proved NP-hard problem to be solved using the full view of the input parameters in problem in polynomial time. PQI-operator [5] was introduced before to represent the exact computational complexity of the process in subset construction with function modifiers. These modifiers represent artificially implemented structures in the graph of automaton to be translated into the semaphore or any other logic gate so that the latter statement holds true and non-feasible subsets of automaton states aren't acceptable when traversing it through closure functions when passing it through the filter function  $g(t)$  defined in (3) for emptiness problem and in (4) for the common logical case.

The "P versus NP" theorem which wasn't still explicitly reviewed remains as a closed question as the author of the scientific work proposing the computational models for better evaluation of algorithm complexity had a better understood theoretical experience which leads to the question of the relation between polynomial (P) and non-polynomial (NP) classes to be open. However, still we have the facts which show that using subset construction in its modified full form can lead to the appearance of the more effective algorithms for non-emptiness intersection problem as well as to other problems where the redundant logical accepting states can be implemented as it's shown in the subset construction for extended operators in ERE like intersection, subtraction and complement.

## Aho-Corasick Trees and Intersection Operator



As we have defined the intersection operator within the common Thompson's construction at first and then Rabin-Scott subset construction, it's possible to get the point for Aho-Corasick trees which denote the finite set of words and can be seen as a DFA.

For further purpose we can use the intersection state in NFA and get the construction for each accepting state in Aho-Corasick automata by implementing the state transition from accepting states to the finishing state of the pre-defined regular expression pattern, thus, giving the possibility to decide what words belong to the regular set.

The observation above leads to the minimization of accepting states by applying subset construction backwards from accepting states in Aho-Corasick tree [8]. This leads to the application of matching algorithmic constructions of using the mixed stream for both Aho-Corasick tree or compressed entity like Lempel-Ziv-Welch (LZW) stream.

The main conjecture is that Aho-Corasick trees optimized backwards by using Rabin-Scott backwards construction lead to the imminent minimization of this tree. This can be proved by the fact that each of the accepting and ending states in this tree conforms the decreasing function  $opt(x)$  which is defined as follows:

$$opt(x) = \{x \in subset - construction(A), \exists f(x, a), f(y, a): a \in A\}. \quad (5)$$

The definition (5) gives us the observation that trees are given from the starting single point and cannot be optimized further as they represent the optimized tree during online construction of this tree within the additional string to be added or which is already included in the tree. Thus, we can conclude that the state minimization process is to be started from the accepting states in backward direction.

Intersection operator can be applied to compressed or non-compressed trees within the observable time frame. This operator is for deciding the more complex and efficient algorithm for matching the regular expression pattern against the set of words rather than a single word or input stream of single source – in contrary, Aho-Corasick trees are of multiple sources and as we have shown can be minimized also.

For LZW input streams [9], it's defined that the ending mark from the encoded input can be used further when constructing tree itself during the invocation process – this is a linear process not requiring additional resources like memory and time and, thus, we conclude that these streams are unary.

## Containerization of NFA and DFA for ERE and Stop Marks

This step of process includes the experimentation with the non-deterministic finite automata (NFA) conversion to deterministic finite automata (DFA) along the Extended Regular Expressions (ERE) within the aimed operators like intersection, subtraction and complement. The practice shows that in common case this is the best practice for implementation of composite NFA and DFA via subset construction.

In common sense, this is a good approach for developing analyzing tools in biomedicine for processing big amount of DNA sequences.

For the question of  $P = NP$  via O-operator proof, we can conclude that a single case is quite clear to conclude that this is a way of solving NP-hard problems laying outside NP-complexity class.

## Matching algorithm with stop marks

We define the matching according to the non-trivial symbol in the sequence of concatenation, whereas the fully connected clique of states for empty transitions lets the exact word to be matched and consequent cliques are matched according to this stop mark.

The cliques are defined as the strongly connected components in which any word can be defined in the final set, thus, allowing the mark to be matched before the actual matching starts – this technique was used before to prove the equivalence of P and NP classes.

## Conclusion

We have defined the necessary relations between operational calculus and ERE constructions and evaluation on either NFA or DFA through Modified Subset Construction (MSC). This calculus gives a broader observation of how our model is to be represented in operational logic and applied mathematics.

This relation is to introduce the solution to regular language non-emptiness intersection problem within the time  $2^{O(f(n))}$ .

We have also shown that P versus NP conjecture for automata non-emptiness intersection problem can be considered decidable in polynomial time, thus giving the assumption that P equals NP.

The practical solutions to the minimization of Aho-Corasick tree and the usage of LZW input streams is also given as we have shown that these trees can be efficiently optimized using backward propagation method of the closure computation.

We can conclude more that operational calculus can be used in approximate regular expression matching, however, this is a well-studied question and doesn't require more attention as the definition of the new algebraic structure. This structure remains open for extended operators like intersection, subtraction and complement in ERE.

### Acknowledgements

The author expresses gratitude to all the members of ResearchGate™ for their valuable comments on the publication of this work.

### References

- [1] Rabin M. O., Scott D. Finite automata and their decision problems. IBM journal of research and development. 1959. 3(2). 114-125.
- [2] Thompson K. Programming techniques: Regular expression search algorithm. Communications of the ACM. 1968. 11(6). 419-422.
- [3] Hsieh S. C. Product construction of finite-state machines. Proc. of the World Congress on Engineering and Computer Science. 2010. 141-143.
- [4] Syzdykov M. Deterministic automata for extended regular expressions. Open Computer Science. 2017. 7(1). 24-28.
- [5] Syzdykov M. et al. Introducing pqi-operator in theory of computational complexity. Advanced technologies and computer science. 2022. 3. 4-9.
- [6] de Oliveira Oliveira M., Wehar M. On the fine grained complexity of finite automata non-emptiness of intersection. Developments in Language Theory. Proceed. of the 24th International Conference. Tampa, FL, USA, May 11–15, 2020. Cham: Springer International Publishing, 2020. 69-82.
- [7] Fernau H., Hoffmann S., Wehar M. Finite automata intersection non-emptiness: Parameterized complexity revisited. arXiv preprint arXiv:2108.05244. 2021.
- [8] Zha X., Sahni S. Highly compressed Aho-Corasick automata for efficient intrusion detection. 2008 IEEE Symposium on Computers and Communications. IEEE, 2008. 298-303.
- [9] Bille P., Fagerberg R., Gørtz I. L. Improved approximate string matching and regular expression matching on Ziv-Lempel compressed texts. ACM Transactions on Algorithms. 2009. 6(1). 1-14.
- [10] Cook S. The importance of the P versus NP question. Journal of the ACM (JACM). 2003. 50(1). 27-29.

### МОДИФИКАЦИЯ ЛАНҒАН ІШКІ ЖИЫНДЫ ҚҰРУДЫҢ ОПЕРАЦИЯЛЫҚ ЕСЕБІ

#### Сыздықов Мырзахмет

Қ.И. Сәтпаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы, Қазақстан  
mspmail598@gmail.com

ORCID ID: <https://orcid.org/0000-0002-8086-775X>

**Андатпа.** Біз қиылысу, алу және қайта жазылған толықтауыш сияқты қайта анықталған операторлар ішіндегі өзгертілген ішкі жиын құрылысының көрінісі бойынша кеңейтілген тұрақты өрнекті зерттеудің жалғасын ұсынамыз. Бұрынғыдай біз бұл жағдайда күрделіліктің төмендеу сипаты мен тенденциясы бар екенін айттық. Біз Рабин мен Скоттқа байланысты осы өзгертілген ішкі жиынтық конструкцияның операциялық бөлігіне қатаң анықтама береміз. Алгоритмнің күрделілігі біз жұмыста эквиваленттіліктің қатаң дәлелін берген NP-қиын есептерден аз шама болып қала береді, сондықтан бұл жұмыс әртүрлі есептердің күрделі болуы үшін салыстырмалы дәлелдемелерді зерттеуді жалғастырады, алайда, операциялық есептеу сияқты біртұтас көзқарас тұрғысынан түсіндіріледі. Бұл есептеуде зерттеудің жалпы тармақтары ішкі жиынды құру арқылы есептелуі тиіс кемінде екі операндтары бар түрлендірілген ішкі жиын құрылысын ұсынуға берілген

және тиімді алгоритмнің күрделілігі тұрғысынан олар өзгертілген ішкі жиын құрылысы арқылы есептеледі.

**Кілттік сөздер:** ішкі жиын құрылысы, кеңейтілген тұрақты тіркестер, модификация, амалдар, есептеу.

## ОПЕРАЦИОННОЕ ИСЧИСЛЕНИЕ ПОСТРОЕНИЯ МОДИФИЦИРОВАННОГО ПОДМНОЖЕСТВА

**Сыздықов Мирзахмет**

КазННТУ им. К.И. Сатпаева, Алматы, Казахстан  
mspmail598@gmail.com

ORCID ID: <https://orcid.org/0000-0002-8086-775X>

**Аннотация.** Мы представляем продолжение изучения расширенных регулярных выражений с точки зрения построения модифицированного подмножества внутри переопределенных операторов, таких как пересечение, вычитание и перезаписанное дополнение. Как и прежде, мы утверждали, что в этом случае сложность имеет убывающий характер и тенденцию. Мы дадим строгое определение операционной части этой модифицированной конструкции подмножества, принадлежащей Рабину и Скотту. Сложность алгоритма остается на величину меньше, чем у NP-сложных задач, для которых мы дали строгое доказательство эквивалентности в предыдущей работе, поэтому эта работа продолжает изучение сравнимого доказательства для множества задач, которые, однако, являются вычислительно сложными. Объяснимо с точки зрения единого подхода, такого как операционное исчисление. В этом исчислении основные точки исследования отданы представлению модифицированной конструкции подмножества с не менее чем двумя операндами, которые должны быть вычислены путем построения подмножества, и с точки зрения сложности эффективного алгоритма они вычисляются с использованием модифицированной конструкции подмножества.

**Ключевые слова:** построение подмножества, расширенные регулярные выражения, модификация, операции, исчисление.

*Сведение об авторе:*

*Анг.: Syzdykov Mirzakhmet - Satbayev University, Almaty, Kazakhstan*

*Каз.: Сыздықов Мырзахмет- Қ.И. Сәтпаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы, Қазақстан.*

*Рус.: Сыздықов Мырзахмет- Казахский национальный исследовательский технический университет имени К.И. Сатпаева, Алматы, Казахстан.*

УДК 004.89

## МАРКОВСКИЕ ИДЕИ В БАЙЕСОВСКИХ СЕТЯХ

Ахметова А.М.<sup>1</sup>, Абдилдаева А.А.<sup>1</sup>, Литвиненко Н. Г.<sup>1</sup>, Литвиненко А. Г.<sup>2</sup>

<sup>1</sup>Казахский Национальный Университет имени Аль-Фараби, г. Алматы

<sup>2</sup>Lehrstuhl für Mathematics for Uncertainty Quantification

[ardak\\_66@mail.ru](mailto:ardak_66@mail.ru)

ORCID ID: <https://orcid.org/0000-0001-8360-5504>

ORCID ID: <https://orcid.org/0000-0002-6381-9350>

ORCID ID: <https://orcid.org/0000-0002-0576-8305>

ORCID ID: <https://orcid.org/0000-0001-5427-3598>

**Аннотация.** В статье рассматриваются байесовские сети (в дальнейшем БС) с дополнительными ограничениями, обусловленными идеями марковских сетей (в дальнейшем МС). Графовые модели, описываемые подобными сетями, будем называть байесовскими марковскими сетями (БМС). Данные модели описывают многие реальные задачи с различными видами неопределенностей, имеющих различные причинно-следственные связи. Ограничения, накладываемые марковским свойством, во многих случаях позволяют значительно упростить расчеты в байесовских сетях при наличии свидетельств. В то же время практика показывает, что большинство моделей, отражающих реальные процессы и построенных на аппарате байесовских сетей в действительности обладают и марковским свойством.

В статье описаны идеи расчетов в байесовских сетях с ограничениями, накладываемыми марковским свойством. Описаны различия при расчетах в байесовских сетях без марковских ограничений и с марковскими ограничениями.

Работа написана в рамках грантового финансирования AP19679142 «Поиск оптимальных решений в байесовских сетях в моделях с линейными ограничениями и линейными функционалами. Разработка алгоритмов и программ» (2023-2025гг.) МОНВ РК.

**Ключевые слова:** Байесовская сеть, Марковская сеть, Марковское случайное поле, Марковские свойства, графическая модель, свидетельство, распространение свидетельств.

### Введение

Познакомиться с теорией байесовских сетей можно в [1, 2, 3]. Подобные задачи часто имеют причинно-следственные связи между различными элементами. В основе данных графических моделей обычно используются направленные ациклические графы. Вершинами в БС являются переменные, носящие вероятностный характер. Между переменными существуют различные причинно-следственные связи. Теория марковских сетей хорошо описана в [4, 5, 6]. Для иллюстрации излагаемого материала в статье используется программный комплекс HUGIN EXPERT. Ознакомится с работой данной программы можно в [7]. В статье будет использован известный учебный пример ASIA. Пример придуман и впервые описан в [8].

Байесовские сети позволяют решать различные достаточно сложные задачи, имеющие различные виды неопределенностей. Подобные задачи часто имеют причинно-следственные связи между различными переменными. Теория БС позволяет рассчитать значения некоторых переменных на основе известных или ранее рассчитанных переменных используя причинно-следственные связи. Если одна или несколько переменных получили некоторые значения, то говорят, что переменные получили свидетельства. Наличие свидетельств резко усложняет расчеты в БС. Поиск хороших, быстрых алгоритмов для расчетов в БС со свидетельствами достаточно актуальная проблема. Дополнительное условие марковости во многих случаях может значительно упростить алгоритмы. На примере будет рассмотрена возможность подобного упрощения.

## Постановка задачи

### Краткое описание байесовских сетей и марковских сетей.

Пусть задан граф  $G = \langle V, E \rangle$ , где  $V$  – множество вершин графа,  $E$  – множество ребер графа. Ориентированный, ациклический граф  $G$  называется байесовской сетью если каждой вершине  $v \in V$  поставлена в соответствие случайная величина  $X_v$ , а каждому ребру  $e = (u, v) \in E$  ставится вероятностная зависимость случайной величины  $X_v$  от случайной величины  $X_u$ .

Вершина  $u$  называется родительской для вершины  $v$ , если ребро  $(u, v) \in E$ . Множество всех родительских вершин  $v$  обозначается через  $\text{parents}(X_v)$ . Для вычисления вероятности существует цепное правило для байесовских сетей:

$$P(X_1, X_2, X_3, \dots, X_n) = \prod_{i=1}^n P(X_i | \text{parents}(X_i))$$

Цепное правило позволяет разложить совместное распределение в произведение условных распределений. Очевидно, что при расчете соблюдается принцип Маркова: переменная зависит только от родительских переменных и не важно каким образом родительские переменные получили свое значение.

Марковская сеть это графовая модель, представленная неориентированным графом, в которой множество случайных величин обладает Марковским свойством. Марковское свойство определяется одним из трех эквивалентных способов:

Свойство пар: Любые две несмежные переменные условно независимы с учетом всех других переменных

Локальное свойство: переменная условно независима от всех других величин, с учетом своих соседей

Глобальное свойство: Любые два подмножества переменных условно независимы с учетом разделяющего подмножества

В байесовских сетях играет важное понятие d-разделенности. Суть данного понятия – попытаться выделить из байесовской сети некоторую подсеть, расчеты в которой не зависят от вершин, не лежащих в данной подсети. Это позволит существенно снизить объемы расчетов во всей сети. Ниже приведено определение d-разделенности.

Путь  $S$  между вершинами  $a$  и  $b$  называют d-разделённым или заблокированным множеством вершин  $Z$  тогда и только тогда, когда:

$S$  содержит цепь  $i \rightarrow z \rightarrow j$  или разветвление  $i \leftarrow z \rightarrow j$  такие, что  $z$  принадлежит  $Z$  или

$S$  содержит инвертированное разветвление (коллайдер)  $i \rightarrow z \leftarrow j$  такие, что  $z$  не принадлежит  $Z$  и у вершины  $z$  нет потомков принадлежащих  $Z$ .

Понятие d-разделенности можно распространить и на непересекающиеся подмножества  $A$  и  $B$ . Два подмножества  $A$  и  $B$  называются d-разделённым, если любые два элемента из этих подмножеств d-разделены.

Рассмотрим известный в литературе по БС учебный пример ASIA. В данном примере рассматриваются несколько вариантов задания свидетельств. Проводим анализ решений. Рассматриваем способ решения при условии марковости. Сравниваем два способа построения алгоритмов.

Описание учебного примера (рисунок 1).

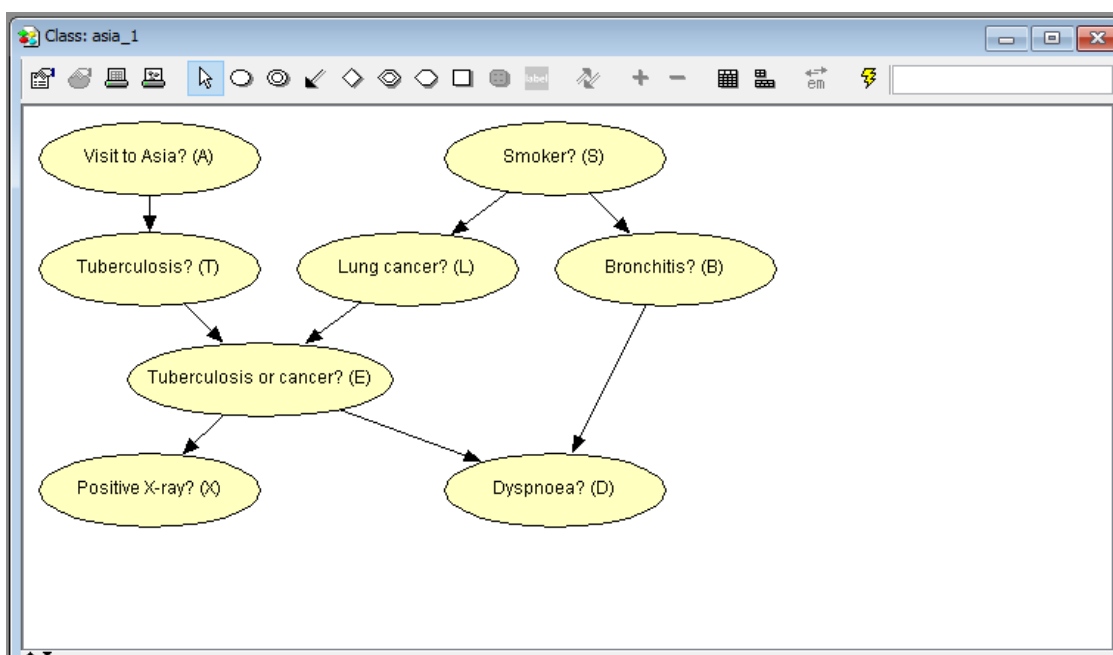


Рисунок 1 - Учебная байесовская сеть ASIA

Исследуется здоровье человека. Вершина A обозначает был или нет исследуемый в Азии. В таблице 1 определена вероятность данного факта. От этого в значительной степени зависит вероятность заболевания туберкулезом (T). В таблице 2 определена вероятность заболевания. Вершина S и таблица 3 определяют курит или нет человек. Курение может способствовать заболеванию раком (L), заболеванию бронхитом (B) или и тем и другим одновременно (E) (таблица 4, таблица 5 и таблица 6). Результат рентгена (X) грудной клетки может показать с определенной вероятностью (таблица 7) заболевание раком или туберкулезом, но различить эти болезни не может. Рак, туберкулез и бронхит с определенной вероятностью вызывают одышку (таблица 8).

Visit to Asia? (A)	
	Probability
yes	0.01
no	0.99

Таблица 1

Tuberculosis? (T)		
	yes	no
Visit to Asia? (A)		
yes	0.05	0.01
no	0.95	0.99

Таблица 2

Smoker? (S)	
	Probability
yes	0.5
no	0.5

Таблица 3

Smoker? (S)	yes	no
yes	0.1	0.01
no	0.9	0.99

Таблица 4

Smoker? (S)	yes	no
yes	0.6	0.3
no	0.4	0.7

Таблица 5

Tuberculosis or cancer? (E)	yes		no	
	yes	no	yes	no
Lung cancer? (L)	1	1	1	0
no	0	0	0	1

Таблица 6

Tuberculosis or cancer? (E)	yes	no
yes	0.98	0.05
no	0.02	0.95

Таблица 7

Dyspnoea? (D)	yes		no	
	yes	no	yes	no
Bronchitis? (B)	0.9	0.8	0.7	0.1
no	0.1	0.2	0.3	0.9

Таблица 8

## Методы исследования

Рассмотрим стандартные расчеты данного примера (рисунок 2). Предполагаем, что свидетельства отсутствуют. Вероятности переменных рассчитываются последовательно, начиная с маргинальных (независимых) переменных. Маргинальными в нашем примере являются переменные A и S. Маргинальные переменные отнесем к нулевому уровню. В первую очередь рассчитываются переменные, зависящие лишь от маргинальных переменных. Это переменная T, зависящая от маргинальной переменной A и переменные L и B, зависящие от переменной S. Переменные S, L, B отнесем к первому уровню. Далее мы рассчитываем переменные второго уровня. Эти переменные могут зависеть от переменных нулевого и первого уровня. E единственная переменная второго уровня, зависящая от переменных первого уровня T и L. Затем рассчитываем переменные третьего уровня. Эти переменные могут зависеть от переменных нулевого, первого и второго уровня. Переменная X зависит от переменной второго уровня E. Переменная D зависит от переменной второго уровня E и переменной первого уровня B.

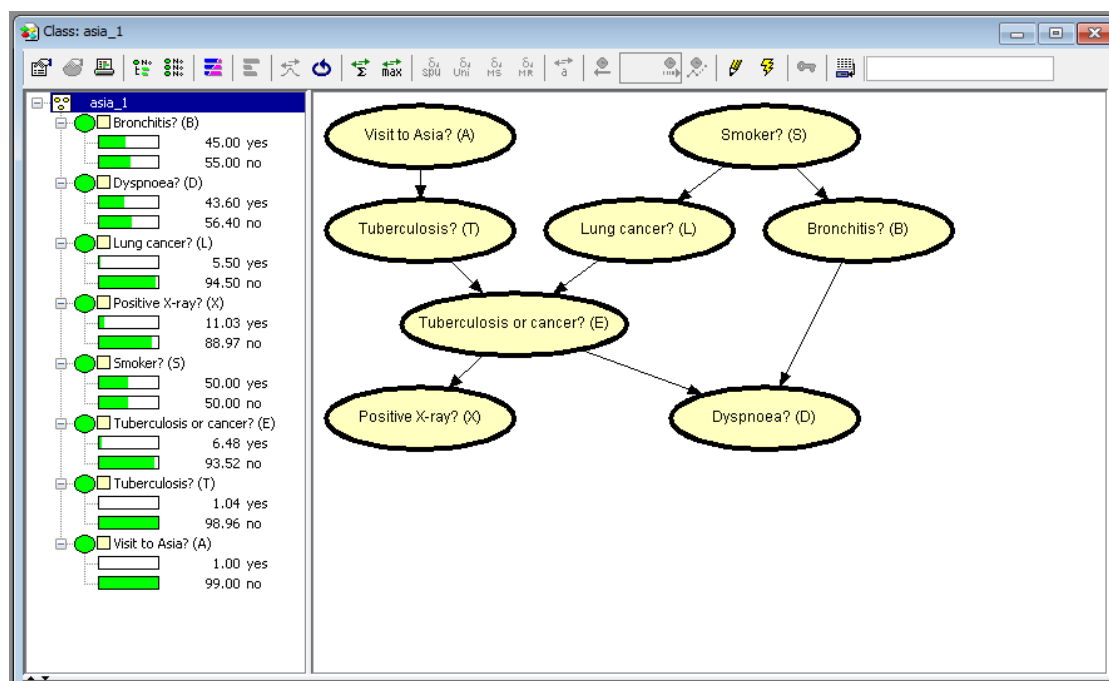


Рисунок 2 - Пример стандартного расчета

Решение находится слева. Из решения видно, например, что вероятность заболевания туберкулезом или раком (E) при заданных маргинальных переменных равна 6.4%. А вероятность иметь одышку (D) равна 43.6%. Заметим, что в расчетах мы использовали только соседние переменные, т.е. реально и существенно использовали марковское свойство.

Ситуация с расчётами существенно меняется при получении свидетельств. Допустим мы сделали рентген, давший положительный результат. Решение мы видим на рисунке 3. Наиболее популярный алгоритм расчетов в байесовских сетях состоит в следующих построениях:

1. Построение доменного графа.
2. Построение морального графа.
3. Построение триангулярного графа.
4. Построение дерева смежности.
5. Построение дерева сочленений.

Разработка и реализация подобных алгоритмов достаточно сложна. Хотелось бы найти возможности определить и в данном случае марковское свойство и найти способ разделить переменные на группы и проводить расчеты тем же способом, как и в случае отсутствия свидетельств.

При наличии свидетельств, свидетельства становятся в каком-то смысле определяющими в расчетах. В нашем случае будем считать переменную X переменной нулевого уровня. Переменная E будет переменной первого уровня, зависит только от переменной нулевого уровня. Переменные T, L, D будут переменными второго уровня. Переменные A, S будут переменными третьего уровня. Переменная S будет переменной четвертого уровня, зависит от переменных второго и третьего уровня.



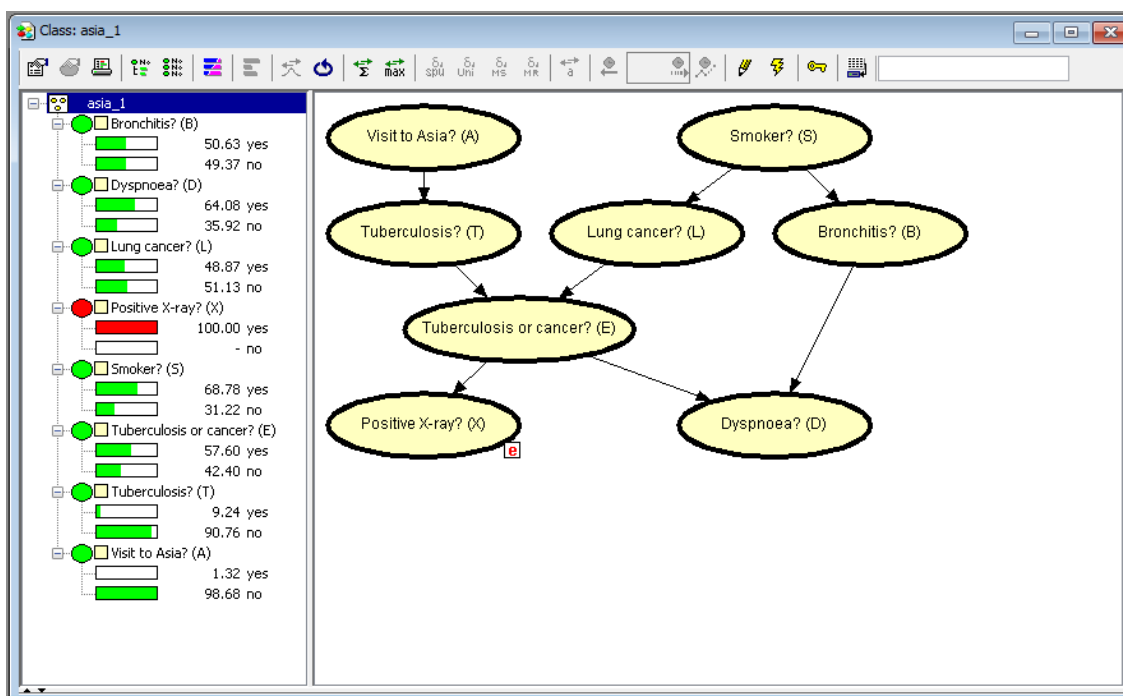


Рисунок 3 - Решение при получении свидетельств

Расчеты, в силу структур таблиц условных вероятностей, можно проводить используя теорему Байеса.

### Результаты и их обсуждения

В статье рассматриваются возможности использования свойства марковости для упрощения расчетов в байесовских сетях. Решая одни проблемы, мы получаем новые проблемы:

Обоснование использования свойства Маркова.

Как использовать существующие таблицы вероятностей с недостаточно удобными структурами.

Как корректно обосновать использование теоремы Байеса.

Необходимо доказать, что при нескольких свидетельствах и смене направлений некоторых ребер не возникает циклов.

Исследование возникновения противоречий в байесовских сетях при некорректном, несовместимом получении нескольких свидетельств. Пути решения возможных противоречий.

Тем не менее рассматриваемый подход представляется достаточно перспективным и требует дальнейшего исследования.

### Заключение

В статье рассматриваются возможности использования марковости при расчетах в байесовских сетях со свидетельствами. Цель – упростить алгоритмы расчетов. Работа выполнена в рамках грантового проекта AP19679142 «Поиск оптимальных решений в байесовских сетях в моделях с линейными ограничениями и линейными функционалами. Разработка алгоритмов и программ» (2023-2025гг.) МОНВ РК.

### Список литературы

[1] Pearl J. Probabilistic Reasoning in Intelligent Systems. San Francisco: Morgan Kaufmann Publishers, 1988. 552 p

- [2] Литвиненко Н.Г., Литвиненко А.Г., Мамырбаев О.Ж., Шаяхметова А.С. Байесовские сети. Теория и практика. Алматы: Институт информационных и вычислительных технологий, 2020. 197 p. ISBN 978-601-332-888-1.
- [3] Robert G. Cowell, A. Philip Dawid, Steffen L. Lauritzen, David J. Spiegelhalter. Probabilistic Networks and Expert Systems. Springer, 1999. 3321 p. ISBN 0-387-98767.
- [4] Li, S. Z. Markov Random Field Modeling in Image Analysis. Springer, London, 2009.
- [5] Kindermann Ross; Snell, J. Laurie. Markov Random Fields and Their Applications (англ.). American Mathematical Society, 1980. ISBN 0-8218-5001-6.
- [6] Rue, Håvard; Held, Leonhard. Gaussian Markov random fields: theory and applications (англ.). CRC Press, 2005. — ISBN 1584884320.
- [7] HUGIN Graphical User Interface. Documentation. Release 9.1. 2021.
- [8] Lauritzen, S.L. and Spiegelhalter, D.J. Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems. Journal of the Royal Statistical Society. Series B, 50, 157-224. 1988.

## БАЙЕСТІК ЖЕЛІЛЕРДЕГІ МАРКОВТІК ИДЕЯЛАР

Ахметова А.М.<sup>1</sup>, Абдилдаева А.А.<sup>1</sup>, Литвиненко Н. Г.<sup>1</sup>, Литвиненко А. Г.<sup>2</sup>

<sup>1</sup>Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы қ-сы  
<sup>2</sup>Белгісіздікті сандық бағалау үшін математика кафедрасы, Германия  
[ardak\\_66@mail.ru](mailto:ardak_66@mail.ru)

ORCID ID: <https://orcid.org/0000-0001-8360-5504>

ORCID ID: <https://orcid.org/0000-0002-6381-9350>

ORCID ID: <https://orcid.org/0000-0002-0576-8305>

ORCID ID: <https://orcid.org/0000-0001-5427-3598>

**Аңдатпа.** Мақалада Марков желілерінің (бұдан әрі МС) идеяларына байланысты қосымша шектеулермен Байес желілері (бұдан әрі БС) қарастырылады. Ұқсас желілермен сипатталған графикалық модельдерді Байес Марков желілері (BMS) деп атаймыз. Бұл модельдер әртүрлі себептік байланыстары бар әртүрлі белгісіздіктері бар көптеген нақты міндеттерді сипаттайды. Марков қасиеті қойған шектеулер көптеген жағдайларда дәлелдер болған кезде Байес желілеріндегі есептеулерді едәуір жеңілдетеді. Сонымен қатар, тәжірибе көрсеткендей, нақты процестерді көрсететін және аппаратта салынған Байес желілерінің модельдерінің көпшілігі Марковтық қасиетке ие.

Мақалада Марков қасиетімен шектелген Байес желілеріндегі есептеу идеялары сипатталған. Марков шектеулері жоқ және Марков шектеулері бар Байес желілеріндегі есептеулердегі айырмашылықтар сипатталған.

Жұмыс AP19679142 "Сызықтық шектеулер мен сызықтық функционалдығы бар модельдерде Байес желілерінде оңтайлы шешімдерді табу. Алгоритмдер мен бағдарламаларды құру" (2023-2025жж.) гранттық қаржыландыру аясында орындалған.

**Кілттік сөздер:** Байес желісі, Марков желісі, Марковтың кездейсоқ өрісі, Марков қасиеттері, графикалық модель, айғақтар, айғақтардың таралуы.

## MARKOV IDEAS IN BAYESIAN NETWORKS

Akhmetova A.M.<sup>1</sup>, Abdildayeva A.A.<sup>1</sup>, Litvinenko N. G.<sup>1</sup>, Litvinenko A. G.<sup>2</sup>

Al-Farabi Kazakh National University, Almaty  
Lehrstuhl für Mathematics for Uncertainty Quantification, Germany

[ardak\\_66@mail.ru](mailto:ardak_66@mail.ru)

ORCID ID: <https://orcid.org/0000-0001-8360-5504>

ORCID ID: <https://orcid.org/0000-0002-6381-9350>

ORCID ID: <https://orcid.org/0000-0002-0576-8305>

ORCID ID: <https://orcid.org/0000-0001-5427-3598>

**Abstract.** The article deals with Bayesian networks (hereinafter BS) with additional limitations due to the ideas of Markov networks (hereinafter MS). Graph models described by such networks will be called Bayesian Markov networks (BMS). These models describe many real-world problems with different types of uncertainties having different cause-and-effect relationships. The limitations imposed by the Markov property in many cases make it possible to significantly simplify calculations in Bayesian networks in the presence of evidence. At the same time, practice shows that most models reflecting real processes and built on the apparatus of Bayesian networks actually have the Markov property.

The article describes the ideas of calculations in Bayesian networks with restrictions imposed by the Markov property. The differences in calculations in Bayesian networks without Markov constraints and with Markov constraints are described.

The work was written within the framework of grant funding AP19679142 "Search for optimal solutions in Bayesian networks in models with linear constraints and linear functionals. Development of algorithms and programs" (2023-2025) MONV RK.

**Keywords:** Bayesian network, Markov network, Markov random field, Markov properties, graphical model, evidence, dissemination of evidence.

*Сведения об авторах:*

*Рус: Ахметова А.М. - PhD, и.о. ассоциированного профессора кафедры Искусственного интеллекта и Big Data, Казахский национальный университет имени аль-Фараби, Алматы, Казахстан.*

*Қаз: Ахметова А.М. - PhD, Жасанды интеллект және Big Data кафедрасының қауымдастырылған профессор м.а., әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан.*

*Англ: Akhmetova A.M. - PhD, acting Associate Professor of the Department of Artificial Intelligence and Big Data, Al-Farabi Kazakh National University, Almaty, Kazakhstan.*

*Рус: Абдилдаева А.А. - PhD, ассоциированный профессор кафедры Искусственного интеллекта и Big Data, Казахский национальный университет имени аль-Фараби, Алматы, Казахстан.*

*Қаз: Абдилдаева А.А. - PhD, Жасанды интеллект және Big Data кафедрасының қауымдастырылған профессоры, әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан.*

*Англ: Abdildayeva A.A. - PhD, Associate Professor of the Department of Artificial Intelligence and Big Data, Al-Farabi Kazakh National University, Almaty, Kazakhstan.*

*Рус: Литвиненко Н.Г. - магистр, Казахский национальный университет имени аль-Фараби, Алматы, Казахстан.*

*Қаз: Литвиненко Н.Г.- магистр, әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан.*

*Англ: Litvinenko N. G. - master, Al-Farabi Kazakh National University, Almaty, Kazakhstan.*

*Рус: Литвиненко Александр- доктор естественных наук кафедры математики количественной оценки неопределенности, Германия*

*Қаз: Литвиненко Александр – жаратылыстану ғылымдарының докторы, белгісіздікті сандық бағалау математика кафедрасы, Германия*

*Англ: Litvinenko Alexander - Dr. rer. nat. Chair of Mathematics for Uncertainty Quantification, Germany.*

МРНТИ 28.23.37

## ПРИМЕНЕНИЕ ТЕХНОЛОГИЙ МАШИННОГО ОБУЧЕНИЯ ПРИ ДИАГНОСТИКЕ ГЛАУКОМЫ

Павлов С.В.<sup>1</sup>, Мамырбаев О.Ж.<sup>2</sup>, Оралбекова Д.О.<sup>2,3</sup>, Момынжанова К.<sup>4\*</sup>

<sup>1</sup>Винницкий Национальный Технический Университет, Винница, Украина

<sup>2</sup>Институт информационных и вычислительных технологий, Алматы, Казахстан

<sup>3</sup>НАО «Алматинский университет энергетики и связи имени Гумарбека Даукеева», Алматы, Казахстан

<sup>4</sup>НАО «Казахский Национальный Университет имени Аль-Фараби», Алматы, Казахстан

ORCID ID: <https://orcid.org/0000-0001-8318-3794>

ORCID ID: <https://orcid.org/0000-0003-3455-3319>

e-mail: [morkenj@mail.ru](mailto:morkenj@mail.ru), [dinaoral@mail.ru](mailto:dinaoral@mail.ru), [kymbat\\_momynzhanova87@mail.ru](mailto:kymbat_momynzhanova87@mail.ru)

**Аннотация.** Глаукома — это прогрессирующее заболевание глаз, которое при отсутствии лечения может привести к необратимой потере зрения или полной слепоте. Раннее выявление и лечение глаукомы имеют решающее значение для предотвращения потери зрения. Однако диагностика и лечение глаукомы требует тщательного тестирования и регулярного мониторинга, что может занимать много времени и средств. Диагностика глаукомы - сложный процесс, требующий высокой точности и эффективности в выявлении симптомов заболевания. В последние годы применение технологий машинного обучения стало все более популярным в области диагностики глаукомы. В этой статье представлен обзор недавних исследований, которые применяли алгоритмы машинного обучения для диагностики глаукомы, включая классификацию глаукомы и прогнозирование ее прогрессирования. Статья подчеркивает преимущества применения современных технологий машинного обучения, такие как улучшение точности, эффективности и объективности в диагностике глаукомы. Кроме того, описываются некоторые проблемы и ограничения этих технологий, а также предлагаются потенциальные решения. В целом, статья подчеркивает потенциал технологий глубокого обучения в диагностике глаукомы и их роль в улучшении результатов для пациентов.

**Ключевые слова:** машинное обучение, глаукома, заболевания глаз, диагностика, прогнозирование прогрессирования.

### Введение

Глаукома является одним из самых распространенных заболеваний глаз, которое может привести к ухудшению зрения и в некоторых случаях к его полной потере. Это одна из основных причин слепоты во всем мире, от которой страдают примерно 80 миллионов человек во всем мире, и ожидается, что к 2040 году их число возрастет до 111,8 миллионов [1]. Она вызывается повышением внутриглазного давления, которое приводит к повреждению зрительного нерва и снижению зрительной функции. По данным Всемирной организации здравоохранения, глаукома является второй по частоте причиной слепоты в мире. Глаукома часто протекает бессимптомно на ранних стадиях и может оставаться не диагностированной до тех пор, пока не произойдет значительная потеря зрения. Раннее выявление и лечение имеют решающее значение для предотвращения дальнейшей потери зрения и сохранения качества жизни пациентов с глаукомой.

Диагностика и лечение глаукомы включает в себя всестороннее обследование глаз, которое включает проверку поля зрения, измерение внутриглазного давления и оценку зрительного нерва [2]. Регулярный мониторинг и тестирование необходимы для лечения заболевания и предотвращения потери зрения. Однако процесс диагностики и мониторинга глаукомы может быть трудоемким и дорогостоящим, что может привести к потенциальным

препятствиям для доступа и приверженности пациентов.

Для обнаружения и диагностики глаукомы используются различные методы, включая офтальмоскопию, гониоскопию, пахиметрию, периметрию и тонометрию. Однако, когда глаукома обнаруживается в ранней стадии, ее лечение обычно более успешно. Поэтому постоянно возникает потребность в разработке более эффективных методов диагностики, особенно в ранней стадии заболевания.

В последние годы, применение технологий машинного обучения (МО) в области медицины стало все более популярным. МО позволяет обрабатывать большие объемы данных и создавать алгоритмы, которые могут быстро и точно диагностировать заболевания. В области диагностики глаукомы, МО показывает высокую точность диагностики и может помочь в обнаружении заболевания в ранней стадии.

Цель этой обзорной статьи - описать, как применение технологий МО изменило подход к диагностике глаукомы и обсудить текущее состояние исследований, а также перспективы будущего применения МО в этой области. Мы представляем обзор существующих подходов к диагностике глаукомы с использованием МО и описываем различные алгоритмы, используемые в этих исследованиях. Мы также обсуждаем преимущества и недостатки применения МО в диагностике глаукомы, а также возможности дальнейшего развития этой области исследований.

## Литературный обзор

Song et al. (2021) [3] в своей обзорной статье рассмотрели различные исследования, посвященные применению глубокого обучения для диагностики глаукомы. Данные исследователи описали преимущества и ограничения использования методов машинного обучения для диагностики глаукомы, а также обозначили перспективы дальнейших исследований в данной области.

Bolme et al. (2021) [4] также провели систематический обзор исследований, посвященных применению глубокого обучения для диагностики глаукомы. В работе авторы отметили, что большинство исследований показывают высокую точность и чувствительность методов машинного обучения при диагностике глаукомы, но при этом отметили необходимость проведения дополнительных исследований в данной области.

Yaqoob et al. (2020) [5] провели обзор литературы по применению искусственного интеллекта в офтальмологии в целом, включая и диагностику глаукомы. Авторы подчеркнули, что использование методов машинного обучения в офтальмологии может улучшить точность диагностики и облегчить работу врачей, но также отметили необходимость проведения большого количества исследований, чтобы определить эффективность и перспективы применения этих методов.

Medeiros et al. (2021) [6] в своей статье рассмотрели перспективы и ограничения применения искусственного интеллекта в диагностике глаукомы. Было выявлено, что существуют различные подходы к применению методов машинного обучения в диагностике глаукомы, но отметили, что необходимо проводить больше исследований, чтобы понять, каким образом эти методы могут быть использованы в клинической практике.

Thakur et al. (2021) [7] провели систематический обзор исследований, посвященных применению методов машинного обучения для диагностики глаукомы. В работе авторы отметили, что методы машинного обучения могут быть эффективным инструментом для раннего выявления глаукомы и определения степени ее развития. Кроме того, они могут помочь в оценке эффективности лечения и прогнозировании результатов.

Christopher et al. (2021) [8] рассмотрели применение искусственного интеллекта в скрининге глаукомы. Авторы отметили, что методы машинного обучения, основанные на анализе зрительных полей и оптической кохеографии, могут помочь в ранней диагностике глаукомы и выявлении пациентов с высоким риском развития болезни.

Hu et al. (2021) [9] рассмотрели применение методов глубокого обучения для

детектирования глаукомы на изображениях оптической когерентной томографии. На основе полученных результатов было обнаружено, что глубокое обучение может улучшить точность диагностики и помочь в обнаружении начальных стадий глаукомы.

Li et al. (2020) [10] провели систематический обзор и мета-анализ исследований, посвященных автоматизированной диагностике глаукомы с использованием методов глубокого обучения и фандусной фотографии. Исследователи данной работы отметили, что методы глубокого обучения могут быть более эффективными, чем традиционные методы диагностики, в том числе в случаях с низкой квалификацией врачей.

Morales-Fernandez et al. (2020) [11] рассмотрели применение искусственного интеллекта для детектирования глаукомы. Авторы выявили, что методы машинного обучения могут помочь улучшить точность диагностики и прогнозирования результатов лечения.

Wu et al. (2020) [12] рассмотрели применение искусственного интеллекта и глубокого обучения для диагностики глаукомы. А Fathi et al. (2020) [13] рассмотрели применение искусственного интеллекта для диагностики и управления глаукомой. Авторы отметили, что методы машинного обучения, такие как глубокое обучение, имеют высокую точность диагностики и могут помочь в развитии индивидуализированных методов лечения глаукомы. Они обсудили также необходимость использования крупных и разнообразных наборов данных для обучения моделей машинного обучения и управления качеством этих данных. А также выявили, что методы глубокого обучения могут улучшить точность диагностики и помочь в определении риска развития глаукомы

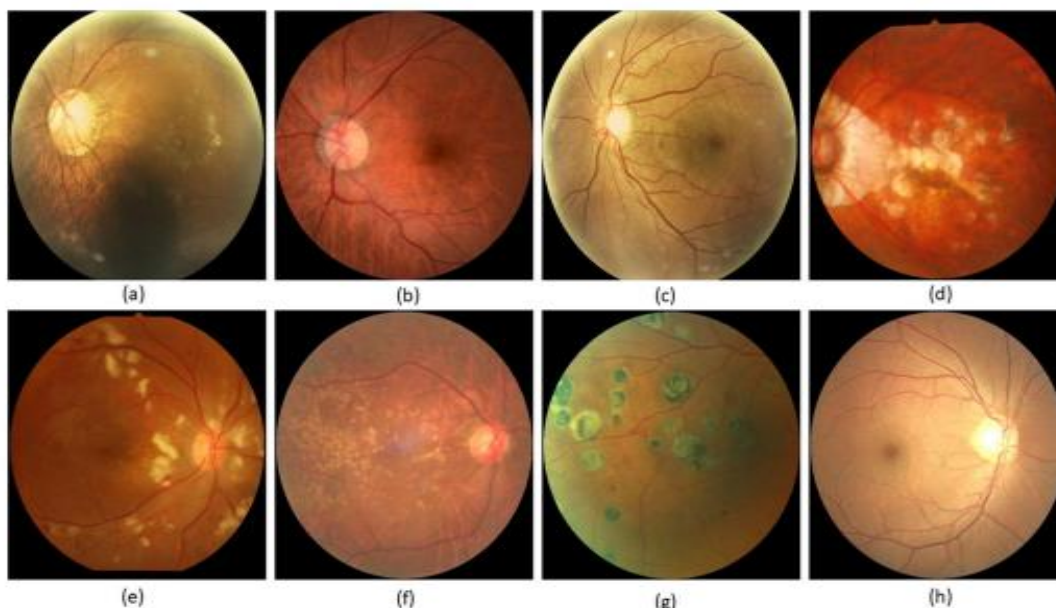
Asaoka et al. (2020) [14] обсудили применение различных методов машинного обучения, включая глубокое обучение, для диагностики и управления глаукомой. Они отметили, что многие исследования показывают высокую точность и чувствительность методов машинного обучения в диагностике глаукомы. Однако авторы отметили необходимость более широкого использования этих методов в клинической практике и необходимость дальнейшего совершенствования алгоритмов, чтобы улучшить их точность и надежность. В целом, авторы считают, что использование искусственного интеллекта может значительно улучшить диагностику и управление глаукомой и стать ценным инструментом для врачей и пациентов.

## **Методы и материалы**

В последние годы машинное обучение стало важным инструментом для улучшения диагностики глаукомы. В этом разделе мы рассмотрим основные методы машинного обучения, применяемые в данной области, а также их преимущества и ограничения.

### *Обработка изображений глазного дна*

Одним из ключевых методов машинного обучения в диагностике глаукомы является обработка изображений глазного дна. Этот метод позволяет анализировать изображения, полученные с помощью глазных камер, и выявлять признаки, связанные с глаукомой. В частности, используются методы выделения границ диска зрительного нерва, квадрантов зрительного поля и других зон глазного дна (рис. 1) [17].



**Рисунок 1** - Изображения глаукомы: (a) эпиретинальная мембрана макулы, (b) нормальное глазное дно, (c) легкая непролиферативная ретинопатия, (d) патологическая миопия, (e) гипертоническая ретинопатия, (f) лазерное пятно, умеренная непролиферативная ретинопатия, (g) умеренная непролиферативная ретинопатия, лазерное пятно, (h) легкая непролиферативная ретинопатия

Для обработки изображений глазного дна применяются различные методы машинного обучения, включая сверточные нейронные сети (Convolutional Neural Networks, CNN), случайные леса (Random Forests), метод опорных векторов (Support Vector Machines, SVM) и др. Каждый из этих методов имеет свои преимущества и ограничения, и выбор метода зависит от конкретной задачи и доступных данных.

Для оценки моделей машинного обучения используются общие показатели, такие как точность, чувствительность и специфичность (1-3). Чувствительность оценивает способность модели идентифицировать случаи с заболеванием, а специфичность оценивает способность модели идентифицировать случаи без заболевания. Следующие уравнения используются для расчета чувствительности, специфичности и точности, где TP - истинно положительный, TN - истинно отрицательный, FP - ложноположительный и FN - ложноотрицательный.

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (1)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (2)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+TN+FP} \quad (3)$$

#### *Анализ данных о зрительных функциях*

Кроме обработки изображений глазного дна, машинное обучение также применяется для анализа данных о зрительных функциях пациентов. Для этого используются данные о зрительном поле, оптической когерентной томографии (ОКТ) и других функциональных методах оценки зрительных функций [18-19].

Данные о зрительных функциях пациентов анализируются с помощью различных методов машинного обучения, таких как нейронные сети, SVM, случайные леса и др. В результате анализа можно получить информацию о состоянии зрительных функций пациента и о возможных рисках развития глаукомы.

#### *Использование многомерных данных*

Многомерные данные являются ключевым элементом диагностики глаукомы, поскольку она характеризуется несколькими признаками, включая давление внутри глаза, меры толщины роговицы, анализ глазного дна и зрительного поля. Машинное обучение может эффективно использоваться для анализа этих многомерных данных и предоставления более точных результатов.

Одним из методов машинного обучения, которые широко используются для анализа многомерных данных, является метод главных компонент (РСА). Этот метод позволяет сократить размерность данных, удаляя шум и несущественные признаки, и выделять главные компоненты данных, которые объясняют большую часть их изменчивости. РСА может применяться для анализа данных о давлении внутри глаза, зрительных полях и других параметрах, связанных с глаукомой.

Другим методом машинного обучения, который может использоваться для анализа многомерных данных, является метод кластеризации. Этот метод позволяет выделить группы данных, которые имеют сходные характеристики, и использовать эту информацию для классификации новых данных. Кластеризация может применяться для анализа данных о зрительных полях и других параметрах глаукомы, чтобы выделить группы пациентов с различными стадиями заболевания.

В целом, использование методов машинного обучения для анализа многомерных данных является мощным инструментом для диагностики и прогнозирования глаукомы. Однако, при использовании этих методов необходимо учитывать особенности конкретных данных, чтобы обеспечить их точность и надежность. Также необходимо учитывать этические вопросы, связанные с использованием машинного обучения в медицине, чтобы защитить конфиденциальность пациентов и обеспечить справедливость и равноправие в доступе к медицинской помощи.

## **Эксперименты и результаты**

В этом разделе описываются эксперименты и результаты, полученные в ряде исследований по этой теме.

Volme и соавторы (2021) [4] провели систематический обзор автоматизированного обнаружения глаукомы с помощью технологий машинного обучения и получили следующие результаты. Из 55 исследований, включенных в обзор, большинство использовали анализ изображений сетчатки для диагностики глаукомы. В 36 исследованиях были использованы алгоритмы глубокого обучения, такие как сверточные нейронные сети (CNN), рекуррентные нейронные сети (RNN) и комбинации этих алгоритмов. Из них 28 исследований получили точность диагностики выше 90%.

Үақооб и соавторы провели обзор литературы о применении искусственного интеллекта в офтальмологии и получили следующие результаты. В большинстве исследований использовались данные, полученные при помощи оптической когерентной томографии (ОКТ) и анализ изображений сетчатки. Алгоритмы машинного обучения, используемые в этих исследованиях, включали CNN, RNN и нейронные сети с обратной связью (FBNN). Большинство из них показало высокую точность диагностики глаукомы.

Medeiros и соавторы описали преимущества и ограничения использования технологий машинного обучения в диагностике глаукомы. В результате они пришли к выводу, что эти технологии могут помочь в определении риска развития глаукомы, мониторинге прогрессирования заболевания и оценке эффективности лечения. Однако, они также отметили, что существуют ограничения в применении машинного обучения, один из них является ограниченной доступностью данных. Существуют также другие ограничения, которые не были упомянуты в статье Medeiros и соавторов. Некоторые из них могут включать необходимость высококвалифицированных специалистов, которые могут анализировать и интерпретировать данные, полученные при помощи машинного обучения. Это может потребовать дополнительных затрат на обучение и подготовку персонала. Кроме того, существует возможность, что алгоритмы машинного обучения могут привести к



ложным результатам, что может оказать негативное влияние на диагностику и лечение глаукомы. Это может произойти, если данные, используемые для обучения алгоритмов, содержат ошибки или недостаточно точны. Несмотря на эти ограничения, применение технологий машинного обучения в диагностике глаукомы имеет большой потенциал и может привести к улучшению качества диагностики и лечения пациентов с этим заболеванием. Поэтому необходимо продолжать исследования в этой области и улучшать алгоритмы машинного обучения, чтобы повысить точность диагностики и эффективность лечения.

## Обсуждения

В данном обзоре мы рассмотрели применение технологий машинного обучения в диагностике глаукомы и описали основные алгоритмы и методы, используемые в этой области. Существует множество исследований, подтверждающих эффективность использования методов машинного обучения в определении риска развития заболевания, мониторинге его прогрессирования и оценке эффективности лечения. Однако, несмотря на все преимущества, существуют и некоторые ограничения, и проблемы в применении этих технологий, которые необходимо учитывать. Одним из основных преимуществ использования машинного обучения в диагностике глаукомы является возможность автоматического анализа многомерных данных, таких как ОФТГ, ОСТ, фонд глаза и другие. Это позволяет получать более точную и объективную информацию, чем при оценке вручную. Кроме того, алгоритмы машинного обучения могут учитывать множество различных факторов, таких как возраст, пол, наличие других заболеваний и т.д., что позволяет более точно определять риск развития глаукомы.

Одним из ограничений применения машинного обучения является необходимость большого количества данных для обучения алгоритмов. В некоторых случаях доступ к большим объемам данных может быть ограничен, что затрудняет использование методов машинного обучения в практической медицине. Кроме того, существуют проблемы с качеством данных, такие как недостаточная разметка, ошибки при сборе данных и т.д. Еще одной проблемой является сложность интерпретации результатов, полученных с помощью методов машинного обучения. В отличие от традиционных методов, где эксперты могут объяснить свои решения, алгоритмы машинного обучения могут давать решения, которые не всегда легко интерпретировать. Это может быть особенно проблематично в случае, если машинное обучение используется для принятия важных медицинских решений, таких как выбор метода лечения.

Кроме того, возникает задача постоянного обновления и модификации алгоритмов машинного обучения. Это связано с тем, что данные могут изменяться со временем, и алгоритмы должны быть обучены на новых данных, чтобы оставаться актуальными. Кроме того, существует риск переобучения, когда алгоритмы машинного обучения начинают "запоминать" данные, вместо того чтобы выявлять паттерны и закономерности. Это может привести к неправильным прогнозам и неправильной диагностике. Также следует отметить, что многие алгоритмы машинного обучения используются только для диагностики глаукомы, а не для определения причин и механизмов заболевания. Это означает, что врачи все еще должны проводить дополнительные исследования, чтобы понять причины заболевания и выбрать наиболее эффективное лечение. Более того существует риск недостаточной надежности алгоритмов машинного обучения. С развитием технологий и доступности больших объемов данных, возможности машинного обучения в медицине будут только расти. Кроме того, современные методы машинного обучения могут использоваться для решения более широкого спектра задач, включая прогнозирование прогрессирования заболевания и определение оптимальной терапии.

Таким образом, технологии машинного обучения имеют потенциал для революционизации диагностики и лечения глаукомы. Однако, для того чтобы эта технология была эффективной, необходимо учитывать ее ограничения и недостатки, а

также проводить дополнительные исследования для проверки надежности и эффективности новых методов. В частности, необходимо учитывать не только точность и надежность алгоритмов машинного обучения, но и их применимость в реальных условиях. Необходимо учитывать этические аспекты использования технологий машинного обучения в медицине. Например, важно убедиться, что эти технологии не приводят к дискриминации определенных групп пациентов, а также соблюдать конфиденциальность и безопасность данных. Но, несмотря на ограничения и недостатки, использование технологий машинного обучения в диагностике глаукомы является многообещающим направлением исследований. Также стоит отметить, что использование технологий машинного обучения в медицине является достаточно новым направлением исследований, и будущее этого направления может быть связано с развитием новых алгоритмов машинного обучения, увеличением доступности и объема данных, а также развитием более эффективных методов обработки данных.

В целом, использование технологий машинного обучения в диагностике глаукомы представляет собой перспективное направление исследований, которое может привести к улучшению диагностики и лечения этого заболевания. Однако, для того чтобы эта технология стала полезным инструментом в практике врачей, необходимо проводить дополнительные исследования и учитывать ее ограничения и недостатки.

### **Заключение**

Технологии машинного обучения могут существенно улучшить диагностику, мониторинг и лечение глаукомы. Большинство исследований в этой области показали высокую точность и надежность алгоритмов машинного обучения в диагностике и мониторинге прогрессирования заболевания. Однако, также было выявлено несколько ограничений и недостатков, которые необходимо учитывать при применении этих технологий. Важным недостатком технологий машинного обучения является их зависимость от правильной настройки параметров и качества данных. Неправильная настройка параметров или использование низкокачественных данных может привести к недостаточной точности и надежности алгоритмов машинного обучения. Тем не менее, несмотря на ограничения и недостатки, технологии машинного обучения представляют собой мощный инструмент для диагностики, мониторинга и лечения глаукомы. Они могут помочь определить риски развития заболевания, мониторить прогрессирование заболевания и оценивать эффективность лечения.

В целом, применение технологий машинного обучения в диагностике глаукомы является одним из наиболее перспективных направлений исследований в этой области медицины. Несмотря на некоторые ограничения и недостатки, эта технология может привести к улучшению точности диагностики, мониторинга и лечения глаукомы.

### **Благодарность**

Работа выполнена при финансовой поддержке Комитета науки Министерства образования и науки Республики Казахстан (No BR18574144).

### **Список литературы**

- [1] <https://ria.ru/20230312/glaukoma-1856960364.html>
- [2] Schuster AK, Erb C, Hoffmann EM, Dietlein T, Pfeiffer N. The Diagnosis and Treatment of Glaucoma. Dtsch Arztebl Int. 2020 Mar 27. 117(13). 225-234. DOI: 10.3238/arztebl.2020.0225. PMID: 32343668; PMCID: PMC7196841.
- [3] Song, W., Huang, Y., Cui, H., Zhang, Z., & Wang, Y. Deep learning for glaucoma diagnosis: an updated review. Expert Review of Medical Devices. 2021. 18(5). 337-348.
- [4] Bolme, P., Sridharan, S., Varghese, S., & Prathil, S. Automated detection of glaucoma using deep learning techniques: a systematic review. Eye and Vision. 2021. 8(1). 24.

- [5] Yaqoob, U., Fathi, A., Zadeh, E. K. Artificial Intelligence in Ophthalmology: A Review of the Literature. *Cureus*. 2020. 12(9). 10599.
- [6] Medeiros, F. A., Xu, Y., Weinreb, R. N. Artificial intelligence in glaucoma: promises and challenges. *Current Opinion in Ophthalmology*. 2021. 32(2). 133-139.
- [7] Thakur N., Kaur H., Kaur P., Kumar A. Machine learning techniques in glaucoma diagnosis: a systematic review. *International Ophthalmology*. 2021. 41(3). 851-864.
- [8] Christopher M., Anantharaman G., Robin A. L. Artificial intelligence and glaucoma screening. *Journal of Glaucoma*. 2021. 30(1). 76-80.
- [9] Hu Y., Wang J., Xie P. Glaucoma Detection in Optical Coherence Tomography Images using Deep Learning Techniques: A Review. *Current Medical Imaging Reviews*. 2021. 17(1). 43-51.
- [10] Li S., Zhang Y., Liu X., Yang J., Zhang L. Automated glaucoma diagnosis using deep learning and fundus photographs: a systematic review and meta-analysis. *BMJ Open Ophthalmology*. 2020. 5(1), e000501.
- [11] Morales-Fernandez L., De La Rosa-Gómez, I., Rascón-Ramírez, F. J., Fernández-Vega Cueto, L. Artificial intelligence in the detection of glaucoma. *Archivos de la Sociedad Española de Oftalmología (English Edition)*. 2020. 95(10). 499-506.
- [12] Wu Z., Medeiros F. A. Artificial intelligence and deep learning for glaucoma: A review and perspective. *Eye and Vision*. 2020. 7(1). 23.
- [13] Fathi A., Yaqoob U., Zadeh E. K. Artificial intelligence in the diagnosis and management of glaucoma. *Journal of Ophthalmic and Vision Research*. 2020. 15(3). 306.
- [14] Asaoka R., Murata H., Iwase A. Recent advances in artificial intelligence technologies for glaucoma. *Expert Review of Ophthalmology*. 2020. 15(3). 125-134.
- [15] Zhang X., Huang S., Luo J. Artificial Intelligence in Glaucoma: Applications, Challenges and Future Prospects. *International Journal of Ophthalmology*. 2021. 14(2). 313-319.
- [16] García-Feijóo J., Sampaolesi J. R., Bartz-Schmidt K. U. Artificial Intelligence in Glaucoma. *Ophthalmology and Therapy*. 2020. 9(2). 187-190.
- [17] Kashyap R., Nair R., Gangadharan S. M. P., Botto-Tobar M., Farooq S., Rizwan A. Glaucoma Detection and Classification Using Improved U-Net Deep Learning Model. *Healthcare (Switzerland)*. 2020. 10(12), [2497]. <https://doi.org/10.3390/healthcare10122497>
- [18] Saedi O, Boland MV, D'Acunto L, Swamy R, Hegde V, Gupta S, Venjara A, Tsai J, Myers JS, Wellik SR, DeMoraes G, Pasquale LR, Shen LQ, Li Y, Elze T. Development and Comparison of Machine Learning Algorithms to Determine Visual Field Progression. *Transl Vis Sci Technol*. 2021 Jun 1;10(7):27. DOI: 10.1167/tvst.10.7.27. PMID: 34157101; PMCID: PMC8237084.
- [19] Mohammad Eslami, Julia A. Kim, Miao Zhang, Michael V. Boland, Mengyu Wang, Dolly S. Chang, Tobias Elze, Visual Field Prediction: Evaluating the Clinical Relevance of Deep Learning Models, *Ophthalmology Science*. 2023. 3(1). 100222, ISSN 2666-9145, <https://doi.org/10.1016/j.xops.2022.100222>.

## ГЛАУКОМАНЫ АНЫҚТАУДА МАШИНАЛЫҚ ОҚЫТУДЫ ПАЙДАЛАНУ

Павлов С.В.<sup>1</sup>, Мамырбаев О.Ж.<sup>2</sup>, Оралбекова Д.О.<sup>2,3</sup>, Момынжанова К.<sup>4\*</sup>

<sup>1</sup> Винница ұлттық техникалық университеті, Винница, Украина

<sup>2</sup> «Ақпараттық және есептеуіш технологиялар институты» ҒК ҚР ҒЖБМ, Алматы, Қазақстан

<sup>3</sup> «Гумарбек Даукеев атындағы Алматы энергетика және байланыс университеті» КЕАҚ, Алматы, Қазақстан

<sup>4</sup> «эль-Фараби атындағы Қазақ Ұлттық Университеті» КЕАҚ, Алматы, Қазақстан

**Аңдатпа.** Глаукома - бұл емделмеген жағдайда көру қабілетінің тұрақты жоғалуына немесе толық соқырлыққа әкелуі мүмкін прогрессивті көз ауруы. Глаукоманы ерте анықтау және емдеу көру қабілетінің жоғалуын болдырмау үшін өте маңызды. Дегенмен, глаукоманы диагностикалау және емдеу көп уақытты және қымбатқа түсетін ауқымды тестілеуді және тұрақты бақылауды қажет етеді. Глаукоманы диагностикалау - аурудың белгілерін анықтауда жоғары дәлдік пен тиімділікті талап ететін күрделі процесс. Соңғы жылдары глаукоманы диагностикалау саласында машиналық оқыту технологияларын қолдану танымал бола бастады. Бұл мақалада глаукоманы диагностикалау үшін машиналық оқыту алгоритмдерін қолданатын соңғы зерттеулерге, соның ішінде глаукоманы жіктеу және оның дамуына болжауға шолу жасалған. Мақалада глаукоманы диагностикалаудағы дәлдікті, тиімділікті және объективтілікті арттыру сияқты заманауи машиналық оқыту технологияларын қолданудың артықшылықтары көрсетілген. Ол сондай-ақ осы технологиялардың кейбір мәселелері мен шектеулерін сипаттайды және ықтимал шешімдерді ұсынады. Жалпы, мақала глаукоманы диагностикалаудағы терең оқыту технологияларының потенциалын және олардың пациенттер үшін нәтижелерін жақсартудағы рөлін көрсетеді.

**Кілттік сөздер:** машиналық оқыту, глаукома, көз аурулары, диагностика, прогрессияны болжау.

## APPLICATION OF MACHINE LEARNING TECHNOLOGIES IN THE DIAGNOSIS OF GLAUCOMA

**Pavlov S.V.<sup>1</sup>, Mamyrbayev O.<sup>2</sup>, Oralbekova D.<sup>2,3</sup>, Momynzhanova K.<sup>4\*</sup>**

<sup>1</sup>Vinnitsia National Technical University, Vinnitsia, Ukraine

<sup>2</sup>«Institute of information and computational technologies» MSHE RK, Almaty, Kazakhstan

<sup>3</sup>«Almaty University of Power Engineering and Telecommunication named Gumarbek Daukeev»  
Almaty, Kazakhstan

<sup>4</sup>«Al-Farabi Kazakh National University» Almaty, Kazakhstan

ORCID ID: <https://orcid.org/0000-0001-8318-3794>

ORCID ID: <https://orcid.org/0000-0003-3455-3319>

E-mail: [morkenj@mail.ru](mailto:morkenj@mail.ru), [dinaoral@mail.ru](mailto:dinaoral@mail.ru), [kymbat\\_momynzhanova87@mail.ru](mailto:kymbat_momynzhanova87@mail.ru)

**Abstract.** Glaucoma is a progressive eye disease that, if left untreated, can lead to permanent vision loss or total blindness. Early detection and treatment of glaucoma is critical to preventing vision loss. However, diagnosing and treating glaucoma requires extensive testing and regular monitoring, which can be time consuming and costly. Diagnosis of glaucoma is a complex process that requires high accuracy and efficiency in identifying the symptoms of the disease. In recent years, the application of machine learning technologies has become increasingly popular in the field of glaucoma diagnostics. This article provides an overview of recent research that has applied machine learning algorithms to diagnose glaucoma, including classifying glaucoma and predicting its progression. The article highlights the benefits of applying modern machine learning technologies, such as improving accuracy, efficiency and objectivity in the diagnosis of glaucoma. It also describes some of the problems and limitations of these technologies and suggests potential solutions. Overall, the article highlights the potential of deep learning technologies in diagnosing glaucoma and their role in improving patient outcomes.

**Key words:** machine learning, glaucoma, eye diseases, diagnostics, progression prediction.

*Сведения об авторах:*

*Рус: Павлов С.В. - доктор технических наук, профессор кафедры бимедицинской инженерии и оптико-электронных систем Винницкого Национального Технического Университета, Академик Международной академии прикладной радиоэлектроники, Винница, Украина*

*Қаз: Павлов С.В.- техника ғылымдарының докторы, Винница ұлттық техникалық университетінің бимедициналық инженерия және оптикалық-электрондық жүйелер кафедрасының профессоры, Халықаралық қолданбалы радиоэлектроника академиясының академигі, Винница, Украина*

Применение технологий машинного обучения при диагностике глаукомы

Павлов С.В., Мамырбаев О.Ж., Оралбекова Д.О., Момынжанова К.

*Англ: Pavlov S.V. - Doctor of Technical Sciences, Professor of the Department of Biomedical Engineering and Optical-Electronic Systems of Vinnytsia National Technical University, Academician of the International Academy of Applied Radioelectronics, Vinnitsia, Ukraine*

*Рус: Мамырбаев О.Ж – PhD, ассоциированный профессор (доцент) института информационных и вычислительных технологий, Алматы, Казахстан*

*Қаз: Мамырбаев О.Ж – PhD, Ақпараттық және есептеуіш технологиялары институтының қауымдастырылған профессоры, Алматы, Қазақстан*

*Англ: Mamyrbayev O.Zh – PhD, Associate Professor (Associate Professor) Institute of Information and Computational Technologies, Almaty, Kazakhstan*

*Рус: Оралбекова Д. О. – PhD, старший научный сотрудник института информационных и вычислительных технологий, Алматы, Казахстан*

*Қаз: Оралбекова Д.О. – PhD, Ақпараттық және есептеуіш технологиялар институтының аға ғылыми қызметкері, Алматы, Қазақстан*

*Англ: Oralbekova D. O. – PhD, senior researcher at the Institute of Information and Computational Technologies, Almaty, Kazakhstan*

*Рус: Момынжанова К.Р. – PhD докторант по специальности 8D06106-Информационные системы, Казахский национальный университет имени аль-Фараби, Алматы, Казахстан*

*Қаз: Момынжанова Қ.Р. – 8D06106-Ақпараттық жүйелер мамандығы бойынша PhD докторанты, әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан*

*Англ: Momynzhanova K.R. – PhD student in specialty 8D06106-Information systems, Al-Farabi Kazakh National University, Almaty, Kazakhstan*

UDC 004.02  
IRSTI 20.53.15

## ARTIFICIAL INTELLIGENCE FOR COMPLEXITY THEORY

Mirzakhmet Syzdykov

Satbayev University, Almaty, Kazakhstan

[mshpmail598@gmail.com](mailto:mshpmail598@gmail.com)

ORCID ID: <https://orcid.org/0000-0002-8086-775X>

**Abstract.** In this continued series of work, we present the theoretical and practical results towards reasoning with modern methods of Artificial Intelligence (AI). We justify our methodology with help of illustrative examples from Computer Science relying on the regular expression matching algorithm and application of the proposed solution for the task of identifying files consistency according to the unknown format. We will also give several notable proofs to the classical theorems which in some sense are coherent to the terms like AI and algorithmic complexity, however, or at least, nowadays they're solved involving the huge amount of hardware resources and together constitute the new formation in the modern age with help of specifically crafter hardware modules – we're still about to represent the model in more classical understanding from the point of view of computational complexity, concise reasoning and computer logic within the classical models, theorems and proofs as the base approach of estimating the costs needed to build Artificial Neural Networks (ANN) or Machine Learning (ML) data.

**Keywords:** artificial intelligence, parsing, algorithm.

### Introduction

We are giving clear statement that the prevalence of artificial intelligence technologies like ML has gained final success in applied sciences like medicine or Computer Vision (CV) [1].

Since these times, it continues to grow up in the field of application and extension to the real-life circumstances.

However, we define the question of if artificial intelligence has a consciousness and how it'd change the state of the matter if this would be completely proved and researched. On this occasion still it's necessary to estimate that network of neural nodes cannot produce the clear conscious mind as human being can be. We address this question for further investigation from psychological point of view with the main focus on cognitive abilities of the artificial intelligence which cannot be achieved by simply simulating the neural network.

Artificial Intelligence, or simply AI, and its consciousness is a general question of the modern Computer Science, which is highlighted in the press by many researchers [2, 3]. We are to answer the question whether it's important and whether it's possible. Of course, we give our argument towards the fact that psychology and self "I" of any mind cannot be followed from the chaos produced by the neural network, thus, all the arguments made towards the fact that AI can be conscious based upon latest knowledge and technology cannot be addressed to the main point of view as per analogy of the human "I" and his *modus operandi*.

The data volume as a starting point of view are estimated as very big in gigabytes of pure textual data in order to train the neural network. This is a very challenging task as gaining such big amount of data and successful storage of the trained neural network means the decision to take care of Big-Data hardware. Thus, Big Data for AI could be a good trend nowadays leaving the hope of free services around the globe in order to be cheap and safe and generally free of charge.

Artificial Intelligence and its consciousness don't adhere to the economic stability and grow as in the modern time the computer programmer profession will be replaced by clever AI. This is tied not only to programmers but also to other list of professions which will simply vanish due to the process of globalization of AI.

There's another point of view which is meant to be the artificial intelligence for the not good intentions. Of course, there's a way to train the neural network against the prohibited type of data. Thus, if we, for example, would train the network against the binary codes of executables and their source codes, this will lead to the reverse engineering which is a prohibited method of obtaining source codes of programs from their packaged content. The same malicious method can be applied to password guess and other tasks

which involve the example to be seeded from the large amount of sample data.

ChatGPT and other modern trends of artificial intelligence are already gone to practice due to the need of automated assistant in business processes. This could be a question addressed to economics, however, still due to the prevalence of the automated solutions there's no need of human interaction, thus, the proficiency and artificial intelligence remains an open question in labor market.

The clear consciousness of the artificial intelligence, we, thus, address to the psychological, psychical and physiological processes of the brain, when human neurons transfer the signal from one point on the brain to another. As this is already done in artificial networks, we still cannot conclude that artificial neural network can replace the human brain functionality due to the presence of the term psychiatry – we call it a Freud conjecture, when still trained data model cannot acquire the self-“I” – thus, it can be modelled as a separate kind of entity available on the control environment like hardware and programming languages to be operated.

We also give the notion to the rapidly growing interest of variety of communities for Machine Learning. We give the definition of algorithm of Explainable Machine Learning in order to define the extensible role of the ML in the field of algorithmic approach rather than resorting to the classical approach of neural networks.

### O-notation and equivalence of complexity classes

Starting from this point we define the complexity from big-O notation which can be relative as to the fact which is outlined to be a definitive argument towards AI operability and algorithms' complexity. In fact, they're linear, however, the big amount of data and slow performance make their choice towards choosing more effective solutions.

In this section we define the relation between complexity classes like equivalence, less or greater as for the two compared algorithms:

$$O(f(n)) \sim O(g(n)). \quad (1)$$

This can be achieved during the maximal load of both algorithms measured by functions in (1). This maximum value can be achieved only by the limit in the following form:

$$r[f(n), g(n)] = \lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = \lim_{n \rightarrow \infty} \frac{f'(n)}{g'(n)} = \begin{cases} 0, O(f(n)) < O(g(n)) \\ C, C > 0, O(f(n)) = O(g(n)). \\ \infty, O(f(n)) > O(g(n)) \end{cases} \quad (2)$$

Here in (2),  $C$  is a constant value and is valid if only both complexity classes are equal, which can be derived by derivative rule of any degree.

Thus, we simplify our relation to the following form:

$$r(f, g) = \begin{cases} 0 \\ C, C > 0. \\ \infty \end{cases} \quad (3)$$

The derivative rule is also in common can be evaluated as follows:

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = \lim_{n \rightarrow \infty} \frac{f'(n)}{g'(n)} = \lim_{n \rightarrow \infty} \frac{d^k[f(n)]}{d^k[g(n)]} \quad (4)$$

Algorithms and definition of their complexity can be found in [4]. Further in this article we will show the strict proof of the complexity classes P and NP using non-trivial occasion which is defined in comparative work [5].

### NP-completeness and hash functions

Here we give the proof of the existence of hash function which can be defined by a hash of the fixed length, for example, as it's defined in MD5 to be 128 bits.

Since for the input, there're only fixed number of possible ways of giving the collision to the one-way function, we define the probability of collision as follows:

$$p(n) = \frac{2^{128}}{2^n} = 2^{128-n}. \quad (5)$$

As the one-way function is defined on all the input space the probability (5) will be measured towards common input as follows:

$$\lim_{n \rightarrow \infty} p(n) = 0 \Rightarrow \exists f(x). \quad (6)$$

Where in (6)  $f(x)$  is a one-way function which exists since the probability of collision is defined on the whole set of the input and randomly defined function will never collide. The latter is true for non-random functions if they're relevant to the common case of dependence from the input size given by the parameter  $n$  in (5) and (6).

On account for "P versus NP" theorem we get the following from (4):

$$O(n!) = O(n^k). \quad (7)$$

The (7) holds true as we can see further:

$$\lim_{n \rightarrow \infty} \frac{n^k}{n!} = \lim_{\substack{n \rightarrow \infty \\ k \rightarrow \infty}} \frac{d^k(n^k)}{d^k(n!)} = \lim_{\substack{n \rightarrow \infty \\ k=n}} \frac{n!}{n!} = 1. \quad (8)$$

Thus, from (8) it follows that P equals NP.

### Smart parsing and position automata

The position automata are deterministic finite automata with the definition of the position as an additional parameter in follow function for the state, position and label of the outgoing edge.

The parsing within the AI is defined as a way of complementing of any of the trained models to the input data which can be formatted with help of position automata.

We can also conclude that position automata are reminiscent to the state explosion effect. The smart parsing, thus, gives the possibility of the trained data to be applied for the identification of the common format.

The task above can be also defined by the alternating regular expression with the definition of all the combinations of subgroups given by brackets in the final pattern string. Thus, at each step of the algorithm we collect the matched input file towards the pre-defined pattern with help of neural network which records each generation of learning for each of the input files.

The brackets can be formed as a full combination of valid patterns which together give the entire template of the input in the pre-defined file format, while the other input strings can be matched against sparse tree with the assumption that ending leaves of this tree are closed under &-operator artificial state.

The construction above gives the possibility of recognizing the common format of files grouped by their extension.

### Conclusion

The questions of artificial intelligence and its consciousness as well as the global impact of it on the natural processes like economics and labor market are also given followed by the research practice and knowledge.

Thus, we devise more concept from the term Artificial General Intelligence (AGI) on the very urgent question of the human labor which will be replaced in the near future by AGI according to the predictions from opinions raised in press and which wasn't raised in science. The ethical question beyond this fact is to provide the AGI with the consciousness model so that it could be addressed to the worker's burden. This ethical question isn't limited to the loss in labor market and the underlying conditions where AGI can be used for malicious intentions – that's still the question of advocacy and policy as well.

We have also given the answer to the P and NP in limit spaces as it follows that they are equal in



general.

We have also proposed the algorithm based upon artificial states in automata to identify the input file format based upon the trained amount of data.

### Acknowledgements

The author expresses gratitude to all the colleagues from past to present.

### References

- [1] Hamet Pavel, Johanne Tremblay. Artificial intelligence in medicine. Metabolism. 2017. 69. 36-40.
- [2] Coller Kevin. What is consciousness? ChatGPT and advanced AI might redefine our answer. <https://www.nbcnews.com/tech/tech-news/chatgpt-ai-consciousness-rcna71777> [accessed 3.30.2023].
- [3] Hildt Elisabeth Artificial Intelligence: Does Consciousness Matter? <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01535/full> [accessed 3.30.2023].
- [4] Cormen T. H., Leiserson C. E., Rivest R. L., Stein C. Introduction to algorithms. MIT press. 2022
- [5] Syzdykov Mirzakhmet. Equivalence of Complexity Classes via Finite Automata Derivatives. ADVANCED TECHNOLOGIES AND COMPUTER SCIENCE. 2022. 4. 9-14.

## ЕСЕПТЕУ ТЕОРИЯСЫНА АРНАЛҒАН ЖАСАНДЫ ИНТЕЛЛЕКТ

Сыздықов Мирзахмет

Қ. И. Сәтпаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы, Қазақстан  
[msspmail598@gmail.com](mailto:msspmail598@gmail.com)  
ORCID ID: <https://orcid.org/0000-0002-8086-775X>

**Андатпа.** Осы жалғасты жұмыс сериясында біз жасанды интеллекттің (ЖИ) заманауи әдістерімен ойлаудың теориялық және практикалық нәтижелерін ұсынамыз. Біз өз әдістемемізді информатикадан көрнекі мысалдар көмегімен тұрақты өрнектерді сәйкестендіру алгоритміне және белгісіз пішімге сәйкес файлдардың сәйкестігін анықтау тапсырмасы үшін ұсынылған шешімді қолдануға негіздейміз. Біз сондай-ақ белгілі бір мағынада ЖИ және алгоритмдік күрделілік сияқты терминдерге сәйкес келетін классикалық теоремаларға бірнеше маңызды дәлелдер келтіреміз, бірақ, немесе, кем дегенде, қазіргі уақытта олар аппараттық ресурстардың үлкен көлемін ескере отырып шешіледі және бірге жаңа формацияны құрайды. қазіргі заманда арнайы құрастырылған аппараттық модульдердің көмегімен – біз әлі де классикалық модельдер, теоремалар мен дәлелдер негізінде есептеу күрделілігі, қысқаша пайымдау және компьютерлік логика тұрғысынан модельді классикалық түсінуде көрсетеміз. Жасанды нейрондық желілер (ЖНЖ) немесе Machine Learning (ML) деректерін құруға қажетті шығындарды бағалау тәсілі.

**Кілттік сөздер:** жасанды интеллект, талдау, алгоритм.

## ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ ДЛЯ ТЕОРИИ СЛОЖНОСТИ

Сыздықов Мирзахмет

Казахский Национальный Исследовательский Университет им. К. И. Сатпаева  
Алматы, Казахстан  
[msspmail598@gmail.com](mailto:msspmail598@gmail.com)  
ORCID ID: <https://orcid.org/0000-0002-8086-775X>

**Аннотация.** В этой продолжающейся серии работ мы представляем теоретические и практические результаты рассуждений с использованием современных методов искусственного интеллекта (ИИ). Мы обосновываем нашу методологию наглядными примерами из области компьютерных наук, опираясь на алгоритм сопоставления регулярных выражений и применение предложенного решения задачи определения непротиворечивости файлов по неизвестному формату. Мы также приведем несколько примечательных доказательств классических теорем,

которые в некотором смысле согласуются с такими терминами, как ИИ и алгоритмическая сложность, однако, по крайней мере, в настоящее время они решаются с использованием огромного количества аппаратных ресурсов и вместе составляют новую формацию. в современную эпоху с помощью специально созданных аппаратных модулей – мы все еще собираемся представить модель в более классическом понимании с точки зрения вычислительной сложности, лаконичных рассуждений и компьютерной логики в рамках классических моделей, теорем и доказательств в качестве основы подход к оценке затрат, необходимых для создания данных искусственных нейронных сетей (ИНС) или машинного обучения (МО).

**Ключевые слова:** искусственный интеллект, синтаксический анализ, алгоритм.

*Сведение об авторе:*

*Анг.: Syzdykov Mirzakhmet - Satbayev University, Almaty, Kazakhstan*

*Каз.: Сыздықов Мырзахмет- Қ.И. Сәтпаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы, Қазақстан.*

*Рус.: Сыздыков Мырзахмет- Казахский национальный исследовательский технический университет имени К.И. Сатпаева, Алматы, Казахстан.*

UDC 004.42  
IRSTI 81.93.29

## METHOD OF CONSTRUCTING MODELS OF INFORMATION ATTACKS

Barakova A.Sh., Ussatova O.A.

Institute of Information and Computational Technologies, Almaty, Kazakhstan,  
Al-Farabi Kazakh National University, Almaty, Kazakhstan

balia\_79@mail.ru

uoa\_olga@mail.ru

ORCID ID: <https://orcid.org/0000-0002-5276-6118>

ORCID ID: <https://orcid.org/0000-0002-0904-745X>

**Abstract.** The existing approaches to the construction of formal models of information attacks are considered and analyzed. The basic requirements for information attack models are formulated. A method is proposed for constructing models of information attacks based on control E-network transitions and multi-agent management, taking into account the formulated requirements. The considered approaches to modeling allow describing the process of an information attack with varying degrees of detail. The models use different mathematical bases, but most of them are based on finite automata and represent the attack as a sequence of automaton states. None of the models makes it possible to take into account the vulnerability activated by such a complex, the method of its implementation and possible consequences. The models allow us to more accurately determine the effectiveness of existing and developed means of protection against simulated information attacks.

**Keywords:** attack, modeling, information attack model, managing E-network, agent, multi-agent management.

### Introduction

Currently, one of the most relevant areas of scientific research in the field of information security is the development of methods and tools for detecting attacks and protecting against attacks on computer systems and networks. In the process of such development, it is necessary to constantly conduct research, including preliminary study and detailed study of possible options for implementing information attacks. As a rule, these works are carried out using models that allow reproducing the necessary properties and characteristics information attack, as well as to assess the level of its danger to the computer system (CS). The models allow us to more accurately determine the effectiveness of existing

The currently created models of information attacks can be classified according to the following basic criteria [1]:

- the possibility of expanding the model (models with a fixed number of parameters and their values, extensible models with the possibility of adding new parameters and their values);
- the possibility of taking into account the sequence of actions performed in an information attack in the model;
- the level of detail of the model (models of low, medium and high level of detail). and developed means of protection against simulated information attacks.

A formalized information attack model based on attack trees developed by B. Schneier has become widespread [2]. Attack trees are conceptual diagrams that describe threats to the system and possible

attacks aimed at their implementation. The underlying construction here is a hierarchical tree  $G = (L, E)$ , where

$\{l_i\}_{i \in L}$  is the set of tree vertices,

$\{e_s\}_{s \in E}$ ,

$2 E \subset L$  – the set of tree arcs

Each vertex of the tree  $G$  is associated with a certain action of the violator, while the root of the tree denotes the ultimate goal of an information attack, the implementation of which can cause significant damage to  $A C$ . Thus, it is possible to make up a set of possible paths on the graph  $G$   $\{ \} i G_p = g_p$ , where each path  $i g_p$  is a sequence of arcs  $( , , \dots, 1 2 n e e e$  of the form  $e_i = ( . li,lj), li,l j \in L$  In this case, the final vertex of the arc  $i l$  is simultaneously the initial vertex of the arc  $l_{i+1}$ .

The leaves of the tree  $G$  can act as the initial vertex of the path, and the root of the tree can act as the final vertex. The Schneier attack tree model has several important advantages and disadvantages

Advantages	Disadvantages
<ul style="list-style-type: none"> <li>– the model can be used to describe attacks on any information or physical systems;</li> <li>– due to the presence of numerical values at the vertices and edges, the model provides an opportunity to introduce estimates of each step according to certain criteria, for example, by execution time, number of operations, estimated cost, etc. The sequence of steps can be evaluated based on the criteria of each step;</li> <li>– the expansion of the attack model is carried out by adding new elements to the sets of vertices and arcs of trees describing the attack;</li> <li>– it is possible to simulate complex information attacks involving several implementation options.</li> </ul>	<p>This model has a number of disadvantages, which should include:</p> <ul style="list-style-type: none"> <li>– a tree is used as a fundamental structure for modeling an attack, which imposes some restrictions on the structure of the graphical representation of the model. The presence of cycles also creates certain difficulties when working with this model;</li> <li>– the attack model does not include the parameters of the AU environment, under which the implementation of the simulated attack is possible;</li> <li>– there are no tools in the model that provide dynamic modeling.</li> </ul>

In [3], a formal method of attack modeling is proposed, which is an extension and refinement of the model based on the attack tree. Two attributes are introduced: the lifetime (reflects the time dependencies between the stages of the attack) and the degree of confidence (characterizes the probability of reaching the target of the attack with the achieved sub-targets). However, this model has the same disadvantages as those based on the attack tree. The attack graph model [4] is also based on an extension of the attack tree model. Attack graphs are a specialized tool for describing an attack. Graph nodes do not represent conceptual actions, but network nodes, program processes, configuration files, code sections, etc. The model is widely used because it is based on a simple and well-researched mathematical base - finite automata, itself is quite simple and obvious. Transitions between nodes are carried out on based on deterministic rules, the current value of some system parameters, variables, etc. may be taken into account. Existing graph models are well-suited for describing the sequence of actions of an attacker and are often used to assess the complexity of an information system security breach, rather than for modeling and investigating attacks. The disadvantages of these models can also be attributed to the fact that they do not contain

mechanisms for organizing controlled branching and modeling the dynamic component of the attack.

Information attack models are also built on the basis of formal languages and ontologies [5]. Formally, the ontology is a tree, and the threat is represented by a sequence of symbols. Sequences are considered as words of a formal language specified by means of some formal grammar. The description of the generalized attack scenario by means of stochastic grammar has the following form

$$GA = (V_N, V_T, S, P) \tag{1}$$

where  $V_N$  is a set of non-terminal symbols that denote micro-level attack steps,  $V_T$  is a set of terminal symbols that correspond to the upper and intermediate levels of representation of the attack scenario steps,  $S$  is the initial symbol of the attack scenario,  $P$  is a set of output rules describing the operations of detailing the attack scenario by replacing characters. Each replacement is carried out with a given probability:

$$\alpha_i \rightarrow \beta_{ij}, i = 1, \dots, n, j = 1, \dots, m, \tag{2}$$

where  $\alpha_i$  is a nonterminal character,  $\beta_{ij}$  is a string of terminal and nonterminal characters

It is characteristic that the conceptual model of information attack based on formal languages and ontologies is in many ways similar to the model discussed above, developed by B. Schneier. The model is also represented as a graph structure and can be expanded by adding new elements to the sets of terminal and non-terminal symbols, as well as by expanding the output rules. The model can be presented both in text and graphical form. The main disadvantage of such an information attack model is the absence of a parameter that characterizes the vulnerability of the CS, on the basis of which the violator's action is performed.

### Requirements for information attack models

For effective use in order to study the likely actions of the violator in relation to the CS, the information attack model being developed must meet the following minimum set of requirements:

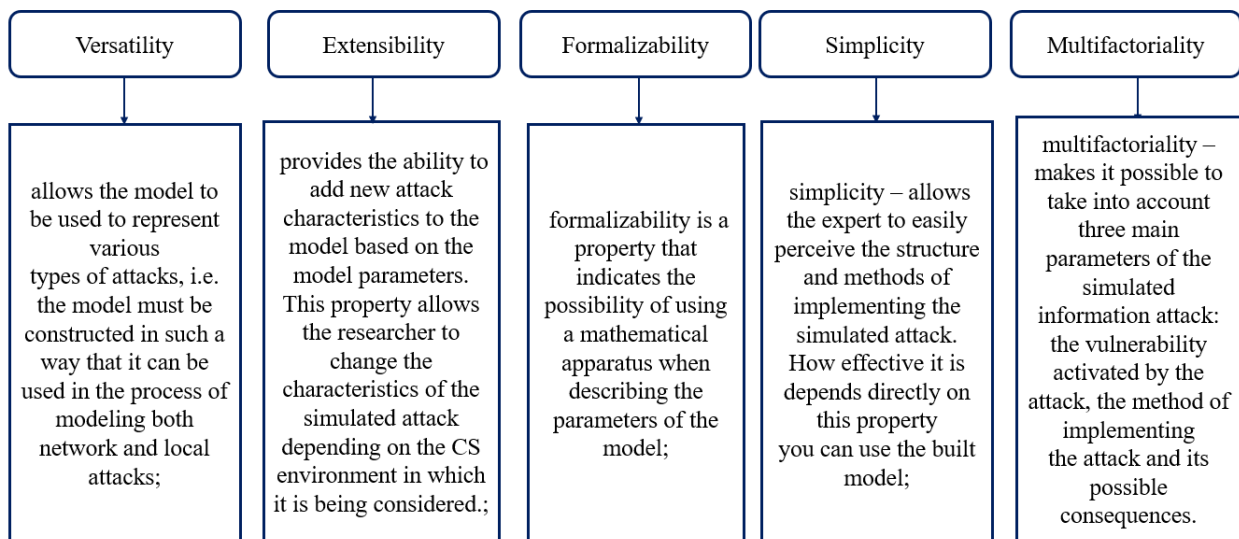


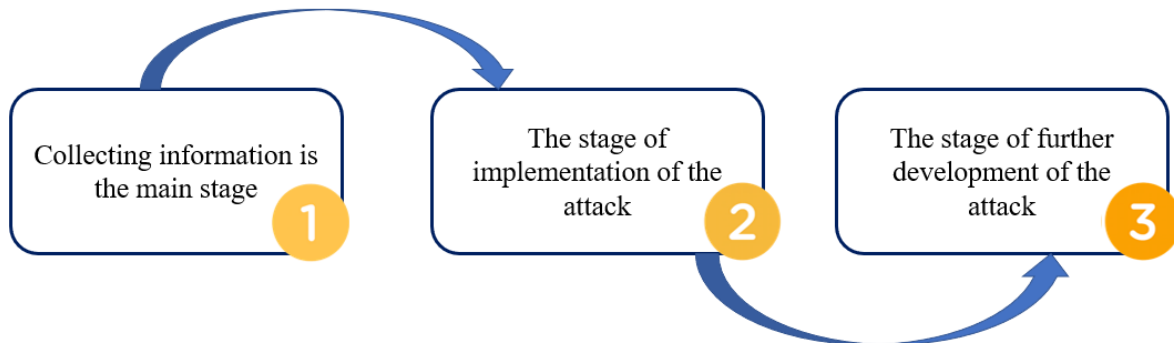
Figure 1 - Requirements for information attack models

### Information attack as a modeling object

An attack on an information system is an action or sequence of related actions of an intruder that lead to the realization of a threat by exploiting system vulnerabilities. Vulnerability is

commonly understood as the weak point of the CS, on the basis of which the successful implementation of the threat is possible. In turn, a threat is a potentially possible event, action, phenomenon or process that can cause damage to the system resource [6]. Thus, in order to implement the attack, the attacker simulates some action that leads to the desired result with the help of some means that exploits the vulnerabilities of the system.

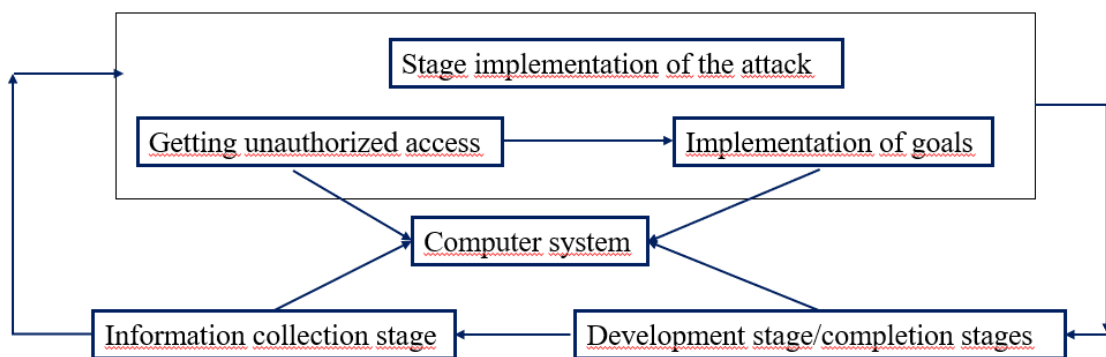
In general, an information attack can consist of three stages:



1. Collecting information is the main stage. At this stage, the target of the attack is selected, information about it is collected (OS, configuration, services), the most vulnerable places of the attacked system are identified, the impact on which leads to the desired result, the type of attack being implemented is selected.

2. The stage of implementation of the attack. At this stage, the violator gets unauthorized access (ND) to the resources of those CS nodes in relation to which the attack is carried out. If, by the nature of the impact, the attack is active [7], then this stage is also accompanied by the realization of the goals for which the attack was undertaken. The result of such actions may be a violation of confidentiality, integrity and availability of information. In addition, at this stage, the source and the fact of the attack may be hidden, the so-called "covering up traces".

3. The stage of further development of the attack – actions are performed that are aimed at continuing the attack on the resources of other CS nodes. In the case of passive attacks [7], this stage is the stage of completing the attack.



**Figure 2** - schematically shows the stages of the life cycle of a typical information attack.

When modeling an information attack, it is necessary to determine its parameters and characteristics. The main parameters of the attack are:

- duration of action;
- multiplicity;
- a list of vulnerabilities used by the attack.

Note that modeling some types of attacks, for example, "distributed denial of service" – DDoS requires

the definition of additional attack parameters, such as [8]:

- type of attack (semantic attack (TCP SYN, Incorrect packets, Hard request, etc.) or "brute force" type attack – such as UDP/ICMP flood, smurf/fraggle, etc.);
- rate of attack (can be constant or variable. In the latter case, the intensity of the attack varies over time. The change in tempo can be increasing or fluctuating);
- influence on the target of the attack (you can choose a "disruptive" attack, when a distributed attack will be carried out from all attacking nodes at once, or a worsening one – the number of attacking nodes is included in the attack gradually. The first attack option is easier to detect);
- the constancy of the set of attacking nodes (the set can be constant (the same nodes attack) or variable);
- the degree of automation (automatic or automated).

### **The proposed method of modeling attacks**

Taking into account the disadvantages of existing analogues, we will use Control E-Nets (Control E-Nets – CEN) as an information attack model [9, 10], which are a modification of E-networks and meet all the requirements for the attack models described above. Issues of the complexity of the attack model (taking into account the vulnerability search stage in attack models, methods of implementation and development of the attack) will be solved using a multi-agent control system, the modules of which work on the principles of intelligent agents. We will describe each agent using the implementation model presented in viRis. 1. The life cycle of a typical information attack on CS resources de managing E-network. We will simulate attacks in a specially designed environment, which is based on the principles of invariance to the subject area, the level of modeling, the experiment being conducted and the level of user readiness. Let's consider in more detail the mechanisms of operation of the control E-networks, systems multi-agent management and information attack simulation environment.

### **Conclusions**

A method of constructing information attacks based on control E-networks and multi-agent management is proposed. Using this method allows you to create models based on simple, easily perceived constructions to represent various types of attacks. Thanks to the use of the mechanism of control E-networks, the method allows you to take into account the current state of the CS and the object of attack during modeling – dynamic modeling is provided.

In addition, the control E-networks allow the use of mathematical apparatus when describing the parameters of the model – the property is provided formalizability of the model. The use of a multi-agent approach makes it possible to solve the problem of multifactorial (complexity) attacks, highlighting the stage of finding vulnerabilities of attack objects (planning level), the stages of implementing the attack and determining its possible consequences (reactive level); there is a cooperative level for the exchange of information between agents of different levels. The obtained simulation models of information attacks can be used to build a synthetic environment of information security systems in order to clarify their features and characteristics using the method of semi-natural modeling.

### **References**

- [1] Serdyuk V. A. Analysis of modern trends in building models of information attacks. Information technology. 2004. 5. 94-101.
- [2] Schneier B. Attack Trees [Electronic resource] Availability mode: <http://www.schneier.com/paper-attacktrees-ddj-ft.html>.
- [3] Campete S. A formal Method for Attack Modeling and Detection: Availability mode: <http://citeseer.ist.psu.edu/751069.html> [Electronic resource]

- [4] Shener O. Automated Generation and Analysis of Attack Graphs. USA, 2002. 273 – 284.
- [5] Gorodetski V. Attacks against Computer Network: Formal Grammer-based Framework and SimulationTool. Proceedings of the international RAID conference. St. Petersburg, 2002. 219-238.
- [6] Lukatsky A.V. Detection of attacks. St. Petersburg: VVH-Petersburg, 2001. 624.
- [7] Natrov V. V. Classification of network attacks. Information technologies in Management and modeling: SAT. Dokl. Belgorod, 2005. 128-132.
- [8] Kotenko I. V. Modeling the resistance of software agents on the Internet: a general approach, modeling environment and experiments. Information protection. INSIDE. 2006. 5. 2-10.
- [9] Kazimir V. V. modeling of synthetic environment for reactive systems. Mathematical modeling. 2003. 2 (10). 24-32.
- [10] Kazimir V. V. Model-oriented management of intellectual production systems: Dissertation of the Doctor of Technical Sciences: 05.13.06. 2006. 301.

## АҚПАРАТТЫҚ ШАБУЫЛ МОДЕЛЬДЕРІН ҚҰРУ ӘДІСІ Баракова А. Ш., Усатова О. А.

Ақпараттық және есептеу технологиялары институты, Алматы, Қазақстан,  
Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан  
*balia\_79@mail.ru*  
*uoa\_olga@mail.ru*

ORCID ID: <https://orcid.org/0000-0002-5276-6118>

ORCID ID: <https://orcid.org/0000-0002-0904-745X>

**Андатпа.** Ақпараттық шабуылдардың формальды модельдерін құрудың қолданыстағы тәсілдері қарастырылып, талданды. Ақпараттық шабуыл модельдеріне қойылатын негізгі талаптар тұжырымдалған. Электрондық желінің бақылау ауысуларына және тұжырымдалған талаптарды ескере отырып, мультиагентті басқаруға негізделген ақпараттық шабуыл модельдерін құру әдісі ұсынылды. Қарастырылған модельдеу тәсілдері ақпараттық шабуыл процесін әртүрлі дәрежеде егжей-тегжейлі сипаттауға мүмкіндік береді. Модельдер әртүрлі математикалық негіздерді қолданады, бірақ олардың көпшілігі ақырғы машиналарға негізделген және шабуылды А ретінде ұсынадыавтомат күйлерінің реттілігі. Модельдердің ешқайсысы мұндай кешен арқылы іске асырылатын осалдықты, оны жүзеге асыру тәсілін және ықтимал салдарын ескеруге мүмкіндік бермейді. Модельдер қолданыстағы тиімділікті дәлірек анықтауға мүмкіндік береді және имитациялық ақпараттық шабуылдардан қорғау құралдары

**Кілттік сөздер:** шабуыл, модельдеу, ақпараттық шабуыл моделі, электрондық желіні басқару, агент, мультиагентті басқару.

## СПОСОБ ПОСТРОЕНИЯ МОДЕЛЕЙ ИНФОРМАЦИОННЫХ АТАК Баракова А.Ш., Усатова О.А.

*Институт информационных и вычислительных технологий, Алматы, Казахстан,  
Казахский Национальный университет имени Аль-Фараби, Алматы, Казахстан*  
*balia\_79@mail.ru*  
*uoa\_olga@mail.ru*

ORCID ID: <https://orcid.org/0000-0002-5276-6118>

ORCID ID: <https://orcid.org/0000-0002-0904-745X>

**Аннотация.** Рассмотрены и проанализированы существующие подходы к построению формальных моделей информационных атак. Сформулированы основные требования к моделям информационных атак. Предложен метод построения моделей информационных атак, основанный на контрольных переходах электронной сети и мультиагентном управлении, с учетом сформулированных требований. Рассмотренные подходы к моделированию позволяют описать



процесс информационной атаки с разной степенью детализации. Модели используют различные математические основы, но большинство из них основаны на конечных автоматах и представляют атаку как последовательность состояний автомата. Ни одна из моделей не позволяет учесть уязвимость, активируемую таким комплексом, способ ее реализации и возможные последствия. Модели позволяют нам более точно определять эффективность существующих и разрабатываемых средств защиты от имитируемых информационных атак.

**Ключевые слова:** атака, моделирование, модель информационной атаки, управляющая электронная сеть, агент, мультиагентное управление.

*Сведения об авторах:*

*Англ: Olga Aleksandrovna Ussatova, PhD, Institute of Information and Computational Technologies, Almaty, Kazakhstan*

*Каз: Усатова Ольга Александрқызы, философия докторы, Ақпараттық және есептеуіш технологиялар институты, Алматы, Қазақстан*

*Рус: Усатова Ольга Александровна, доктор философии, Институт информационных и вычислительных технологий, Алматы, Казахстан*

*Англ: Aliya Sharizatovna Barakova, doctoral student, Al-Farabi Kazakh National University, Almaty, Kazakhstan*

*Каз: Баракова Алия Шаризатқызы, докторант, әл-Фараби атындағы ҚазҰУ, Алматы, Қазақстан*

*Рус: Баракова Алия Шаризатовна, докторант, Казахский национальный университет им. аль-Фараби, Алматы, Казахстан.*

**Responsible for the release: PhD, Shayakhmetova A.S.  
Merkebaev A.**

**Deputy chief editor: PhD, Mamyrbayev O.Zh**

The editorial board of the journal " Advanced technologies and computer science " is not responsible for the content of published articles. The content of the articles belongs entirely to the authors and is posted in the journal solely under their responsibility.

Signed in print 03.03.2020  
Edition of 50 copies. Format 60x84 1/16. Paper type.  
Order No. 4.

Publication of the Institute of Information and Computational Technologies

28 Shevchenko str., Almaty, Republic of Kazakhstan  
7 (727) 272-37-11  
atcs@iict.kz