# Thematic areas for the Master State Examination in elective courses

## Data Analysis and Processing

1. Data types, data preprocessing, data features. Feature Selection (principle, Entropy, Gini index).
2. Frequent pattern mining (basic principles, methods, variants, implementation).
3. Clustering algorithms (representative-based clustering, hierarchical clustering). Density-based clustering, cluster validation), advanced clustering methods (CLARANS, BIRCH, CURE).
4. Decision Trees (principle, algorithm, metrics used for proper splitting attribute selection, pruning).
5. Probability classification (Bayes theorem, Naive Bayes theorem).
6. Support Vector Machines (principle, algorithm, kernel trick).
7. Neural networks (basic principle, learning methods, activation function).
8. Assessment of classification algorithms (error-rate, precision, recall, F-measure).
9. Regression (linear and non-linear regression, regression trees, model quality assessment).
10. Network types. Graph and adjacency matrix as network representation. Data structures for representation of different types of networks, advantages and disadvantages (adjacency matrix, adjacency list, and adjacency tree), operation complexity, hybrid representation.
11. Topological properties of networks, characteristic values and their distribution (degree, path length, radius, clustering coefficient), centrality types.
12. Global properties of networks (small world, scale-freeness, growing and preferential attachment). Power law and its interpretation in real-world network environment. Assortativity.
13. Network models and their properties (Edrös-Rényi, Watts-Strogatz, Barabasi-Albert).
14. Communities. Global and local approaches. Modularity.
15. Advanced network models – multilayer networks, community based models, temporal networks.
16. Resilience of networks, propagation of phenomena in networks. Propagation and maximization of influence in networks. Prediction of links. Sampling.
17. Multilayer networks, their types and representations. Methods of analysis and visualization of multilayer networks, projection, flattening.
18. Local and global properties of multilayer networks, centrality types and random walks. Methods of community detection in multilayer networks.
19. Algorithm for pattern matching on strings (one/multiple sample searching, regular expression searching, and approximate search).
20. Information retrieval (IR) (models – Boolean, vector, lexical analysis, stemming and lemmatization, stop-words, index construction, query evaluation, precision, relevancy, recall, F-measure).
21. Linear algebra in IR (dimensionality reduction methods, matrix factorization, latent semantics, hypertext documents analysis, PageRank).
22. Neural networks text processing (word embedding, text classification, text generation, …).
23. Describe the architecture of convolutional neural networks, layers used, principle of operation, basic types of architectures.
24. Describe the architecture of recurrent neural networks, types of neurons, principle of operation.

Topics are covered by these courses: Machine Learning, Network Science I, Network Science II, Methods of Analysis of Textual Data, Deep Learning.

## Database Systems

1. Relational data model, the SQL language, a procedural extension of SQL; function dependencies, normal forms.
2. Transactions and isolation levels in SQL, recovery, ACID, concurrency control: locking, multi-versioning.
3. Physical implementation of database systems; tables (the heap table, clustered table, hash table) and indices (the B-tree, bitmap index), materialized views, data partitioning.
4. Paging of the query result, compression of tables and indices, column and row storages.
5. Query execution plan, logical and physical operations, random and sequential accesses, tuning of the query processing.
6. Algorithms of the join operation.
7. Physical implementation of data structures and algorithms of the query processing; optimization of accesses to the main memory and disk; design and implementation of the cache buffer.
8. Operators of the query execution plan, statistics of data in DBMS, cost-based optimization.
9. CAP theorem, NoSQL DBMS, BASE, replication, MongoDB, CRUD operations.
10. Multidimensional data structures, handling multidimensional data in DBMS.

These topics cover the following courses: Advanced Database Systems, Physical Implementation of Database Systems, Algorithms of Query Evaluation in Database Systems.

## Computer graphics and image analysis

1. Color systems in computer graphics, nonlinearity of graphical output (gamma correction), composition of raster images (alpha channel), HDR.
2. Affine and projective space. Affine and projective transformations and their mathematical notation. Modeling and imaging transformations in computer graphics.
3. Methods of photorealistic images synthesis, recursive ray tracing, radiometry, rendering equation, Monte Carlo approaches in lighting calculation, acceleration methods.
4. Graphics pipeline and implementation of its individual steps, illumination models and shading algorithms, visibility (hidden surface removal), real-time global illumination techniques, brief characteristics of the OpenGL standard.
5. Image and video compression, principles of image manipulation in spatial and frequency domain.
6. Basic methods of image modification and segmentation (filtering, thresholding, edges, areas, corners).
7. Basic methods of object recognition, feature recognition. Universal feature descriptors (e.g. HOG), trainable classifiers (e.g. SVM).
8. Deep neural networks (e.g. convolutional, description of individual types of layers).
9. Reconstruction of 3D objects from 2D images (basic principles).

These topics are discussed in Computer Graphics I, Computer Graphics II, Digital Image Processing, and Image Analysis I.

## Computer Systems and Networks

1. Routing in autonomous systems – OSPF, and IS-IS protocols; optimization.

2. Routing between autonomous systems – AS definition, BGP, MP-BGP, routing policies.
3. Switched local area networks with redundancy – STP variants, link aggregation, VLAN security.
4. Ensuring quality of service in computer networks. Multimedia data transfers.
5. Multicasts in LAN and WAN.
6. IPv6 – basic concepts and mechanisms, supporting and routing protocols.
7. Attacks on computer systems and networks, their detection and prevention.
8. Basic data types and addressing methods in machine code. Principles for combining ASM and C source code. Ways of passing arguments to and returning values from functions.
9. Instruction set properties required to implement calculations with large numbers. When, and why, it is necessary to differentiate between signed and unsigned data types in the instruction set.
10. Basic characteristics and differences between processes and threads, their advantages, disadvantages and implementation limitations.
11. I/O operations in the OS: blocking and non-blocking calls, synchronous and asynchronous operation.
12. Interprocess communication methods, their properties, and appropriateness of use.
13. Basic use of secure communication using SSL.
14. PC boot process: BIOS, UEFI, network boot – principles, description of necessary services.
15. The specifics of mail server configuration, the mechanism of aliases and canonical names, basic and advanced approaches to ensure the mail server security.
16. Types of operating system kernels. Operating system core drivers, description of driver creation, specifics of its development, ways of communication between the driver and the user environment.

Topics are covered in following subjects: Routed and Switched Networks, Programming in Operating Systems, Assembly Languages, Computer Attacks and Defence, and UNIX Systems Management.

## Software engineering

1. Requirements engineering discipline. Discipline activities, artifact, models. Classification, prioritization, management, traceability, requirements dependency. Characteristics of “well-done” requirements. Analytical mechanism, analytical patterns.
2. Use cases -suggested structure, guidelines for a scenario writing level of use case, extend, extensions, relations among use-cases. Usage during software development.
3. Design of architecture and detail design. Views on architecture. Sources of design advice (Similar system, Reference model, Architectural styles, Design patterns, Design Conventions, Design principles). Key issues in Software Design. Object oriented design – inheritance and composition; dynamic and static dispatch; Liskov substitution principle; Demeter Law; Dependency Inversion.
4. UML – properties of language, description of diagrams, usage of diagrams when create model of requirements and analysis and design.
5. Design patterns – GoF, Patterns of Enterprise Application Architecture (M. Fowler). Enterprise Integration Patterns (G. Hohpe).
6. Web Services, Service Oriented Architecture. Architectonic Style REST. Microservice Architecture.
7. The importance of testing. Testing terminology. Test process. Test planning. Verification vs. Validation. Expected Results. Incident Management.

8. Testing through the software lifecycle. Levels of Testing (V-model). Testing in each level. Testing techniques. Configuration management.
9. Software process. RUP, SCRUM, XP – description and comparison.
10. Declarative programming language, substance of functional and logic programming.

Themes are covered by lectures: Software Engineering I, Software Engineering II, Software Engineering III, Enterprise Application Development.

## Theoretical Computer Science

1. Computational complexity of algorithms. Advanced methods of analysis of complexity of algorithms: recursive algorithms, amortized analysis, average-case analysis.
2. General methods of design of efficient algorithms (divide and conquer, dynamic programming, greedy algorithms for optimization problems, ...).
3. Mathematical models of algorithms (Turing machines, RAM machines, Minsky machines). Algorithmically undecidable problems, halting problem. Semi-decidable problems, Post's theorem. Rice's theorem and its consequences for automated verification of properties of programs.
4. Complexity classes, in particular, classes PTIME, NPTIME, PSPACE, NPSPACE, polynomial hierarchy, EXPTIME, EXPSPACE, LOGSPACE, NLOGSPACE, and their relations. Examples of practical problems from the above mentioned classes.
5. Polynomial reducibility between problems and other types of reducibilities. NP-hard and NP-complete problems. PSPACE-complete problems. The question of relations of classes PTIME, NPTIME, PSPACE, LOGSPACE, NLOGSPACE.
6. Randomized algorithms, e.g., for primality testing. Properties necessary for a practical use. Applications in cryptography.
7. Problems of discrete optimization. Approximation algorithms for NP-hard optimization problems. Approximation ratio, the class of (well) approximable problems. Examples, in particular, the travelling salesman problem (TSP).
8. Parallel algorithms. Models of computation for parallel algorithms (PRAM). Analysis of computational complexity of parallel algorithms. Class NC. PTIME-complete problems. Relationship between parallel algorithms and circuits.
9. Distributed algorithms. Models of computation for distributed algorithms. Communication complexity.
10. Semantics of programming languages: formalisms for a description of semantics (operational semantics, denotational semantics). Methods for proving correctness of programs.
11. Deductive reasoning, a definition of a valid inference.
12. Propositional logic, syntax and semantics of the language, proving in propositional logic.
13. First-order predicate logic, syntax and semantics of the language (interpretations, models, satisfiability).
14. Semantic methods of proofs in first-order predicate logic and Aristotle's logic.
15. General resolution method in first-order predicate logic.
16. Proof calculi: definitions and properties of calculi (soundness, completeness, undecidability).
17. Natural deduction in first-order predicate logic.
18. Theories of arithmetic, incompleteness of arithmetic, Gödel's incompleteness theorems.
19. Algebraic theories, theories of relations and lattices.

The topics are covered by courses: Mathematical Logic, Theoretical Computer Science, Advanced Logic, Selected Topics of Theoretical Computer Science.

Last actualization 20.10.2023