POZNAN UNIVERSITY OF TECHNOLOGY

EUROPEAN CREDIT TRANSFER AND ACCUMULATION SYSTEM (ECTS)

COURSE DESCRIPTION CARD - SYLLABUS

Course name

Infomation Theory [N1Inf1>TINF]

Course

Field of study

Computing

Area of study (specialization)

–

Level of study

first-cycle

Form of study

part-time

Year/Semester

4/7

Profile of study

general academic

Course offered in

Polish

Requirements

elective

Number of hours

Lecture

12

Laboratory classes

12

Other

0

Tutorials

0

Projects/seminars

0

Number of credit points

2,00

Coordinators

dr hab. inż. Robert Susmaga
robert.susmaga@put.poznan.pl

Lecturers

Prerequisites

Basic knowledge regarding: a) data structures (one- and two-dimensional arrays, lists, trees), b) probability theory and statistics (probability, including conditional probability, random variables, probability distributions and probability distribution functions, mean values, expected values, variance), c) calculus (logarithmic function, exponential function, function derivative). Basic skills regarding designing, creating and testing computer programs (in a programming language of one’s choice) that implement simple processing of static (vectors and matrices) and dynamic (lists, trees) data structures. (recommended) A fair amount of cognitive curiosity and not less perseverance in pursuing the goals of personal development.

Course objective

The objective of the course is to present a selection of aspects of the Information Theory, one of the most fundamental theories underlying theoretical Computer Science of modern-day. The Information Theory deals with representing, storing and communicating information expressed in the form of symbols. Owing to the fact that many important applications of this theory reach far beyond the core of Computing Science, the presented selection of aspects will be confined to the most fundamental ones, mainly those related to Data Compression, in particular: Lossless Data Compression. The fruits of rapid development of notions in the Information Theory have soon turned out to be incredibly useful and to have a great deal of practical value. In the modern-day these solutions show up in virtually all imaginable computer systems in existence, ones that could hardly survive without the ubiquitous multimedia content, the popularity and versatilty of which has been consistently and unwaveringly influenced by the accessibility of data compression methods. Detailed objectives of the course include sharing skills and knowledge sufficient for: a) undestanding the fundamental ideas underlying the Information Theory, especialy those pertaining to Data Compression, b) identifying, formulating and solving basic problems in the Information Theory and Data Compression, c) designing and creating computer programs that successfully implement the presented methods and algorithms.

Course-related learning outcomes

Knowledge:
The students:
1. have a basic, ordered and well-grounded knowledge essential for important areas of computer science such as algorithmics -- [K1st_W2]
2. know and understand the basic techniques, methods, algorithms, and tools used for solving computer problems related to Information Theory and its applications to Data Compression -- [K1st_W4]
3. know and understand the basic techniques, methods, algorithms, and tools used for solving computer problems as well as problems in computing science, including an automated recognition of patterns in data of different types and their processing -- [K1st_W5]
4. have a basic knowledge of key directions and the most important successes of, making use of the achievements of other scientific disciplines, like Information Theory to provide solutions with a high practical impact -- [K1st_W5]

Skills:
The students:
1. are able to collect information from the appropriate sources of different nature, perform its critical analysis, interpretation and synthesis as well as comprehensively draw and justify formulated conclusions regarding the information, especially in the context of Information Theory and its applications to Data Compression -- [K1st_U1]
2. can efficiently plan and carry out experiments, including computer measurements and simulations related to various aspects of Information Theory, interpret the obtained results and draw conclusions based on the experimental outcomes -- [K1st_U4]
3. can retrieve, analyse and transform different types of data and carry out data synthesis to knowledge and conclusions useful for solving a variety of problems that arise in computer science -- [K1st_U10]
4. can -- following a pre-defined specification -- design and create an IT system by first selecting and then using available methods, techniques and computer tools (including programming languages) --- [K1st_U8]
5. can adapt existing algorithms as well as formulate and implement novel ones, including algorithms typical for different streams of Computing Science, using at least one well-known tool -- [K1st_U9]

Social competences:
The students:
1. understand that knowledge and skills quickly become outdated and perceives the need for constant additional training and raising one's qualifications -- [K1st_K1]
2 are aware of the importance of scientific knowledge and research related to computer science and Computing Science in solving problems that are essential for the functioning of individuals, firms, organizations as well as the entire society -- [K1st_K2]

Methods for verifying learning outcomes and assessment criteria

Learning outcomes presented above are verified as follows:

Learning outcomes presented above are verified as follows:
Formative assessment (laboratory clasess): evaluation of the solutions to the assigned programming problems (as they arise).
Final assessment:
-- (laboratory clasess): evaluation of the solutions to the assigned programming problems,
-- (lectures): evaluation of the results of a written test (30--45 min) with both mutiple choice, short answer and (small) computational questions (mostly: micro-problems to be solved in writing).

Programme content

Information and information measure, the measure of
Hartley. The measure of Shannon (entropy), entropy as the measure of information. Applications of the measures of Hatley and Shannon. Multi-dimensional aspects of entropy: joint entropy; conditional entropy and mutual
information, and their applications in data analysis.
Two types of compression: lossless and lossy. Two techniques of compression: code-based and dictionary-based.
Code-based compression: Shannon-Fano encoding and Huffman encoding.
Dictionary-based compression: Lempel-Ziv-Welch algorithm.

Course topics

The course includes, but is not limited to, the following.
Fundamentals of:
-- calculus (fundamental properties and graphs of log(x) and x*log(x)),
-- probability theory (discrete random variables): probability and conditional probability, probability distributions.
Information and an information measure. Basic properties of the information measure. The measure of Hartley. The measure of Shannon (entropy): basic mathematical properties (graphs and extrema in two- and three-dimensional cases). The entropy as the measure of information. Applications of the measures of Hatley and Shannon to text, audio and video transmissions.
Three types of data processing: encoding/decoding, encrypting/decrypting and hashing. Compression (discrete data): idea and objectives. Compression as a form of coding. The two basic types of compression (lossless and lossy). The two basic techniques of compression (code-based and dictionary-based).
Encoding/decoding (discrete data): the idea and objectives. Codes: definitions and examples, tree representations of codes, prefix codes. Kraft's inequality. Data encoding aimed at compressing.
Code-based lossless compression (discrete data): Shannon-Fano encoding and Huffman encoding: the idea, code trees, optimal codes, algorithms, examples, properties. Entropy in code-based compression.
Dictionary-based lossless compression: Lempel-Ziv-Welch method: the idea, representations, dictionaries, algorithms, examples, properties.
Modern day compression systems. The future of compression.
(Optionally) Multi-dimensional aspects of entropy (joint entropy, conditional entropy, mutual information).
(Laboratory classes)
Simple programs operating on scalar and vector arguments. Probability distributions, probabilistic approximations to languages.
Applications of selected code-based methods in compression: Shannon-Fano coding and Huffman coding.
Applications of selected dictionary-based methods in compression: Lempel-Ziv-Welch.

Teaching methods

Lectures: multimedia presentations (theoretical elements, explanations, examples, exercises).
Laboratory classes: designing and creating (in a programming language of one's choice) programs that solve the assigned problems (which illustrate the ideas and notions presented during the lectures).

Bibliography

Basic
1. D.J.C. MacKay: Information Theory, Inference, and Learning Algorithms, Cambridge University Press,

Cambridge, UK, 2003.
2. T.M. Cover, J.A. Thomas, Elements of Information Theory, 2nd Edition, Wiley and Sons, Hoboken, New Jersey, 1991.
3. K. Sayood (red.): Lossless Compression Handbook, Academic Press, Elsevier Science, San Diego, California, 2003.
Additional
1. Lecture notes (slide show presentations)
2. K. Sayood: Introduction to Data Compression, 3rd Ed., Morgan Kaufmann Publishers, San Francisco, California, 2006.
3. A. Drozdek: Wprowadzenie do kompresji danych WNT, Warszawa, 1999 (in Polish).
4. A. Przelaskowski: Kompresja danych. Podstawy, metody bezstratne, kodery obrazów, BTC, Legionowo, 2005 (in Polish)

Breakdown of average student's workload

	Hours	ECTS
Total workload	50	2,00
Classes requiring direct contact with the teacher	24	1,00
Student's own work (literature studies, preparation for laboratory classes/ tutorials, preparation for tests/exam, project preparation)	26	1,00