

Міністерство освіти і науки України
Львівський національний університет імені Івана Франка

Основи теорії інформації та кодування

Навчальний посібник

Львів
2023

УДК 519.7:004(075.8)

О 75

Автори:

І. А. Прокопишин, Р. Є. Рикалюк, В. Ф. Чекурін, К. А. Червінка

Рецензенти:

д-р техн. наук, проф. В. Б. Дудикевич

(Національний університет "Львівська політехніка", м. Львів);

д-р фіз.-мат. наук, проф. Т. С. Нагірний

(Інститут механічної інженерії, Зеленогурський Університет,
м. Зелена Гура, Польща);

д-р техн. наук, проф. Р. М. Пасічник

(Західноукраїнський національний університет, м. Тернопіль)

Рекомендовано до друку Вченою радою

Львівського національного університету імені Івана Франка

(протокол № 49/6 від 29.06.2023 року)

О 75 **Основи теорії інформації та кодування** : навч. посібник / [І. А. Прокопишин, Р. Є. Рикалюк, В. Ф. Чекурін, К. А. Червінка]. – Електрон. вид. – Львів : ЛНУ ім. Івана Франка, 2023. – 156 с.

ISBN 978-617-10-0820-5 (електрон. вид.)

Розглянуто головні поняття теорії інформації, ефективного кодування даних, моделювання та кодування графічних зображень, звуку та відео, кодування з захистом від завад та шифрування, а також теоретичні основи побудови електронних схем комп'ютерів.

Для студентів математичних та інших спеціальностей, які вивчають комп'ютерні науки.

УДК 519.7:004(075.8)

ISBN 978-617-10-0820-5 (електрон. вид.)

© Прокопишин І. А., Рикалюк Р. Є.,
Чекурін В. Ф., Червінка К. А., 2023

© Львівський національний університет
імені Івана Франка, 2023

ЗМІСТ

СПИСОК АБРЕВІАТУР.....	5
ПЕРЕДМОВА.....	8
Розділ 1. ПРЕДМЕТ КОМП'ЮТЕРНИХ НАУК	10
1.1. Інформатика. Кібернетика. Комп'ютерні науки	10
1.2. Інформаційні технології та інформаційні системи	13
Запитання та завдання.....	14
Розділ 2. БАЗОВІ ПОНЯТТЯ ТЕОРІЇ ІНФОРМАЦІЇ.....	15
2.1. Інформація	15
2.2. Повідомлення, сигнали, дані. Загальна схема передавання інформації	17
2.3. Математичні моделі сигналів.....	18
2.4. Кодування даних	24
2.5. Представлення та кодування даних у комп'ютерах.....	25
2.6. Представлення чисел у комп'ютерах	28
Запитання та завдання.....	31
Розділ 3. КІЛЬКІСНА ОЦІНКА ІНФОРМАЦІЇ ДИСКРЕТНИХ ДЖЕРЕЛ	32
3.1. Кількість інформації випадкової події.....	32
3.2. Ентропія дискретного джерела інформації.....	34
Запитання та завдання.....	39
Розділ 4. ЕФЕКТИВНЕ КОДУВАННЯ.....	40
4.1. Побудова двійкових префіксних кодів за допомогою бінарних дерев	40
4.2. Теорема Шеннона про ефективне кодування.....	43
4.3. Алгоритм Шеннона–Фано	44
4.4. Алгоритм оптимального кодування Гаффмана.....	48
4.5. Арифметичне кодування	50
Запитання та завдання.....	52
Розділ 5. МЕТОДИ СТИСНЕННЯ ДАНИХ БЕЗ ВТРАТ ІНФОРМАЦІЇ	53
5.1. Стиснення даних	53
5.2. Словникові методи стиснення	55
5.3. Архівація даних	56
Запитання та завдання.....	57
Розділ 6. МОДЕЛЮВАННЯ, КОДУВАННЯ ТА ВІДТВОРЕННЯ КОЛЬОРІВ	58
6.1. Світлові величини	58
6.2. Основи колориметрії.....	60
6.3. Колірні моделі	64
6.4. Кодування кольорів.....	67
6.5. Управління кольором.....	68
Запитання та завдання.....	69
Розділ 7. ЦИФРОВІ ГРАФІЧНІ МОДЕЛІ ТА ФОРМАТИ.....	71
7.1. Растрові зображення	71
7.2. Векторні моделі зображень	75
7.3. Алгоритми стиснення зображень	76
7.4. Растрові графічні формати	80

7.5. Векторні та комбіновані формати.....	83
Запитання та завдання.....	86
Розділ 8. КОДУВАННЯ ЗВУКУ ТА ВІДЕО	87
8.1. Перетворення звукового сигналу у цифровий	87
8.2. Звукові формати	89
8.3. Моделювання відеоданих	92
8.4. Базові методи стиснення відеоданих.....	93
8.5. Головні відеостандарти.....	95
Запитання та завдання.....	95
Розділ 9. КОДУВАННЯ З ЗАХИСТОМ ВІД ЗАВАД	97
9.1. Загальні поняття	97
9.2. Теоретико-множинний аналіз завадостійкого кодування.....	98
9.3. Кодова відстань	100
9.4. Лінійні систематичні коди.....	102
9.5. Код Геммінга.....	105
Вправи	106
Запитання та завдання.....	107
Розділ 10. ОСНОВИ КРИПТОГРАФІЇ.....	108
10.1. Методи захисту інформації. Криптологія.....	108
10.2. Короткий історичний огляд.....	110
10.3. Симетричні криптографічні системи	112
10.4. Криптосистеми з відкритим ключем.....	117
10.5. Стандарти шифрування даних	119
10.6. Гешувальні функції.....	120
10.7. Цифровий підпис.....	122
10.8. Адміністрування ключами	123
10.9. Криптоаналіз. Теоретична та практична стійкість шифру.....	125
Запитання та завдання.....	128
Розділ 11. КРИПТОСИСТЕМА <i>RSA</i>	130
11.1. Подільність чисел. Алгоритм Евкліда.....	130
11.2. Обчислювальна складність алгоритмів.....	132
11.3. Кільце зведених лишків за модулем n	133
11.4. Функція Ейлера та її властивості.....	135
11.5. Алгоритм <i>RSA</i>	136
11.7. Надійність алгоритму криптування <i>RSA</i>	138
Запитання та завдання.....	139
Розділ 12. ТЕОРЕТИЧНІ ОСНОВИ ПОБУДОВИ КОМП'ЮТЕРІВ.....	140
12.1. Архітектура універсального комп'ютера.....	140
12.2. Теоретичні основи побудови цифрових електронних схем.....	142
Запитання та завдання.....	149
СПИСОК ЛІТЕРАТУРИ.....	150
ПРЕДМЕТНИЙ ПОКАЖЧИК	153

СПИСОК АБРЕВІАТУР

AIC	–	автоматизована інформаційна система
АЦП	–	аналого-цифровий перетворювач
ДКП	–	дискретне косинус-перетворення
ЕОМ	–	електронно-обчислювальна машина
ЕПТ	–	електронно-променева трубка
ІС	–	інформаційна система
МК	–	матриця квантування
МКО	–	міжнародна комісія з освітленості
НСД	–	найбільший спільний дільник
ОП	–	оперативна пам'ять
ОС	–	операційна система
ТВЧ	–	телебачення високої чіткості
ЦАП	–	цифро-аналоговий перетворювач
AAC	–	<i>Advanced Audio Coding</i>
AC-3	–	<i>family of Audio Compression technologies developed by Dolby Laboratories</i>
AES	–	<i>Advanced Encryption Standard</i>
ANSI	–	<i>American National Standard Institute</i>
ASCII	–	<i>American Standard Code for Information Interchange</i>
AT&T	–	<i>American Telephone and Telegraph</i>
AVC	–	<i>Advanced Video Coding – video compression standard</i>
BD	–	<i>Blu-ray Disc</i>
BMP	–	<i>Bitmap Picture – bitmap image storage format developed by Microsoft</i>
CCITT	–	<i>Comité Consultatif International Téléphonique et Télégraphique</i>
CD	–	<i>Compact Disc</i>
CD-DA	–	<i>Compact Disc Digital Audio – standard for recording sound on CD</i>
CD-R	–	<i>Compact Disc-Recordable</i>
CD-RW	–	<i>Compact Disc-ReWritable</i>
CIE	–	<i>Commission Internationale de l'Eclairage</i>
CIELAB	–	<i>color space developed by CIE in 1976 based on CIE XYZ</i>
CIE XYZ	–	<i>color space created by the CIE in 1931</i>
CIF	–	<i>Common Interchange Format – video standard with scan of 352x288 pixels</i>
CMM	–	<i>Color Matching Module</i>

Список аббревиатур

<i>CMS</i>	– <i>Color Management System</i>
<i>CMY</i>	– <i>subtractive color model, in which the main colors are Cyan, Magenta and Yellow</i>
<i>CMYK</i>	– <i>subtractive color model, in which the main colors are Cyan, Magenta, Yellow and black</i>
<i>DES</i>	– <i>Data Encryption Standard – symmetric block encryption algorithm</i>
<i>DjVu</i>	– <i>lossy image compression technology and the corresponding format</i>
<i>DVD</i>	– <i>Digital Video Disc</i>
<i>ENIAC</i>	– <i>Electronic Numerical Integrator and Computer</i>
<i>EPS</i>	– <i>Encapsulated PostScript – simplified version of PostScript</i>
<i>FPX</i>	– <i>FlashPix – bitmapped computer graphics file format</i>
<i>GIF</i>	– <i>Graphics Interchange Format – 8-bit bitmap graphics format</i>
<i>HDTV</i>	– <i>High-Definition Television – high-definition television standard with scan of 1920x1080 pixels</i>
<i>Full HD</i>	– <i>see HDTV</i>
<i>HSB</i>	– <i>color intuitive model based on three color characteristics: Hue, Saturation, and Brightness</i>
<i>HSV</i>	– <i>see HSB</i>
<i>ICC</i>	– <i>International Color Consortium</i>
<i>ISO</i>	– <i>International Organization for Standardization</i>
<i>JPEG</i>	– <i>name of full-color image compression algorithm and corresponding format, developed by the Joint Photographic Expert Group</i>
<i>LZ</i>	– <i>prefix for family of data compression algorithms (LZ77, LZ78, LZW etc.) – the initials of Abraham Lempel and Jacob Ziv</i>
<i>MD</i>	– <i>prefix for family of Message-Digest Algorithms (MD1–MD5 etc.)</i>
<i>MIDI</i>	– <i>Musical Instrument Digital Interface</i>
<i>MP3</i>	– <i>coding format for digital audio</i>
<i>MPEG</i>	– <i>Moving Pictures Experts Group – line of digital television standards</i>
<i>NTSC</i>	– <i>National Television System Committee – standard for analog television broadcast</i>
<i>PAL</i>	– <i>Phase Alternating Line – color encoding system for analogue television</i>
<i>PCM</i>	– <i>Pulse Code Modulation</i>
<i>PDF</i>	– <i>Portable Document Format</i>
<i>PGP</i>	– <i>Pretty Good Privacy – encryption program</i>
<i>PKE</i>	– <i>Public Key Encryption</i>
<i>PNG</i>	– <i>Portable Networks Graphic – raster-graphics file format</i>
<i>PS</i>	– <i>PostScript – page description language and the corresponding format</i>

Список аббревиатур

- QCIF* – *Quartet Common Interchange Format – video standard with scan of 176x144 pixel*
- RA* – *Real Audio – audio data format that uses several audio codecs*
- RGB* – *additive color model, in which the main colors are Red, Green and Blue*
- RIFF* – *Resource Interchange File Format – standard for storing any data structures*
- RLE* – *Run Length Encoding*
- RSA* – *Rivest–Shamir–Adleman – public-key cryptosystem developed by Ron Rivest, Adi Shamir and Leonard Adleman*
- SECAM* – *Séquentiel de couleur à mémoire – analog color television system*
- SHA* – *prefix for family of Secure Hash Algorithm standards, SHA-1– SHA-3*
- SI* – *Système International d’Unités*
- TIFF* – *Tagged Image File Format – format for storing raster graphic images*
- UHDTV* – *Ultra High Definition Television – digital ultra high definition television standards*
- VOC* – *Voice File – 8-bit audio format*
- WAV* – *Wave Form Audio File – 16-bit audio format developed by Microsoft*
- WCS* – *Windows Color System*
- WMA* – *Windows Media Audio – series of audio codecs and their corresponding audio coding formats developed by Microsoft*
- WMF* – *Windows Meta File – universal vector graphics file format for Windows applications*
- XOR* – *eXclusive OR – logical and bitwise operation modulo two addition*
- XYZ* – *see CIE XYZ*
- YCbCr* – *color model that uses a luminance component and two color difference components*
- YUV* – *color model that uses a luminance component and two color difference components*

ПЕРЕДМОВА

Сьогодні важливим чинником у розвитку суспільства є інформаційні технології, що проникають у всі сфери людської діяльності – в державне управління, економіку, науку, освіту, культуру та побут.

Отож доволі строгий, проте доступний виклад основ теорії інформації та кодування, як одних з фундаментальних галузей комп'ютерних наук, є своєчасним і потрібним для підготовки спеціалістів.

Посібник містить дванадцять розділів.

Перший розділ, вступний, окреслює предмет комп'ютерних наук. У ньому означено поняття інформаційної системи та інформаційної технології.

У другому та третьому розділах викладено головні поняття теорії інформації та питання кількісної оцінки інформації дискретних ймовірнісних джерел. Зокрема, розглянуто математичні моделі сигналів, перетворення аналогових сигналів у цифрові, теорему про відліки, способи кодування тексту та чисел у комп'ютерах, ентропію простого та стаціонарного джерела інформації.

У четвертому та п'ятому розділах викладено основи теорії ефективного кодування та її застосування до стиснення даних. Розглянуто побудову префіксних кодів за допомогою бінарних дерев, теорему Шеннона для стаціонарних джерел, методи ефективного кодування Шеннона–Фано, Гаффмана та арифметичного кодування, а також приклади їхнього використання. Описано методи стиснення даних без втрати інформації, зокрема, словникові методи стиснення і метод Лемпела–Зіва–Велча.

У розділах 6, 7 та 8 розглянуто основи моделювання та представлення у комп'ютерах графічних, звукових та відеоданих. Викладено основи колориметрії, колірні моделі та кодування кольорів. Розглянуто базові графічні моделі та формати, методи кодування звуку та відеозображень, а також методи їхнього стиснення.

У розділі 9 вивчатимемо проблеми завадостійкого кодування. На строгому математичному рівні розглянуто загальні питання побудови лінійних систематичних кодів, зокрема, коду Геммінга.

У десятому та одинадцятому розділах викладено основи криптографії. Детально описано симетричні криптосистеми та криптосистеми з відкритим ключем. Розглянуто стандарти шифрування даних, гешувальні функції, цифровий підпис, викладено алгоритм *RSA* та його надійність.

У дванадцятому розділі розглянуто теоретичні основи побудови електронних схем комп'ютерів, які ґрунтуються на апараті математичної логіки.

Для розуміння головних теоретичних положень, викладених у посібнику, достатньо знань з курсу вищої математики для ЗВО. Посібник містить значну кількість прикладів, які ілюструють теоретичний матеріал. Наприкінці кожного розділу подано запитання для самоконтролю.

Під час підготовки посібника автори використали усі джерела, наведені у списку літератури.

Автори висловлюють глибоку вдячність професорам В. Б. Дудикевичу, Т. С. Нагірному та Р. М. Пасічнику за зауваження та пропозиції, надані ними під час рецензування посібника.

Електронне навчальне видання

ПРОКОПИШИН Іван Анатолійович
РИКАЛЮК Роман Євстахович
ЧЕКУРІН Василь Феодосійович
ЧЕРВІНКА Костянтин Андрійович

Основи теорії інформації та кодування

Навчальний посібник

Редактор
Ірина Лоїк

Комп'ютерне верстання
Іван Прокопишин

Дизайн обкладинки
Василь Роган

Формат 60x84/8. Умовн. друк. арк. 18,1. Зам. № 8Е

Видавець і виготовлювач:
Львівський національний університет імені Івана Франка,
вул. Університетська, 1, м. Львів, 79000 Львів.

Свідоцтво про внесення суб'єкта видавничої справи до Державного реєстру видавців,
виготівників і розповсюджувачів видавничої продукції:
Серія ДК № 3059 від 13.12.2007.