

МЕДИЧНА ІНФОРМАТИКА

*Навчальний посібник
для студентів медичних університетів*

**Міністерство охорони здоров'я України
Харківський національний медичний університет**

МЕДИЧНА ІНФОРМАТИКА

*Навчальний посібник
для студентів медичних університетів*

**Харків
ХНМУ
2020**

УДК 61:004(075.8)

М 42

Затверджено Вченою радою ХНМУ.

Протокол № 8 від 19.09.2019.

Рецензенти:

Тиманюк В. А. – канд. фізико-мат. наук, проф. (НФУ)

Берест В. П. – канд. фізико-мат. наук, доц. (ХНУ ім. В. Н. Каразіна)

Автори:

Кнігавко В. Г., Зайцева О. В., Бондаренко М. А., Батюк Л. В., Рукін О. С.

М 42 Медична інформатика : навч. посібник для студентів мед. ун-тів / В. Г. Кнігавко, О. В. Зайцева, М. А. Бондаренко та ін. – Харків : ХНМУ, 2020. – 64 с.

Навчальний посібник створено відповідно до оновленої робочої навчальної програми дисципліни «Медична інформатика». Структура навчального посібника дає можливість його використання на практичних заняттях з даної дисципліни.

УДК 61:004(075.8)

© Харківський національний
медичний університет, 2019

© Кнігавко В. Г., Зайцева О. В.,
Бондаренко М. А., Батюк Л. В.,
Рукін О. С., 2020

ТЕМА 1. Основні поняття медичної інформатики. Комп'ютер у діяльності майбутнього лікаря

Навчальна дисципліна «Медична інформатика» викладається з метою ознайомити студентів з використанням інформаційно-комунікаційних технологій (ІКТ) в галузі охорони здоров'я, навчити обробляти медико-біологічні дані за допомогою ІКТ і забезпечувати розвиток інформаційної компетентності в майбутніх лікарів.

Предметом вивчення дисципліни «Медична інформатика» є інформаційні процеси, що передбачають використання ІКТ в галузі охорони здоров'я.

1.1. Основні поняття медичної інформатики

Медична інформатика (МІ) – це наукова дисципліна, що займається дослідженням процесів одержання, передачі, обробки, зберігання, поширення, відображення інформації з використанням інформаційної техніки й технології в медицині й охороні здоров'я.

Основною метою МІ є *оптимізація інформаційних процесів* у медицині за рахунок використання *комп'ютерних технологій*, що забезпечує підвищення якості охорони здоров'я населення.

Інформація – це отримана в процесі переробки даних сукупність знань (*нових*, раніше не відомих фактів).

З поняттям інформації пов'язані такі поняття, як дані, повідомлення, сигнал, канал зв'язку.

Дані – це інформація, представлена у формалізованому вигляді та призначена для обробки її технічними засобами, наприклад ЕОМ.

Повідомлення – це впорядкована сукупність сигналів, здатних переносити інформацію.

Сигнал являє собою будь-який процес, що впливає на сенсорні системи.

Каналом зв'язку називається середовище, по якому передаються сигнали. Наприклад, при усній розмові сигналом є мова, а каналом зв'язку – повітря, у нервовій системі сигналом є нервові імпульси, а каналами – нервові волокна.

Сигнали можуть бути **дискретними й безперервними**. Прикладом дискретного сигналу є передача цифр імпульсами струму, прикладом безперервного – зміна напруги в ланцюзі, що відповідає зміні температури.

Усяке повідомлення складається з комбінації невеликої кількості простих сигналів певної фізичної природи. Повний набір таких сигналів називають **алфавітом**, один сигнал із цього набору – **буквою** алфавіту.

Опис якого-небудь повідомлення за допомогою певного алфавіту називається **кодуванням**, переклад цього повідомлення на інший алфавіт називається **перекодуванням**, розшифрування повідомлення – **декодуванням**.

Будь-яку інформацію можна закодувати за допомогою алфавіту із **двох знаків (0; 1)**. Такий код називається **двійковим**. Це найпоширеніший спосіб кодування в сучасних інформаційних системах. Існують й інші способи кодування, у тому числі створені природою. Наприклад, кодування інформації в ДНК клітини реалізується за допомогою 4 різних «букв» алфавіту (А, Г, Т, Ц) – азотистих основ (аденін, гуанін, тимін, цитозин), а в первинній структурі білків – за допомогою 20 амінокислот.

1.2. Кількість інформації

Одне з основних понять теорії інформації – **кількість інформації** (I). Чим більше різних можливостей має подія, тим більшу інформацію про нього несе повідомлення. Кількість інформації про подію A змінюється у відношенні, зворотному **ймовірності** $P(A)$ цієї події, тобто чим більше ймовірність події, тим менше міститься інформації в повідомленні про те, що ця подія відбулася, і навпаки.

Залежність кількості інформації від імовірності події описується **формулою Хартлі** (введена американським ученим Ральфом Хартлі в 1928 р.):

$$I(A) = -\log_2 P(A).$$

Кількість інформації є позитивною величиною. Оскільки ймовірність будь-якої події $P(A) \leq 1$, величина $\log_2 P(A)$ завжди негативна, тому у формулі Хартлі стоїть знак мінус.

Якщо $P(A) = 1$, то така подія обов'язково повинна відбутися, і повідомлення про цю подію не несе інформації (кількість інформації дорівнює нулю): $I(A) = 0$. Якщо ймовірність події дуже мала, тобто $P(A) \rightarrow 0$, то кількість інформації прямує до нескінченності: $I(A) \rightarrow \infty$.

Одиниця виміру кількості інформації – **біт**.

1 біт – це кількість інформації, що міститься в повідомленні про те, що відбулася подія, імовірність якої дорівнює $\frac{1}{2}$:

$$I(A) = -\log_2 \frac{1}{2} = -(\log_2 1 - \log_2 2) = -(0 - 1) = 1 \text{ (біт)}.$$

Іншими словами, **1 біт** – це інформація, яка міститься в повідомленні про те, що відбулася одна із двох рівноймовірних подій.

Наведемо приклад обчислення кількості інформації, що міститься в повідомленні про подію. Знайдемо кількість інформації в повідомленні про те, що при киданні грального кубика випала цифра 5. Імовірність цієї події $P(5) = \frac{1}{6}$.

Відповідно,

$$I(5) = -\log_2 \frac{1}{6} = -(\log_2 1 - \log_2 6) = \log_2 6 = \frac{\ln 6}{\ln 2} \approx 2.58 \text{ (біт)}.$$

Похідні одиниці кількості інформації: **1 байт = 8 біт**, 1 кбайт = 2^{10} байт = 1024 байт, 1 Мбайт = 1024 кбайт, 1 Гбайт = 1024 Мбайт і т.д.

Якщо події не є рівноймовірними, важливо знати **середню кількість інформації** (\bar{I}), що припадає на одне повідомлення. Ця величина є математичним очікуванням величини $I(A)$ й обчислюється за формулою **Шеннона** (формула отримана американським математиком Клодом Шенноном в 1948 р.):

$$\bar{I} = -\sum_{i=1}^n P_i \log_2 P_i,$$

де n – число всіх можливих подій, про яких приходять повідомлення, P_i – імовірність i -ї події ($i = 1, 2, \dots, n$). Середня кількість інформації також називається **інформаційною ентропією** (H) і є мірою невизначеності інформації в повідомленні.

Наведемо приклад. У кошику лежать 2 білих і 5 чорних куль. Знайти середню кількість інформації в повідомленні про колір вибраної навмання кулі.

Імовірність навмання вийняти білу кулю $P(A) = \frac{2}{7}$, чорну кулю $P(B) = \frac{5}{7}$. Застосовуючи формулу Шеннона, отримуємо

$$\bar{I} = -\sum_{i=1}^n P_i \log_2 P_i = -\left(\frac{2}{7} \log_2 \frac{2}{7} + \frac{5}{7} \log_2 \frac{5}{7}\right) = 0.863 \text{ (біт)}.$$

1.3. Інформаційні технології. Комп'ютер

Інформаційні технології – це комп'ютеризовані способи обробки, зберігання, передачі й використання інформації.

Сукупність пристроїв, призначених для автоматичної або автоматизованої обробки даних, називається **обчислювальною технікою**, конкретний набір взаємодіючих між собою пристроїв і програм називається **обчислювальною системою**.

Центральним пристроєм більшості обчислювальних систем є **комп'ютер**. Як правило, комп'ютер містить наступні пристрої:

- **арифметико-логічний пристрій**, що виконує арифметичні й логічні операції;
- **пристрій управління**, який організує процес виконання програм;
- **запам'ятовувальні пристрої** для зберігання програм і даних;
- **зовнішні (периферійні) пристрої** для **введення-виведення** інформації.

У сучасних комп'ютерах арифметико-логічний пристрій і пристрій управління об'єднані в один пристрій – **центральний процесор**.

Запам'ятовувальні пристрої (пам'ять) комп'ютера підрозділяється на кілька видів:

1. **Оперативна пам'ять** (оперативний запам'ятовувальний пристрій – ОЗП, random access memory – RAM). Зберігає дані й програми, з якими працює комп'ютер у поточний момент часу. Має невеликий обсяг і високу швидкодію. При вимиканні комп'ютера інформація, що перебуває в ОЗП, стирається.

2. **Довгочасна (зовнішня) пам'ять**. Має невисоку швидкодію, але більшу ємність, та інформація, що зберігається в ній, не стирається при вимиканні комп'ютера. До цього виду пам'яті належать носії на магнітних дисках (наприклад, жорсткі диски), компакт-диски (CD, DVD), флеш-накопичувачі і т. д.

3. **Кеш-пам'ять** (cache memory) – спеціальна надшвидка оперативна пам'яті, яка зберігає копії найбільше часто використовуваних ділянок оперативної пам'яті. У сучасних комп'ютерах звичайно знаходиться в складі центрального процесора.

4. **Постійна пам'ять** (постійний запам'ятовувальний пристрій – ПЗП, read only memory – ROM), у яку внесені дані при виготовленні комп'ютера, причому ці дані не можуть бути змінені.

5. **CMOS-пам'ять** (за назвою структури застосовуваних мікросхем CMOS, complementary metal-oxide-semiconductor), призначена для зберігання параметрів конфігурації комп'ютера. Для зберігання інформації звичайно використовує незалежне живлення – батарею або акумулятор.

6. Інші види пам'яті (наприклад, відеопам'ять).

Периферійні пристрої (пристрою введення-виведення) призначені для введення й виведення інформації.

Основними пристроями введення даних у комп'ютер є **клавіатура й миша**. Пристроями введення також є **сканер**, відеокамера і т. д. До пристроїв введення належать **аналого-цифрові перетворювачі (АЦП)** для введення інформації у вигляді електричних сигналів, наприклад для запису звуку за допомогою мікрофона або для одержання інформації з датчиків.

Основні пристрої **виведення** – **монітор і принтер**. Комп'ютер також здатний виводити інформацію у вигляді електричних сигналів на які-небудь виконавчі пристрої через **цифро-аналогові перетворювачі (ЦАП)**. Виведення звуку також здійснюється за допомогою ЦАП.

До пристроїв введення-виведення належать також **мережні пристрої** (мережні адаптери, модеми, пристрої бездротового зв'язку Wi-Fi і т. д.). Вони використовуються для обміну інформацією між комп'ютерами по різних каналах зв'язку.

1.4. Програмне забезпечення. Файлова система

Програмне забезпечення (ПО) – сукупність програм, які забезпечують працездатність комп'ютера, використовуються програмістами й користувачами комп'ютера при розв'язанні різноманітних задач.

Програмне забезпечення підрозділяється на три види.

1. **Системне ПО**. До цього виду належать **операційні системи, оболонки й допоміжні програми (утиліти)**.
2. **Прикладне ПО загального призначення**. До цього виду належить ПО широкого застосування: системи програмування, системи управління базами даних (СУБД), текстові редактори й видавничі системи, універсальні офісні пакети і т. д.
3. **Прикладне ПО спеціального призначення**. До цієї групи належать програми спеціального призначення: бухгалтерські програми, спеціалізовані бази даних, програми фінансового аналізу, системи автоматичного проектування, статистичні програми, експертні системи й багато інші.

На сучасних комп'ютерах інформація, як правило, зберігається у вигляді **файлів**. **Файл** – це **одиниця зберігання даних** довільного обсягу, що має унікальне власне **ім'я**. Файли надають спосіб зберігати інформацію на диску й зчитувати її знову. При цьому від користувача сховані такі деталі, як спосіб і місце зберігання інформації.

Файлова система призначена для зберігання даних на дисках і забезпечення доступу до них. Файлова система визначає **спосіб організації** даних на диску або іншому носії.

При роботі з файлами звичайно вводяться механізми **структурування**. Як правило, вони являють собою **ієрархічні відношення**. Для організації таких відносин вводиться поняття **каталогу (directory)**, що у теперішній час часто називається **папкою (folder)**. Каталог містить інформацію про дані, організовані у вигляді файлів. Усередині каталогів можуть міститися інші каталоги, які називають **підкаталогами (subdirectory)**. За їхньою допомогою ми одержуємо можливість будувати майже нічим не обмежену ієрархію.

ТЕМА 2. Інформаційні ресурси системи охорони здоров'я

2.1. Інформаційні ресурси мережі Інтернет і доступ до них

Інформаційні ресурси в широкому сенсі розуміння цього терміна – це сукупність даних, організованих для ефективного одержання достовірної інформації. Тобто це масиви документів в інформаційних системах: бібліотеках, архівах, банках і базах даних та інших видах інформаційних систем.

Найпоширенішим на сьогоднішній день способом одержання доступу до інформаційних ресурсів є **сервіси** (або служби) **мережі Інтернет**.

Інформаційні сервіси Інтернету надають користувачам можливість доступу до певних інформаційних ресурсів, що зберігаються в Інтернеті. Такими ресурсами є або файли стандартних форматів, або різного роду документи, які можна переглянути, роздрукувати, зберегти.

До основних інформаційним службам відносять **сервіс передачі файлів** (здійснюваний за допомогою протоколу FTP – file transfer protocol) і **Всесвітню павутину** (WWW – world wide web, протоколами якої є **НТТР** – hypertext transfer protocol, і **НТТРС** – hypertext transfer protocol secure – для захищених з'єднань). На частку Всесвітньої павутини припадає більш ніж 80 % мережних інформаційних запасів світу. Таке лідерство WWW має за рахунок адаптованої для Інтернету форми відображення інформації – гіпертексту.

Гіпертекст, на відміну від звичайного тексту, містить у собі систему посилань, за допомогою яких можна зручно й інтуїтивно зрозуміло структурувати дуже великі за обсягом масиви інформації. Інформаційною базою служби WWW є мережа документів (**Web-сторінок**), що зберігаються на Web-серверах Інтернету та зв'язані між собою гіперпосиланнями. Сукупність взаємозалежних сторінок, що належать якійсь одній особі або організації, називається **Web-сайтом**.

Web-сервер – це програма або служба, запущена на якому-небудь комп'ютері, завданням якої є надання доступу до даних, які на ньому розміщені, по протоколах НТТР і НТТРС. Web-сервером називають і програмне забезпечення подібного типу, і комп'ютер, на якому воно встановлене. У світі існують мільйони Web-серверів, і на кожному з них може міститися один або кілька сайтів. Більшість Web-серверів у світі працюють під керуванням операційних систем Linux і Unix, а найбільш популярним Web-сервером на сьогоднішній день є Apache. Сайти й сторінки лежать у певній папці на комп'ютері із запущеним Web-сервером, і коли ми набираємо адресу Web-сторінки, то ми просто відкриваємо файли в цій папці.

2.2. Технології пошуку інформації в мережі Інтернет

Проблема забезпечення зручного пошуку інформації для користувачів Мережі вирішується у двох напрямках. Перший напрямок – це створення Інтернет-каталогів, друге – пошукових машин (або пошукових систем).

Каталог Інтернет-ресурсів – це ієрархічний каталог посилань на різні інформаційні ресурси, що постійно оновлюється й поповнюється та містить безліч окремих Web-сторінок, Web-сайтів або їх категорій (груп) з коротким описом їх змісту.

Каталоги прості у використанні, спосіб пошуку інформації з каталогу передбачає «рух униз по сходах», тобто рух від більш загальних категорій до більш конкретних. На головній сторінці сайту каталогу Інтернет-Ресурсів розташований тематичний рубрикатор (рис. 1 а-б). Клацанням по назві рубрики Ви попадаєте в список підрозділів, що стосуються обраної тематики, і далі маєте можливість переглянути список сайтів і ознайомитися з їхнім вмістом. Такий спосіб пошуку інформації найбільш зручний, коли Ви утрудняєтеся чітко сформулювати мету свого пошуку, або прагнете скласти загальне уявлення про тему.

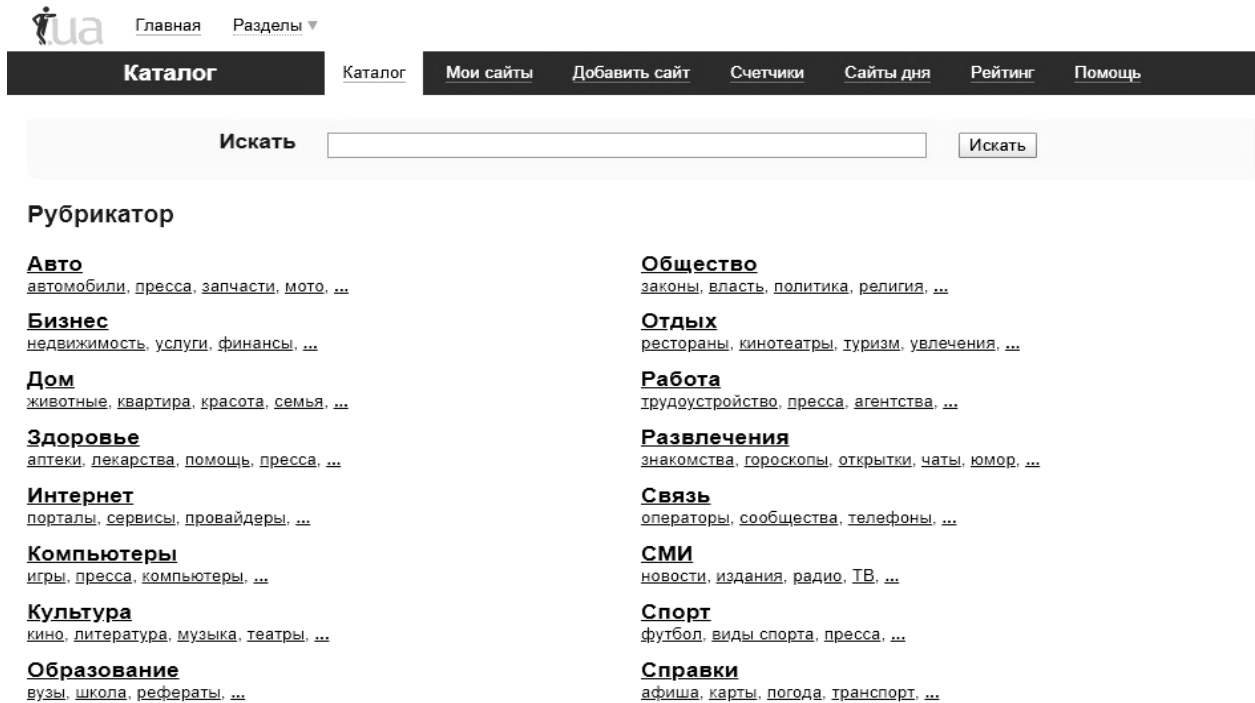


Рис. 1а. Інтерфейс каталогу Інтернет-Ресурсів i.ua (catalog.i.ua)

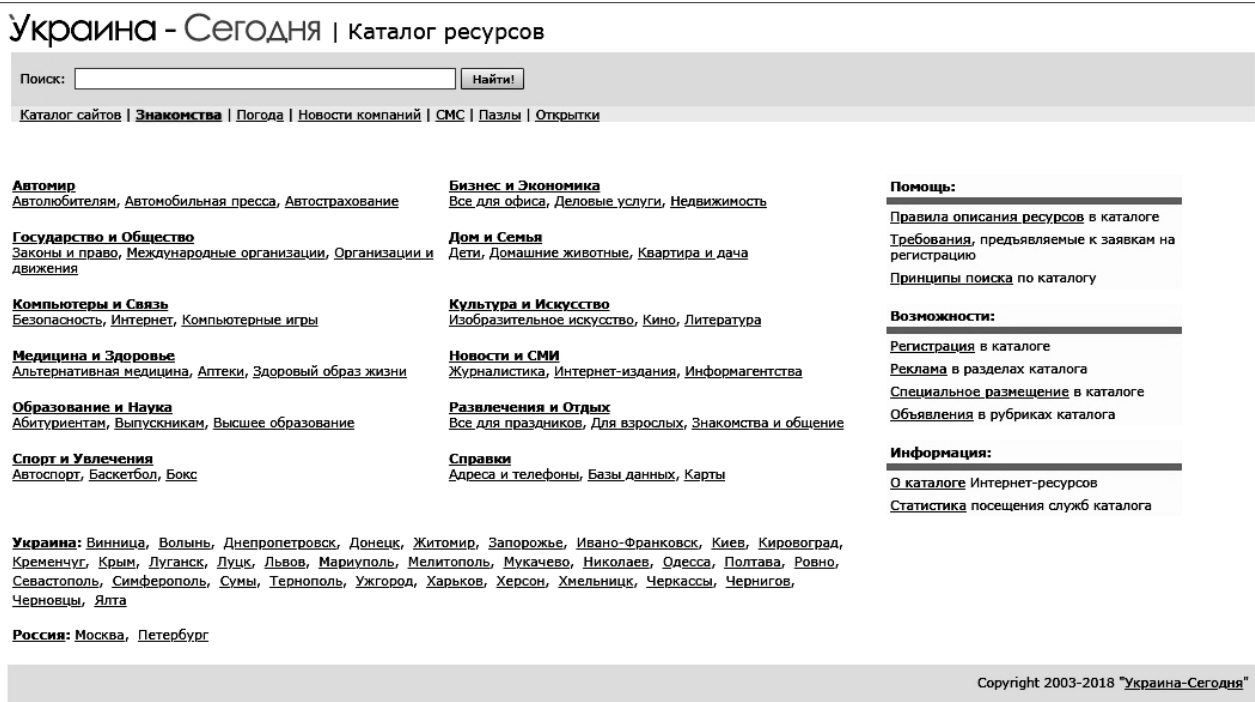


Рис. 1б. Інтерфейс каталогу Інтернет-Ресурсів www.ukraine-today.net

Однією з переваг тематичних каталогів є те, що пояснення до посилань надаються самим авторами сайту, включеного до каталогу, і повністю відбивають його зміст, що дає Вам можливість точніше визначити, наскільки відповідний зміст сайту до мети Вашого пошуку. Однак є й істотний недолік: у каталоги заносяться посилання на сервери, сайти або великі розділи сайтів, тому посилання на всі сторінки, де могла б зустрітися потрібна Вам інформація, у каталог не потрапляють. Для детального, тонкого пошуку каталоги не підходять.

Приклади деяких тематичних українських і російськомовних каталогів загального призначення: catalog.i.ua, favorites.com.ua, www.ukraine-today.net і ін.

Кожний з каталогів Інтернет-Ресурсів загального призначення містить розділ, присвячений медицині та здоров'ю. Наприклад, i.ua – <http://catalog.i.ua/catalog/7/>.

Пошукові системи (пошукові машини, або машини Web-пошуку) – це сервери з величезною базою даних Інтернет-адрес, які автоматично звертаються до Web-сторінок по всіх цих адресах, вивчають вміст цих сторінок, формують і прописують ключові слова зі сторінок у свою базу даних (індексують сторінки). Більше того, роботи пошукових систем переходять за посиланнями, що зустрічаються на сторінках, і переіндексують їх. Оскільки майже будь-яка Web-сторінка має безліч посилань на інші сторінки, то в результаті подібної роботи пошукова машина зрештою теоретично може обійти всі сайти в Інтернет. Тому пошукові машини є найбільш придатним інструментом для детального, тонкого пошуку інформації за ключовими словами.

Найпопулярніша сьогодні пошукова система – **Google** (www.google.com). Не менш відомі Yahoo (www.yahoo.com), Bing (www.bing.com).

Технологія використання пошукової системи проста. Користувач набирає ключову фразу й активізує пошук, тим самим одержує добірку документів за сформульованим запитом. Цей список документів ранжується пошуковою машиною за певними критеріями так, щоб угорі списку виявилися ті документи, які найбільше відповідають запиту користувача. Кожна з пошукових машин використовує різні критерії ранжування документів, як під час аналізу результатів пошуку, так і під час формування індексної бази даних Web-сторінок. Тому якщо вказати в рядку пошуку різних пошукових систем однаковий запит, можна отримати різні результати пошуку.

Головний недолік пошукових систем – відносно низький відсоток відповідності інформації в списку результатів пошуку. Тобто далеко не кожна адреса в списку результатів пошуку відповідає запиту користувача. Іноді запит користувача сформульований занадто широко, і результат пошуку неосяжно великий (десятки тисяч сторінок), і тоді пошукова система неправильно інтерпретує ключові слова.

Багато з названих вище сайтів поєднують у собі функції пошукової машини й каталогу, надаючи при цьому доступ і до інших ресурсів Інтернет: електронна пошта, обмін швидкими повідомленнями в соціальних мережах, доступ і управління віддаленими (хмарними) сховищами даних і т. д. (рис. 2). Такі сайти називаються **Web-Порталами**.



Рис. 2. Інтерфейс Web-Порталу Yahoo (www.yahoo.com)

2.3. Медичні Інтернет-Ресурси

Питання, пов'язані з охороною здоров'я, знаходять своє відображення на безлічі сайтів. У мережі Інтернет можна знайти матеріали, що представляють інтерес для пацієнтів, практикуючих лікарів, організаторів охорони здоров'я, науковців і т. д. Існують спеціальні ресурси, цікаві кожній окремій групі, але є й такі, які необхідні широкому колу користувачів. Можна визначити два основні напрямки класифікації медичних ресурсів мережі Інтернет: по типу відвідувачів і по меті відвідування.

За типом очікуваних відвідувачів можна виділити наступні групи ресурсів:

- для пацієнтів (ресурси, що пропонують довідкову медичну інформацію про різні захворювання, їх симптомах, способах профілактики; про лікарів і установи, що надають відповідну медичну допомогу);
- для лікарів (спеціалізована медична інформація для фахівців практичної охорони здоров'я й науковців);
- для фахівців з організації охорони здоров'я (законодавчі й нормативні акти, довідкові матеріали, необхідні для організації роботи й підготовки звітності; аутсорсингові ресурси – медичні послуги, надані іншими установами й комерційними фірмами за договорами);
- для фахівців фінансово-господарчих служб і підприємців, чия діяльність пов'язана з охороною здоров'я (устаткування, медикаменти, видаткові матеріали, інструментарій, засоби зв'язку, транспорт і т. д.);
- для фахівців кадрових служб і пошуку роботи (сервіси, що дозволяють переглядати резюме фахівців-медиків, здійснювати пошук вакансій і т. п.).

За метою відвідування Інтернет-ресурсу виділяються такі групи ресурсів:

- для пошуку спеціалізованої інформації;
- для пошуку медичних послуг;
- для пошуку лікувальних і профілактичних засобів;
- для навчання;
- для бізнесу й забезпечення діяльності лікувально-профілактичних установ;
- для пошуку роботи й співробітників;
- для спілкування.

Приклади спеціалізованих медичних сайтів:

- <http://www.moz.gov.ua> – офіційний сайт Міністерства охорони здоров'я України;

- <http://www.medicusamicus.com> – «Medicus Amicus» – сайт для лікарів і фармацевтів;
- <http://www.morion.ua> – сайт компанії «МОРІОН», що спеціалізується в сфері розвитку й впровадження комп'ютерних технологій у медицині й фармації, видавничого бізнесу, дослідження й аналізу фармацевтичного ринку України;
- <http://www.health-ua.org> – «Здоров'я України» – медичний портал;
- <http://www.compendium.com.ua> – On-line компендіум;
- <http://www.medicinform.net> – «Медична інформаційна мережа» – пошукова машина й каталог, новини.

2.4. База даних медичних публікацій Medline

Будь-яке наукове дослідження починається з пошуку відомих результатів, отриманих раніше іншими дослідниками, в обраній галузі знань. Такий пошук дозволяє провести аналіз і визначити новизну дослідження, його актуальність, оцінити значимість отриманих результатів, вибрати актуальний напрям для розробки. Оскільки більшість наукових результатів публікується в спеціалізованих журналах, то якісний пошук передбачає перегляд усіх періодичних видань по обраній і баз даних, що містять інформацію про наукові статті та монографії, що публікуються у світі. Однією з найвідоміших бібліографічних систем медичних публікацій є Medline. Ця база даних була створена в Національній медичній бібліотеці Національного інституту здоров'я США на початку 80-х років ХХ століття. У ній зібрані реферати й бібліографічні дані всіх публікацій з кінця 1960-х років і по сьогоднішній день із більш ніж 4 000 світових наукових журналів. До переваг бази даних Medline можна віднести наступне:

- можливість швидко підібрати бібліографічні дані про статті на задану тему;
- реферати статей дозволяють одержати загальне уявлення про зміст, матеріали й результати публікації;
- для близько 80 % наявних у бібліографічному каталозі бібліотеки журнальних публікацій доступні повнотекстові версії;
- дозволяє провести пошук за пов'язаними темами.

База даних Medline і програма PubMed, що обслуговує її, вільно доступні на сервері <http://www.ncbi.nlm.nih.gov/pubmed> або <http://pubmed.gov>.

Пошук публікацій здійснюється за ключовими словами і фразами. Для ключових слів на сторінці пошуку є рядок введення. Однак введення ключових слів і фраз здійснюється тільки англійською мовою. Запит обробляється після натискання на кнопку Search (рис. 3). Результат пошуку – це список публікацій, у яких зустрічаються задані ключові слова й фрази.

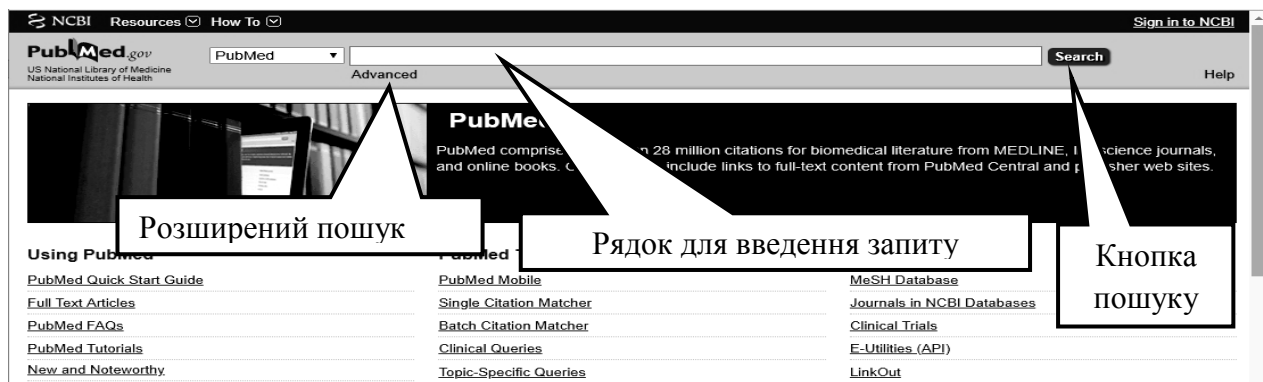


Рис. 3. Інтерфейс початкової сторінки програми Pubmed

ТЕМА 3. Створення та ведення медичної документації

3.1. Класифікація програм для роботи з текстами

Для створення документації, у тому числі й медичної, у цей час використовуються програми для роботи з текстами. Ці програми діляться на текстові редактори, текстові процесори й видавничі системи.

Текстові редактори призначені для створення й зміни текстових даних взагалі й текстових файлів зокрема. Вони дозволяють переглядати вміст текстових файлів і виконувати над ними різні дії – вставку, видалення й копіювання тексту, контекстний пошук і заміну, сортування рядків, перегляд кодів символів і конвертацію кодувань, друкування і т. п.

Для серйозної роботи з документами використовуються **текстові процесори**. Текстові процесори відрізняються від текстових редакторів додатковою можливістю *форматування тексту*, тобто його оформлення (використанням методів вирівнювання тексту, застосуванням декількох шрифтів і т. п.). Сучасні текстові процесори, крім форматування шрифтів і абзаців та перевірки орфографії, містять можливості, раніше властиві лише настільним видавничим системам, у тому числі створення таблиць і вставку графічних зображень. У цей час більшість текстових редакторів є текстовими процесорами. Найбільш популярні текстові процесори – Microsoft Word, LibreOffice Writer, WordPerfect.

Видавнича система – комплект устаткування для підготовки оригінал-макета видання, готового для передачі в друкарню. До складу видавничої системи входить один або кілька персональних комп'ютерів з необхідним програмним забезпеченням для створення макета оформлення, розпізнання, набору й верстки тексту, редагування зображень, додрукарської підготовки оригінал-макета.

3.2. Можливості текстового процесора

У процесі підготовки документа в розпорядженні користувача є набір засобів і процедур організації роботи з уведення, редагування й форматування тексту та вбудованих об'єктів.

До основних функціональних можливостей для роботи з документом належать:

- використання шрифтів різних розмірів і накреслень символів і різних способів їх виділення;
- установка параметрів абзацу;
- задання міжрядкових інтервалів;
- перевірка правопису, граматики й добір синонімів;
- пошук і заміна символів, слів і фрагментів тексту;
- автоматичний перенос слів на новий рядок;
- автоматична нумерація сторінок;
- друк верхніх і нижніх заголовків сторінок (колонтитулів);
- установка виносок;
- побудова змістів, покажчиків;
- набір тексту в декілька колонок;
- створення таблиць, малюнків і побудова діаграм;
- перегляд документів перед друком;
- установка розмірів паперового носія й параметрів друку;
- скасування й повторення попередніх дій користувача;

- вставка полів з інформацією стандартного типу (дата, час, авторські дані і т. д.);
- створення макрокоманд і гіпертекстових посилань;
- включення в документ різних об'єктів (файлів, формул та ін.);
- імпорт документів, створених в інших програмах і т. д.

Текстовий процесор пропонує також широкий вибір засобів надання документів привабливого зовнішнього вигляду: автоформатування, застосування стилів, бібліотеки стилів і шаблонів документів. За допомогою шаблонів можна автоматизувати процес підготовки стандартних документів, таких, як факсимільні повідомлення, стандартна ділова переписка й документація.

Зупинімося докладніше на деяких із зазначених можливостей.

1. Вибір шрифту. Текстовий процесор дозволяє вибрати будь-який з наявних на комп'ютері шрифтів, зазначити його розмір, або кегль (розмір шрифту звичайно задається в пунктах, 1 пункт дорівнює $1/72$ дюйма = 0,3528 мм), застосовувати різні способи виділення (напівжирний шрифт, курсив, підкреслення, закреслювання і т. д.), задавати колір шрифту та тла. Також може задаватися ширина символів (без зміни висоти) та інтервал між символами.

2. Форматування абзацу. Форматування абзацу містить у собі вказання методу вирівнювання (по лівому або правому краю, по центру, по ширині); задання відступів від лівого й правого краю сторінки, абзацного відступу першого рядка; задання міжрядкових інтервалів і інтервалів між абзацами; налаштування інших властивостей абзацу (наприклад, заборона відриву першого або останнього рядка при переході на наступну сторінку).

3. Стилi. Стилем називається набір параметрів форматування, який застосовують до тексту документа, щоб швидко змінити його зовнішній вигляд. Стилi дозволяють застосувати одночасно всю групу атрибутів форматування (звичайно форматування шрифту й абзацу) до тексту документа. Текстові процесори містять набір вбудованих стилів для створення звичайного тексту, заголовків, списків і т. п., а також дозволяють змінювати існуючі стилі та створювати нові. Застосування стилів дозволяє досягти однаковості у форматуванні різних частин одного документа або різних документів одного типу, а також дає можливість при необхідності швидко переформатувати весь текст документа або окремі його елементи.

4. Створення таблиць. Текстові процесори дозволяють вставляти в документ таблиці. Створення таблиць починається із завдання кількості рядків і стовпців у таблиці. Таблиця складається із комірок, у які можна вводити текст і вставляти об'єкти (символи, формули).

Надалі з таблицею можна робити різні дії: поєднувати й розбивати комірки, рядки, стовпці; формувати текст у комірках, задавати вирівнювання тексту (горизонтальне й вертикальне); змінювати розміри комірок (висоту й ширину), у тому числі підбирати розміри автоматично за вмістом; задавати зовнішній вигляд границь комірок, фрагментів таблиці або всієї таблиці (тип лінії, її товщина, колір, видимість), тло комірок і колір тексту.

5. Створення закладок і гіперпосилань у документі. Текстові процесори дозволяють створювати в документі закладки, тобто вказувати місця в документі, на які можна швидко переходити за посиланням (як на веб-сторінці). Поси-

ланням (гіперпосиланням) є фрагмент тексту, який зв'язується із установленою раніше закладкою. Гіперпосилання автоматично виділяється спеціальним стилем (звичайно іншим кольором і підкресленням). Клацання по посиланню переводить курсор у те місце, де була створена закладка. Гіперпосилання зручно використовувати для створення змістів і покажчиків на конкретні місця в документі.

6. Створення графічних об'єктів. До графічних об'єктів належать лінії (прямі, криві, ламані, стрілки) і фігури (прямокутник, коло, еліпс, трикутник, елементи блок-схем, винесення і т. д.). Для кожної фігури задаються її розміри (висота, ширина), положення на сторінці (відносно полів сторінки, щодо абзацу, безпосередньо в тексті), кут повороту, кольори й шаблони границь і тла, прозорість, додаткові властивості (наприклад, типи стрілок). Готову фігуру можна скопіювати в інше місце в документі. Також є можливість вирівнювати фігури одна щодо одної, відносно сторінки й по сітці, групувати кілька фігур в один об'єкт для подальшого копіювання, розміщати всередині фігур текстові написи.

Крім вбудованих графічних об'єктів, текстові процесори дозволяють вставляти в документ інші типи об'єктів – картинки, діаграми, формули і т. д.

7. Створення шаблону документа. Шаблон – це модель для створення нового документа. У шаблоні зберігаються елементи, що становлять основу документа:

- постійний текст, графічні об'єкти разом з форматуванням;
- поля для введення інформації (текстові, зі списками, дата і т. д.);
- параметри друкованої сторінки документа;
- список доступних стилів;
- макрокоманди (послідовність дій, що автоматизують роботу з документом).

При завантаженні шаблону автоматично створюється новий документ, що містить копію всієї присутньої в шаблоні інформації (текст, графіка, таблиці і т. д.) з форматуванням, а також пропонується введення даних у поля за допомогою діалогових вікон. Шаблони зручно використовувати для створення форм (бланків) стандартних документів.

ТЕМА 4. Бази даних.

Системи управління базами даних

4.1. Основні поняття

Сучасна медицина неможлива без використання баз даних. Наприклад, базою даних локального характеру може бути база обліку пацієнтів, яка асоціюється з поняттям «реєстратура». Тут можуть зберігатися всі історії хвороб, результати аналізів, ЕКГ, рентгенограми й інша інформація, яка може бути доступна лікареві в будь-який момент часу без сторонньої участі. Електронний спосіб ведення обліку пацієнтів дозволяє оперативно вирішувати проблеми передачі інформації в іншу медичну установу (у зв'язку з переїздом або направленням на лікування хворого), забезпечувати захист від несанкціонованого доступу.

Базою даних (БД) називають поєднану сукупність структурованих даних, що належать до певної предметної галузі.

Системою управління базами даних (СУБД) називають комплекс програмних і мовних засобів, необхідних для створення баз даних, підтримки їх в актуальному стані й організації пошуку в них необхідної інформації.

Структуризація даних – це запровадження угод про способи представлення даних. Створюючи базу даних, користувач прагне *впорядкувати* інформацію про різні ознаки об'єктів і швидко отримати вибірку даних з довільним поєднанням ознак. Зробити це можливо, тільки якщо дані *структуровані*.

4.2. Класифікація БД

Існують різні способи класифікації БД, зокрема, за технологією обробки даних, по способу доступу до даних, по ступеню універсальності.

1. **За технологією обробки** даних бази даних поділяються на централізовані й розподілені.

Централізована база даних зберігається в пам'яті однієї обчислювальної системи. Якщо ця обчислювальна система є компонентом комп'ютерної мережі, можливий розподілений доступ до такої бази. Такий спосіб використання баз даних часто застосовують у локальних мережах.

Розподілена база даних складається з декількох частин, які, можливо, перетинаються або навіть дублюють одна одну, збережених на різних комп'ютерах. Робота з такою базою здійснюється за допомогою системи управління розподіленою базою даних.

2. **За способом доступу** до даних бази даних поділяються на автономні бази даних з локальним доступом і централізовані бази даних з віддаленим (мережним) доступом.

Автономні локальні БД є найбільш простими. Вони зберігають свої дані в локальній файлової системі на тому комп'ютері, на якому встановлені. СУБД, що здійснює до них доступ, перебувають на тому ж самому комп'ютері.

Системи **централізованих БД із віддаленим (мережним) доступом** передбачають різні архітектури подібних систем – файл-сервер і клієнт-сервер.

Файл-сервер. Архітектура систем БД із віддаленим (мережним) доступом передбачає виділення однієї з машин мережі як центрального сервера файлів. На такій машині зберігається спільно використовувана централізована БД. Файли бази даних відповідно до користувацьких запитів передаються на робочі станції, де в основному й проводиться обробка.

Клієнт-сервер. У цій концепції передбачається, що, крім зберігання централізованої бази даних, центральна машина (**сервер** бази даних) повинна забезпечувати виконання основного обсягу обробки даних. Запит на дані, що надається клієнтом (робочою станцією), породжує пошук і вибір даних на сервері. Здобуті дані, але не файли, транспортуються мережею від сервера до клієнта.

Специфікою архітектури клієнт-сервер є використання **мови запитів SQL** (Structured Query Language – мова структурованих запитів). Це універсальна мова, призначена для створення й виконання запитів, обробки даних як у власній базі даних програми, так і з базами даних, створеними іншими програмами, що підтримують SQL. Запит мовою SQL складається з одного або декількох операторів, що слідують один за одним і розділені крапкою з комою.

3. **За ступенем універсальності** розрізняють два класи СУБД – системи загального призначення й спеціалізовані системи.

4.3. Інформаційні одиниці БД

Одиницею інформації, що зберігається в БД, є **таблиця**.

Кожна таблиця являє собою сукупність **рядків і стовпців**, де **рядки** відповідають **екземпляру об'єкта** (конкретній події або явищу), а **стовпці** – **атрибутам** (ознакам, характеристикам, параметрам) об'єкта, події, явища. У термінах БД стовпці таблиці називаються **полями**, а її рядки – **записами**.

Об'єктами обробки СУБД є наступні інформаційні одиниці БД.

Поле – елементарна одиниця логічної організації даних, яка відповідає неподільній одиниці інформації – реквізиту.

Запис – сукупність логічно пов'язаних полів.

Таблиця – упорядкована структура, що складається з кінцевого набору однотипних записів.

Первинний ключ – поле (або група полів), що дозволяє однозначним образом визначити кожний рядок у таблиці. Первинний ключ повинен мати дві властивості:

- **однозначна ідентифікація** запису: запис повинен однозначно визначатися значенням ключа;
- **відсутність надмірності**: ніяке поле не можна вилучити із ключа, не порушуючи при цьому властивості однозначної ідентифікації.

4.4. Моделі організації БД

Між полями й записами існують певні зв'язки. Залежно від характеру цих зв'язків існують три види моделей організації баз даних: ієрархічна, мережна й реляційна.

1. **Ієрархічна модель** БД являє собою сукупність елементів, які розташовані у порядку їх підпорядкування від загального до часткового й утворюють перевернене за структурою дерево (**граф**). Вузол цього графа – це сукупність атрибутів даних, що описують деякий об'єкт. Кожний вузол на більш низькому рівні пов'язаний тільки з одним вузлом, що перебуває на більш високому рівні.

2. **Мережна модель** БД – за тих же основних понять (рівень, вузол, зв'язок) кожен елемент може бути пов'язаний з будь-яким іншим елементом.

3. **Реляційна модель** БД є в цей час найпоширенішою. В основі цієї моделі бази даних лежить поняття відношення (relation). Відношення представляються у вигляді двовимірних таблиць.

Відношення (таблиця) представляється в комп'ютері у вигляді файлу даних. Рядок таблиці відповідає запису у файлі даних, а стовпець – полю. У теорії реляційних баз даних рядки називають кортежами, а стовпці – атрибутами. Список імен атрибутів відношення називається схемою відношення. У кожному відношенні виділяють один атрибут, який називають ключовим або просто ключем. Ключовий атрибут повинен бути унікальним, тобто він повинен однозначно визначати (ідентифікувати) кортежі.

Над відношеннями (таблицями) можуть виконуватися різні **операції**, подібно до виконання арифметичних операцій (наприклад, об'єднання таблиць). Це дає можливість одержувати з одних відношень, збережених у комп'ютері, інші відношення.

Реляційна модель баз даних використовується для створення електронних медичних карт.

4.5. Етапи розробки БД

1. **Постановка завдання.** На цьому етапі формується завдання зі створення БД. У ньому докладно описується склад бази, призначення й мета її створення, а також перелічується, які види робіт передбачається здійснювати в цій базі даних.

2. **Аналіз об'єкта.** На цьому етапі розглядається, з яких об'єктів може складатися БД, які властивості цих об'єктів. Усі ці відомості можна розташовувати у вигляді окремих записів і таблиць.

3. **Вибір моделі БД.** На цьому етапі необхідно вибрати вид моделі організації БД. Далі необхідно скласти її схему із зазначенням зв'язків між таблицями або вузлами.

4. **Вибір способів представлення інформації** й програмного інструментарію.

5. **Створення бази даних.** У процесі створення бази даних можна виділити деякі стадії, типові для будь-якої СУБД:

- запуск СУБД, створення нового файлу бази даних;
- створення початкової таблиці або таблиць;
- створення екранних форм, тобто графічного інтерфейсу для введення й відображення даних;

- заповнення БД.

6. **Робота зі створеною базою даних.** Робота із БД містить у собі наступні дії:

- пошук даних;
- сортування даних;
- відбір даних;
- виведення на друк;
- зміна й доповнення даних.

Усі ці дії, як правило, виконуються шляхом подачі запитів у СУБД. **Запит** (query) – це засіб вибору необхідної інформації з бази даних.

Запити в сучасних СУБД, як правило, формуються двома способами – за допомогою графічного інтерфейсу (так званого конструктора запитів) або мовою SQL.

Основні види запитів:

- запит на вибірку, який повертає дані з однієї або кількох таблиць і відображає їх у вигляді таблиці;
- запит з параметрами – запит, що відображає діалогове вікно для введення даних;
- запити на зміну записів, зокрема, на додавання, видалення або оновлення записів;
- запити на створення або видалення таблиць.

ТЕМА 5. Медичні інформаційні системи. Електронні медичні карти

5.1. Основні поняття

Інформаційна система (ІС) – це організаційно впорядкована сукупність документів (масивів документів) та інформаційних технологій (у тому числі, з використанням засобів обчислювальної техніки та зв'язку), що реалізують інформаційні процеси.

Інформаційні системи, що використовуються в медицині й охороні здоров'я, називаються **медичними інформаційними системами (МІС)**.

Існують різні підходи до класифікації МІС. Система охорони здоров'я є багаторівневою структурою й побудована за ієрархічним принципом. Відповідно до цього принципу медичні інформаційні системи поділяються на:

- МІС **базового** рівня;
- МІС рівня **установ**;
- МІС **територіального** рівня.

5.2. МІС базового рівня

Метою МІС базового рівня є комп'ютерна підтримка роботи безпосередньо *медичного працівника* (лікаря-клініциста, гігієніста, лаборанта й ін.).

Групи МІС базового рівня:

- медичні інформаційно-довідкові системи;
- медичні консультативно-діагностичні системи;
- медичні апаратно-програмні комплекси;
- автоматизоване робоче місце (АРМ) лікаря.

Медичні інформаційно-довідкові системи призначені для пошуку й видачі медичної інформації за запитом користувача. Інформаційні масиви таких систем (бази й банки даних) містять медичну *довідкову інформацію* різного характеру. Це й наукова інформація з різних медичних дисциплін, і довідкова статистична й технологічна інформація широкого профілю, і обліково-документальна інформація.

Системи цього класу не здійснюють обробку інформації, але забезпечують швидкий доступ до необхідних відомостей.

Медичні консультативно-діагностичні системи (МКДС) призначені для *діагностики патологічних станів* (включаючи прогноз і формування рекомендацій зі способів лікування) при захворюваннях різного профілю й для різних категорій хворих. Вхідною інформацією для таких систем слугують дані про симптоми захворювання, які вводять у комп'ютер у діалоговому режимі або у форматі інформаційних карт.

У загальному випадку МКДС містить:

- базу даних (БД);
- базу знань (БЗ);
- механізм (машину) логічного висновку (МЛВ);
- інтерфейс із користувачем.

База даних призначена для зберігання сукупності фактів, конкретних даних про об'єкти у сфері діяльності МКДС.

База знань містить знання, що належать до конкретної прикладної галузі, у тому числі окремі факти, правила, а також, можливо, евристики, що відносяться до розв'язання задач у цій прикладній галузі.

Механізм логічного висновку з використанням правил і методів БЗ перетворює конкретну інформацію про об'єкт до виду, що відповідає до призначення МКДС (діагноз, план дій і т. п.).

Інтерфейс із користувачем забезпечує безперебійний обмін інформацією між користувачем і системою; він також дає користувачеві можливість спостерігати за процесом розв'язку задач, що протікають у МЛВ.

За способом реалізації МЛВ розрізняють **імовірнісні** МКДС і **експертні** МКДС, тобто в основі МЛВ можуть лежати методи теорії ймовірностей або методи штучного інтелекту.

Медичні апаратно-програмні комплекси (МАПК) призначені для інформаційної підтримки та/або автоматизації діагностичного й лікувального процесу, здійснюваного за умови безпосереднього контакту з організмом хворого або об'єктом дослідження.

По призначенню МАПК можуть бути розділені на ряд класів:

- системи для проведення функціональних і морфологічних досліджень;
- моніторні системи;
- системи управління лікувальним процесом;
- системи лабораторної діагностики;
- системи для медико-біологічних досліджень.

Автоматизоване робоче місце (АРМ) лікаря – комп'ютерна інформаційна система, що призначена для автоматизації всього технологічного процесу лікаря відповідної спеціальності та забезпечує інформаційну підтримку під час прийняття діагностичних і тактичних (лікувальних, організаційних і ін.) лікарських рішень. Усі розглянуті вище інформаційні системи клінічного рівня можуть і повинні входити в структуру АРМ, забезпечуючи автоматизацію всього технологічного процесу медика.

5.3. МІС рівня лікувально-профілактичних установ (ЛПУ)

Інформаційні системи цього рівня призначені для роботи з **інформаційними потоками ЛПУ**. Вони представлені наступними групами:

- ІС консультативних центрів;
- банки інформації медичних установ і служб;
- персоніфіковані реєстри;
- скринінгові системи;
- інформаційні системи ЛПУ;
- ІС для НДІ й вузів.

ІС консультативних центрів призначені для забезпечення функціонування відповідних підрозділів та інформаційної підтримки лікарів під час консультування, діагностики та прийняття рішень при невідкладних станах.

Банки інформації медичних установ і служб містять зведені дані про якісний і кількісний стан працівників установи, прикріпленого населення; основні статистичні відомості, характеристики районів обслуговування й інші необхідні відомості.

Персоніфіковані реєстри (бази й банки даних) – різновид інформаційно-довідкових систем, що містять інформацію на прикріпленій або спостережуваний контингент на основі формалізованої історії хвороби або електронної медичної карти.

Скринінгові системи призначені для проведення долікарського профілактичного огляду населення, а також для лікарського скринінгу (формування груп ризику й виявлення хворих, що потребують допомоги фахівця).

Інформаційні системи ЛПУ – інформаційні системи, які побудовані на об'єднанні всіх інформаційних потоків у єдину систему і забезпечують автоматизацію різних видів діяльності установи.

ІС для НДІ та ВНЗ призначені для інформатизації процесу навчання, науково-дослідної роботи й управлінської діяльності науково-дослідних інститутів (НДІ) та вищих навчальних закладів (ВНЗ).

5.4. МІС територіального рівня

МІС територіального рівня забезпечують управління спеціалізованими й профільними медичними службами, поліклінікою, стаціонарною та швидкою допомогою населенню на рівні територій (міста, області, держави). Такі МІС можуть бути й загальносвітовими (глобальними).

На цьому рівні МІС представлені наступними групами:

- **адміністративно-управлінські** МІС для керівників територіальних медичних служб;
- **статистичні** МІС для роботи зі зведеною інформацією по території;
- **МІС спеціалізованих служб** і напрямів: швидкої допомоги й надзвичайних ситуацій, лікарського забезпечення, реєстри (фтизіатрія, психіатрія, інфекційні хвороби і т. д.);
- **комп'ютерні телекомунікаційні мережі**, що створюють єдиний інформаційний простір у сфері охорони здоров'я.

5.5. Електронна медична карта

Електронна медична карта (електронна історія хвороби; електронний паспорт пацієнта; electronic medical record – EMR) – це медична карта пацієнта медичного закладу в електронній формі.

Електронна історія хвороби може використовуватися як у поліклінічних, так і в стаціонарних установах, з урахуванням характеру й особливостей надання медичної допомоги в них. Складається та зберігається в автоматизованій інформаційній базі даних медичної установи. Карта містить дані *історії хвороб* пацієнта, інформацію про зроблені *щеплення* і його бажання стати *донором*.

Електронні медичні карти повинні поступово замінити традиційні картки медичного страхування. Однією з перших країн Євросоюзу, яка з жовтня 2011 р. впровадила електронні медичні карти, стала Німеччина.

Електронна медична карта є єдиним інформаційним ресурсом, що дозволяє оперувати особистими даними пацієнтів, а також обмінюватися такими даними з іншими медичними установами для складання, обліку й зберігання медичної інформації. Медична документація (інформація) з картами може передаватися до компетентних організацій: страхових компаній, органів контролю за наданням медичної допомоги, правоохоронних органів і т. д.

В Україні електронна медична карта почала працювати в тестовому режимі з 1 березня 2019 р. Вона доступна сімейним лікарям, терапевтам і педіатрам тих лікувальних закладів, які підключені до електронної системи охорони здоров'я й мають контракт із Національною службою здоров'я України.

Електронна медична карта є інформаційною системою, яка спирається на технологію *реляційних баз даних*. Такий спосіб зберігання інформації дозволяє зручно, в автоматичному режимі, відбирати дані за деякою ознакою або набором ознак, упорядковувати їхнє відображення за різними стовпцями таблиці, наприклад, за датою, прізвищем, діагнозом.

Електронна форма медичної карти полегшує вирішення багатьох завдань:

- документування (накопичення, надійне зберігання, можливість зручного перегляду) довільної медичної інформації про пацієнта із прив'язкою до календарної дати;
- пошук (фільтрацію за набором ознак) необхідної інформації;
- відстеження залежності окремих діагностичних параметрів від часу;
- дослідження ефективності роботи.

Така інформаційна система передбачає наявність зручних *засобів введення й відображення інформації*. Для введення текстової й цифрової інформації надаються *форми* (бланки), що відображаються на екрані комп'ютера. За можливості формалізації даних (формування скінченного або доповнюваного списку значень параметра, що вводиться) вибір значення робиться з наданого списку. У систему можуть додаватися різноманітні *електронні довідники* (класифікатор діагнозів, лікарських засобів, адресна книга і т. д.). Передбачаються засоби *імпорту даних апаратних обстежень*. Багато сучасних медичних приладів (томографи, флюорографи, ультразвукові сканери, електрокардіографи і т. д.) формують результат безпосередньо в електронному вигляді. Існують міжнародні стандарти, наприклад, DICOM (Digital Imaging and Communications in Medicine), для приладової діагностичної інформації. Для захисту даних від випадкової втрати в результаті аварії обладнання або навмисного злому сховища даних такі інформаційні системи розміщуються на добре захищених серверах, що мають засобу резервного копіювання даних і журналювання стану.

ТЕМА 6. Прийоми роботи з медичною інформацією за допомогою табличного процесора

6.1. Основні поняття

При обробці різних видів інформації (у т. ч. медичної) **таблиці** й діаграми є зручними візуальними формами представлення даних. Таблиці й діаграми дозволяють зберігати й аналізувати інформацію, приймати рішення відповідно до результатів обробки даних.

Програмні системи, використовувані для обробки даних у формі таблиць, називаються **табличними процесорами**, або **електронними таблицями**. До цього класу програм належать Microsoft Excel, LibreOffice Calc, Quattro Pro і багато інші. У даному розділі буде розглядатися табличний процесор LibreOffice Calc.

Після завантаження табличного процесора (або створення нового файлу в ньому) на екрані з'являється **електронна таблиця** (Worksheet). В одному файлі може міститися кілька таких таблиць, що звичайно називаються **аркушами**.

Як і в СУБД, у табличних процесорах будь-яка таблиця складається з **рядків** і **стовпців**. У більшості табличних процесорів, у тому числі в Calc, стовпці позначаються однією, двома або трьома латинськими буквами (А, В, ..., Z, АА, АВ, ... ZZ, ААА, ...), а рядки – номерами.

Стовпці Рядки	А	В	С	Д	Е	F	G
1							
2							
3							
4							
5							

6.2. Комірка і її характеристики

Дані в таблицях зберігаються в **комірках**. Кожна комірка характеризується адресою, форматом і статусом.

Адреса комірки, як правило, записується комбінацією імені стовпця й номера рядка, наприклад, А1, В20, АХ157.

Формат комірки визначає відображення вмісту комірки в таблиці. Він містить у собі параметри форматування числових значень (а також дати й часу), шрифт, вирівнювання, зовнішній вигляд границь і колір тла.

Статус комірки дозволяє захистити комірку від зміни, сховати формулу, сховати комірку при відображенні або під час друку і т. д.

Вміст комірки. Комірка в LibreOffice Calc може містити три типи даних: текстові, числові, формули.

Текстові дані можуть використовуватися для заголовків, пояснень і т. д. і являють собою рядок тексту довільної довжини.

Числові дані можуть бути представлені у вигляді цілих чисел (наприклад, 1; -6), десяткових чисел (наприклад, 10,6; -0,8) або чисел в експонентній формі (5,6E-4).

Формули являють собою інструкції для обчислень і складаються з операндів, пов'язаних символами арифметичних операцій (наприклад, «+», «-» і т. д.). **Операндами** можуть бути числа (цілі, десяткові, в експонентній формі), адреси комірок і функції (арифметичні й тригонометричні, дати й часу і т. д.).

При введенні в комірку формули обов'язково повинні **починатися зі знака рівності (=)**.

З поняттям формули тісно пов'язане поняття **значення комірки**. Значення комірки є результатом операцій, що виконуються табличним процесором на підставі вмісту. Наприклад, якщо вмістом комірки є формула $= 5 * 2 + 1$, то її значенням буде число 11. Значення комірки відображається в самій комірці, а її вміст – у рядку введення.

6.3. Відносні й абсолютні адреси комірок

Адреси комірок у формулах (називані звичайно **посиланнями**) можуть бути відносними й абсолютними. Відмінність між відносними й абсолютними посиланнями є істотною в разі копіюванні формули з однієї комірки в іншу.

За замовчуванням, LibreOffice Calc використовує **відносні** посилання. При копіюванні формули з відносним посиланням таке посилання автоматично *коригується* відповідно до нового місця розташування, тобто залишається незмінним *відносно розташування* комірки з формулою й комірки, на яку вказує посилання. При цьому посилання, що міститься в скопійованій або переміщеній формулі, належать до нової комірки. Відносні посилання зручно використовувати для застосування формули методом автозаповнення до великого масиву даних, наприклад, до рядка, стовпця або прямокутного діапазону комірок.

При копіюванні формули з **абсолютним** посиланням LibreOffice Calc копіює абсолютне посилання в тій же формі, як воно виглядає у вихідній формулі. Абсолютні посилання використовуються, зокрема, для вставки у формули констант, що зберігаються в окремих фіксованих комірках, або даних з певного рядка або стовпця.

У формулах LibreOffice Calc абсолютне посилання створюється додаванням *знака долара «\$»* перед адресою. Він може стояти перед іменем стовпця (буквами), перед номером рядка (цифрами) або в обох місцях, тобто кожний елемент посилання може бути абсолютним або відносним *незалежно* від іншого:

Посилання	Стовпець А	Рядок 1
\$A\$1	Не змінюється	Не змінюється
A\$1	Змінюється	Не змінюється
\$A1	Не змінюється	Змінюється
A1	Змінюється	Змінюється

Наведемо приклад. Нехай у комірці В3 поміщена формула з посиланням на комірку А1. Потім ця формула копіюється або переміщується на одну комірку вправо й униз, тобто в комірку С4. При цьому, якщо формула містить відносні посилання, то вони також зсунуться: $A \rightarrow B, 1 \rightarrow 2$, а абсолютні посилання залишаться незмінними. Результат копіювання залежно від типу посилання буде таким:

Вихідна формула в В3	Формула після копіювання в С4
=A\$1	=A\$1
=A\$1	=B\$1
=\$A1	=\$A2
=A1	=B2

Посилання в електронних таблицях можуть указувати не тільки на одну комірку, але й на **діапазон комірок**. Діапазоном комірок може бути:

- рядок або її частина;
- стовпець або його частина;
- кілька рядків або стовпців;
- прямокутна область.

Діапазон комірок задається зазначенням *двох адрес однакового типу* (стовпців, рядків або комірок) через символ «:» у такий спосіб:

- кілька рядків (наприклад, з 1 по 5 включно): **1:5**;
- один рядок (наприклад, рядок 2): **2:2** (саме так, щоб відрізнити від числового значення!);
- частина рядка: **A1:D1**;
- кілька стовпців (від стовпця В до стовпця F включно): **B:F**;
- один стовпець (наприклад, стовпець D): **D:D**;
- частина стовпця: **B2:B5**;
- прямокутна область (вказуються адреси комірок, розташованих у лівому верхньому й правому нижньому куті області): **B1:G5**.

Посилання при описі діапазонів комірок, як і для окремих комірок, можуть бути як відносними, так і абсолютними. У багатьох випадках, наприклад при використанні «Майстра функцій», діапазони комірок можна вказувати, виділяючи їх безпосередньо мишею в таблиці.

6.4. Функції в табличних процесорах

Для проведення обчислень та інших видів обробки даних табличні процесори забезпечені набором вбудованих **функцій**. Функції поділяються на категорії: математичні, статистичні, фінансові, логічні, текстові, для роботи з датою й часом і т. д.

Кожна функція має унікальне **ім'я** й один або кілька **аргументів**, розділених символом «;». Аргументами функцій можуть бути:

- фіксовані значення (числові або текстові);
- посилання на комірки (відносні або абсолютні);
- діапазони комірок;
- інші функції (вкладені функції).

Для вставки функцій у формулу можна ввести ім'я функції та її аргументи вручну або скористатися спеціальним діалоговим вікном «**Майстер функцій**», яке дозволяє знайти функцію за категорією, за описом або за іменем, переглянути опис функції зі списком її аргументів і ввести аргументи в призначені для цього поля. На малюнку показаний приклад використання «Майстра функцій» для введення логічної функції IF (якщо).

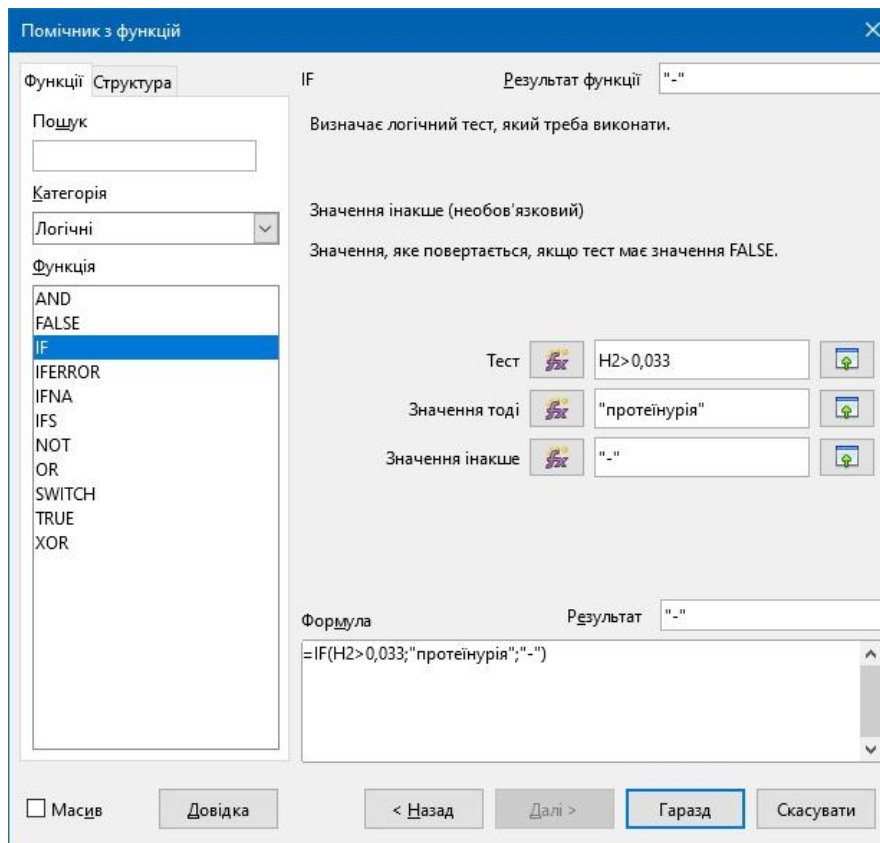


Рис. 4. Вікно «Майстер функцій»

6.5. Побудова діаграм

Сучасні табличні процесори дозволяють відображати дані з таблиць **графічно** у вигляді **діаграм**. В LibreOffice Calc існують різні типи діаграм: стовпчаста й стрічкова, кругова, точкова, у вигляді ліній (графік) і т. д.

Для створення діаграми призначений спеціальний інструмент **«Майстер діаграм»**. Щоб побудувати діаграму, необхідно вказати наступні відомості:

- один або кілька **діапазонів** комірок із даними (вимоги до діапазонів залежать від типу діаграми);
- **тип** діаграми;
- розташування даних у діапазонах (ряди даних по вертикалі або по горизонталі), наявність або відсутність підписів у таблиці;
- кількість, порядок і властивості рядів даних (якщо їх кілька);
- інші елементи діаграми – заголовки, підписи осей, зовнішній вигляд сітки, легенда.

Засоби побудови діаграм в LibreOffice Calc також дозволяють проводити обробку даних безпосередньо на діаграмі, зокрема, знаходити середні значення, будувати лінії тренду (регресії), зазначати похибки і т. д.

ТЕМА 7. Методи біостатистики.

Статистичний аналіз медико-біологічних даних

7.1. Основні поняття

Математична статистика – це наука, що розробляє математичні методи систематизації й використання статистичних даних для наукових і практичних виводів. Ця мета досягається вирішенням двох основних завдань.

1. **Визначення способів збору та групування** статистичних відомостей, отриманих у результаті спостережень або в результаті спеціально поставлених експериментів.

2. **Розробка методів аналізу** статистичних даних залежно від цілей дослідження. До методів аналізу даних належать:

- **оцінка** невідомої ймовірності події; оцінка невідомої функції розподілу; оцінка параметрів розподілу, вид якого відомий; оцінка залежності випадкової величини від однієї або декількох випадкових величин і ін.;
- **перевірка статистичних гіпотез** про вид невідомого розподілу або про величину параметрів розподілу, вид якого відомий.

Випадкова величина – величина, яка приймає в результаті експерименту одне з множини можливих значень, причому поява того або іншого значення цієї величини являє собою випадкову подію.

Дискретною випадковою величиною називається випадкова величина з кінцевою або зліченною множиною можливих значень. Як правило, дискретною випадковою величиною описується кількість об'єктів (подій, явищ), порядковий номер об'єкта в списку й т. п.

Неперервною (безперервною) випадковою величиною називається випадкова величина, яка може набувати будь-якого зі значень, що належать до інтервалу (або інтервалів), у якому вона існує. Неperервна випадкова величина завжди має нескінченну множину значень. Більшість фізичних величин (маса, довжина, концентрація і т. д.) є неперервними.

Першим кроком статистичного аналізу є **класифікація типу даних**, тобто віднесення їх до тієї або іншої **шкали вимірів**.

Шкали вимірів класифікують:

- 1) за типом випадкової величини – **неперервні** (температура, показник гемоглобіну в крові) і **дискретні** (результат захворювання, група крові).
- 2) за дозволеними операціями зі значеннями – **номінальні**, **порядкові**, **інтервальні**, **абсолютні**.

Номінальна шкала (шкала найменувань) використовується для угруповання об'єктів за якісною ознакою (наприклад, за кольором, статтю, групою крові). Ця шкала дає можливість порівнювати об'єкти за такою ознакою (тобто для неї визначені тільки співвідношення еквівалентності $x = y$ й $x \neq y$). Номінальна шкала не передбачає впорядкування значень і кількісних взаємозв'язків між значеннями.

Порядкова (рангова) шкала є впорядкованою послідовністю значень і дозволяє не тільки встановити факт рівності або нерівності вимірюваних об'єктів, але й визначити характер нерівності у вигляді суджень: «більше-менше», «вище-

нижче», «гірше-краще» і т. п. Для порядкової шкали визначені співвідношення $x = y, x \neq y, x < y, x > y$, але не визначені арифметичні операції (додавання, віднімання і т. д.). Приклади порядкових шкал: оцінка успішності (незадовільно, задовільно, добре, відмінно), підсумок лікування пацієнта (видужання, поліпшення стану, без поліпшення, погіршення стану, летальний наслідок).

Інтервальна шкала (шкала різниць) дозволяє описувати властивості об'єкта кількісно шляхом порівняння ознаки об'єкта з еталоном. Для інтервальних шкал мають сенс усі вищезгадані операції порівняння й арифметичні операції додавання та віднімання. Початок відліку для такої шкали задається довільно. Приклад інтервальної шкали – шкала Цельсія, на якій початок відліку обраний умовно по температурі танення льоду.

Абсолютна шкала (шкала відношень) – це інтервальна шкала, у якій є додаткова властивість – природня й однозначна присутність нульової точки. Наприклад, число людей в аудиторії. Це єдина з чотирьох шкал, що має абсолютний нуль. Нульова точка характеризує відсутність вимірюваної якості. За допомогою таких шкал можуть вимірюватися фізичні величини – маса, довжина, сила і т. д. Прикладом абсолютної шкали є шкала Кельвіна, в якій температура відлічується від абсолютного нуля. Операції множення й ділення мають сенс тільки для абсолютної шкали.

7.2. Розподіли випадкових величин

Для визначення випадкової величини необхідно вказати **закон розподілу** цієї величини, який може бути заданий у вигляді таблиці, формули або графіка.

Закон розподілу дискретної випадкової величини X задається у вигляді таблиці, у якій перелічуються всі можливі значення цієї величини x_i й відповідні ймовірності появи цих значень $P(x_i)$.

Закон розподілу безперервної випадкової величини X задається **функцією розподілу** $F(x) = P(-\infty < X \leq x)$, яка дорівнює ймовірності того, що випадкова величина X прийме значення, менше або рівне x . Другий спосіб задання розподілу безперервної випадкової величини – **щільність імовірності** $f(x) = \frac{dF(x)}{dx}$.

З поняттям функції розподілу тісно зв'язаний поняття квантиля. **Квантиль** (x_α) – це значення, яке задана випадкова величина X не перевищує з фіксованою ймовірністю α . Для неперервних випадкових величин α -квантилем (квантилем рівня α) розподілу $F(x)$ є розв'язок рівняння $F(x_\alpha) = \alpha$. Квантиль можна розглядати як функцію, зворотну функції розподілу, тобто аргумент функції розподілу, для якого її значення дорівнює α . Квантилі нормального розподілу широко використовуються в статистиці для побудови інтервальних оцінок характеристик розподілу.

Квантилі з рівнями, кратними 0,25, називаються **квартілями**:

- квантиль рівня 0,25 – нижній (перший) квантиль,
- квантиль рівня 0,5 – **медіана** (другий квантиль),
- квантиль рівня 0,75 – верхній (третій) квантиль.

Мінімальне й максимальне значення випадкової величини іноді називають нульовим і четвертим квантилями відповідно.

Закон розподілу випадкової величини є найбільш повною характеристикою цієї випадкової величини, однак найчастіше достатньо інформації, яка задається за допомогою **числових характеристик** випадкової величини. Серед них найбільше часто використовуються: математичне очікування, дисперсія, середнє квадратичне (стандартне) відхилення.

Математичне очікування $M(X)$ випадкової величини X має сенс середнього значення цієї величини й обчислюється за допомогою формул

$$M(X) = \sum_{i=1}^N x_i P(x_i); \quad M(X) = \int_{-\infty}^{+\infty} x f(x) dx$$

для дискретних і неперервних випадкових величин відповідно.

Дисперсія $D(X)$ характеризує величину розкиду значень x_i випадкової величини X навколо її математичного очікування $M(X)$. Дисперсія обчислюється як математичне очікування квадрата відхилення значень випадкової величини від $M(X)$:

$$D(X) = \sum_{i=1}^N (x_i - M(X))^2 P(x_i); \quad D(X) = \int_{-\infty}^{+\infty} (x - M(X))^2 f(x) dx.$$

Середнє квадратичне відхилення (стандартне відхилення) $\sigma(X)$ дорівнює кореню квадратному з дисперсії: $\sigma(X) = \sqrt{D(X)}$.

Основні види розподілів, що використовуються у математичній статистиці: для дискретних величин – біноміальний (розподіл Бернуллі) і розподіл Пуассона, для неперервних – нормальний розподіл (розподіл Гаусса).

Біноміальний розподіл, або **розподіл Бернуллі** – розподіл числа появ деякої події (A) у серії з n незалежних випробувань, якщо в кожному із цих випробувань імовірність події p постійна. Імовірність появи значення m ($0 \leq m \leq n$) має вигляд

$$P(m) = C_n^m p^m q^{n-m},$$

де $q = 1 - p$, $C_n^m = \frac{n!}{m!(n-m)!}$.

Для цього розподілу $M(X) = np$, $D(X) = npq$.

Розподіл Пуассона (закон рідкісних подій) – граничний випадок розподілу Бернуллі при $n \rightarrow \infty$, $p \rightarrow 0$ і при постійному добутку $np = a$.

$$P(m) = \frac{a^m}{m!} e^{-a}.$$

Для розподілу Пуассона $M(X) = D(X) = a$.

Більшість неперервних фізичних величин у природі мають розподіл, близький до нормального.

Нормальний розподіл (розподіл Гаусса) задається формулою

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}},$$

де a та σ – параметри розподілу. Для нормального розподілу $M(X)=a$, $D(X)=\sigma^2$.

Нормальний розподіл з параметрами $a=0$ і $\sigma=1$ називається **стандартним**

і має вигляд $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

7.3. Вибірковий розподіл. Графічне відображення розподілу

Нехай потрібно вивчити кількісну ознаку **генеральної сукупності**, тобто найбільшої сукупності, елементи якої мають хоча б одну загальну властивість. Припустимо, що з теоретичних міркувань удалося встановити, який саме розподіл має досліджувана ознака. Природно, виникає задача оцінки параметрів, якими визначається цей розподіл.

Для оцінки параметрів генеральної сукупності із цієї сукупності шляхом проведення експериментів отримується деякий набір з n значень випадкової величини. Ця сукупність значень називається **вибіркою** обсягу (об'єму) n .

Розподіл значень випадкової величини у вибірці можна представити у вигляді таблиці, що містить усі значення величини x_i (звичайно відсортовані за зростанням) і числа спостережень кожного значення (частоти) f_i . Такий вигляд розподілу називають **варіаційним рядом**.

Статистичний розподіл, представлений варіаційним рядом, можна наочно зобразити **графічно**, зокрема, у вигляді полігона, гістограми й кумуляти.

Полігон і гістограма варіаційного ряду є графіками емпіричної щільності ймовірності. Відзначимо, що **сума площ** усіх стовпців, що утворюють гістограму, дорівнює **одиниці**.

Кумулята являє собою графік емпіричної функції розподілу випадкової величини.

Статистичні графіки дають можливість оцінити вид розподілу й перейти до знаходження параметрів цього розподілу. Зокрема, якщо розподіл близький до нормального, необхідно знайти вибіркові оцінки математичного очікування й дисперсії. Вони виражаються формулами:

$$\hat{M}(X) = \bar{x} = \frac{1}{n} \sum_{i=1}^k f_i x_i ; \quad \hat{D}(X) = \hat{S}^2 = \frac{1}{n-1} \sum_{i=1}^k (x_i - \bar{x})^2 \cdot f_i .$$

Крім цих величин, для вибіркового розподілу (особливо якщо він не відповідає нормальному) часто шукають такі характеристики, як мода (значення, що найчастіше зустрічається у вибірці), медіана, квантілі.

7.4. Перевірка статистичних гіпотез

Однією із завдань математичної статистики є перевірка статистичних гіпотез, які формулюються під час вивчення деякої вибірки значень x_1, x_2, \dots, x_n випадкової величини X .

Статистична гіпотеза (позначимо її H) – це гіпотеза про передбачуваний вид досліджуваного розподілу ймовірностей або про значення параметрів цього розподілу. Наприклад, можуть перевірятися гіпотези про незалежність двох випадкових величин, про рівність параметрів розподілів і ін.

Алгоритм перевірки статистичної гіпотези.

1. За вибіркою x_1, x_2, \dots, x_n обчислюється спеціальна функція величин x_1, x_2, \dots, x_n , яка називається **статистикою** (T): $T = T(x_1, x_2, \dots, x_n)$.

2. Формулюється **критерій статистичної гіпотези**, тобто правило, що дозволяє відкинути або прийняти гіпотезу H на підставі значень вибірки x_1, x_2, \dots, x_n . Критерій визначає *критичну область значень* статистики T .

3. Гіпотеза H відкидається, якщо значення T належить критичній області, і ухвалюється в протилежному випадку.

Описане правило прийняття або відкидання гіпотези не визначає однозначно правильність або помилковість гіпотези. Тут можливі чотири випадки.

1. Гіпотеза H вірна та ухвалюється згідно із критерієм. Імовірність того, що гіпотеза H вірно ухвалена, називається *довірчою ймовірністю* (α).

2. Гіпотеза H вірна, але відкидається згідно із критерієм. Цей випадок називається *помилкою першого роду*. Імовірність помилки першого роду (P) називається *рівнем значущості* критерію й дорівнює $P = 1 - \alpha$.

3. Гіпотеза H невірна, але ухвалюється згідно із критерієм. Цей випадок називається *помилкою другого роду*. Імовірність цієї помилки позначається β .

4. Гіпотеза H невірна й відкидається згідно із критерієм. Імовірність цього випадку дорівнює $1 - \beta$ й називається *статистичною потужністю*.

Перераховані випадки можна звести в таблицю.

Результат застосування критерію	Гіпотеза H	
	Дійсна	Неправильна
Гіпотеза ухвалена	Гіпотеза вірно ухвалена	Гіпотеза невірно ухвалена (<i>помилка другого роду</i>)
Гіпотеза відкинута	Гіпотеза невірно відкинута (<i>помилка першого роду</i>)	Гіпотеза вірно відкинута

Перевірка гіпотези про параметри нормального розподілу.

У разі перевірки статистичних гіпотез із відомим видом розподілу випадкової величини діють у такий спосіб: у таблицях розподілу статистики T знаходять критичне значення T_0 , що залежить від заздалегідь заданого рівня значущості P (у медико-біологічних дослідженнях він звичайно береться рівним 0,01 або 0,05), і перевіряють нерівність $T \leq T_0$. Якщо нерівність виявляється вірною, то гіпотеза H ухвалюється. Якщо з'ясується, що $T \geq T_0$, то гіпотеза H відкидається.

Цим методом перевіряється гіпотеза про те, що незалежні результати спостережень x_1, x_2, \dots, x_n підкоряються нормальному розподілу із середнім значенням $a = a_0$ за відомої дисперсії σ^2 . Для перевірки цієї гіпотези виконують наступні дії.

1. Знаходиться середнє вибіркове

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

2. Обчислюється статистика T за формулою

$$T = \sqrt{n} \frac{\bar{x} - a_0}{\sigma}$$

3. При заданому рівні значимості P за відповідними таблицями нормального розподілу знаходиться критичне значення T_0 .
4. Якщо виявиться, що $T > T_0$, то висунута гіпотеза про те, що вибірка взята з генеральної сукупності із середнім значенням a_0 , відкидається, у протилежному випадку гіпотеза ухвалюється.

Перевірка гіпотези про значимість відмінності двох середніх вибірових

Така гіпотеза перевіряється аналогічним методом. Порівнюються середні значення (вибірові середні) двох вибірок, що мають нормальний закон розподілу. Відповідний критерій називається **критерієм Стьюдента**.

Нехай є дві вибірки $X_1\{x_{1i}\}$ й $X_2\{x_{2i}\}$ обсягами відповідно n_1 й n_2 . Для цих вибірок виконують наступні обчислення:

1. Обчислюються середні значення \bar{x}_1 , \bar{x}_2 і дисперсії S_1^2 , S_2^2 .
2. Статистика Стьюдента обчислюється за допомогою формули

$$T = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

3. Розподіл Стьюдента має додатковий параметр, названий числом ступенів свободи. У випадку $n_1 = n_2 = n$ число ступенів свободи обчислюється за формулою

$$v = n - 1 + \frac{2n - 2}{\frac{S_1^2}{S_2^2} + \frac{S_2^2}{S_1^2}}$$

4. По таблицях критичних значень критерію Стьюдента для заданого значення рівня значимості P (або довірчої ймовірності $\alpha = 1 - P$) і знайденого числа ступенів свободи v шукається критичне значення T_c .
5. Якщо $T > T_c$, гіпотеза про значиму відмінність вибірових середніх ухвалюється із заданим рівнем значущості, в протилежному випадку – відкидається.

7.5. Кореляційна залежність між випадковими величинами. Коефіцієнти кореляції Пірсона та Спірмена

Кореляційна залежність (кореляція) – статистичний взаємозв'язок двох або більше випадкових величин, при якій зміни значень однієї або декількох із цих величин приводять до систематичної зміни значень іншої або інших величин. Математичною мірою кореляції двох випадкових величин є **коефіцієнт кореляції R** .

Метод обчислення коефіцієнта кореляції залежить від виду шкали, до якої належать змінні. Так, для виміру змінних з кількісною шкалою використовується коефіцієнт кореляції Пірсона. Якщо, щонайменше, одна із двох змінних має порядкову шкалу або не є нормально розподіленою, необхідно використовувати рангову кореляцію Спірмена.

Лінійний коефіцієнт кореляції Пірсона $R(X, Y)$ двох випадкових величин X і Y обчислюється за формулою

$$R(X, Y) = \frac{M \left[\begin{matrix} x - M(X) & y - M(Y) \end{matrix} \right]}{S(X)S(Y)} = \frac{M(XY) - M(X)M(Y)}{S(X)S(Y)}.$$

Коефіцієнт кореляції може набувати значень в інтервалі $[-1; 1]$. Якщо випадкові величини є незалежними, коефіцієнт кореляції дорівнює нулю. Чим ближче значення $|R|$ до одиниці, тим більше підстав уважати кореляційну залежність між випадковими величинами лінійною функцією. Знак коефіцієнта кореляції показує характер цієї функції: при $R > 0$ вона зростаюча (позитивна кореляція), при $R < 0$ – спадна (негативна кореляція).

Коефіцієнт рангової кореляції Спірмена (r_s) застосовується для порівняння двох величин з довільним законом розподілу або для порівняння якісних показників з порядковою шкалою. Він обчислюється в такий спосіб.

1. Значення x_i й y_i випадкових величин X і Y сортуються (окремо) за зростанням.
2. Кожному з них надається **ранг** R_{xi}, R_{yi} , тобто порядковий номер у відсортованій послідовності. Якщо серед значень якої-небудь із величин є повторювані, їх ранги усереднюються.
3. Для кожної пари значень (x_i, y_i) обчислюється різниця рангів $R_{xi} - R_{yi}$.
4. Коефіцієнт кореляції Спірмена обчислюється за допомогою формули

$$r_s = 1 - \frac{6 \sum_{i=1}^n (R_{xi} - R_{yi})^2}{n^3 - n},$$

де n – обсяг вибірки.

Як і для коефіцієнта кореляції Пірсона, значення коефіцієнта Спірмена змінюється від -1 (послідовності рангів повністю протилежні) до $+1$ (послідовності рангів повністю збігаються). Нульове значення показує, що випадкові величини незалежні.

7.6. Непараметричні методи перевірки статистичних гіпотез.

Критерій Вілкоксона й Манна – Уїтні

Критерій Вілкоксона (W), або критерій знакових рангів Вілкоксона – непараметричний статистичний тест (критерій), що використовується для перевірки відмінностей між двома вибірками парних або незалежних вимірів за рівнем якої-небудь кількісної ознаки, вимірюваної в неперервній або в порядковій шкалі.

Суть методу полягає в тому, що зіставляються абсолютні величини вираженості зсувів (різниці значень двох величин) у тому або іншому напрямку.

Для застосування критерію Вілкоксона необхідно:

1. Обчислити різниці $y_i - x_i$ між індивідуальними значеннями в другому й першому вимірах; знайти їхні абсолютні величини $|y_i - x_i|$ й знаки $\text{sgn}(y_i - x_i)$. Нульові різниці виключаються з вибірки.

2. Відсортувати абсолютні величини різниць по зростанню, призначивши їм ранги R_i .

3. Обчислити статистику $W = \sum_{i=1}^N \text{sgn}(y_i - x_i)R_i$.

4. Визначити по таблицях критичні значення $W_c(N)$ для даного обсягу вибірки. Якщо $|W| > W_c$ – зсув в одну сторону вірогідно переважає, тобто результати двох вимірів є різними.

Існує інший варіант цього критерію (T -критерій Вілкоксона). Статистикою є менша із сум рангів, обчислених окремо для різних знаків зсувів. У цьому варіанті гіпотеза про відмінність ухвалюється при $T < T_c$.

Критерій Манна – Уїтні (U) є аналогом критерію Вілкоксона для двох незалежних вибірок. Для застосування критерію необхідно:

1. Скласти єдиний ранжований ряд з обох вибірок, що зіставляються (обсягами n_1 та n_2), розставивши їх елементи за зростанням та приписавши меншому значенню менший ранг.

2. Розділити ранжований ряд на два ряди, що складаються відповідно зі значень першої й другої вибірок. Підрахувати окремо суму рангів, що припадають на частку елементів першої вибірки, і окремо – на частку елементів другої вибірки. Визначити більшу із двох рангових сум (R_x), що відповідає вибірці з n_x елементами.

3. Визначити значення U -критерію Манна – Уїтні за формулою

$$U = n_1 n_2 + \frac{n_x(n_x + 1)}{2} - R_x.$$

4. У таблиці для вибраного рівня значущості P визначити критичне значення критерію для даних n_1 і n_2 . Якщо отримане значення U менше за табличне або дорівнює йому, то визнається наявність істотної відмінності між рівнем ознаки в розглянутих вибірках.

ТЕМА 8. Кластерний аналіз у медичних дослідженнях

8.1. Основні поняття

Кластерний аналіз (*cluster analysis*) – багатовимірна статистична процедура, що виконує збір даних, які містять інформацію про вибірку об'єктів, й потім упорядковує об'єкти в порівняно однорідні групи. Інакше кажучи, кластерний аналіз – це сукупність методів і алгоритмів класифікації об'єктів.

Загальне питання, що задається дослідниками в багатьох областях, полягає в тому, як *організувати* спостережувані дані в наочні структури – таксономії. Математично *таксономією* є деревоподібна структура класифікацій певного набору об'єктів. Угорі цієї структури – об'єднуюча єдина класифікація – кореневий таксон, який відноситься до всіх об'єктів даної таксономії. Таксони, що розташовані нижче кореневого, є більш специфічними класифікаціями, які належать до підмножин загального набору класифікованих об'єктів.

Даний вид аналізу можна використовувати в біології (для класифікації організмів), психології, медицині й у багатьох інших сферах діяльності людини. Наприклад, в галузі медицини кластеризація захворювань, методів лікування захворювань або симптомів захворювань приводить до широко використовуваних таксономій. В галузі психіатрії правильна діагностика кластерів симптомів, таких як параноя, шизофренія і т. д., є вирішальною для успішної терапії.

Застосування кластерного аналізу припускає наступні **етапи**:

1. Відбір вибірки для кластеризації. Мається на увазі, що доцільно кластеризувати тільки кількісні дані.
2. Визначення множини ознак, тобто змінних, за якими будуть оцінюватися об'єкти у вибірці.
3. Обчислення значень тієї або іншої міри подібності (або відмінності) між об'єктами.
4. Застосування методу кластерного аналізу для створення груп подібних об'єктів.
5. Перевірка вірогідності результатів кластерного розв'язку.

Мета кластеризації:

1. Розуміння даних шляхом виявлення кластерної структури. Розбивка вибірки на групи схожих об'єктів дозволяє спростити подальшу обробку даних і прийняття рішень, застосовуючи до кожного кластера свій метод аналізу.
2. Стискання даних. Якщо вихідна вибірка надмірно велика, то можна скоротити її, залишивши по одному найбільш типовому представникові від кожного кластера.
3. Виявлення новизни. Виділяються нетипові об'єкти, які не вдається приєднати до жодного із кластерів.

8.2. Визначення міри подібності між об'єктами

Кластерний аналіз заснований на об'єднанні об'єктів у досить великі кластери шляхом визначення деякої **міри подібності (відстані)** між об'єктами. Результатом такої кластеризації є *ієрархічне дерево*.

Об'єднання, або метод деревоподібної кластеризації, використовується при формуванні кластерів з урахуванням відстаней між об'єктами. Ці відстані мо-

жуть визначатися в одновимірному або багатовимірному просторі. Завданням дослідників є вибір правильного методу обчислення відстані.

Існує кілька різних **методів обчислення відстані**: евклідова відстань, квадрат евклідової відстані, манхеттенська відстань, степенева відстань, відстань Хеммінга. У задачах кластерного аналізу в біології й медицині найбільше часто використовуються евклідова відстань і відстань Хеммінга.

Найбільш прямий спосіб обчислення відстаней між об'єктами в багатовимірному просторі полягає в обчисленні евклідових відстаней.

Евклідова відстань $p_E(X, Y)$ між об'єктом X з ознаками x_1, x_2, \dots, x_n й об'єктом Y з ознаками y_1, y_2, \dots, y_n визначається за формулою:

$$p_E(X, Y) = \sqrt{\sum_{k=1}^n (x_k - y_k)^2}.$$

Евклідова відстань є аналогом геометричної відстані між об'єктами у випадку двовимірного або тривимірного простору. Однак алгоритм кластеризації не «турбується» про те, чи є надані для аналізу «відстані» справжніми геометричними відстанями або деякими іншими мірами відстані.

Відстань Хеммінга $d(X, Y)$ – число позицій, у яких відповідні символи двох послідовностей символів однакової довжини відрізняються.

Наприклад, зрівняємо дві послідовності символів: 1011101 і 1001001. Як бачимо, ці послідовності мають 2 позиції, у яких символи не збігаються. Це значить, що відстань Хеммінга $d(10\underline{1}1\underline{1}01, 100\underline{1}00\underline{1}) = 2$. Аналогічним образом можна обчислити відстані Хеммінга для десяткових чисел, слів і інших послідовностей символів. Наприклад: $d(\underline{7}36\underline{2}9604, \underline{5}35\underline{2}2204) = 4$; $d(\underline{М}арина, \underline{Г}алина) = 2$.

Для нуклеїнових кислот (ДНК і РНК) можливість гібридизації двох полі-нуклеотидних ланцюгів з утворенням вторинної структури – подвійної спіралі – залежить від ступеня комплементарності нуклеотидних послідовностей обох ланцюгів. При збільшенні відстані Хеммінга кількість водневих зв'язків, утворених комплементарними парами основ, зменшується і, відповідно, зменшується стабільність подвійного ланцюга. Починаючи з деякого граничного значення відстані Хеммінга, гібридизація стає неможливою.

У разі еволюційного розходження гомологічних ДНК-послідовностей відстань Хеммінга є мірою, за якою можна судити про час, що пройшов з моменту розходження гомологів, наприклад, про тривалість еволюційного відрізка, що розділяє гени-гомологи й ген-попередник.

8.3. Методи об'єднання кластерів

На першому кроці, коли кожний об'єкт являє собою окремий кластер, відстані між цими об'єктами визначаються обраною мірою. Однак, коли зв'язуються разом кілька об'єктів, виникає питання про те, як слід визначити відстані між кластерами. Інакше кажучи, необхідно правило об'єднання або зв'язку для двох кластерів. Існують різні **методи об'єднання кластерів**.

1. **Метод найближчого сусіда** (одиначного зв'язку) – ступінь подібності оцінюється за відстанню між найближчими об'єктами кластерів.

2. **Метод найбільш віддалених сусідів** (повного зв'язку) – ступінь подібності визначається найбільшою відстанню між об'єктами в різних кластерах.

3. **Центроїдний метод** – відстань між кластерами визначається відстанню між їхніми «центрами ваги». Різновид центроїдного методу – зважений центроїдний метод (медіана), у якому використовуються вагові коефіцієнти для врахування розмірів кластерів.

4. **Метод середнього зв'язку** (незважене попарне середнє) – відстань визначається як середнє арифметичне всіх попарних відстаней між представниками розглянутих груп. Для кластерів, що суттєво різняться за розмірами, застосовується метод зваженого попарного середнього, який ураховує розміри кластерів.

ТЕМА 9. Формальна логіка у вирішенні задач діагностики і профілактики захворювань

9.1. Поняття про знання

Сучасні інформаційні технології (ІТ) усе більшою мірою зближують медицину з точними науками. Інформаційні технології дозволяють автоматизувати процеси діагностики, прогнозування й вибору методів лікування, дають можливість визначити закономірності перебігу хвороби, зіставляючи одночасно багато її ознак, забезпечують ефективну обробку великої кількості інформації.

У цей час медицині відомо вже більш 10 тисяч хвороб і близько 100 тис. симптомів, здатних проявлятися в різних комбінаціях. У зв'язку з таким величезним потоком інформації з кожним роком процес постановки діагнозу ускладнюється, що породжує лікарські помилки. Тут на допомогу лікареві приходять інформаційні технології у вигляді, наприклад, **систем штучного інтелекту (СШІ)**.

Під **штучним інтелектом (ШІ)** розуміється здатність автоматичних або автоматизованих систем брати на себе функції інтелекту людини, наприклад, приймати оптимальні рішення на основі аналізу зовнішніх впливів і з урахуванням раніше отриманого досвіду.

СШІ підрозділяються на кілька напрямків: експертні системи, системи розпізнавання образів, робототехніка, системи для спілкування з ЕОМ природньою мовою.

У цей час **область практичного застосування СШІ являє собою задачі, які важко формалізувати**. Для них характерні наступні риси:

- задача не може бути визначена в числовій формі (потрібне символічна вистава);
- алгоритмічний розв'язок задачі невідомий (хоча, можливо, і існує) або не може бути використаний через обмежені ресурси (пам'яті комп'ютера, швидкодії);
- мета поставленої задачі не може бути виражена в термінах точно визначеної цільової функції або не існує точної математичної моделі задачі.

Усе це саме характерно для задач медичної діагностики, лікування й профілактики захворювань.

Основне поняття, використовуване в ШІ – поняття «знання».

Знання – це сукупність тверджень про світ, властивості об'єктів, закономірності процесів і явищ, а також правила логічного виведення одних тверджень із інших і правила використання їх для прийняття рішень. Головна відмінність знань від даних полягає в їхній структурності та активності: поява в базі знань нових фактів або встановлення нових зв'язків між ними може стати джерелом змін у прийнятті рішень.

Знання можна класифікувати декількома способами.

1. **Фактичні та стратегічні знання**. Фактичні знання – основні закономірності предметної області, що дозволяють вирішувати конкретні виробничі, наукові й інші задачі, тобто факти, поняття, взаємозв'язки, оцінки, правила, евристики. Стратегічні знання – стратегії прийняття рішень у предметній області.

2. **Факти та евристики**. Факти вказують на *добре відомі* в тій або іншій предметній галузі обставини. Такі знання ще називають *текстовими*, маючи на увазі достатню їхню освітленість у спеціальній літературі й підручниках. Евристики ґрунтуються на *індивідуальному досвіді* фахівця (експерта) у предметній

галузі, накопиченому в результаті багаторічної практики. Ця категорія знань нерідко відіграє вирішальну роль при побудові інтелектуальних програм.

3. **Декларативні та процедурні** знання. Під декларативними знаннями мають на увазі знання типу «*A – це B*». Такі знання характерні для баз даних. Як правило, це *факти* такого типу, як «скарлатина – інфекційне захворювання». Декларативні знання характеризують те, над чим треба виконувати дії. До процедурних знань належать відомості про способи *перетворення* декларативних знань, тобто *дії* для отримання результату.

9.2. Властивості знань

Знання характеризуються рядом **властивостей**, що відрізняють їх від традиційних моделей даних. Перелічимо ці властивості.

1. **Структурованість**. Знання складаються з окремих інформаційних одиниць, між якими можна встановити *класифікуючі відношення*: рід – вид, клас – елемент, тип – підтип, частина – ціле й т. п. Інформаційні одиниці можуть за необхідності розчленовуватися на більш дрібні й поєднуватися в більші за принципом матрьошки.

2. **Внутрішня інтерпретованість**. Разом з інформаційною одиницею, що представляє власне елемент даних, у пам'яті комп'ютера зберігається *система імен*, пов'язана з такою інформаційною одиницею. Наявність системи імен дозволяє інформаційній системі «знати», що зберігається в її пам'яті, і, отже, уміти відповідати на запити про вміст пам'яті, які можуть породжуватися в процесі виконання програм у самій системі або надходити ззовні від користувачів або інших систем.

3. **Зв'язність**. Між інформаційними одиницями можливе встановлення найрізноманітніших *відношень*, що відображають зв'язки явищ і фактів. Коли між інформаційними одиницями в пам'яті системи виникає система відношень, фрагментами цієї структури починають визначатися нові інформаційні одиниці.

4. **Семантична метрика**. На множині інформаційних одиниць, збережених у пам'яті, запроваджуються деякі *шкали* (відношення), що дозволяють оцінити їхню *семантичну близькість* (тобто силу *асоціативного* зв'язку між ними). Це дозволяє знаходити в інформаційній базі знання, *близькі до вже знайдених*.

5. **Активність**. Ця властивість підкреслює принципову відмінність знань від даних. Виконання тих або інших дій у СШ ініціюється *станом бази знань*. *Поява нових фактів* і зв'язків може *активізувати* систему, тобто та або інша структура декларативних знань виявляється активатором для процедурних. Активність бази знань дозволяє СШ формувати мотиви, ставити цілі й будувати процедури для їхнього виконання.

9.3. Логічна модель представлення знань. Алгебра логіки

Центральним питанням побудови систем, заснованих на знаннях, є вибір форми представлення знань. Найбільш поширеними є наступні *моделі представлення знань*:

- логічні моделі;
- продукційні моделі;
- мережні моделі;
- фреймові моделі.

Найпоширенішою формою представлення знань визнається **логічна модель**. Основним апаратом (системою правил) при роботі з логічними моделями є апарат **алгебри логіки**. Завдання алгебри логіки – оптимізація логічних виразів, тобто приведення виразів до вигляду, що містить найменше число аргументів і операцій над ними.

Висловлюванням називається речення, зміст якого можна оцінити як істинне або хибне. Висловлювання позначаються логічними змінними A, B, C, \dots, Z . Логічні змінні можуть отримувати тільки два значення – «**істина**» і «**хибність**», які називаються **значеннями істинності**. В алгебрі логіки вони відповідають числовим значенням **1** і **0**.

Складні висловлювання утворюються із простих висловлювань за допомогою *логічних зв'язок* (або *логічних операцій*).

Основні логічні операції.

1. **Заперечення** (інверсія, логічне НЕ). Позначається \bar{X} або $\neg X$. Змінює значення істинності висловлювання на протилежне.
2. **Диз'юнкція** (логічне додавання, логічне АБО). Позначається $X \vee Y$ або $X + Y$. Істинна, коли істинна *хоча б одна* зі змінних X або Y .
3. **Кон'юнкція** (логічне множення, логічне І). Позначається $X \wedge Y$ або $X \& Y$. Істинна, коли істинні *обидві* змінні X і Y . Операції диз'юнкції та кон'юнкції визначені аналогічним чином і для більшого числа змінних.
4. **Логічне АБО-НЕ** (функція Пірса). Позначається $X \downarrow Y = \overline{X \vee Y}$. Істинна, коли обидві змінні хибні.
5. **Логічне І-НЕ** (функція Шеффера). Позначається $X | Y = \overline{X \wedge Y}$. Хибна тільки тоді, коли обидві змінні істинні.
6. **Еквівалентність** (рівнозначність). Позначається $X \sim Y$. Істинна, коли обидві змінні істинні або обидві хибні.
7. **Нерівнозначність** (додавання за модулем два, **виключне АБО**, XOR). Позначається $X \oplus Y$. Істинна, якщо істинна *тільки одна* із двох змінних X або Y .
8. **Імплікація**. Позначається $X \rightarrow Y$ (« X спричиняє Y », «якщо X , то Y »). Хибна, коли висловлювання X істинно й Y хибно, істинна в інших випадках.

Властивості логічних функцій і їх вираження через функції кон'юнкції, диз'юнкції та заперечення наведені в **таблиці істинності**.

X	Y	$X \vee Y$	$X \wedge Y$	$X \downarrow Y$	$X Y$	$X \sim Y$	$X \oplus Y$	$X \rightarrow Y$
				$\bar{X} \wedge \bar{Y}$	$\bar{X} \vee \bar{Y}$	$(\bar{X} \wedge \bar{Y}) \vee \vee(X \wedge Y)$	$(\bar{X} \wedge Y) \vee \vee(X \wedge \bar{Y})$	$\bar{X} \vee Y$
0	0	0	0	1	1	1	0	1
0	1	1	0	0	1	0	1	1
1	0	1	0	0	1	0	1	0
1	1	1	1	0	0	1	0	1

Наведемо приклади обчислення значення істинності логічного виразу:

- 1) при $X = 0, Y = 1, Z = 1$ вираз $X \vee (Y \wedge Z) = 1$;
- 2) при $X = Y = Z = 0$ вираз $X \rightarrow (Y \sim Z) = 1$.

Складні висловлювання, записані не словами, а за допомогою логічних змінних і знаків логічних операцій, називають **формулами**. Алгебра логіки *не розглядає конкретний зміст* висловлювань, а займається аналізом і синтезом формул і вивченням відносин між формулами.

Наприклад, складне висловлювання «якщо 30 ділиться на 2 і на 3, то 30 ділиться на 6» можна записати у вигляді формули $A \wedge B \rightarrow C$. Наведена формула буде відповідати не тільки конкретно цьому висловлюванню, але й множині всіх інших висловлювань, що мають подібну структуру.

9.4. Експертні системи

Експертна система (ЕС) – інформаційна комп'ютерна система, здатна частково замінити фахівця-експерта у вирішенні проблемної ситуації. Експертні системи належать до класу систем штучного інтелекту, у яких реалізується логіка прийняття рішення досвідченим фахівцем. Експертні системи широко застосовуються в медицині для підтримки прийняття рішень в галузі діагностики, прогнозування, лікування, управління, навчання і т. д.

Існує багато **видів медичних експертних систем**, серед яких можна виділити наступні:

- *експертні системи інтерпретації даних*, які визначають зміст даних, зокрема, даних медичних спостережень і дослідів;

- *експертні системи діагностики*, під час якої визначається характер відхилення стану об'єкта від норми (на підставі цього його зараховують до відповідної до категорії);

- *експертні системи моніторингу*, орієнтовані на безперервну інтерпретацію даних у реальному часі й сигналізацію про вихід тих або інших параметрів за припустимі границі, зокрема експертні системи медичного моніторингу в палатах реанімації;

- *експертні системи прогнозування* логічно будують імовірнісні висновки про майбутній хід подій, виходячи з поточної ситуації, з урахуванням усіх обставин. У медицині з їхньою допомогою прогнозують перебіг хвороби при різних протоколах лікування, визначаючи найкращий протокол у кожному конкретному випадку;

- *експертні системи навчання* визначають недоліки під час вивчення певної дисципліни, збираючи й аналізуючи дані про «слабкі місця», а потім дають необхідні пояснення й рекомендації, згідно з якими вибираються вправи, необхідні для поліпшення підготовки майбутнього лікаря;

- *експертні системи планування* визначають оптимальні плани дій об'єктів, здатних виконувати деякі функції;

- *експертні системи проектування* готують документацію для створення об'єктів із заздалегідь визначеними властивостями, що містить навіть готові креслення й відповідний опис.

Для побудови ЕС можуть бути використані підходи **формальної й неформальної (нейромережевої)** логіки. Специфіка кожного із цих підходів наведена в таблиці.

Тип ЕС Характеристика	ЕС на базі формальної логіки	ЕС на базі неформальної логіки – нейронні мережі
Джерело знань	Формалізований досвід експерта, виражений у вигляді логічних тверджень, правил і фактів, безумовно прийнятих системою	Сукупний досвід експерта-учителя, що відбирає приклади для навчання + власний досвід нейронної мережі, що навчається на цих прикладах
Характер знань	Формально-логічне «лівопівкульне» знання у вигляді правил	Асоціативне «правопівкульне» знання у вигляді зв'язків між нейронами мережі
Розвиток знань	У формі розширення сукупності правил і фактів (базис знань)	У формі донавчання на додатковій послідовності прикладів, з уточненням границь категорій і формуванням нових категорій
Роль експерта	Задає на основі правил повний обсяг знань експертної системи	Відбирає характерні приклади, не формулюючи спеціально обґрунтування свого вибору
Роль штучної системи	Пошук ланцюжка фактів і правил для доказу судження	Формування індивідуального досвіду у формі категорій, одержуваних на основі прикладів, і категоризація образів

9.5. Структура експертної системи

1. **База знань** – сукупність усіх наявних відомостей про проблемну галузь, для якої призначена дана експертна система, записаних за допомогою певних формальних структур представлення знань (набору правил, фреймів, семантичних мереж та ін.).

2. **Підсистема отримання знань** автоматизує процес заповнення й поповнення експертної системи знаннями експерта, тобто забезпечує базу знань усією необхідною інформацією з даної конкретної предметної галузі.

3. **Машина логічного висновку (МЛВ)** – формально-логічна система у вигляді програмного модуля, яка на основі правил і методів бази знань перетворює конкретну інформацію про об'єкт до виду, що відповідає призначенню експертної системи (діагноз, план дій і т.п.).

4. **Робоча пам'ять** (база даних) призначена для зберігання вихідних і проміжних фактів розв'язуваної в поточний момент задачі.

5. **Диспетчер** визначає порядок функціонування експертної системи, планує порядок постановки й досягнення цілей.

6. **Інтерфейс** забезпечує спілкування користувача з експертною системою в зручній для нього формі; дозволяє передавати системі інформацію, що складає вміст бази даних, звернутися до системи з питанням або за поясненням.

7. **Блок пояснень** роз'ясняє, як система одержала розв'язок даної задачі (або чому вона цього розв'язку не знайшла) і які знання вона при цьому використала. Інакше кажучи, блок пояснень створює звіт про виконану роботу

ТЕМА 10. Прийняття рішень у медицині

10.1. Основні поняття

У різних медичних задачах (збір інформації про хворого, діагностика, вибір тактики лікування) лікар стикається із загальною проблемою – проблемою прийняття рішень. При цьому з кожним роком ростуть вимоги до точності діагнозу і його вірогідності, тобто істинності.

Прийняття рішень має на увазі особливий процес людської діяльності, спрямований на вибір найбільш прийняттого варіанта рішення проблеми. Прикладом може служити процес ухвалення рішення про тип (формі) захворювання по відомій вихідній інформації (результати аналізів, зовнішні прояви хвороби).

Виділяють наступні **основні етапи** процедури ухвалення рішення:

1. Визначення мети.
2. Формування множини альтернатив (визначення множини припустимих рішень).
3. Формування оцінки, що дозволяє порівнювати альтернативи (задача оцінювання).
4. Вибір найкращого рішення з множини припустимих рішень (задача оптимізації).

Прийняття рішень, по суті, є не що інше, як **вибір**. Прийняти рішення – значить вибрати конкретний варіант дій з деякої множини варіантів, які прийнято називати **альтернативами**. Множина альтернатив залежить від наявної бази знань і від самої проблемної ситуації.

Рішення – альтернатива (альтернативи), що задовольняє правилам, які містяться в системі переваг.

Система переваг – правила, критерії, за допомогою яких порівнюються альтернативи та приймаються рішення. Ці критерії, виражені математично, називаються **цільовими функціями**.

Наслідок прийняття рішення – подія (результат), можливість появи якої продиктована цим рішенням.

Загальну задачу прийняття рішень (задачу вибору) можна сформулювати в такий спосіб.

Нехай X – множина альтернатив (рішень), Y – множина можливих наслідків (результатів). Передбачається існування причинного зв'язку між вибором деякої альтернативи x_l й настанням відповідного результату y_l . Крім того, передбачається наявність механізму оцінки якості вибору – звичайно шляхом оцінювання якості результату. Потрібно вибрати найкращу альтернативу, для якої відповідний результат має найкращу оцінку якості.

10.2. Класифікація задач прийняття рішень

Виходячи зі зв'язків між рішеннями та результатами, прийнято наступну класифікацію задач прийняття рішень.

1. **Детермінована задача**. Кожна альтернатива x_l приводить до єдиного результату y_l .

2. **Недетермінована задача.** Протилежний випадок, коли кожній альтернативі відповідає не єдиний результат (недетермінована задача) розпадається на два типи:

- задача прийняття рішення в умовах **ризик**у (імовірнісної **визначеності**) – кожній альтернативі x_i відповідає функція щільності ймовірностей на безлічі результатів Y ;
- задача прийняття рішень в умовах **стохастичної** (імовірнісної) **невизначеності**, коли зазначена щільність імовірності невідома.

В умовах імовірнісної невизначеності зв'язків альтернатива-результат, у свою чергу, існують два типи задач:

- задача прийняття рішень в умовах **пасивної взаємодії** особи, що приймає рішення (ОПР) і зовнішнього середовища, тобто зовнішнє середовище поводить пасивно відносно ОПР;
- задача прийняття рішень в умовах **конфлікту (гри)**. У цій ситуації зовнішнє середовище поводить *активно* відносно ОПР, що проявляється діями іншої особи.

Задачі прийняття рішень також підрозділяються на статичні й динамічні. Якщо в ході прийняття рішення ОПР *не здобуває й не втрачає інформацію*, то прийняття рішення можна розглядати як *миттєвий* акт. Відповідні задачі називаються **статичними**. Напроти, якщо ОПР у ході прийняття рішення *одержує або втрачає інформацію*, то така задача називається **динамічною**.

10.3. Етапи прийняття рішення

Усі вимоги, сформульовані в реальних завданнях і записані у вигляді математичних виразів, становлять так звану **математичну постановку задачі**. Процес математичної постановки задачі й наступного її розв'язання можна представити у вигляді ряду етапів.

1. **Вивчення об'єкта** – аналіз особливостей функціонування об'єкта. На цьому етапі виявляються фактори, що впливають на об'єкт, і визначається ступінь їх впливу; вивчаються характеристики об'єкта при різних умовах; вибираються критерії для оптимізації (*цільові функції*).

2. **Описове моделювання** – визначення й фіксація основних зв'язків і залежностей між характеристиками процесу або явища.

3. **Математичне моделювання.**

4. **Вибір або створення методу розв'язання.** На цьому етапі створюється множина *припустимих рішень*, тобто можливих наборів шуканих змінних, які задовольняють умовам-обмеженням задачі.

5. **Розв'язання задачі.** Рішенням задачі є той набір значень із множини припустимих рішень, для якого *цільова функція досягає свого найбільшого або найменшого значення*. Задачі, які описують поведінку реальних об'єктів, як правило, мають багато змінних і багато залежностей між ними. Тому в розумний термін вони можуть бути вирішені тільки за допомогою комп'ютера.

6. **Аналіз отриманого рішення.** Аналіз рішення буває **формальним** і **змістовним**. При *формальному* (математичному) аналізі перевіряють відповідність отриманого рішення побудованій математичній *моделі* (чи правильно введені

вихідні дані, чи правильно функціонують програми комп'ютера і т. д.). При *змістовному* аналізі перевіряють відповідність отриманого рішення тому *реальному об'єкту*, який моделювали. У результаті змістовного аналізу в модель можуть бути внесені зміни, і весь процес повторюється.

7. Аналіз стійкості рішень. Для перевірки стійкості рішення у вихідні дані вносяться зміни в межах можливих похибок або інтервалів існування ознак, а потім досліджується поведінка рішення аналітичними або чисельними методами.

10.4. Чутливість і специфічність діагностичного тесту

Під час розробки й застосування методів і засобів діагностики завжди виникає питання про **надійність** цих методів. Надійність тесту, який використовується для відділення здорових людей від хворих, можна охарактеризувати за допомогою таких характеристик тесту, як чутливість і специфічність.

Задача про надійність діагностичного методу формулюється в такий спосіб. Дана довільна сукупність пацієнтів, які можуть перебувати в одному із двох станів стосовно деякого захворювання – *норма* або *патологія*. Кожному із цих станів відповідає своя функція розподілу результатів діагностичного тесту. Для кожного пацієнта потрібно зробити найкращий вибір між цими двома станами, тобто фактично поставити діагноз «норми» або «патології» на підставі цього тесту.

За співвідношенням результату методу діагностики, що перевіряється, та дійсного стану пацієнта можливі 4 різних наслідки:

Результати тесту	Дійсний стан	
	Патологія	Норма
Позитивні	Істинно позитивні (true positive, <i>TP</i>)	Хибно позитивні (false positive, <i>FP</i>)
Негативні	Хибно негативні (false negative, <i>FN</i>)	Істинно негативні (true negative, <i>TN</i>)

- 1) *TP* – істинно позитивний діагноз (визначається діагностикою методом «золотого стандарту»);
- 2) *FP* – хибно позитивний діагноз (наявність позитивного тесту, де хвороби немає), названий також *помилкою другого роду*;
- 3) *TN* – істинно негативний діагноз (визначається діагностикою по методу «золотого стандарту»);
- 4) *FN* – хибно негативний діагноз (одержання негативного тесту при наявності діагностованої хвороби), названий *помилкою першого роду*.

Чутливість тесту – це частка позитивних результатів, які правильно ідентифіковані як позитивні (тобто ймовірність визнати істинно хворого хворим за результатами тесту). Чутливість відображає властивість тесту з великою точністю *прийняти дійсну гіпотезу*, тобто уникнути помилки першого роду.

$$\text{Чутливість} = \frac{N(TP)}{N(TP) + N(FN)}.$$

Іншими словами, чутливість – це відношення кількості істинно позитивних результатів тесту до загальної кількості пацієнтів з даним захворюванням.

Специфічність тесту – це частка негативних результатів, які правильно ідентифіковані як негативні (тобто ймовірність визнати істинно здорового здоровим за результатами тесту). Специфічність характеризує властивість тесту з великою точністю *відхилити неправильну гіпотезу*, тобто уникнути помилки другого роду.

$$\text{Специфічність} = \frac{N(TN)}{N(TN) + N(FP)}.$$

Специфічність – це відношення кількості істинно негативних результатів до загальної кількості пацієнтів, що не мають даного захворювання.

Обидва показники можуть отримувати значення від 0 до 1 і часто виражаються у відсотках (від 0 до 100 % відповідно).

Слід звернути увагу, що показники чутливості й специфічності відображають *різні властивості* діагностичного тесту й *не залежать один від одного*, наприклад, тест може мати близьку до 100 % чутливість і малу специфічність або навпаки.

10.5. Розрахунки ймовірності наявності хвороби при позитивному тесті

Використовуючи характеристики чутливості й специфічності тесту й частоту випадків захворювання, можна розрахувати ймовірність наявності хвороби за умови позитивного тесту.

Розрахунки проводяться за формулою *Байєса*. Теорема Байєса (або формула Байєса) – одна з основних теорем елементарної теорії ймовірностей, яка дозволяє визначити ймовірність якої-небудь події за умови, що відбулася *інша статистично взаємозалежна* з нею подія.

Нехай у хворого підозрюється певне захворювання, тобто є гіпотеза H про наявність саме цього захворювання. Для даного захворювання відома первісна (*апріорна*) ймовірність $P(H)$ його появи в пацієнта. Пацієнт проходить діагностичний тест, позитивний результат якого позначимо D . Потрібно знайти ймовірність захворювання при позитивному результаті тесту $P(H | D)$.

Тоді за формулою Байєса ймовірність $P(H | D)$ того, що людина дійсно хвора при позитивному результаті тесту (*апостеріорна* ймовірність) дорівнює

$$P(H | D) = \frac{P(D | H)P(H)}{P(D | H)P(H) + P(D | \bar{H})(1 - P(H))},$$

де $P(H)$ – ймовірність наявності хвороби (апріорна ймовірність);

$1 - P(H)$ – ймовірність відсутності хвороби;

$P(D | H)$ – ймовірність того, що тест позитивний для даного діагнозу, тобто чутливість тесту;

$P(D | \bar{H})$ – ймовірність хибно позитивного діагнозу (ймовірність позитивного результату тесту в здоровій людини, тобто ймовірність помилки другого роду).

10.6. Задача диференціальної діагностики захворювань

Теорема Байєса дозволяє вибрати одну з декількох діагностичних гіпотез, ґрунтуючись на обчисленні ймовірностей хвороб за ймовірностями виявлених у хворих симптомів. Найпоширеніша задача такого типу – задача диференціальної діагностики між двома захворюваннями з одним спільним симптомом.

У клінічній практиці відомі симптоми, наявність яких однозначно визначає захворювання. З іншого боку, зустрічаються симптоми, що виключають той або інший діагноз. Однак найчастіше основні симптоми, що визначають клініку, можуть зустрічатися з деякою частотою при різних захворюваннях. У таких випадках технологія прийняття рішень використовує пов'язане з теоремою Байєса поняття **шансів**.

Розглянемо задачу про диференціальну діагностику між захворюваннями H_1 та H_2 , що мають спільний симптом D . Нехай середньостатистична ймовірність захворювання H_1 становить $P(H_1)$, ймовірність захворювання H_2 – $P(H_2)$.

Відношення ймовірностей $\frac{P(H_1)}{P(H_2)}$ показує **початкові шанси** на користь діагнозу

H_1 без урахування додаткових умов:

$$C_0 = \frac{P(H_1)}{P(H_2)}.$$

Задача полягає в тому, щоб знайти **кінцеві шанси** C_1 діагнозу H_1 за наявності симптому D , тобто співвідношення

$$C_1 = \frac{P(H_1 | D)}{P(H_2 | D)}.$$

Теорема Байєса стверджує, що кінцева ймовірність гіпотези $P(H | D)$ пропорційна до її початкової ймовірності $P(H)$, помноженої на її **правдоподібність** $P(D | H)$:

$$P(H | D) \sim P(H)P(D | H).$$

Відповідно, для обчислення кінцевих шансів необхідно знайти *відношення двох правдоподібностей*, або **відношення правдоподібності** (likelihood ratio, LR), тобто відношення ймовірностей $P(D | H_1)$, $P(D | H_2)$ наявності симптому D при захворюваннях H_1 і H_2 .

$$LR = \frac{P(D | H_1)}{P(D | H_2)}.$$

Кінцеві шанси C_1 діагнозу H_1 за наявності характерного симптому D обчислюються як початкові шанси C_0 , помножені на відношення правдоподібності LR :

$$C_1 = C_0 \cdot LR = \frac{P(H_1)}{P(H_2)} \cdot \frac{P(D | H_1)}{P(D | H_2)}.$$

ТЕМА 11. Математичне моделювання медико-біологічних процесів

11.1. Основні поняття

Одним з найважливіших методів дослідження в різних галузях науки, у тому числі й у медицині, є метод моделювання.

Модель – штучно створений людиною об'єкт будь-якої природи, який замінює або відтворює досліджуваний об'єкт так, що вивчення моделі здатне надавати нову інформацію про об'єкт. Іншими словами, модель – це якийсь *новий* об'єкт, який відбиває *істотні особливості* досліджуваного об'єкта, явища або процесу.

Модель – це не тільки відбиття наших знань про досліджуваний об'єкт, але й *джерело нових знань*. Дослідження моделі дозволяє оцінити поведінку модельованого об'єкта в нових умовах або за різних впливів, які на реальному об'єкті перевірити неможливо (дослідження на людях) або важко (дорогі об'єкти або негативні наслідки експериментів).

Об'єкт дослідження в біології й медицині – живий організм, який є досить складною системою. Тому дослідник неминуче вибирає *спрощену* точку зору, яка підходить для розв'язку конкретно поставленої задачі. Вибір моделі визначається цілями дослідження.

Моделі підрозділяються на три класи:

- речовинні (матеріальні);
- енергетичні;
- інформаційні.

Речовинні моделі – моделі, які відтворюють *структуру* об'єкта та *взаємозв'язки* його частин. Прикладом таких моделей у медицині можуть служити різні протези, за зовнішнім виглядом схожі на реальні частини тіла, які вони заміщають.

Енергетичні моделі використовуються для моделювання *функціональних відносин* у досліджуваних об'єктах. Ці моделі за зовнішнім виглядом не нагадують модельовані об'єкти, але їх метою є *виконання функцій* цих об'єктів. Приклади таких моделей у медицині – апарат штучної нирки або штучного дихання. Властивості речовинних і енергетичних моделей можуть поєднуватися. До таких моделей належать біокеровані протези, штучний кришталік ока, останні розробки в галузі штучного серця.

Інформаційні моделі – це описи об'єкта. У медико-біологічних дослідженнях донедавна для опису роботи біологічних систем використовували переважно словесні моделі. Однак за допомогою словесних моделей важко чітко викласти закономірності роботи досліджуваного об'єкта. Тому все частіше застосовуються **математичні моделі**, які використовують кількісні співвідношення між параметрами досліджуваної біосистеми. Ці моделі мають найбільше значення в медико-біологічних дослідженнях.

Математичні моделі підрозділяються на **детерміновані** й **імовірнісні**. У детермінованих моделях змінні й параметри передбачаються заданими й описуються детермінованими функціями. В імовірнісних моделях змінні й параметри є випадковими функціями або випадковими величинами. Детерміновані математичні моделі найчастіше являють собою *систему алгебраїчних або диференціальних рівнянь*. Імовірнісні моделі будуються за результатами експерименталь-

ного визначення динамічних характеристик об'єктів на основі *методів математичної статистики*.

Розглянемо кілька прикладів математичних моделей, які застосовуються у медицині й біології.

11.2. Математична модель «хижак-жертва»

Уперше в біології математичну модель періодичної зміни кількості антагоністичних видів тварин запропонував італійський математик Віто Вольтерра зі співробітниками. Ця модель була розвитком ідеї, наміченої в 1924 р. Альфредом Лоткою у книзі «Елементи фізичної біології». Тому ця класична модель відома як «**модель Лотки-Вольтерри**».

Завдання моделювання формулюється в такий спосіб. У деякому екологічно замкнутому районі живуть тварини двох видів (наприклад, рисі й зайці). Зайці (**жертви**) харчуються рослинною їжею, яка є в достатній кількості (у рамках цієї моделі не враховується обмеженість ресурсів рослинної їжі). Рисі (**хижаки**) можуть харчуватися тільки зайцями. Потрібно визначити, як буде змінюватися чисельність жертв і хижаків із часом.

Позначимо кількість жертв через N , а кількість хижаків через M . Величини N і M є функціями часу t . У даній моделі врахуємо такі фактори:

- 1) природне розмноження жертв;
- 2) природне вимирання жертв;
- 3) зменшення кількості жертв внаслідок поїдання їх хижаками;
- 4) природне вимирання хижаків;
- 5) збільшення кількості хижаків за рахунок розмноження при наявності їжі.

Необхідно вивести *рівняння*, у які входили б усі позначені фактори і які описували б динаміку, тобто зміну кількості жертв і хижаків із часом.

Нехай за якийсь час Δt кількість жертв зміниться на ΔN , а кількість хижаків зміниться на ΔM .

Зміна кількості жертв ΔN за час Δt визначається першими трьома факторами з наведеного вище списку.

Збільшення кількості жертв внаслідок їхнього **природного розмноження** (ΔN_1) пропорційне кількості жертв N , що існують у цей момент, з коефіцієнтом пропорційності A :

$$\Delta N_1 = AN\Delta t.$$

Зменшення кількості жертв внаслідок **природного вимирання** (ΔN_2) також пропорційне їх кількості N в цей момент, з коефіцієнтом пропорційності B (причому $B < A$):

$$\Delta N_2 = -BN\Delta t.$$

Знак мінус показує зменшення чисельності.

В основі рівняння, яке описує **зменшення кількості жертв внаслідок поїдання їх хижаками** (ΔN_3), лежить ідея про те, що чим частіше відбуваються їхні зустрічі, тим швидше зменшується кількість жертв. Частота зустрічей хижака з жертвою пропорційна і кількості жертв N , і кількості хижаків M . Позначивши коефіцієнт пропорційності через C , запишемо:

$$\Delta N_3 = -CMN\Delta t.$$

З урахуванням усіх трьох факторів для зміни кількості жертв отримуємо рівняння

$$\Delta N = AN\Delta t - BN\Delta t - CMN\Delta t,$$

або, у диференціальній формі,

$$\frac{dN}{dt} = AN - BN - CMN.$$

Зміна кількості хижаків ΔM визначається останніми двома факторами зі списку.

Збільшення числа хижаків внаслідок **природного розмноження** (ΔM_1) залежить від кількості хижаків M і кількості їжі (жертв) N з коефіцієнтом пропорційності Q :

$$\Delta M_1 = QMN\Delta t.$$

Зменшення числа хижаків **внаслідок природного вимирання** (ΔM_2) пропорційне кількості хижаків M у цей момент із коефіцієнтом P :

$$\Delta M_2 = -PM\Delta t.$$

З урахуванням цих двох факторів маємо

$$\Delta M = QMN\Delta t - PM\Delta t,$$

або, у диференціальній формі,

$$\frac{dM}{dt} = QMN - PM.$$

У результаті ми знаходимо систему із двох диференціальних рівнянь:

$$\begin{cases} \frac{dN}{dt} = AN - BN - CMN \\ \frac{dM}{dt} = QMN - PM \end{cases}$$

Розв'язок цієї системи рівнянь дає періодичну залежність чисельності хижаків $M(t)$ і жертв $N(t)$ від часу t , як показано на графіку.

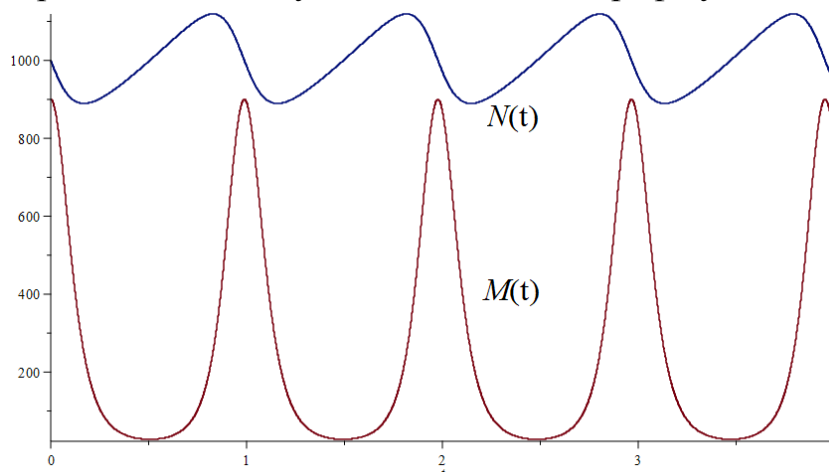


Рис. 5. Поведінка моделі «хижак-жертва»

11.3. Математична модель імунної реакції

Імунітет – це складний комплекс реакцій організму на вторгнення *антигенів* – чужорідних об'єктів: молекул, вірусів, клітин, тканин та ін. Специфічна імунна реакція на молекулярному рівні починається з того, що спеціалізовані

плазматичні клітини виробляють у великій кількості білкові молекули – *анти-тіла*, які нейтралізують антигени.

Антитіла мають конформацію, комплементарну до певної ділянки поверхні антигену. Тому антитіло з'єднується з антигеном, як ключ із замком, і утворений при цьому комплекс зазнає лізису ферментами.

Розглянемо модель роботи імунного апарату під час тривалого інфекційного захворювання. Ця модель використовується в клінічній практиці при лікуванні вірусного гепатиту й гострої пневмонії.

Взаємодія антигенів і імунних сил організму в цій математичній моделі носить характер, схожий на поведінку системи «хижак-жертва». «Жертвою» тут є чужорідний агент, який у моделі буде описуватися кількістю відповідного антигену X , а «хижаком» – антитіла в кількості Y , утворені плазматичними клітинами в кількості Z . Усі три змінні змінюються в часі t .

У даній моделі враховуються наступні процеси й фактори.

1. Розмноження антигенів зі швидкістю AX .
2. Природний розпад антигенів зі швидкістю $-CX$.
3. Природний розпад антитіл зі швидкістю $-LY$.
4. Природна загибель плазматичних клітин зі швидкістю $-NZ$.
5. Взаємодія антиген-антитіло в реакції аглютинації, частота якого пропорційна ймовірності зустрічі відповідного антитіла з антигеном. Ця взаємодія приводить до зменшення кількості антигенів і антитіл зі швидкостями $-BXY$, $-KXY$ відповідно.
6. Продукція антитіл плазматичними клітинами та надходження їх у кров зі швидкістю DZ .
7. Утворення плазматичних клітин зі швидкістю, що залежить від концентрації антигену X складним чином: $MF(X)$.

Тут A, B, C, D, K, L, M, N – коефіцієнти пропорційності.

У наведеній моделі коефіцієнти A й M вважаються залежними від температури, а функція $F(X)$ має вигляд $F(X) = \frac{X}{Q + X}$, де Q – деякий коефіцієнт.

Знаходячи швидкості зміни з часом кількості антигенів $\frac{dX}{dt}$, антитіл $\frac{dY}{dt}$ і плазматичних кліток $\frac{dZ}{dt}$ з урахуванням усіх наведених вище факторів, одержуємо систему трьох диференціальних рівнянь:

$$\begin{cases} \frac{dX}{dt} = AX - BXY - CX; \\ \frac{dY}{dt} = DZ - KXY - LY; \\ \frac{dZ}{dt} = \frac{MX}{Q + X} - NZ. \end{cases}$$

Дослідження математичної моделі полягає в розв'язанні виведеної системи диференціальних рівнянь із відомими коефіцієнтами $A, B, C, D, K, L, M, N, Q$ і за початкових умов (у момент часу $t=0$) $X(0), Y(0), Z(0)$. Важливо відзначити, що одна й та ж модель за різних початкових умов або коефіцієнтів дає абсолютно різну динаміку процесу. При цьому можливі чотири основні форми перебігу інфекційного захворювання:

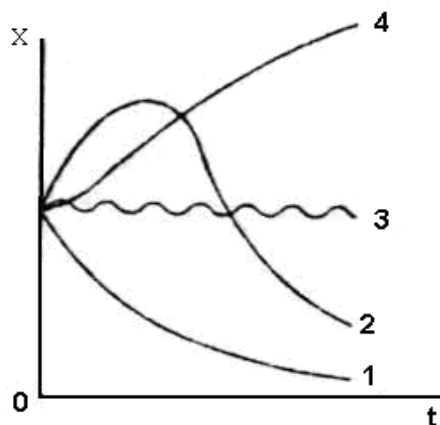


Рис. 6. Можливі форми перебігу інфекційного захворювання

1. Субклінічна форма – перебігає без фізіологічних розладів в організмі й без зовнішніх проявів. Засоби імунного захисту легко знищують антигени, не даючи їм розмножуватися до небезпечної межі.

2. Гостра форма – у цьому випадку організм атакується невідомим антигеном у більших кількостях. Спочатку відбувається його посилене розмноження. Коли ж імунна система створить проти нього достатню кількість антитіл, кількість антигенів різко зменшиться.

3. Хронічна форма – встановлюється динамічна рівновага в кількостях антигенів і антитіл (подібно до моделі «хижак-жертва»). Виникає стійкий стан хвороби.

4. Летальна форма – імунна відповідь занадто запізнюється, і велика кількість антигенів приводить до незворотних змін в організмі.

11.4. Математична модель поширення інфекційного захворювання в населеному пункті

Процес поширення інфекційного захворювання в найпростішому випадку можна описати системою із трьох диференціальних рівнянь першого порядку:

$$\begin{cases} \frac{dX}{dt} = -QAXY, \\ \frac{dY}{dt} = QAXY - \frac{1}{R}Y, \\ \frac{dZ}{dt} = \frac{1}{R}Y, \end{cases}$$

де Q – кількість жителів населеного пункту; A – середня кількість жителів, яких щодня інфікує кожний хворий; R – середня тривалість захворювання, дні; X – кількість здорових людей, які не перенесли дане захворювання й не мають імунітету до нього; Y – кількість хворих людей; Z – кількість тих, хто перехворів і набув імунітету.

Наведена модель дає можливість розрахувати швидкість поширення захворювання серед населення, максимальну кількість одночасно хворих, загальну кількість осіб, що перенесли хворобу.

11.5. Фармакокінетичні моделі

Фармакокінетика вивчає розподіл в організмі досліджуваної біологічно активної речовини та зміну її концентрації з часом. До біологічно активних речовин належать, зокрема, лікарські препарати. Закони зміни концентрації лікарського препарату в організмі різні залежно від способів і параметрів його введення й виведення.

Для спрощення моделювання у фармакокінетиці організм умовно поділяють на **камери**, тобто на частини, у яких досліджуваний препарат розподілений рівномірно.

Найпростішою фармакокінетичною моделлю є **лінійна однокамерна модель**. Вона є адекватною для опису поведінки лікарських препаратів, що вводяться в кров ін'єкційно.

У **лінійній** фармакокінетичній моделі швидкість виведення лікарського препарату з організму $\frac{dM}{dt}$ пропорційна першому ступеню маси M препарату в камері. Вважаємо, що препарат рівномірно розподілений по всьому об'єму V організму (така модель називається **однокамерною**).

Диференціальне рівняння однокамерної лінійної фармакокінетичної моделі має вигляд:

$$\frac{dM}{dt} = -kM,$$

де k – **константа елімінації**, тобто коефіцієнт видалення препарату з організму.

Розв'язок цього диференціального рівняння має вигляд

$$M = M_0 e^{-kt},$$

де M_0 – маса препарату в момент часу $t = 0$, тобто безпосередньо після введення.

Аналогічна залежність від часу справедлива й для концентрації препарату $C = \frac{M}{V}$:

$$C = C_0 e^{-kt},$$

де C_0 – концентрація препарату в початковий момент часу.

Як видно, концентрація лікарського препарату в крові безперервно знижується за спадним експоненційним законом на зразок закону радіоактивного розпаду. За аналогією з періодом напіврозпаду у фармакокінетиці існує поняття **періоду напіввиведення** ($T_{1/2}$), який дорівнює

$$T_{1/2} = \frac{\ln 2}{k}$$

і означає проміжок часу, за який з організму буде виведена половина початкової маси препарату. Ця характеристика часто вказується в інструкціях із застосування лікарських препаратів.

Таким чином, модель дозволяє обчислити залежність концентрації препарату від часу. У реальній медичній практиці вона дає можливість планувати приймання певних доз препарату таким чином, щоб його концентрація перебувала в заданих межах.

ТЕМА 12. Мультимедійні технології

12.1. Основні поняття

Мультимедійні технології – технології, що дозволяють надавати інформацію користувачеві одночасно в різних формах (текст, графіка, анімація, звук, відео) в інтерактивному режимі. Буквально слово «мультимедіа» (multimedia) означає «багато носіїв інформації».

Мультимедійні продукти можна поділити на кілька категорій залежно від того, на які групи споживачів вони орієнтовані:

- комп'ютерні ігри;
- бізнес-програми;
- освітні програми;
- програми, призначені для самостійного створення мультимедійних продуктів.

Технологія мультимедіа містить у собі спеціальні апаратні й програмні засоби. Мультимедійні комп'ютери повинні мати конфігурацію, що дозволяє одержувати гарне зображення графічних об'єктів, демонструвати відео високого дозволу, прослуховувати звук, що відтворюється з якістю не гірше якості звучання аудіоприладів. До такого комп'ютера може підключатися множина різних пристроїв:

- принтер і плотер – пристрої для виведення тексту й іншої графічної інформації на папір;
- 3D-принтер – пристрій, що використовує метод пошарового створення фізичного об'єкта за цифровою тривимірною моделлю;
- сканер – пристрій для введення зображень із паперу або слайдів;
- проектор – для показу зображення (відео або презентацій) на великому екрані;
- пристрої для підключення до локальних або глобальних комп'ютерних мереж (мережні адаптери, модеми, пристрої бездротового зв'язку);
- звукові колонки й інші аудіосистеми – для відтворення звуку;
- TV-тюнер – для приймання теле- або радіопередач;
- джойстик – маніпулятор у вигляді укріпленої на шарнірах ручки із кнопками, який використовується для управління об'єктами в комп'ютерних іграх; інші види маніпуляторів;
- відеокамера;
- технічні засоби для створення віртуальної реальності;
- зовнішні накопичувачі для зберігання даних.

Мультимедіа-технології можуть використовуватися як на окремому комп'ютері (локально), наприклад, відтворення фільму з жорсткого диска або з DVD-диска, так і шляхом завантаження або безпосереднього програвання потоку з Інтернету.

Потокові технології пред'являють дві ключові вимоги до обчислювальної системи:

- 1) високі швидкості передачі даних;
- 2) можливість відтворення в режимі реального часу.

Високі швидкості передачі даних обумовлені природою візуальної й акустичної інформації. Органі зору й слуху людину здатні обробляти за секунду величезні обсяги даних, тому їм необхідно поставляти інформацію з тією швидкістю, яка забезпечить прийнятний рівень якості сприйняття. Наприклад, швидкість передачі даних під час відтворення музики – 128...320 кбіт/с, відео високого дозволу – 10...25 Мбіт/с і вище.

Друга вимога, що накладається мультимедійними продуктами на систему, полягає в необхідності доставки даних у режимі реального часу. Графічна складова відеофільму складається з послідовності кадрів, що передаються з певною частотою (як правило, 25, 30 або 60 кадрів у секунду). Кадри повинні доставлятися через точні інтервали часу, щоб зображення не «смикало». У випадку аудіосигналу вимоги ще більш жорсткі. Чутливість людського вуха перевищує чутливість ока, тому під час відтворення звуку відхилення в часі доставки навіть у кілька мілісекунд буде помітним.

Нерівномірність часу доставки даних, тобто розкид мінімального й максимального часу проходження пакета даних від середнього часу проходження пакета, називається **джитером**. Наприклад, посилаються 100 пакетів. Мінімальний час проходження пакета – 395 мс, середнє – 400 мс, максимальне – 405 мс. У цьому випадку ($405-400=5$; $400-395=5$) джитер становить 5 мс. Для забезпечення високої якості відтворення слід утримувати джитер у певних межах.

Параметри середовища передачі, необхідні для прийнятного відтворення мультимедіа в реальному часі, називають **параметрами якості обслуговування**. До них належать:

- 1) пропускна здатність (швидкість передачі даних);
- 2) затримка при передачі пакета;
- 3) джитер;
- 4) імовірність втрати пакета або біта.

Наприклад, мережний оператор може пропонувати службу, що гарантує середню пропускну здатність, рівну 4 Мбіт/с, 99% затримок при передачі в інтервалі від 105 до 110 мс (джитер у цьому випадку складе 2,5 мс) і ймовірність втрати біта, рівну 10^{-10} , що є прийнятними параметрами для передачі фільму у форматі MPEG-2.

12.2. Мультимедійні файли

Будь-який мультимедійний продукт містить у собі наступні основні компоненти:

- текстові;
- графічні;
- звукові;
- відео й анімацію;
- інші типи, наприклад, інтерактивні тривимірні представлення.

Цим компонентам відповідають певні типи й формати файлів. Типовий мультимедійний продукт складається, як правило, не з одного файлу, а з *великої кількості файлів* різних типів, що мають певну структуру.

12.3. Кодування та стиснення звуку

Звукові хвилі, які за природою є механічними хвилями, приймаються мікрофоном і перетворюються ним на електричний сигнал. Потім цей сигнал переводиться в цифрову форму за допомогою **аналого-цифрового перетворювача** (АЦП), тобто АЦП приймає на вході електричну напругу й формує двійкове число на виході. АЦП робить відліки (виміри) прийнятого електричного сигналу не безперервно, а *дискретно*. Частота, з якою АЦП робить відліки, називається **частотою дискретизації**.

Мінімальна частота дискретизації АЦП може бути знайдена за **теоремою Котельникова** (теоремою Найквіста–Шеннона): якщо звукова хвиля не є чисто синусоїдальною, а являє собою суму декількох синусоїдальних хвиль із частотами від 0 до f , тоді для подальшого повного відновлення сигналу досить вимірювати значення сигналу із частотою дискретизації $2f$. Це твердження вперше було висунуто Найквістом у 1928 році, а частота $2f$ названа **частотою Найквіста**.

Наприклад, для передачі звуку у всьому чутному діапазоні (від 16 до 20 000 Гц) необхідно використовувати частоту дискретизації не менш ніж 40 000 Гц. Найбільш часто в мультимедійних системах використовуються значення частоти дискретизації 44 100 і 48 000 Гц.

Оцифровані відліки ніколи не бувають точними, оскільки АЦП може вимірювати вхідний сигнал не точно, а з якимось кроком за рівнем сигналу – кроком квантування. Помилка, що виникає в результаті неточної відповідності оцифрованого (квантованого) сигналу до вихідного сигналу, називається **шумом квантування**. При недостатній кількості бітів, якими представляється кожний відлік сигналу, цей шум може бути настільки великим, що буде помітний на слух як спотворення вихідного сигналу або як сторонні шуми.

Аудіо-CD диски містять звуковий сигнал, оцифрований із частотою дискретизації 44 100 Гц, у результаті чого вони можуть зберігати звуки з частотами до 22 кГц. Кожний відлік являє собою 16-бітне (2-байтове) ціле двійкове число, пропорційне амплітуді сигналу.

Необхідна для відтворення звуку **пропускна здатність** системи розраховується як добуток частоти дискретизації й кількості біт в одному відліку. У наведеному прикладі вона становить 705,6 кбіт/с для монофонічного сигналу й 1,411 Мбіт/с для стереофонічного (два канали).

Існують різні **алгоритми стиснення** цифрового звуку, які дозволяють використовувати для передачі звуку тієї ж якості суттєво меншу швидкість передачі даних (пропускну здатність). Найбільш популярним алгоритмом стиснення звуку є алгоритм MP3, який дозволяє стискати оцифрований звук приблизно в 10 разів.

Цифровий звук легко обробляти на комп'ютері. Існує багато програм для персональних комп'ютерів, що дозволяють користувачам записувати, відтворювати, редагувати, міксувати і зберігати звук. Сьогодні весь професійний звукозапис і редагування звуку здійснюється в цифровому виді.

12.4. Кодування зображення

Сітківка ока людини має *інерційні властивості*, тобто яскраве зображення, що швидко з'явилося на сітківці, залишається на ній кілька мілісекунд, перш ніж згаснути. Якщо послідовність однакових або близьких зображень з'являються й зникають із досить високою частотою, то око людини не помітить, що воно дивиться на дискретні зображення. Частота, при якій око не помічає миготіння світла (зображення), повинна бути не нижче ніж 50 Гц. Усі відеосистеми використовують цей принцип для створення зображень, що рухаються.

Щоб зрозуміти, як працюють відеосистеми, найкраще почати із простого старомодного чорно-білого телебачення. Для перетворення двовимірного зображення в одновимірну залежність напруги від часу камера швидко *сканує* електронним променем зображення, розбиваючи його на горизонтальні лінії (*лінії розгортки*, або рядки) і записуючи під час просування інтенсивність світла. *Сканування кадра* відбувається аналогічно до читання книги – зліва направо у кожному рядку та зверху донизу в кадрі. Закінчивши сканування, промінь повертається у вихідну точку, і процес повторюється. У різних країнах використовувалися різні стандарти, що описують параметри сканування (кількість рядків, частота кадрів і т. д.).

Сучасні відеосистеми являють собою матричні рідкокристалічні монітори й цифрові камери із ПЗЗ-матрицями (прилад із зарядовим зв'язком). У цих пристроях немає електронно-променевої трубки й скануючого променя, однак принцип передачі відеосигналу залишається таким самим.

Для того, щоб передати плавний рух об'єкта, вистачає частоти в 25 кадрів на секунду. Однак при такій частоті кадрів (частоті розгортки) глядачі сприймають зображення як миготливе. Це пов'язане з тим, що сітківка ока встигне відновитися, перш ніж з'явиться новий кадр. Щоб виправити цей недолік, необхідно збільшувати частоту кадрів, що потребувало б збільшення обсягів збереженої й переданої інформації. Замість цього було обрано інше рішення. Лінії розгортки показуються на екрані не підряд, а через одну – спочатку всі непарні, а потім усі парні. Така техніка називається **черезрядковою розгорткою**. При цьому зображення на екрані промальовується із удвічі більшою частотою, тобто 50 разів на секунду. Експерименти показали, що за такої послідовності відображення рядків на екрані люди не помічають мерехтіння зображення навіть при передачі 25 кадрів на секунду.

Крім черезрядкової, у наш час існує також **прогресивна розгортка**, в якій лінії розгортки відображаються підряд (одна за одною), а плавність зображення досягається високою частотою кадрів. Вона застосовується в популярному стандарті передачі зображень HDTV (телебачення високої чіткості).

У кольоровому відео використовується той же принцип розгортки, що й у чорно-білому, з тою різницею, що замість одного кольору зображення представляється **трьома кольорами** – червоним, зеленим і синім (RGB – red, green, blue). Комбінації цих трьох кольорів виявляється достатньо для передачі будь-якого кольору завдяки особливостям будови людського ока.

Дотепер ми розглядали аналогове відео. Обговоримо тепер **цифрове відео**. Найпростіша форма представлення цифрового відео полягає в послідовності

кадрів, що складаються із прямокутної сітки елементів малюнка, названих пікселями. **Піксель** – найменший логічний елемент двовимірного цифрового зображення або фізичний елемент матриці дисплеїв, що формують зображення. У кольоровому телебаченні достатньо використовувати по 8 біт на кожний із трьох кольорів RGB, що дає в сумі 24 біта на піксель. Хоча числом, що складається із 24 біт, можна позначити близько 16 млн кольорних відтінків, людське око не в змозі навіть розрізнити таку величезну кількість відтінків, не говорячи вже про більші кількості.

Для гладкої передачі руху, як і в аналоговому відео, у цифровому відео необхідно відобразити щонайменше 50 кадрів на секунду.

12.5. Стиснення зображень і відеосигналу

Оцінимо пропускну здатність каналу зв'язку, необхідну для передачі нестисненого відеосигналу. Для прикладу візьмемо цифровий відеосигнал, що містить 50 кадрів у секунду з роздільною здатністю 1920×1080 пікселів, у якому кожний піксель містить 24 біта. Перемноживши всі ці показники, одержуємо пропускну здатність близько 2,3 Гбіт/с, що багаторазово перевищує пропускну здатність типових комп'ютерних мереж. Очевидно, що про передачу мультимедійної інформації в нестисненому вигляді не може бути й мови. За останні кілька десяти років було розроблено багато **методів стиснення**, що роблять можливою передачу мультимедійної інформації.

Для всіх методів стиснення потрібно два алгоритми: один для стиснення даних у джерела інформації, а інший – для відновлення даних у її одержувача. У літературі ці алгоритми називають відповідно алгоритмами *кодування* й *декодування*.

Для **стиснення нерухливих зображень** із неперервно змінним кольором (наприклад, фотографій) використовується стандарт **JPEG**.

Для **стиснення відеосигналу** застосовують стандарти **MPEG** (MPEG-1, MPEG-2, MPEG-4). Перший етап стиснення полягає в тому, що кадри відеосигналу кодуються стандартом JPEG. Додаткового стиску можна досягти, використовуючи перевагу того факту, що послідовні кадри часто бувають майже ідентичні.

У сценах, у яких камера й задній план нерухливі й невелика кількість об'єктів повільно переміщується, майже всі пікселі в сусідніх кадрах будуть ідентичні. У цьому випадку просте віднімання кожного кадра від попереднього й обробка різниці алгоритмом JPEG дасть досить гарний результат. Однак цей метод погано підходить до сцен, у яких камера повертається або наїжджає на об'єкт, що знімається. Для таких сцен використовуються різні способи компенсувати цей рух камери. Після стиснення цифрове відео складається з кадрів різних типів, що містять як повні зображення, так і різницеву інформацію.

12.6. Мультимедійні презентації

Одним із самих популярних застосувань мультимедійних технологій є створення мультимедійних презентацій.

Мультимедійна презентація – це спосіб представлення (презентації) інформації за допомогою мультимедійних технологій. Мультимедійна презентація може містити в собі текстові матеріали, фотографії, малюнки, слайд-шоу,

звукове оформлення й дикторський супровід, відеофрагменти й анімацію, тривимірну графіку. Мультимедійні презентації створюються за допомогою спеціальних програм, наприклад, Microsoft PowerPoint, LibreOffice Impress.

У цей час мультимедійні презентації найчастіше використовуються для представлення доповідей, реклами і т. д., що пов'язане з наступними їхніми перевагами.

Компактність. Мультимедійні презентації зберігаються на цифрових носіях: дисках, USB-накопичувачах. Такі носії відзначаються малою вагою й розмірами. При цьому на одному носії можна зберігати кілька десятків різних презентацій.

Мобільність. Записана на носій презентація може бути використана скрізь, де є комп'ютер. Ще один варіант – це ноутбук, за допомогою якого можна провести презентацію в будь-яких умовах.

Обсяг. В одній мультимедійній презентації можна помістити солідний обсяг різноманітної інформації: аудіо, відео, графічної, текстової. У зручній формі клієнтові, слухачеві доноситься маса інформації.

Наочність. Сучасні інформаційні технології дозволяють працювати із графікою й відео на найвищому рівні.

Вплив на емоції. Комбінація в мультимедійній презентації відео й звукового ряду дозволяє здійснювати комплексний вплив на емоції. Грамотно підібрана палітра, відео-образи й музичний супровід здатний впливати на підсвідомість, зачіпаючи глибинні рівні, до яких складно добратися словами.

Економічність. Запис мультимедійних презентацій на цифрові носії коштує дуже недорого. Фактично, це обходиться дешевше друку гідного буклету або книги. Крім того, мультимедійні презентації можна використовувати багаторазово, немає потреби створювати нові презентації на кожний випадок.

Самостійна робота студентів

Список тем

1. Кодування, класифікація, стандартизація й алгоритмізація медичних задач.
2. Візуалізація медико-біологічних даних. Обробка й аналіз медичних зображень і біологічних сигналів.
3. Доказова медицина.
4. Медичні апаратно-програмні комплекси. Апарати й системи заміщення втрачених функцій людини.
5. Нанотехнології в медицині.
6. Системний аналіз у медичних дослідженнях.
7. Біологічна, медична й фізіологічна кібернетика.
8. Етичні й правові принципи управління інформацією в системі охорони здоров'я.

Список рекомендованої літератури

Основна (базова)

1. Медична інформатика : підручник для студентів медичних ВНЗ / за ред. В. Г. Книгавка. – Харків : ХНМУ, 2015. – 288 с.
2. Медицинская информатика : учебник для студентов 2-го курса мед. вузов : в 2 ч. / под ред. В. Г. Книгавко. – Харьков : ХНМУ, 2016. – Ч. 1. – 150 с. – Ч. 2. – 197 с.
3. Handbook of Biomedical Informatics. Електронний ресурс:
https://en.wikipedia.org/wiki/Book:Handbook_of_Biomedical_Informatics
4. Lubliner David J. Biomedical Informatics: An Introduction to Information Systems and Software in Medicine and Health / David J. Lubliner // Auerbach Publications. – 2015. – 434 p.
5. Nanette B. Health Information Management Technology: An Applied Approach / B. Nanette // American Health Information Management Association. – 2016. – 5th ed. – 686 p.
6. Mervat Abdelhak. Health Information: Management of a Strategic Resource, / Mervat Abdelhak, Mary Alice Hanken // Saunders. – 2015. – 5th edition. – 800 p.
7. Editors J. H. Handbook of Medical Informatics / J. H. Editors, V. Bommel, M. A. Musen // Електронний ресурс
<http://www.mieur.nl/mihandbook>; <http://www.mihandbook.stanford.edu>
8. Mark A. Musen B. Handbook of Medical Informatics / Mark A. Musen B. // Електронний ресурс
<ftp://46.101.84.92/pdf12/handbook-of-medical-informatics.pdf>
9. Edward H. Biomedical Informatics, 2014 / H. Edward, J. Shortliffe, J. Cimino // Електронний ресурс
<http://www.rhc.ac.ir/Files/Download/pdf/nursingbooks/Biomedical%20Informatics%20Computer%20Applications%20in%20Health%20Care%20and%20Biomedicine-2014%20-%20CD.pdf>

Допоміжна

10. Hebda T. L. . Handbook of Informatics for Nurses & Healthcare Professionals / T. L. Hebda, P. Czar // Kindle Edition. – 2012. – 5th Edition. – 624 p.
11. Medical Informatics: Computer Applications in Health Care and Biomedicine, 2011 // Електронний ресурс
<https://books.google.com.ua/books?id=WYvaBwAAQBAJ&pg=PA321&lpg=PA321&dq=book++medical+informatics&source=bl&ots=VjPvStLtIk&sig=b39YVoBlts31QSJkUf4bnAjTqfY&hl=uk&sa=X&ved=0ahUKEwiqkeTdpIzQAUGWSwKHTyIBfw4ChDoAQhHMAc#v=onepage&q=book%20%20medical%20informatics&f=false>
12. Чалий О.В. Основи медичної інформатики та обчислювальної техніки / О. В. Чалий, В. А. Д'яков, І. І. Хаїмзоню –Київ : Вища школа, 1993.
13. Інформаційні технології в охороні здоров'я і практичній медицині : у 10 кн. Кн.5. Оброблення клінічних і експериментальних даних у медицині : навч. посібник / О. П. Мінцер, Ю. В. Вороненко, В. В. Власов. – Київ : Вища школа, 2003. – 350 с.

14. Медицинская информатика : учебник / И. Е. Булах, Ю. Е. Лях, В. П. Марценюк, И. И. Хаимзон. – Киев : ВСИ «Медицина», 2012. – 424 с.
15. Medical Informatics=Медична інформатика : підручник / І. Є. Булах, Ю. Є. Лях, В. П. Марценюк, І. Й. Хаимзон. – Київ : ВСИ «Медицина», 2012. – 368 с.
16. Медична інформатика в модулях : практикум / І. Є. Булах, Л. П. Войтенко, М.Р. Мруга та ін. ; за ред. І. Є. Булах. – Київ : Медицина, 2012. – 208 с.
17. Медична інформатика підручник для студентів II курсу мед. спеціальностей / І. Є. Булах, Ю. Є. Лях, В. П. Марценюк, І.І. Хаїмзон. – Тернопіль : ТДМУ, «Укрмедкнига», 2008. – 316 с.
18. Булах І. Є. Інформаційні технології у психології та медицині : підручник / І. Є. Булах, І. І. Хаїмзон. – Київ : ВСВ «Медицина», 2011. – 216 с.
19. Інформатика в таблицях і схемах: ПК і його складові, операційна система Windows, інтернет, основні та допоміжні пристрої, системне та прикладне програмне забезпечення, моделювання та програмування / Л. І. Білоусова, Н. В. Олефіренко. – Харків : Торсінг плюс, 2014. – 111 с.
20. Основи інформатики. Microsoft Office 2013 (Word, PowerPoint на практиці) : навч. посібник / М. М. Дрінь, Н. В. Романенко ; М-во освіти і науки України, Чернів. нац. ун-т ім. Ю. Федьковича. – Чернівці : Чернів. нац. ун-т, 2014. – 75 с.
21. Інформатика та інформаційні технології : практикум для орг. роботи студентів на практ. та лаборатор. заняттях / Ю. Ю. Білак, В. О. Лавер, Ю. В. Андрашко, І. М. Лях ; М-во освіти і науки України, ДВНЗ «Ужгор. нац. ун-т», Ф-т інформ. технологій, каф. інформатики та фіз.-мат. дисциплін. – Ужгород : Аутдор-шарк, 2015.
22. Глинський Я. М. Інформатика : практикум з інформ. технологій / Я. М. Глинський. – Тернопіль : Підруч. і посіб., 2014. – 302 с.
23. Мінцер О.П. Інформатика та охорона здоров'я / О. П. Мінцер // Медична інформатика та інженерія. – 2010. – № 2. – С. 8–21
24. Основы компьютерной биостатистики. Анализ информации в биологии, медицине и фармации статистическим пакетом MedStat / Ю.Е. Лях, В.Г. Гурьянов, В.Н. Хоменко и др. – Донецк, 2006. 211 с.
25. Медичні інформаційні системи в діагностиці, лікуванні і прогнозуванні цукрового діабету / [С. М. Злепко та ін.] ; М-во освіти і науки, молоді та спорту України, Вінниц. нац. техн. ун-т. – Вінниця : ВНТУ, 2013. – 187 с.
26. Журибеда О. Системи керування базами даних : посібник / О. Журибеда. – Київ : Перше вересня, 2017. – 163 с.
27. Обвінцев О. В. Інформатика та програмування : курс на основі Python: матеріали лекцій: навч. посіб. / О. В. Обвінцев ; Київ. нац. ун-т ім. Т. Шевченка. – Київ : Основа, 2017. – 247 с.
28. Інформаційні технології: проблеми та перспективи / [Н.Г. Аксак та ін.] ; за заг. ред. В. С. Пономаренка ; М-во освіти і науки України. – Харків : Рожко С. Г., 2017. – 446 с.
29. Устенко С. А. Мережні інформаційні технології : навч. посібник для вищ. навч. заклад. / С. А. Устенко, І. В. Устенко ; М-во освіти і науки України, Миколаїв. нац. ун-т ім. В.О. Сухомлинського. – Миколаїв : Швець В. М., 2016. – 321 с.

30. Симбірська Л. М. Інформаційні системи й технології : навч. посібник для самот. вивчення. / Л. М. Симбірська, Г. Д. Симбірський, А. І. Левтеров. – Харків : ХНАДУ, 2016. – 129с.

31. Гойко О.В. Практичне використання пакета STATISTICA для аналізу медико-біологічних даних : навч. посібник / О. В. Гойко. – Київ , 2004. – 76 с

32. Добрянський Д. О. Використання телемедицини у клінічній практиці : навч.-метод. посібник / Д. О. Добрянський, О. П. Мінцер, В. В. Краснов. – Київ : Українсько-Швейцарська програма «Здоров'я матері та дитини», 2011. – Ч. 1. – 86 с.

33. Добрянський Д. О. Використання телемедицини у клінічній практиці : навч.-метод. посібник / Д. О. Добрянський, О. П. Мінцер, В. В. Краснов. – Київ : Українсько-Швейцарська програма «Здоров'я матері та дитини», 2011. – Ч. 2 (для викладача). – 94 с.

34. Європейський стандарт комп'ютерної грамотності. Ч. 1. Базові концепції інформаційних технологій. Опрацювання текстів. Первинна обробка інформації / О. З. Готра, О. В. Бойко, О. З. Ткачук, Л. Б. Лотоцька ; за ред. О. З. Готри. – Львів : ЛНМУ ім. Данила Галицького, 2007. – 78 с.

35. Готра О. З. Європейський стандарт комп'ютерної грамотності. Ч. 2. Сучасні технології обробки, аналізу та представлення інформації. Використання програм підготовки презентацій для представлення даних / О. З. Готра, Л. Б. Лотоцька, О. З. Ткачук ; за ред. О. З. Готри. – Львів : ЛНМУ ім. Данила Галицького, 2007. – 94 с.

36. Готра О. З. Європейський стандарт комп'ютерної грамотності. Ч. 3. Інформація і комунікація. Пошук та передача інформації. Використання технології баз даних для обробки та аналізу інформації / О. З. Готра, Л. Б. Лотоцька, І. С. Собчук ; за ред. О. З. Готри. – Львів : ЛНМУ ім. Данила Галицького, 2007 – 102 с.

37. Реєстрація, обробка та контроль біомедичних електронних сигналів : нав. посібник / В. Вуйцик, З. Ю. Готра, О. З. Готра та ін. – Львів : Ліга-Прес, 2009. – 308 с.

Інформаційні ресурси

1. <http://repo.knmu.edu.ua/handle/123456789/162> (сайт ХНМУ)

2. Handbook of Biomedical Informatics

https://en.wikipedia.org/wiki/Book:Handbook_of_Biomedical_Informatics

3. Societies:

www.amia.org, <http://imia-medinfo.org/wp/>, www.himss.org, www.tmi.or.th

4. U.S. Office of the National Coordinator for Health IT:

<http://www.healthcareitnews.com>

5. Journals in the Field Biomedical Informatics:

6. Healthcare Informatics

www.healthcare-informatics.com

7. Journal of the American Medical Informatics Association:

www.jamia.org

8. www.uacm.kharkov.ua (Українська асоціація "Комп'ютерна Медицина")

9. <http://www.ecdl.org/> (Офіційний сайт організації ECDL Foundation)
10. <https://support.office.com/uk-ua/> (Довідкові та навчальні матеріали пакету Microsoft Office)
11. <http://windows.microsoft.com/uk-ua/windows/help> (Довідкові та навчальні матеріали з роботи в операційній системі Microsoft Windows)

Періодичні видання:

12. Медична інформатика та інженерія. Науково-практичний журнал – http://archive.nbuiv.gov.ua/portal/chem_biol/Mii/index.html
13. Клиническая информатика и телемедицина. Журнал Украинской ассоциации «Компьютерная медицина» – <http://uacm.kharkov.ua/rus/index.shtml?r-klininfo-ujournal.htm>
14. Український журнал телемедицини та медичної телематики – http://archive.nbuiv.gov.ua/portal/chem_biol/Ujtm/index.html

ЗМІСТ

ТЕМА 1. Основні поняття медичної інформатики. Комп'ютер у діяльності майбутнього лікаря	3
ТЕМА 2. Інформаційні ресурси системи охорони здоров'я	7
ТЕМА 3. Створення та ведення медичної документації	12
ТЕМА 4. Бази даних. Системи управління базами даних	15
ТЕМА 5. Медичні інформаційні системи. Електронні медичні карти	18
ТЕМА 6. Прийоми роботи з медичною інформацією за допомогою табличного процесора	22
ТЕМА 7. Методи біостатистики. Статистичний аналіз медико- біологічних даних	26
ТЕМА 8. Кластерний аналіз у медичних дослідженнях	34
ТЕМА 9. Формальна логіка у вирішенні задач діагностики і профілактики захворювань	37
ТЕМА 10. Прийняття рішень у медицині	42
ТЕМА 11. Математичне моделювання медико-біологічних процесів	47
ТЕМА 12. Мультимедійні технології	53
Самостійна робота студентів	58
Список рекомендованої літератури	59

Навчальне видання

Кнігавко Володимир Гілярійович
Зайцева Ольга Василівна
Бондаренко Марина Анатоліївна
Батюк Лілія Василівна
Рукін Олексій Сергійович

МЕДИЧНА ІНФОРМАТИКА

**Навчальний посібник
для студентів медичних університетів**

Відповідальний за випуск М. А. Бондаренко



Комп'ютерний набір О. С. Рукін
Комп'ютерна верстка О. Ю. Лавриненко

Формат А4. Ум.-друк. арк. 2,1. Зам. № 19-33857.

**Редакційно-видавничий відділ
ХНМУ, пр. Науки, 4, м. Харків, 61022
izdatknmurio@gmail.com**

Свідоцтво про внесення суб'єкта видавничої справи до Державного реєстру видавництв, виготівників і розповсюджувачів видавничої продукції серії ДК № 3242 від 18.07.2008 р.