МОДУЛЬ 1
1 ФУНДАМЕНТАЛЬНІ ПОНЯТТЯ

1. Загальна теорія систем.
2. Діалектичні принципи, що застосовуються в загальній теорії систем.
3. Системний аналіз його основні етапи.
4. Постулати загальної теорії систем.
5. Роль системних представлень у практичній діяльності
6. Поняття системи.
7. Цілісність системи.
8. Елементи системи.
9. Види зв’язки між елементами системи.
10. Структурні і функціональні схеми системи.
11. Поняття властивості системи.
12. Типи властивостей системи.
13. Стан системи.
14. Типи станів.
15. Трансформація станів і форми їх опису.
16. Класифікація систем.
17. Складність та подовження системи.
18. „Метасистема”.

1 . 1 Предмет і область визначення теорії систем

Ядром курсу «Основи теорії систем і управління» є загальна теорія систем (ЗТС).

Загальна теорія систем (ЗТС) - науковий напрямок, пов'язаний з розробкою сукупності філософських, методологічних, конкретно-наукових і прикладних проблем аналізу і синтезу складних систем довільної природи. Часто загальну теорію систем називають *«системологією»*. Останню визначають як науку про системи.

Найбільш важливою рисою ЗТС є її міждисциплінарний характер. Основою для цього є аналогічність (ізоморфізм) процесів, що протікають у системах різного типу, різної природи. Це дає можливість переносити знання з однієї галузі в іншу.

Кожне з трьох слів, що входять у назву цього наукового напрямку, має своє визначене значення.

«Загальна» означає, що ЗТС має дедуктивний характер і поєднує інші теорії. Під дедуктивним розуміють, що більшість умовиводів даної теорії про елементи множин і відносини між ними виконують на основі знання загальних властивостей усіх множин.

«Теорія» - підклас висловлень, що вважаються правдивими. Теорія є системою узагальненого знання, пояснення тих чи інших боків дійсності. Теорія є духовним, уявним відображенням і відтворенням реальної дійсності. Структуру теорії формують принципи, аксіоми, гіпотези, закони, судження, положення, поняття, категорії, факти.

Теорія складається з відносно твердого ядра і його захисного поясу. У ядро входять основні принципи. Захисний пояс теорії містить допоміжні гіпотези, що конкретизують ядро.

Під назвою *„теорія систем"* поєднані математичні поняття і методи, що використовуються для вивчення широкого кола явищ і процесів.

«Система» - безліч взаємодіючих елементів (компонентів) і відносин між ними, які в цілому виконують певну функцію. Елемент системи - структурна одиниця, що має риси, які виражають головну якість системи. Відносини характеризують зв'язки між елементами (наприклад, менше ніж..., включене в..., транзитивно..., тотожньотаін.).

Визначень поняття «система» може бути безліч, і це залежить від прийнятого рівня «абстрагування». Під *абстрагуванням* розуміють уявне відволікання від несуттєвих властивостей, зв'язків, відносин і виділення деяких сторін, що цікавлять дослідника. Воно, як правило, здійснюється у два етапи. На першому етапі визначаються несуттєві властивості, зв'язки та ін. На другому - досліджуваний об'єкт заміняється іншим, більш простим, що представляє собою спрощену модель об'єкта, але зберігає головне в простому. Розрізняють такі рівні абстрагування:

- символічний чи лінгвістичний;
- теоретико-множинний;
- абстрактно-алгебраїчний;
- топологічний;
- логіко-математичний;
- теоретико-інформаційний;
- динамічний;
- евристичний.

Метою ЗТС є вивчення законів, які діють у системах довільної природи на всіх рівнях організації.

У загальній теорії систем виділяють два напрямки.

У першому напрямку ЗТС розглядають як *метатеорію* різних системних концепцій, розробок спеціалізованих теорій систем. Даний напрямок аналізує розвиток ЗТС як деякої філософської концепції, що включає в себе такі поняття, як системні принципи, системний підхід, системний аналіз та ін.

Другий напрямок - теорія систем узагалі, що представляє собою деякий математичний апарат, який описує закономірності формування і розвитку будь-яких систем. Теорію систем взагалі називають *абстрактною теорією систем (АТС)*. Іноді абстрактну теорію систем розглядають як молоду галузь кібернетики.

ЗТС орієнтована на застосування **діалектичних принципів**: цілісності, системності, релятивності, універсальності і розвитку.

Під *принципом* розуміють правило, яке виникає в результаті суб'єктивно осмисленого досвіду людей.

Принцип цілісності. Відповідно даного принципу система виявляє свою цілісність як конкретний об'єкт, у якого є граничні і якісні властивості.

Принцип системності орієнтований на цілісне представлення досліджуваних об'єктів.

Принцип релятивності системи вказує, що будь-яку безліч предметів можна розглядати як систему і як несистему.

Принцип універсальності системи вказує, що завжди можна знайти такий аспект, стосовно якого щось можна описувати як систему.

Принцип розвитку вказує на безперервне співвіднесення й узгодження зовнішньої і поточної внутрішньої детермінант системи в період її існування.

Системний підхід - поняття, що підкреслює значення комплексності, широти охоплення і чіткої організації в дослідженні, проектуванні й плануванні. Системний підхід спирається на відомий діалектичний закон взаємозв'язку і взаємозумовленості явищ у світі й суспільстві. Він вимагає розглядати досліджувані явища й об'єкти не тільки як самостійну систему, але й як підсистему деякої великої системи. Найбільш повно сутність системного підходу сформульована В.Г. Афанасьєвим, який визначив ряд взаємозалежних аспектів, що у сукупності й єдності складають системний підхід. Такими аспектами є:

- системно-елементний, що відповідає на запитання, з чого (яких компонентів) утворена система;
- системно-структурний, який розкриває внутрішню організацію системи, спосіб взаємодії утворюючих її компонентів;
- системно-функціональний показує, які функції виконує система та компоненти, що її утворюють;
- системно-комунікаційний розкриває взаємозв'язок даної системи з іншими як по горизонталі, так і по вертикалі;
- системно-інтегративний, що показує механізми, фактори збереження, удосконалення і розвитку системи;
- системно-історичний (еволюційний) відповідає на запитання як, яким чином виникла система, які етапи проходила у своєму розвитку, які її історичні перспективи.

Системний аналіз - процедура вирішення взаємозалежних одна з іншою проблем.

Аналіз у широкому значенні - це метод пізнання за допомогою розчленовування чи розкладання предметів дослідження на складові частини. Основна мета системного аналізу - забезпечення цілісного, всебічного підходу до вирішення складної проблеми. У його основі лежить поняття системи. На базі цього поняття виконують урахування всіх обставин. При цьому використовують кількісне порівняння всіх альтернатив для того, щоб уможливити свідомий вибір найкращого в даній ситуації вирішення. Оцінку альтернатив виконують за різними критеріями, наприклад, за критерієм вимірності, ефективності, надійності та ін. Головну роль у системному аналізі відіграють теоретичні побудови, засновані на таких загальних поняттях, як «зв'язок», «відношення», «властивість», «процес» та ін.

Системний аналіз можна розглядати як наукову дисципліну, що на основі системно організованих, структурно взаємозалежних і функціонально взаємодіючих евристичних процедур, методологічних засобів, математичного апарата, програмного забезпечення та обчислювальних можливостей комп'ютерних систем і мереж забезпечує в умовах концептуальної невизначеності отримання і накопичення інформації про досліджуваний предмет. Така інформація потрібна для наступного формування знань про цей предмет як єдиного, цілісного об'єкта з позицій побудови мети дослідження й ухвалення раціонального вирішення в умовах різнорідних багатофакторних ризиків.

Основні етапи системного аналізу:

- визначення цілей системи і встановлення їхньої ієрархії до початку процесу ухвалення рішення;
- структуризація: виділення об'єкта дослідження і середовища його існування;
- розробка математичних моделей, що відбивають зміст цілей;
- визначення обмежень і вимог, що накладаються на систему середовищем;
- розробка різних (альтернативних) способів досягнення цілей;
- оцінка варіантів вирішень, яка базується на всьому прийнятому комплексі критеріїв;
- вибір кращого варіанту.

Загальна теорія систем спирається на три постулати.

Відповідно до *першого постулату* функціонування систем будь-якої природи може бути описано на основі розгляду формальних структурно-функціональних зв'язків між окремими елементами систем.

Другий постулат показує, що організація системи може бути визначена на основі спостережень, проведених ззовні за допомогою фіксованих станів тільки тих елементів системи, що безпосередньо взаємодіють з її оточенням.

Третій постулат показує, що організація системи цілком визначає її функціонування і характер взаємодії з навколишнім середовищем.

Постулати ЗТС дають можливість вирішувати дві задачі:

1) визначення організації системи виходячи з характеристик взаємодії із зовнішнім середовищем;
2) визначення характеристик взаємодії виходячи з організації системи.

Теоретичною частиною загальної теорії систем є кібернетика, теорія інформації, теорія ігор, теорія прийняття рішень, топологія, факторний аналіз, теорія множин, теорія осередків, теорія графів, теорія мереж, теорія автоматів, теорія масового обслуговування та ін.

Прикладною частиною ЗТС є системотехніка, дослідження операцій, інженерна психологія, ергономіка.

Через те, що системні дослідження носять загальний, міждисциплінарний характер, тобто стосуються утворення, розвитка, функціонування, синтезу будь-яких систем, цілком правомірним є таке світоглядне запитання: чи не заміняють вони філософію, чи не є вони новою загальною методологією науки?

Деякі теоретики системного методу відповідають на це запитання позитивно, деякі негативно. Для внесення ясності в цю проблему необхідно розглянути три її аспекти.

По-перше, прихильники позитивної відповіді занадто вузько розуміють функції філософського знання, зводячи їх до однієї - бути Методологією наукового дослідження.

По-друге, не можна теоріям ставити за провину ті світоглядні характеристики, що даються їхніми творцями.

По-третє, слід рішуче відмовитись від метафізичної альтернативи: діалектика чи системний підхід. У процесі системних досліджень виникає багато філософських проблем, але це не дає підстави ототожнювати системний підхід з філософією. Системний метод, подібно математичним методам відноситься до загальнонаукових методів дослідження, що існує у взаємозв'язку, з одного боку з конкретно науковими, з іншого боку - із загальними (філософськими) методами. Тобто, багато понять теорії систем - загальнонаукові поняття.

Таким чином, системний підхід, з одного боку, дозволяє застосовувати ряд загальнофілософських положень для вирішення приватних задач, а з іншого боку - збагачує саму філософію за рахунок розвитку конкретних наук.

Загальна теорія систем зародилася в 30-х роках XX сторіччя і в 50-і роки сформувалася в самостійний науковий напрямок. Ще в 1945 році при організації сектора філософських питань природознавства в Інституті філософії Академії наук СРСР академік СІ. Вавілов сформулював мету і завдання роботи сектора, серед яких головне місце займали проблеми цілісності й системності.

Засновниками ЗТС вважають біолога Л. Берталанфі, фахівців із математичних проблем у галузі біології і психології Р. Жерара, А. Рапопорта, А. Хола, Р. Фейджина, Л. Хартлі, економіста К. Боуддінга, фахівця з лінгвістики М. Месаровича, бельгійського фізика И. Прігожина. Великий внесок у розробку ЗТС внесли радянські вчені А.І. Ракітов, В.М. Глушков, В.Г. Афанасьєв, І.В. Блауберг, Є.Г. Юдін, П.К. Анохін, М.П. Бусленко, А.І. Берг, В.М. Бурков, Ю.М. Горянський, Л.А. Колісников, І.Ф. Зразків, В.М. Садовський, В.В. Павлов, Г.П. Мельников, Ю.А. Шредер та інші.

У 1954 році у СІЛА було організовано «Суспільство досліджень в області загальної теорії систем» (Society for General Systems Research). З 1956 року це суспільство видає щорічник «General Systems». У 1959 році при Кейсівському технологічному інституті (США) був створений «Центр системних досліджень». У 1963 році корпорація «Інтернейшенал бізнес машинз корпорейшен» організувала Інститут системних досліджень. З 1969 року в СРСР видається щорічник Системні дослідження». У ці роки був створений теоретичний базис математичного і методологічного інструментарію формалізації й автоматизації на базі ЕОМ процедур вирішення реальних організаційних і технічних системних проблем у різних сферах практичної діяльності. На даному етапі досліджень у 70 роках XX сторіччя сформувалась методологічна криза, яка проявилась у невідповідності існуючої методології глобальній проблематиці. Це привело до розробки багаторівневої, ієрархічної структури методології системного аналізу глобальної проблематики.

Теперішнім часом методологічні та інструментальні засоби системології дозволяють досліджувати властивості складних об'єктів довільної природи, а також причини їхнього виникнення, динаміку Адаптації і розвитку. Завдяки цьому можна більш ефективно здійснювати управління будь-якими складними об'єктами для забезпечення їхньої стійкості і безпеки, прогнозувати еволюції систем, розробляти і конструювати системи будь-якої природи.

Системологія, яка інтерпретована в термінах конкретної науки, може виконувати методологічні функції, використовуватись як засіб «діалектичної обробки» створюваного конкретною наукою знання, тобто як засіб «організації» ноосфери.

1.2. Система та її компоненти

Поняття *«система»,* як і всяке інше поняття, - ідеальний об'єкт, що відбиває у свідомості деяку групу (клас) матеріальних об'єктів (явищ), які виділяють у процесі пізнання з матеріального світу. Поняття - це думка, що відбиває істотні і необхідні ознаки предмета 'чи явища. Усяке поняття як сукупність істотних ознак об'єкта, "Об'єднаних в одній думці, є результат логічної *абстракції (лат.* abstrahere - відволікати) від всіх об'єктів даного роду. Ми відволікаємо їхні істотні ознаки і потім поєднуємо ці ознаки в одну цільну думку, чи в одну ідею, про даний об'єкт.

Існує безліч визначень поняття системи. Найбільш поширені з них:

- система - цілісна взаємозалежна безліч об'єктів;
- система - цілісна безліч об'єктів (елементів), пов'язаних міжсобою взаємними відносинами;
- система - порядок (план, класифікація), згідно якого розташовується група понять для утворення єдиного ставного цілого;
- система - сукупність взаємозалежних, певним чином організованих і взаємодіючих елементів;
- система - організована безліч структурних елементів, взаємопов'язаних і виконуючих певні функції;
- система - комплекс виборчо залучених компонентів, у яких взаємодія і взаємини здобувають характер взаємосприяння компонентів на одержання фіксованого корисного результату;
- система - сукупність взаємозалежних елементів, відособлена від середовища і взаємодіюча з нею як ціле.

Якщо розглянути усі визначення системи, тоді можна виділити дві тенденції: одна з них спрямована на отримання формулювання поняття системи у найбільш загальному вигляді, що відбиває структуру складних понять, їхню членованість, абстрактність; друга спрямована на підкреслення у понятті «система» практичної спрямованості об'єктів. У другій тенденції фігурують мета, продукція, результат.

Перша тенденція використовує вищий рівень абстрагування. Тому у визначення системи включені й різного роду *конгломерати,* суміші, властивості яких адитивні і зводяться до простої суми властивостей його складових. Так, наприклад, кілька молекул газу теж система, оскільки наявність хоча б двох молекул, що вдаряються, вже означає їхню механічну взаємодію. Але такі системи ще не мають *інтегративних,* цілісних властивостей. Головна ознака конгломерату полягає в тому, що при включенні до нього чи виключенні з нього компонентів ні сам конгломерат, ні його компоненти не перетерплюють скільки-небудь помітних якісних змін. Конгломерат лише збільшується чи зменшується в розмірах. Кожний з компонентів конгломерату автономний, його зміна залежить лише від нього самого, тому що зв'язки між компонентами носять зовнішній, украй хитливий характер. Яскравим прикладом конгломератів є транспортні потоки малої щільності. Вилучення з їхнього складу чи включення до нього частини транспортних засобів не приводить до помітних якісних змін потоку. Конгломерати часто називають сумативними системами.

Друга тенденція у визначенні поняття системи показує її доцільність, прагнення до досягнення визначеної мети. Наявність мети виступає у ролі системоутворюючого фактора, що поєднує компоненти в єдине ціле, перетворює конгломерат у цілісну систему. Головною відмінністю цілісної системи від сумативної є наявність у цілісній системі інтегративних властивостей. Системи, де в якості системоутворюючого фактора виступає мета, називають *цілеспрямованими.* Визначення таких систем характерно для нижчого рівня абстрагування. Конгломерати і цілеспрямовані системи - крайні точки системності об'єктивного світу.

На проміжних рівнях абстрагування визначення системи відображують ступінь її цілісності. Цілісна система не зводиться до механічної суми частин, що її утворюють. Так, звертаючись до молекул газу, можна сказати, що при наявності достатньої кількості молекул для повного прояву статистичних закономірностей, система молекул стає термодинамічно цілісною системою. Вона чітко виявляє нові інтегративні властивості (температуру, тиск), яких окремі молекули не мають. Аналогічне явище спостерігається у транспортному потоці. При збільшенні щільності потоку залежність руху одного автомобіля від іншого збільшується, рух автомобілів стає зв'язаним. Перехідний інтервал за часом між автомобілями від вільного до зв'язаного руху складає 6-9 с. У зв'язаному стані транспортний потік здобуває якісно нові цілісні властивості. Такі властивості називають *емерджентними* (породженими). Їх немає у окремих елементів системи. Вони виникають у результаті взаємодії цих елементів. Красивий приклад прояву породженої властивості привів М. Арбиб: „Нехай є деякий цифровий автомат S, що перетворить будь-яке ціле число на його вході в число, на одиницю більше вхідного. Якщо з'єднати два таких автомати послідовно в кільце, то в новій системі виявиться нова властивість: вона генерує зростаючі послідовності на виходах А и В, рис. 1.1. Одна з цих послідовностей складається тільки з парних, інша - тільки з непарних чисел.”

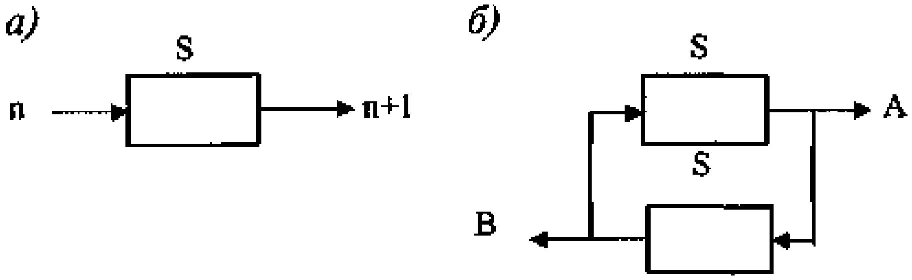

Рис. 1.1. Ілюстрація властивості емерджентності:
***а)* вихідний цифровий автомат; *б)* з'єднання двох автоматів.**

Слід зазначити, що нічого містичного в породженні нових властивостей немає. Нові властивості виникають завдяки конкретним зв'язкам між конкретними елементами. Інші зв'язки дадуть інші властивості. Наприклад, паралельне з'єднання тих же автоматів нічого не змінює в арифметичному відношенні, але збільшує надійність обчислень, якщо на вихід надходить сигнал тільки від справного автомата.

Властивість емерджентності є проявом внутрішньої цілісності системи.

Відзначаючи відміну цілісної системи від системи сумативної, необхідно мати на увазі, що абсолютної границі між ними не існує. У процесі еволюції матеріального світу в результаті дії інтегративних сил (скажемо, сил притяжіння у фізичних і хімічних процесах) адитивні системи здобувають характер цілісних систем, і, навпаки, у наслідок дії сил дезорганізації, наприклад зростання ентропії, цілісні системи нерідко розпадаються і перетворюються в системи сумативні. Перетворення сумативних систем у цілісні може відбуватись різним чином: способом зміцнення зв'язків вже існуючих компонентів; за допомогою розвитку взаємодії компонентів, раніше не взаємодіючих; способом утворення нових компонентів та їхніх зв'язків чи перегрупування наявних та ін. Перетворення цілісних систем у сумативні теж може відбуватись різним чином: за допомогою ослаблення чи руйнування обмежених зв'язків компонентш; у результаті вилучення чи втрати компонентів та ін. Прикладом також може бути транспортний потік. Так, зменшення щільності потоку веде до зменшення його зв'язності і перетворенню у вільний потік, тобто конгломерат.

Якщо кожна частина системи зв'язана з іншою частиною таким чином, що зміна в одній частині викликає зміну у всіх інших частинах та у всій системі, то вважають, що система поводиться *когерентно* чи як *ціле.* Іншу крайність складає безліч частин, зовсім це пов'язаних між собою, так що зміна в кожній частині залежить винятково від цієї частини. Така поведінка називається незалежністю чи *фізичною адитивністю.*

Цілісність (чи когерентність) і незалежність (чи адитивність) - це не дві окремі властивості, а крайні ступені тієї ж самої властивості.

Формально, у теоретико-множинному смислові, цілісність системи може бути оцінена за допомогою міри системності *М,,* запроонованої Г.П. Мельником:

$$M_\tau = \frac{|F_b \cap F_t|}{|F_b \cup F_t|},$$

де F_t - множина необхідних функціональних станів;

F_b -множина можливих функціональних станів;

$\cap$ - операція пересічення множин;

$\cup$ - операція об'єднання множин.

Поняття «система» відрізняється від поняття «комплекс». *Комплекс* означає в перекладі з латинського «зв'язок», «сполучення». У світі все взаємозалежно, але одного взаємозв'язку для розуміння системи недостатньо, тут потрібно виділити конкретну специфіку взаємодії елементів сукупності. Комплексність тільки грань системності.

Неодмінними приналежностями системи є елементи, компоненти, частини, саме те, з чого безпосередньо утворене ціле і без чого воно неможливе.

Звичайно під *частиною* системи розуміють щось просторово віддільне від системи. Наприклад, дерево можна розчленувати на безліч окремих шматків, кожний з який називають частиною. Таке представлення про частину системи носить механістичний характер. Більш обґрунтованим є представлення частини системи у вигляді її компонента.

Компоненти - це взаємодіючі структури цілісної системи, що підкоряються тим же законам, що і вся система. Звичайно як компоненти виступають підсистеми. ***Підсистема*** - це система в системі більш високого порядку ***(надсистемі).***

Підсистеми утворені компонентами нижчого порядку стосовно системи в цілому. Усякий розподіл цілого на частини, усякий розподіл системи на підсистеми є відносним, деякою мірою умовним. Наприклад, гальмову систему автомобіля можна віднести або до ходової частини, або до підсистеми керування. Іншими словами, границі між підсистемами умовні, відносні. Це стосується і границь між самою системою і навколишнім середовищем.

Елемент (елементарний) означає кінцевий, неподільний, останній. Елемент - це поріг членування в межах даної якості системи й елементарний його носій. Зрозуміло, елемент неподільний не взагалі, а тільки в рамках даної якості. Членування елемента виводить дослідника в якісно іншу систему.

Відносини між елементами, компонентами, підсистемами і системами реалізуються через *зв'язки* між ними. Зв'язки можуть бути енергетичними, речовинними, інформаційними, внутрішніми і зовнішніми, прямими і зворотними.

Елементи, компоненти, підсистеми і системи в цілому графічно позначають квадратами, ромбами, прямокутниками, окружностями та ін. Зв'язки - суцільними, пунктирними лініями і стрілками, рис. 1.2. Кожен елемент, компонент та ін. мають вхід і вихід.

Рис. 1.2. Позначення елементів і зв'язків

Вхід - місце прикладання зовнішньої дії. Зовнішню дію називають стимулом, подразником, вхідним сигналом.

Вихід - місце зняття вихідної характеристики. Вихідну характеристику називають вихідним сигналом, реакцією системи на зовнішню дію.

Часто вихідну характеристику визначають як зміну вихідного сигналу в часі.

Через входи із зовнішнього середовища у певні моменти часу в систему надходить речовина, енергія чи інформація; в інші моменти часу результати процесів їхнього перевтілення надходять у зовнішнє середовище через виходи.

Через співвідношення вхідних і вихідних величин можливі різні схеми їхньої взаємодії. Найбільш типовими є такі чотири схеми (рис. 1.3):

1) одномірно-одномірна (однофункціональна) - один вхідний сигнал і одна вихідна характеристика;
2) одномірно-багатомірна - один вхідний сигнал і декілька вихідних характеристик;
3) багатомірно-одномірна - декілька вхідних сигналів і одна вихідна характеристика;
4) багатомірно-багатомірна - декілька вхідних сигналів і декілька вихідних характеристик.

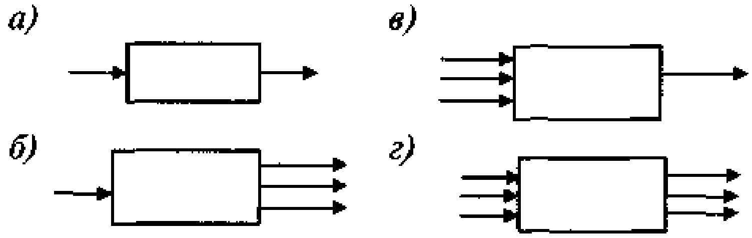

Рис. 1.3. Схеми взаємодії:
***а)* одномірно-одномірна; *б)* одномірно-багатомірна;**
***в)* багатомірно-одномірна; *г)* багатомірно-багатомірна.**

Елементи і компоненти системи, їхні входи і виходи по різному пов'язані між собою. Існують такі види зв'язків: незамкнені, замкнені і складні.

Основні *незамкнені зв'язки:* прямий послідовний (простий) зв'язок *(a);* паралельний розподільчий зв'язок *(b);* паралельний з'єднуючий зв'язок *(c);* послідовний зв'язок між системами A, B, C, D *(d);* непрямий зв'язок між системами A і C, A і D, B і D *(d);* паралельний зв'язок, що розгалужується *(e)* (рис. 1.4).

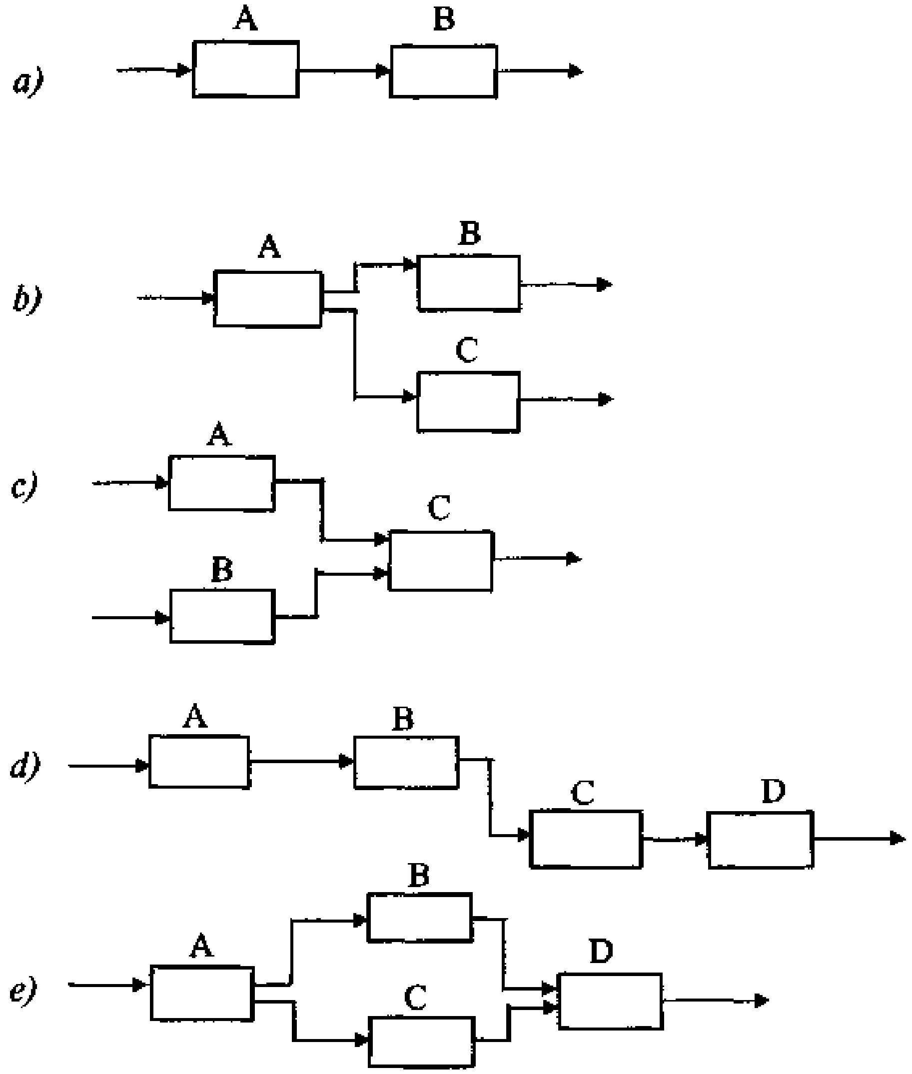

Рис. 1.4. Види незамкнених зв'язків у системах

Зв'язки типу *c)* і *e)* характеризують випадок, коли з'єднуються виходи багатьох однотипних елементів. Такого роду з'єднання називають масовими, а системи, що містять такі з'єднання - *масовими системами.*

Основні *замкнені зв'язки* формуються за допомогою зворотного зв'язку. *Зворотний зв'язок* - це зв'язок між виходом і входом того самого елемента чи системи. Він може здійснюватись безпосередньо або через інші елементи системи. Зворотний зв'язок, що зменшує дію вхідного впливу на вихідну величину, називають негативним, а той, що збільшує цей вплив - позитивним. Негативний зворотний зв'язок сприяє відновленню рівноваги в системі, порушеної зовнішньою дією, а позитивний - підсилює відхилення від рівноважного стану відносно до його величини в системі без такого зворотного зв'язку. До складу замкнених зв'язків включають: власний зворотний зв'язок *(а)*; прямий зворотний зв'язок *(b);* непрямий зворотний зв'язок *(c, d)* (рис. 1.5).

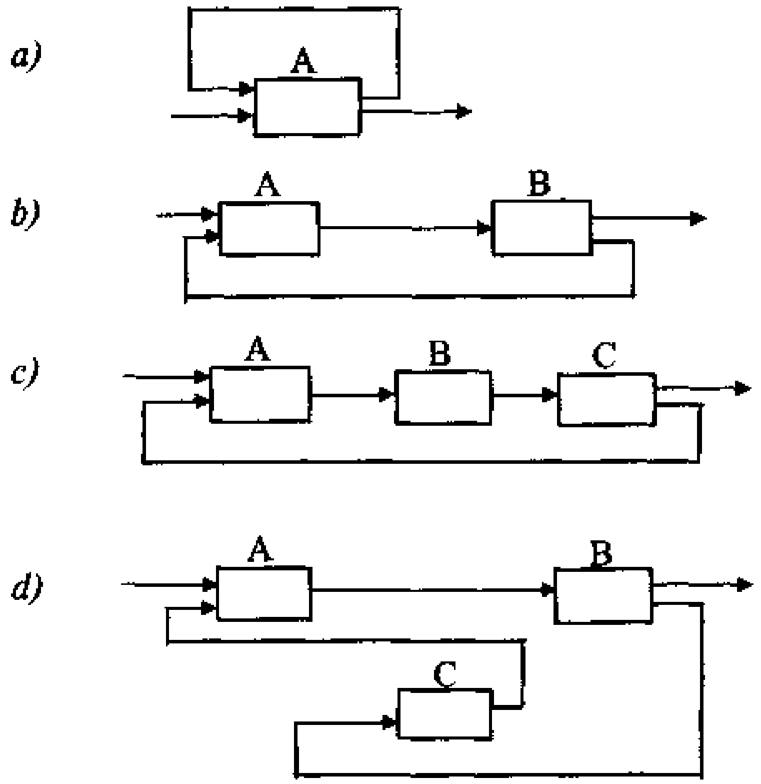

Рис. 1.5. Основні види замкнених зв'язків у системах

Складні зв'язки. У складних системах виникає безліч комбінацій зв'язків між окремими елементами і підсистемами. Найчастіше зустрічаються: зворотний паралельний розподільчий зв'язок *(a)*; зворотний паралельний сполучний зв'язок *(b);* послідовний паралельний зв'язок *(c)* (рис. 1.6).

Аналіз елементів, компонентів і зв'язків між ними дозволяє встановити з чого складається система. Але при дослідженні системи важливо розкрити, як улаштована, організована система, як і в ім'я чого вона діє. Тут перед дослідником постає питання про структуру і функції системи.

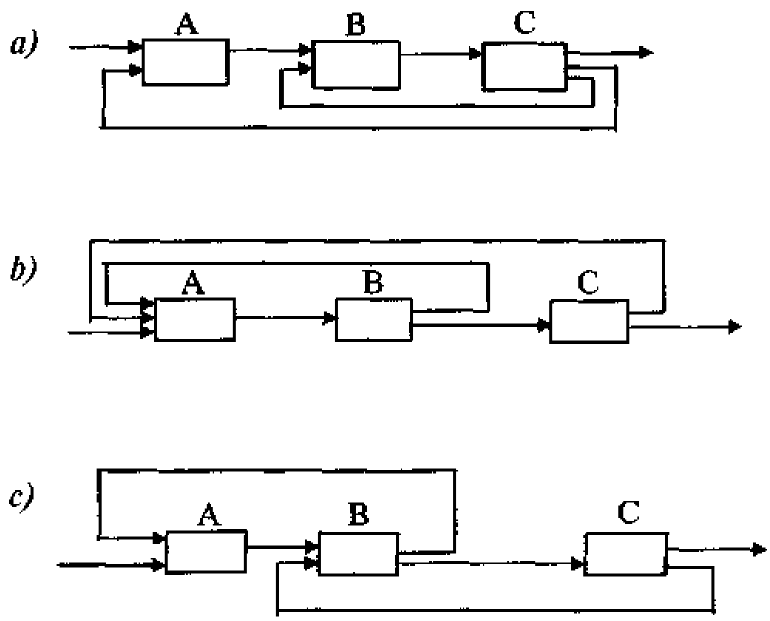

Рис. 1.6. Види складних зв'язків у системах

Структура - внутрішня організація системи, що представляє собою специфічний спосіб взаємозв'язку, взаємодії утворюючих його компонентів. Структура - це упорядкованість, організованість системи. Поняття структури вживається і в іншому, більш широкому смислі як сукупність необхідних і достатніх для досягнення мети відносин між елементами.

Поняття *«функція»* вживається в самих різних значеннях. Вона може означати здатність до діяльності і саму діяльність. Такий підхід до визначення функції розглядає її як спосіб прояву активності, життєдіяльності системи і її компонентів. Часто функцію визначають як реакцію системи на зовнішній вплив. Але такий підхід справедливий тільки для однофункціональних систем.

У широкому значенні функція - це форма, спосіб прояву активності, життєдіяльності системи та її компонентів. Функції характерні й системі, й її компонентам, причому функцією системи є інтегративний результат утворюючих її компонентів.

Поняття функції близько до поняття мети, вони тісно пов'язані одне з одним. *Мета* - це майбутній стан системи, стан, до якого вона прагне. Звичайно мета даної системи ставиться системою більш високого порядку (надсистемою), до якої входить дана система. З позицій системи більш високого порядку, мета даної системи є її функція. Тоді *функція* - це *призначення* системи чи її компонентів, що реалізується через спрямоване перевтілення вхідних сигналів у вихідні характеристики. Функція - це частка участі системи у функціонуванні надсистеми. Надсистема впливає на систему таким чином, що властивості системи як функціонального об'єкта змінюються в напрямку посилення здатності сприяти ефективному функціонуванню надсистеми.

Формально функція представляється залежністю вихідної характеристики від вхідних сигналів.

Для графічного представлення систем використовують структурні і функціональні схеми.

Структурна схема - схематичне зображення взаємодії між елементами, компонентами, підсистемами і зовнішнім середовищем. У структурній схемі вказуються всі елементи системи, усі зв'язки між елементами усередині системи і зв'язки певних елементів з навколишнім середовищем. Часто структурні схеми представляють у вигляді графів. Граф складається з позначень елементів, що називаються вершинами, і позначень зв'язків між ними, що називаються ребрами (іноді дугами). Графи можуть зображувати будь-які структури, якщо не накладати обмежень на пересіченість ребер. Деякі типи структур мають особливості, важливі для практики, вони виділені з інших і отримали спеціальні назви. Так, в організаційних системах часто Зустрічаються *лінійні, деревоподібні, матричні, мережні* структури іструктури зі *зворотними зв'язками* (рис. 1.7).

Функціональна (феноменологічна) схема - графічне представлення функцій системи чи її компонентів без відображення їхнього внутрішнього структурування. Компоненти функціональної схеми не вимагають того, щоб бути цілком подібними компонентам прототипу, тому що метою створення такої схеми є тільки опис функцій прототипу. Кожний з компонентів системи, власне кажучи, є «чорною скринькою». Поняття ***чорної скриньки*** образно підкреслює повну відсутність відомостей про її внутрішню будову.

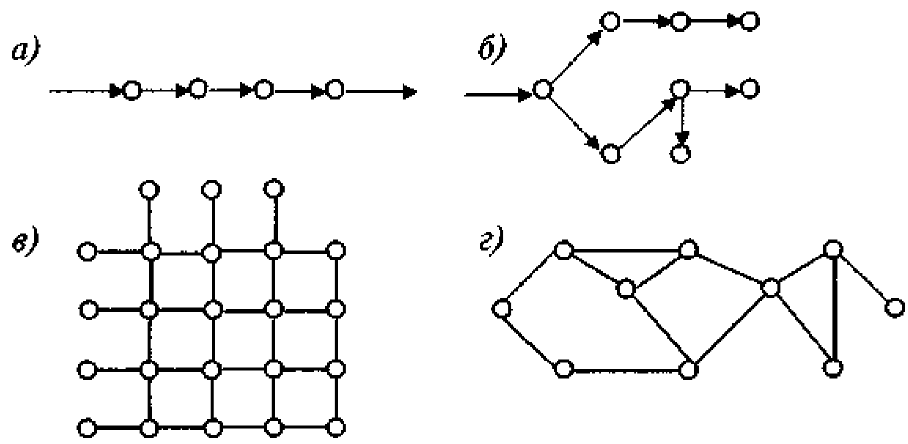

Рис. 1.7. Графи різних структур:
а) **лінійна структура;** *б)* **деревоподібна структура;**
в) **матрична структура;** *г)* **мережна структура**

Чорна скринька - основа макропідходу до аналізу систем, тобто таких методів вивчення системи, що ґрунтуються на властивостях співвідношення вхід-вихід.

1.3. Властивості систем

Властивість - здатність системи виявляти ті чи інші сторони в процесі взаємозв'язку і взаємодії. Ця здатність обумовлюється внутрішньою природою системи, її будовою, структурою. Властивість можна розглядати як окремий випадок відношення. У відношенні бере участь не менше двох об'єктів, а властивістю ми називаємо деякий атрибут одного об'єкта. Іншими словами, властивість - це одномісне згорнуте відношення.

Розрізняють три типи властивостей:

1) властивості, обумовлені структурою і функціями системи;

2) властивості, що формують здатність до самозбереження системи;

3) властивості, що характеризують тактику і стратегію поводження системи при досягненні мети. Такі властивості часто називають функціональними чи цільовими.

До першої групи властивостей відносять ***властивості, породжені відносинами в предметах,*** уже представлених як системи. Ці властивості часто називають *системоутворюючими.* У їхнє число включають: *структурну і функціональну складність, організованість, розчленованість, взаємну автономність елементів, варіативність, елементарність, іманентність, надійність, однорідність, завершеність, мінімальність та ін.*

Головними в сполуці цих властивостей є ***структурна і функціональна*** *складності.*

Існує кілька концепцій визначення складності:

- логічна концепція, заснована на аналізі властивостей предикатів, які характеризують систему;

- алгоритмічна концепція, що визначає складність як довжину алгоритму відтворення системи;

- обчислювальна концепція, що прив'язує алгоритмічну складність до засобів обчислень;

- статистична концепція, яка характеризує складність через міру розрізнення розподілів імовірностей;

- теоретико-множинна концепція, що ототожнює складність системи з числом її елементів;

- теоретико-інформаційна концепція, яка пов'язує складність системи з її ентропією.

У рамках теоретико-інформаційної концепції в якості міри ***складності*** *У.Р.* Ешбі запропоновано використовувати розмаїття, що кількісно оцінюється числом можливих станів системи «n». Надалі Л. Хартлі було показано, що для оцінки складності зручно використовувати логарифмічну міру

$$H_m = \log n,$$

де H_m - міра складності (максимальна ентропія системи).

Оцінка складності за числом станів не показує, в якому **з** можливих станів система знаходиться та в який вона перейде в наступний момент часу.

Якщо покласти, що P_i є імовірність того, що система знаходиться в і-тому стані, тоді для оцінки стану системи можна використовувати формулу К. Шеннона:

$$H = -\sum_{i=1}^{n} P_i \cdot \log P_i,$$

де H - поточна невизначеність системи.

Для закритої системи її організованість (визначеність) чи абсолютна організація може бути оцінена за формулою:

$$Q = H_m - H,$$

де Q - абсолютна організація системи.

Якщо ліву і праву частини даного рівняння розділити на H_m, ми одержимо оцінку відносної організації системи у формі Г. Ферстера:

$$R = 1 - \frac{H}{H_m}, \text{ де } R = \frac{Q}{H_m}.$$

Розчленованість системи характеризує її здатність поділятись на складові.

Властивість ***взаємної автономності елементів*** системи проявляється в тому, що кожному її елементу притаманні властивості системи в цілому.

Варіативність - це прояв здатності змінюватись, проходячи ряд станів, перетворюватись **в** іншу систему.

Елементарність оцінюється відношенням структури системи до її елементів. Кожен елемент елементарної системи не є системою з такою ж структурою як і система.

Властивістю, яка визначається у тому ж відношенні, що й елементарність, можна вважати ***іманентність.*** Дана властивість проявляється в тому, що системоутворююче відношення охоплює тільки елементи даної системи і не поширюється поза нею.

Надійність - здатність зберігати свою системоутворюючу властивість при елімінації будь-якої кількості елементів, крім одного.

Властивість ***однорідності*** може відноситись і до елементів і до самої структури системи. У першому випадку, коли системи складаються з однорідних елементів, говорять про елементну (чи субстрактну) гомогенність. У випадку однорідної структури система може бути структурно (чи функціонально) гомогенною. Однорідність, звичайно, розуміють в тому ж самому плані, в якому утворена система, стосовно системоутворюючій властивості.

Завершеність системи виявляється в тому, що вона не допускає приєднання нових елементів без руйнування цієї системи.

На відміну від завершеності ***мінімальність*** указує не на можливість приєднання, а на можливість вилучення елементів. Якщо структура системи така, що система руйнується при вилученні хоча б одного елемента, то система мінімальна.

Властивості, які формують здатність системи до самозбереження, часто називають зберігаючими. У їхній структурі виділяють:

а) здатність зберігати стаціонарний нерівновагий стан незалежно від умов зовнішнього середовища та умов функціонування системи, що змінюються;

б) здатність утримувати суттєві перемінні внутрішнього середовища в допустимих межах (стосовно до технічних систем цю властивість називають *гомеостазом,* а до біосистем - ***гомеостазисом);***

в) здатність пристосовуватись (адаптуватись) до зміни речовинного, енергетичного та інформаційного середовища (процес адаптації розглядається як зміна властивостей і поведінкових реакцій системи, спрямована на збереження її цілісності і гомеостазу на всіх системно-структурних рівнях організації);

г) здатність динамічно реагувати на зміни і впливи навколишнього середовища за рахунок структурно-функціональних перебудов (властивість динамічності);

д) властивість якісної неоднорідності, що виявляється в тому, що в рамках однієї й тієї ж функціональної системи спільно і злагоджено працюють підсистеми з якісно різними керуючими сигналами (речовинними, енергетичними, інформаційними);

є) властивість тимчасової неоднорідності, яка проявляється в тому, що в одній функціональній системі взаємодіють підсистеми з різними постійними часу (повільні, швидкі, надшвидкі);

ж) властивість структурно-функціональної організованості, що виявляється у високій стійкості структури і функцій системи;

з) властивість структурно функціональної стохастичності, яка проявляється в різновиді реакцій у відповідь на одні й ті ж дії середовища;

к) властивість структурної дискретності, що виявляється в можливості поділу системи на дискретні елементи, компоненти, підсистеми;

л) властивість функціональної безперервності, тобто варіабельність кількісних параметрів у межах однієї і тієї ж дискретності.

До ***властивостей, що характеризують тактику і стратегію поводження системи,*** відносять:

а) властивість спонтанної чи внутрішньої активності;

б) властивість цілеспрямованої активності.

Властивість ***спонтанної активності*** є найважливішим для будь-якої біологічної системи. Визначаючи організм як спонтанно активну систему, Л. Берталанфі писав, що «навіть при постійних зовнішніх умовах і при відсутності зовнішніх стимулів організм являє собою не пасивну, а істотно активну систему. Про це свідчать, зокрема, функція нервової системи і поведінка. Внутрішня активність, а не реакції на стимули лежить в основі цих процесів. Можна також показати, що це поняття справедливе і для еволюції нижчих тварин, і для розвитку перших рухів ембріонів, і утробного плоду».

Властивість спонтанної активності ще більше відноситься до людини, що обумовлено її соціальною природою. Дана властивість лежить в основі мотивації цілеположення і поводження, а також становлення різних форм пристосування. Ясно, що оскільки людина може виконувати рухи чи акти, а самі поняття руху чи акта невіддільні від поняття мети, тим самим властивість спонтанної активності визначає властивість бути багатоцільовим. При цьому під «багатоцільовим» розуміють можливість досягнення будь-якої мети з деякого набору цілей. Вибір цілей і порядок їхнього досягнення носить спонтанний характер.

Спонтанна активність мінімальна в зоні комфорту. Так, наприклад, водій, керуючи автомобілем, мимоволі качає кермове колесо, відшукуючи положення на проїзній частині дороги, при якому сумарне психічне примушення, що формується за рахунок впливу об'єктів середовища праворуч і ліворуч від напрямку руху, мінімальне. У випадку відшукання зазначеного мінімуму амплітуда качань стає мінімальною. Аналогічне явище спостерігається в льотчика, що качає літак за рахунок мимовільного руху штурвала керування.

Вище вказувалось, що будь-яка дія чи рух невіддільний від поняття мети. Тому у випадку усвідомленого вибору мети спонтанна активність стає ***цілеспрямованою.*** У системах з цілеспрямованою активністю спонтанна активність підлегла закону диференціальної чутливості.

Спонтанна і цілеспрямована активність системи забезпечують її готовність підтримувати функціонування надсистеми.

1.4. Стан системи

Стан системи - це якась (внутрішня) характеристика системи, значення якої в даний момент часу визначає поточне значення вихідної величини.

Представлення про стан системи пов'язується із широким колом показників і характеристик, які визначають її функціонування і реакції на різні зовнішні впливи. Стан системи - точно визначена умова чи властивість, яка може бути пізнана, чи повторена знову.

Стан системи в момент часу t_0 є такий набір зведень про поводження системи, якого разом з деяким можливим вхідним впливом, заданим при $t_0 \le t \le t_j$ досить для однозначного визначення вихідного сигналу для $t_0 \le t \le t_j$ при кожному $t_j \le t_0$. Звичайно при *канонічному* представленні системи її *стан визначається як найменший набір чисел, який необхідно задати в даний момент часу, щоб можна було в рамках математичного опису системи передбачити її поводження в будь-який майбутній момент часу.* Так, наприклад, для опису стану виробничого підприємства як мінімум необхідно задати числа, що оцінюють наявні потужності q_1 трудові ресурси q_2, запаси матеріалів q_3 та ін. Дані числа називають координатами стану чи зображуючими точками елемента (системи) у k-мірному просторі станів. Такий простір називають ***гіперпростором.***

При зміні стану системи змінюються її координати, вони стають змінними величинами і представляються у вигляді $q_1(t_1), q_2(t_2), ..., q_n(t_n)...$ Ці величини прийнято називати характеристиками станів системи.

Залежно від мети дослідження до уваги приймають тільки вагомі характеристики. При зміні мети досліджень склад вагомих характеристик може змінюватись. Чим ретельніше будуть відібрані вагомі характеристики, тим точніше можливий опис змін стану системи і тим ефективніше вплив на неї.

Агрегація координат стану формує вектор стану Q.

Для практичного аналізу динаміки станів системи важливі три типи станів: нульовий стан, стан, що установився і стан рівноваги.

Нульовим станом називають деякий стан θ, для якого

$$0 = \psi(\theta, 0, t)$$
$$\text{при всіх } t_0 \le t < \infty$$

Іншими словами, нульовий стан має таку властивість: якщо система знаходиться в нульовому стані

$$Q(t_0) = \theta$$

і вхідний вплив є нульовим

$$X(t) = 0, \quad t_0 \le t < \infty,$$

тоді вихідний сигнал системи також виявляється нульовим

$$Y(t) = 0, \quad t_0 \le t < \infty,$$

де X, Y - вектори входів і виходів відповідно.

Помітимо, що нульовий стан необов'язково єдиний.

Стан, що встановився, якщо він існує, є такий єдиний стан y, у який система приходить при нульовому вхідному впливі незалежно від початкового стану.

Стан рівноваги є деякий стан η, у якому система залишається при нульовому вхідному впливі

$$\Phi(\eta,0,t)=0 \text{ при будь-яких } t_0 \geq t < \infty$$

Зміна стану реального об'єкта неминуче пов'язана з переносом і перевтіленням речовини, енергії чи інформації, і не може відбуватись миттєво. Процес переходу з одного стану в інший називають ***перехідним процесом,*** а сам перехід - ***трансформацією*** (перевтіленням) стану.

У кібернетиці деякий початковий стан, на який чиниться вплив з метою трансформації, називають ***операндом*** (операнд - те, що випробує дію). Діючий фактор називають ***оператором.*** Новий стан системи, що виникає під дією оператора, називають ***образом*** (образ - те, у що перетворюється операнд). Наприклад, нагрівається вода, перетворюється в пару. Вода - операнд, тепло - оператор, новий стан - пар. Ця зміна описується формулою

вода → пар

і називається переходом. Стрілкою вказується напрямок переходу.

Трансформація станів може бути представлена такими способами:

а) словесний опис;
б) загальна форма;
в) матрична форма;
г) кінематичний графік;
д) математичні формули;
є) логічні формули.

Прикладами ***словесного опису*** можуть бути технологічні карти виконання робіт. Наприклад, виконання робіт при влаштуванні шару дорожнього одягу з ґрунту, укріпленого бітумом у межах однієї за-хватки включає такі операції: транспортування ґрунту, розрівняння ґрунту, транспортування води, роздрібнення і зволоження ґрунту, транспортування бітуму, змішування ґрунту з бітумом, ущільнення ґрунту, обробленого бітумом. На наступній захватці всі операції повторюються.

Даний словесний опис характеризує трансформацію станів машинно-дорожньої ланки з влаштування шару дорожнього одягу.

Загальна форма трансформації описує процес виконання дорожньо-будівельних робіт як процес переходу від безлічі первісних станів до безлічі наступних, що здійснюються деяким часом.

Приймемо позначення: транспортування ґрунту - а; розрівняння ґрунту - *б*; транспортування води - *в*; роздрібнення і зволоження ґрунту *г*; транспортування бітуму - *д*; змішування ґрунту з бітумом – *е*; ущільнення ґрунту, обробленого бітумом - *ж*.

Тоді загальна форма запису прийме вигляд:

а б в г д е ж
↓
б в г д е ж а,

де стрілка означає напрямок переходу.

Верхній рядок букв у даному записі представлений операндами, нижній - образами. Остання операція позначена буквою *ж* переходить на операцію *а,* що відповідає переходу машинно-дорожньої ланки на нову захватку.

У даному записі кожен елемент нижнього рядка зустрічається серед елементів верхнього рядка. Такий перехід відноситься до *замкненого виду.* Якщо безліч наступних станів буде містити й інші стани порівняно з вихідними, то таке перевтілення відносять до *незамкненого виду.* Так, незамкненим є перевтілення

а б в г д е ж
↓
б в г д е ж.

Замкненим перевтіленням можна моделювати дії, що мають деяку тривалість і циклічно повторюються. Незамкнені перевтілення моделюють дії чи діяльність, розвиток яких від деякого стану невідомий, зведення про них закінчуються.

У тому випадку, коли кожному операнду відповідає один образ, перевтілення називають *однозначним.* У складі однозначних перевтілень виділяють дві групи - *взаємооднозначні* та *однозначні лише в один бік.*

Взаємооднозначними є такі перевтілення, в яких початкові і наступні стани можна однозначно перетворювати в обох напрямках. Наприклад,

а б с д
↓
б с д а.

Якщо кожному елементу з безлічі операндів відповідає один і тільки один елемент із безлічі образів, але не кожному образу відповідає один і тільки один операнд, тоді таке перевтілення є однозначним лише в один бік. Наприклад,

а б с д
↓
б с д б.

Тут образу *б* відповідають вихідні стани *аід.*

Однозначними перевтіленнями моделюють розвиток детермінованих систем.

У тих випадках, коли в системі якому-небудь початковому стану відповідає кілька наступних станів, тоді таке перевтілення називають *неоднозначним.*

Неоднозначне перевтілення може бути записане у вигляді

а   *б ...*
↓
б + с   *с + д ...*

З'єднання + між декількома наступними станами в одному переході використані у виключному смислові. Наприклад, операнд *а* перейде або в образ *б* або в образ *с* з якоюсь імовірністю. Приведений запис є неоднозначним перевтіленням з імовірністю умовно (1/3, 2/3), (3/4, 1/4).

Неоднозначними перевтіленнями можна моделювати розвиток імовірнісних систем.

Розрізняють ще один вид перевтілень - *тотожні* перевтілення. При цьому перевтіленні не відбувається ніяких змін, і кожен образ збігається зі своїм операндом.

Якщо всі операнди різні, то це перевтілення є взаємооднозначним:

$$\begin{array}{c} а\ б\ с\ д \\ \downarrow \\ а\ б\ с\ д. \end{array}$$

При описі трансформації способом ***матричної форми*** переходи від одного стану до іншого позначають одиницями на полях матриці (табл. 1.1).

Таблиця 1.1. Матриця станів машинно-дорожньої ланки

		Початкові стани				
		a	b	c	d	e
Наступні стани	a	-	-	-	-	1
	b	1	-	-	-	-
	c	-	1	-	-	-
	d	-	-	1	-	-
	e	-	-	-	1	-

У широкому смислові матрицею розміру mхn називають прямокутну таблицю чисел, розташованих у m рядках і n стовпцях. У матричній формі можна записувати однозначні і неоднозначні перевтілення.

Спосіб запису за допомогою ***кінематичного графіка*** показано на рис. 1.8. Послідовність переходів на кінематичному графіку зображується стрілками між початковим і наступним станом.

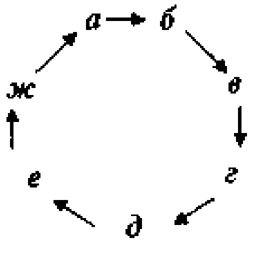

Рис. 1.8. Кінематичний графік

Математичні і ***логічні формули*** використовують для зображення характеристики кількісної зміни трансформації (але не відображення якісної зміни стану, як це було зроблено в чотирьох представлених способах запису).

Загальною формулою такого запису може бути $n' = f(n)$, де *n* - початковий стан, *n'* - наступний стан. Цей спосіб запису дозволяє пояснити основну операцію при трансформації. Розміри зміни початкового стану можуть бути враховані величиною коефіцієнта трансформації *k*. Тоді, якщо $n' = kn$ і $n'' = kn'$, підставляючи значення *n',* одержимо $n'' = kkn$. З цього можна зробити висновок, що, знаючи *k* і початковий стан *n* для деякого інтервалу часу *t,* можна обчислити наступний стан $n = k'n$.

Прикладом застосування математичних формул для опису трансформації стану системи є математичний опис транспортних потоків на основі використання математичного апарату марковських випадкових процесів. Під марковським процесом (процесом без післядії) розуміють процес, що має такі властивості: для кожного моменту часу t_0 імовірність будь-якого стану системи в майбутньому (при $t > t_0$) залежить тільки від її стану в сьогоденні і не залежить від того, коли і яким чином система прийшла в цей стан.

Марковським ланцюгом називають таку випадкову послідовність подій, при якій у ланцюжку подій імовірність переходу з будь-якого стану S_i в кожний S_j не залежить від того, коли і як система прийшла в стан S_i. При використанні ланцюгів Маркова ставиться завдання визначення імовірності перебування системи в якому-небудь стані.

Рух потоку автомобілш представляється як марковський процес переходу потоку зі стану S_1 у стани S_2, S_3, S_4, ... S_n. Імовірності станів $p_1(t), p_2(t),..., p_n(t)$ визначаються як результат рішення системи диференціальних рівнянь. При складанні диференціальних рівнянь користуються графом станів (рис. 1.9).

На графі позначають щільність потоку подій λ_{ij} що переводять систему із стану S_i у S_j. Для системи, показаної на рис. 1.9, диференціальні рівняння мають вигляд:

$$\frac{dp_1(t)}{dt} = \lambda_{21}p_2(t) + \lambda_{31}p_3(t) - (\lambda_{12} + \lambda_{13})p_1(t);$$

$$\frac{dp_2(t)}{dt} = \lambda_{12}p_1(t) - (\lambda_{21} + \lambda_{23} + \lambda_{24})p_2(t);$$

$$\frac{dp_3(t)}{dt} = \lambda_{13}p_1(t) + \lambda_{23}p_2(t) - (\lambda_{31} + \lambda_{34})p_3(t);$$

$$\frac{dp_4(t)}{dt} = \lambda_{24}p_2(t) + \lambda_{34}p_3(t).$$

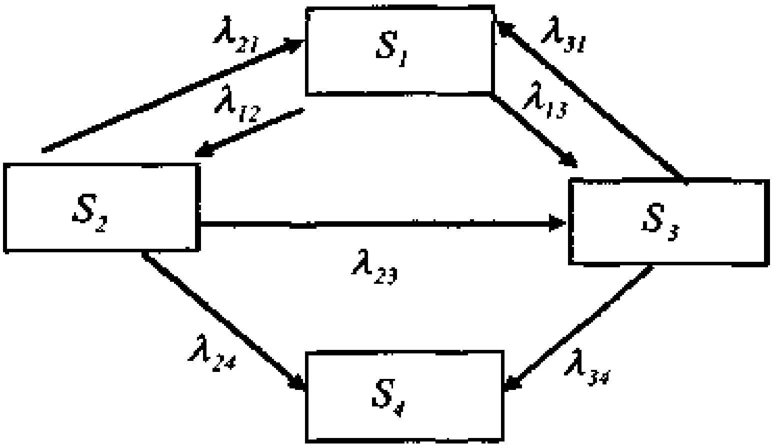

Рис. 1.9. Граф станів потоку автомобілів

При вивченні й описі перевтілень, що відбуваються у системах, конструктивним є їх опис мовою, властивою символічній логіці з елементами Бульової алгебри.

Бульова алгебра - це алгебраїчна система, що залежно від обставин може бути інтерпретована або як система подій, або як система висловлень. У ролі подій розглядають майбутні стани системи. Під *висловленням* у логіці розуміють оповідальне речення, що має властивість бути класифіковане або як щире, або як помилкове, але не як те й інше разом.

Між подіями і висловленнями існує деяка аналогія. Це дозволяє обслуговувати логіку і теорію імовірностей тим самим формальним апаратом. Такий апарат був створений Дж. Булєм і згодом розвинутий С.М. Бернштейном і А.І. Колмогоровим.

Причинно-наслідковий зв'язок подій чи логічний зв'язок висловлень описують формулами, які мають вигляд нерівностей. Наприклад, нерівність $x < y$ виражає велику вірогідність події y порівняно з подією x, і ця ж нерівність може виражати більшу правдоподібність висловлення y порівняно з висловленням x.

Кожна логічна перемінна, що відображує подію чи висловлення, може здобувати два протилежні значення, які позначають нулем чи одиницею (1 - подія відбулася, висловлення щире; 0 - подія не відбулася, висловлення помилкове). 1 і 0 розглядають як найбільший і найменший елементи Бульової алгебри. Крім цих елементів використовують інші, які називають ***сентенціональними зв'язками.*** У висловленнях це слова «не», «і», «чи», «якщо..., тоді...», «тоді і тільки тоді, коли...». З їхньою допомогою з простих речень, що не містять зв'язки, можна утворити складні речення.

Речення, змінене словом «не», називають ***запереченням.***

З'єднання двох речень за допомогою зв'язки «і» називають ***кон'юнкцією*** цих двох речень. З'єднання двох речень у складне за допомогою зв'язки «чи» називають ***диз'юнкцією*** двох речень. Складне речення, побудоване з двох простих за допомогою зв'язки «якщо..., тоді...», називають ***імплікацією.*** Складне речення, утворене за допомогою зв'язки «тоді і тільки тоді, коли...», називають ***еквіваленцією.***

Для позначення зв'язок використовують такі символи: - для «не», $\wedge$ для „і”, $\vee$ для „чи”, $\rightarrow$ для „якщо..., тоді...”, $\leftrightarrow$ для „тоді і тільки тоді, коли...”.

1.5 Класифікація систем

Для класифікації систем широко використовують такі ознаки:
1) путь прояву цілісності;
2) субстанціональну природу системи;
3) тип елементів;
4) тип відносин між елементами усередині системи і зовнішнім середовищем;
5) обумовленість взаємодії;
6) системоутворюючі властивості.

Залежно від ***путі прояву цілісності*** системи розділяють на зовнішні і внутрішні.

Внутрішня система - це цілісне утворення, до якого можна застосувати членування, представляючи цю систему у вигляді деякої структури складових частин.

Зовнішня система - це клас об'єктів загальної природи, об'єднаних деякою цілісною сутністю. Елементи такої системи можуть не мати ні просторової, ні тимчасової спільності, ні навіть генетичного зв'язку. Важлива лише спільність природи утворюючих систему об'єктів.

Субстанціональна природа системи визначається її сутністю, характером і походженням. Відкіля система - чи вона існує в об'єктивному світі, чи у свідомості людини.

За цією ознакою виділяють чотири класи систем.

Перший клас систем - ті, що існують в об'єктивній дійсності, неживій і живій природі, суспільстві.

Другий клас - системи концептуальні, ідеальні. Елементами таких систем є поняття. Іноді ні системи називають абстрактними.

Третій клас - системи штучні, які спроектовані, сконструйовані і створені людиною.

Четвертий клас - змішані системи, в яких органічно злиті елементи, що є продуктом природи, і елементи, «придумані», створені людиною. Прикладом таких систем є система «людина - автомобіль».

За типами елементів розрізняють технічні, біологічні, біотехничні, фізичні, хімічні, політичні, економічні системи та ін.

За типом відносин між елементами усередині системи і зовнішнім середовищем розрізняють відкриті і закриті системи, лінійні і нелінійні, ієрархічні, керовані системи, цілеспрямовані, адаптивні, системи, що самоорганізуються та ін.

Еквівалентні за типом відносин системи поєднуються в класи, усередині класів формуються підкласи. Таку еквівалентність називають ізоморфізмом. Тому системи в класах і підкласах ізоморфні.

У *закритих системах* надходження з зовнішнього середовища речовини, енергії та інформації не приводить до зміни стану системи. Закрита система розглядається як абсолютно відособлена, що не має зовнішніх входів і виходів. У реальній дійсності такі системи не існують. Але часто виявляється зручним користатись такою абстракцією, обриваючи на деякому кроці зовнішні зв'язки.

У *відкритих системах* надходження ззовні речовини, енергії чи інформації змінюють систему. Елементи відкритих систем можуть бути віднесеш або до зовнішньої системи, або до зовнішнього середовища.

Лінійність чи *нелінійність* системи визначається її статичною характеристикою. Під *статичною характеристикою* системи розуміють зв'язок між величиною зовнішнього впливу $x(t)$ на систему (величиною вхідного сигналу) і максимальною величиною (амплітудою) вихідної характеристики y_m. Якщо $y_m=f(x)$ лінійна, то і система лінійна. Нелінійність статичної характеристики і наявність запізнювання в реагуванні є ознакою нелінійності системи.

Поняття «лінійності» означає наявність деякого виду пропорційності між вхідними і вихідними перемінними. Точне математичне визначення лінійності пов'язане з нульовим станом системи і нульовим вхідним впливом.

Система називається *лінійною щодо нульового стану,* якщо вона задовольняє таким двом умовам:

1.*Однорідність.* Якщо y_1 є реакція на нульовий стан при довільному вхідному впливі x, тоді cy є реакція на нульовий стан при вхідному впливі cx, де с - довільна константа.

2. *Адитивність.* Якщо y_1 є реакція на нульовий стан при довільному вхідному впливі x_1, а y_2 є реакція на нульовий стан при довільному вхідному впливі x_2, тоді y_1+y_2 є реакція на нульовий стан при вхідному впливі x_1+x_2.

Систему називають *лінійною щодо нульового вхідного впливу,* якщо виконуються дві умови:

1.*Однорідність.* Якщо реакція на нульовий вхідний вплив для будь-якого довільного початкового стану Q_0 є y_0, тоді реакція на нульовий вхідний вплив cx_0 є cy_0.

2.*Адитивність.* Якщо y_1 є реакція на нульовий вхідний вплив при будь-якому довільному початковому стані Q_1, а y_2 - реакція на нульовий вхідний вплив при будь-якому довільному початковому стані Q_2,тоді $y_1 + y_2$ - реакція на нульовий вхідний вплив при початковому стані $Q_1 + Q_2$.

Для лінійної системи важлива наявність властивості *декомпозиції.*

Система має *властивість декомпозиції,* якщо вона задовольняє такій умові: якщо y_0 - реакція на нульовий вхідний вплив для довільного початкового стану і y_u - реакція на нульовий стан для довільного початкового вхідного впливу, тоді результуюча реакція на той же початковий стан і вхідний вплив є $y_0 + y_u$.

Таким чином, система може бути названа *лінійною,* якщо вона лінійна щодо нульового стану, лінійна щодо вхідного впливу і задовольняє умові декомпозиції.

До лінійної системи застосовується *принцип суперпозиції.* Відповідно з цим принципом при впливові на систему декількох вхідних сигналів кожен з них фільтрується, начебто ніякі інші сигнали на систему не діють. Загальний вихідний сигнал за принципом суперпозиції утвориться в результаті підсумовування її реакції на кожен вхідний сигнал.

За обумовленістю дії розрізняють системи з детермінованою дією і системи з випадковою (імовірнісною, стохастичною) дією.

У *детермінованій* системі складові її елементи і зв'язки між ними взаємодіють точно передбаченим способом. У цих системах при фіксованих зовнішніх умовах і способі керування перехід з одного стану в інший цілком визначений.

У *випадковій* (імовірнісній, стохастичній) системі складові її елементи і зв'язки між ними взаємодіють таким чином, що не можна зробити точного, детального передбачення її поводження. Така система завжди залишається невизначеною, і передбачення про її майбутнє поводження ніколи не виходить з рамок імовірнісних категорій, за допомогою яких це поводження описується.

Стосовно детермінованої системи можна говорити про статичність чи динамічність.

Для *статичної* системи середні арифметичні значення вихідного сигналу на різних відрізках часу не виходять за допустимі межі, обумовлені точністю методики виміру досліджуваного показника.

У *динамічній* системі середнє арифметичне значення вихідного сигналу на різних відрізках часу змінюється, оскільки в такій системі відбувається зміна станів її елементів. Елементи системи, що змінюються, розглядаються як змінні величини. Якщо ці змінні величини допускають їхній вимір і представлення у вигляді конкретних чисел, то можна допустити конкретну оцінку стану системи. Ця оцінка відбиває кількість інформації, що міститься в системі, тобто те, що можна довідатись про неї.

Розрізняють *безперервні* і *дискретні* динамічні системи. У першому випадку процес перевтілення вхідного сигналу у вихідну характеристику розглядається в часі як безперервний, у другому - тільки у фіксовані (дискретні) моменти.

Усе вищевикладене показує, що динамічними можна назвати системи, в яких відбуваються зміни у часі.

Випадкові системи поділяються на *стаціонарні, нестаціонарні* й *ергодичні.* Такий розподіл систем заснований на різній залежності від часу основних статистичних характеристик.

Поняття стаціонарності різне у вузькому і широкому смислах.

У вузькому смислові система стаціонарна, якщо відсутня зміна в часі математичного чекання і дисперсії вихідного сигналу, чи зміна залежності між членами ряду.

У широкому смислові у стаціонарній системі математичне чекання не залежить від часу, а автокореляційна функція залежить тільки від величини інтервалу $\tau = t_2 - t_1$.

Система, стаціонарна у вузькому смислові, завжди стаціонарна в широкому смислові. Зворотне твердження не справедливе. Для процесів, підпорядкованих нормальному закону розподілу, поняття стаціонарності у вузькому і широкому смислах співпадають.

У різних системах властивість стаціонарності виражена неоднаково. В одних процесах стаціонарність зберігається протягом тривалого часу, в інших - лише на коротких відрізках часу. Розрізняють *ергодичні* і *неергодичні* стаціонарні системи. Для ергодичних систем усереднення за часом однієї реалізації вихідної характеристики досить великої тривалості рівносильне усередненню за ансамблем реалізацій вихідних характеристик. Для ергодичної системи автокореляційна функція прагне нуля при необмеженому збільшенні $\tau = t_2 - t_1$. У неергодичних системах ця умова не виконується.

За системоутворюючими властивостями розрізняють: прості, складні, дуже складні системи, метасистеми, розчленовані і нерозчленовані, елементарно автономні та елементарно неавтономні, варіативні і неваріативні, елементарні і неелементарні, іманентні і неіманентні, суцільнонадійні і несуцільнонадійні, однорідні і неоднорідні, завершені і незавершені, мінімальні і немінімальні системи та ін.

Простими прийнято вважати системи, що не мають розгалуженої структури (не можна виділити ієрархічні рівні), з невеликою кількістю взаємозалежних і взаємодіючих елементів, що виконують найпростіші функції. Ці системи легко піддаються опису.

Складними вважають системи з розгалуженою структурою і значною кількістю взаємозалежних і взаємодіючих елементів, що виконують більш складні функції. Високий ступінь зв'язності елементів у складних системах приводить до того, що зміна якого-небудь одного елемента чи зв'язку спричиняє зміну багатьох інших елементів системи. У складних системах можлива наявність декількох різних структур, декількох різних цілей. Але конкретний стан складної системи може бути описано.

Дуже складними називають такі системи, стан яких за тими чи іншими причинами неможливо докладно і точно описати.

Метасистеми - це надскладні системи, для яких сучасний рівень знань недостатній для проникнення в суть зв'язків системи або вони незрозумілі.

У терміні «метасистема» використовується грецький префікс «мета». У перекладі з грецького він має три значення:

1) «мета X» називається те, що спостерігається (має місце) після X, тобто є передумовою мета X;

2) вираження «мета X» показує, що X міняється і служить загальною назвою цієї зміни;

3) «мета X» використовується як назва того, що вище X в тому смислові, що воно більш високо організовано, має більш високий логічний тип чи розглядається в більш широкому смислові.

«Метасистема» включає всі три смисли цього поняття. Отже:

1) метасистема може бути визначена тільки після того, як визначені інші типи систем;

2) метасистема описує зміну - заміну однієї системи іншою;

3) метасистема вище інших систем.

Комбінуючи різниці систем за обумовленістю дії і за ступенем складності, виділяють дев'ять типів систем, яким можна дати такі визначення:

1. *Проста система* з детермінованою дією. Ця система містить мало елементів і взаємних зв'язків, легко описується, її динамічні властивості легко передбачати.

2. *Складна система з детермінованою дією* має розгалужену структуру, багато елементів із складними зв'язками, доступна в описі. Зміну її станів можливо передбачати. У подібних системах кожне відхилення від заздалегідь передбаченої дії є помилкою, що свідчить про псування системи.

3. *Дуже складні системи з детермінованою дією* практично не піддаються опису, хоча і зустрічаються в житті.

4. *Прості системи з квазідетермінованою дією.* Ці системи функціонують у детермінованому режимі лише в окремі періодично повторювані відрізки часу. На цих відрізках вони легко описуються Поводження таких систем на інших відрізках часу можна передбачити з залученням теорії імовірностей.

5. *Складні системи з квазідетермінованою дією.* Ці системи ще піддаються опису в детермінованому режимі функціонування. Але розвиток таких систем точно передбачити неможливо.

6. *Дуже складні системи з квазідетермінованою дією.* На детермінованих ділянках функціонування таких систем можливо приблизно описати їхнє поводження, на імовірнісних - неможливо ні описати, ні передбачити точно їхній розвиток.

7. *Прості системи з імовірнісною дією.* Можливе передбачення їхнього поводження із залученням теорії імовірностей.

8. *Складні системи з імовірнісною дією.* Ці системи піддаються опису в статиці, фотографічному опису. Розвиток таких систем точно передбачити неможливо.

9. *Дуже складні системи з імовірнісною дією.* Ці системи практично описати неможливо.

Слід розрізняти поняття *«великої»* і *«складної»* системи. У понятті *«великої системи»* фіксується співвідношення різновидів системи, що спостерігається, і можливостей її спостереження і дослідження. При виділенні такої системи задається одне відношення між її елементами. Якщо ж при виділенні системи задається не одне, а безліч відносин між елементами і відповідно утворюється не одна, а безліч структур, то така система характеризується неоднорідністю, різноякісністю виділених елементів і зв'язків, структурною розмаїтістю. Таку систему називають складною. Поєднання величини і складності формує *великі прості* і *великі складні* системи.

Вище показано, що складність системи кількісно може бути оцінена за допомогою максимальної ентропії системи. Відносна організація характеризує обумовленість поводження системи.

Враховуючи, що складність і відносна організація не пов'язані між собою лінійно, представляється можливим використовувати ці характеристики для побудови класифікаційної діаграми систем (рис. 1.10).

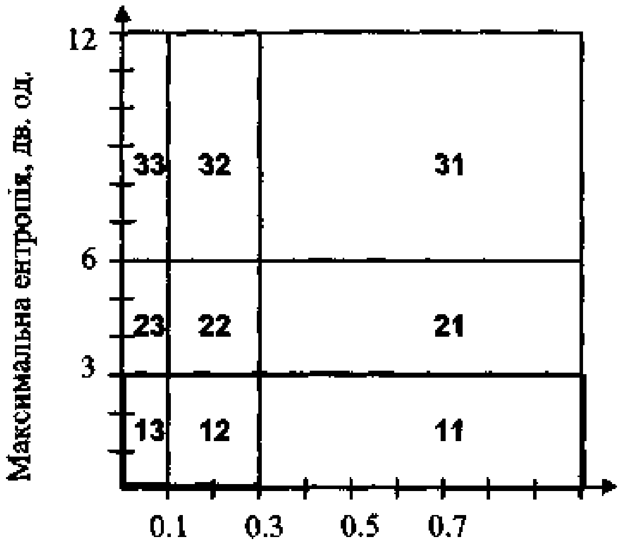

Рис. 1.10. Класифікаційна діаграма:
11-13 — проста детермінована, квазідетермінована, імовірнісна системи;
21-23 — складна детермінована, квазідетермінована, імовірнісна системи;
31-33 — дуже складні системи

При цьому для класифікації систем за складністю можна користуватись шкалою С. Біра:

- прості системи ($0 < H_m \leq 3$);
- складні системи ($3 < H_m \leq 6$);
- дуже складні системи ($6 < H_m \leq 12$);
- метасистеми ($H_m > 12$).

У свою чергу, для класифікації систем за обумовленістю дії можна користатись шкалою Ю.Г. Антомонова:

- детерміновані системи ($0,3 < R \leq 1$);
- квазідетерміновані системи ($0,1 < R \leq 0,3$);
- імовірнісні системи ($0 < R \leq 0,1$).

Розподіл систем відносний. Ця відносність виявляється в наступному: кожна система може характеризуватись і вивчатись з різних позицій, що визначається точкою зору дослідника; кожна система може бути представлена як елемент більш загальної суперсистеми (системи більш високого рангу, порядку). У той же час елементи чи групи елементів даної системи у визначених умовах можна розглядати як системи.

Групи елементів, що розглядаються як системи більш низького рангу, виділяються. При цьому враховують відносно стійкий порядок внутрішніх відносин між елементами системи і внутрішню структуру системи. Таким чином, класифікація систем завжди прив'язується до деякого рівня, тобто до їхнього *ієрархічного* розташування в структурі організаційної побудови систем.

В ієрархії систем первинним прийнято вважати такий елемент чи таку сукупність елементів системи, що не допускають їхнє подальше розчленовування без втрати основної якості всієї системи, враховуючи обрану дослідником точку зору.

Система другого рівня (порядку) поєднує дві і більше первинні системи; третього рівня (порядку) - дві і більше системи другого порядку та ін. При виділенні систем другого, третього і подальшого порядків виходять з таких принципових положень:

- поділ системи на внутрішні підсистеми здійснюється так, щоб загальна цілеспрямованість функціонування всієї системи зберігалась;
- виділення внутрішніх підсистем здійснюється, враховуючи виникнення деяких особливих характеристик для кожного з виділених рівнів;
- кількість виділених рівнів має бути мінімальною, але не повинна утруднювати (ускладнювати) вивчення систем кожного рівня.

Ієрархічна побудова систем як методичний прийом дозволяє успішно вирішувати багато практичних питань, пов'язаних з удосконаленням керування.

Цілеспрямований розвиток *складних* і *дуже складних імовірнісних систем* має потребу в керуванні. Якщо на систему здійснюється цілеспрямований вплив і система реагує на нього, то такі системи називають *керованими системами.*

У кожній керованій системі можна виділити дві підсистеми - керовану і керуючу. Враховуючи такий розподіл саме *керування* можна охарактеризувати, як такий цільовий вплив виконавця (керуючої підсистеми) на об'єкт функціонування (керовану підсистему), при якому керована система переходить з безлічі різних можливих її станів у такий стан, при якому досягається визначена необхідна (що задається) мета.

2 ДИНАМІКА СТАНІВ СИСТЕМ

1. 19 Фазова траєкторія і простір.
2. 20 Норма та норматив стану.
3. 21 Можливі режими функціонування системи.
4. 22 Чутливість та стійкість системи.
5. 23 Інваріантність у послідовності станів системи.
6. 24 Принцип, групи принципів функціонування системи.
7. 8. 26 Критерії самоорганізації системи.
9. 27 Сутність методу деформації еталонного зв’язку при оцінці сил діючих на систему.
10. 28 Жорсткість і гнучкість зв’язків. Примушення зв’язку.
11. 29 Умови поєднання елементів у системі.
12. 30 Критичне число елементів у системі та критична жорсткість зв’язків.
13. 31 Еволюція системи.
14. 32 Причинно-наслідкові відносини між елементами системи.
15. 33 Умови статистичної і динамічної рівноваги системи та адекватності.

2.1 Перевтілення систем

Розглядаючи послідовну зміну станів системи і визначаючи відповідні їм точки (точки, що відображують) у просторі станів, можна одержати траєкторію її поведінки. Дану траєкторію називають ***фазовою траєкторією,*** а простір станів ***фазовим простором.*** Число вимірювань фазового простору дорівнює числу перемінних (координат), що визначають стан системи. Різниця між кількістю перемінних і кількістю зв'язків між ними визначає кількість ступенів свободи системи, вона показує кількість незалежних перемінних.

У реальних системах не всі координати можуть змінюватись у необмежених границях. Область простору станів, у якій може знаходитись точка, що відображує, ***називається областю допустимих станів.*** Але навіть у межах області допустимих станів не завжди будь-яка точка зображує можливий стан системи. Таку властивість має лише ***безперервний простір станів,*** що відповідає такій системі, координати якої можуть приймати будь-яке значення (у допустимих межах). Але існують системи, що називаються дискретними, в яких координати можуть приймати лише кінцеве число фіксованих значень. Простір станів таких систем також є дискретним. У цьому випадку точка, що відображує, може займати лише кінцеве число положень.

У межах області допустимих станів системи розміщено стан, що найбільшою мірою відповідає (адекватний) меті, завданням і умовам функціонування системи. Такий стан називають ***нормою стану.*** Іншими словами, норма стану є функціональний оптимум. *Під оптимальним функціонуванням системи розуміють проходження всіх процесів з найбільш можливою злагодженістю, надійністю й економічністю.*

Часто норму розуміють як деяке середнє, найбільш імовірне, що найчастіше зустрічається, звичне, очікуване та ін. Але таке розуміння норми виявляє ряд недоліків і ближче до розуміння ***нормативу.***

Нормативи звичайно встановлюються в розрахунку на середню, «стандартну» людину. Нормативи або декретуються, або встановлюються договірним способом, але в тому чи іншому випадку нормативи залежать від рівня пізнання об'єктивних норм. Норми об'єктивні і розкриваються перед людиною залежно від рівня його знань. Нормативи найчастіше лише якоюсь мірою відповідають об'єктивним нормам, але в інших випадках вони можуть бути нав'язані суспільству за тими чи іншими причинами (чи мотивами).

Норму можна розглядати як точку, до якої тяжіє система, до якої вона прагне повернутися кожного разу, коли зовнішній вплив припиняється.

Рух системи (зміна її стану) може відбуватись як під дією зовнішніх впливів, так і в результаті процесів, що відбуваються усередині самої системи. Впливи на систему в математичному смислові можуть відбуватись під дією на координати системи, а також за допомогою зміни її параметрів.

Рух системи, її поводження може здійснюватись у трьох режимах - ***рівноважному, перехідному*** і ***періодичному.***

Якщо стан і перевтілення системи пов'язані між собою так, що перевтілення не змінює стан, тоді режим поведінки *є **рівноважним.*** Припустимо далі, що в системі відбуваються деякі перевтілення таким чином, що вона постійно знаходиться у стані рівноваги. На якомусь етапі зовнішні збуджуючі впливи зміщують систему зі стану рівноваги в який-небудь сусідній стан. Але подальші перевтілення відбуваються таким чином, що система знову повертається в стан, характерний як рівноважний. Звідси, період, коли система знаходиться в сусідньому стані (від періоду фіксованої рівноваги до періоду повернення в цей стан) може вважатись ***перехідним періодом*** у режимі поведінки системи ***(перехідний режим системи).***

Перехідні процеси описують за допомогою систем диференціальних (чи кінцево-різницевих) рівнянь, що зв'язують координати стану з інтенсивностями його входів і виходів,

$$\dot{Q}(t) = \Phi(Q(t), Y(t)),$$
$$Y(t) = \psi(Q(t), X(t)).$$

де Q - вектор стану;

$\dot{Q}$ - похідна вектора стану;

X - вектор входів;

У - вектор виходів;

t - час.

Перше рівняння задає динамічні властивості системи. Друге - дозволяє за поточним станом знайти вихідні характеристики системи. Якщо перехідний період повторюється з визначеною регулярністю, тоді такий режим поведінки системи називають *періодичним.* У періодичному режимі не обов'язково щоб зміщення стану рівноваги системи в який-небудь сусідній стан відбувалося під дією однакових зовнішніх збуджуючих впливів, з однаковим характером перевтілень. Важливо, що система періодично повертається до стану рівноваги.

Для характеристики здатності системи повертатись до стану рівноваги чи змішувати свій стаціонарний стан при тривалих відхиленнях значень перемінних зовнішнього середовища використовують поняття її *чутливості.*

Чутливість системи - це чутливість вектора стаціонарних (рівноважних) значень координат стану до вектора стаціонарних значень входів.

У стаціонарному режимі Q=0. Для систем, описуваних лінійними диференціальними рівняннями:

$$\dot{Q} = AQ + BX;$$
$$Y = CQ + DX.$$

де A, B, C, D - матриці, складені з коефіцієнтів рівнянь. Рівняння статичної рівноваги представляється у вигляді

$$AQ + BX = 0.$$

Тому стаціонарне значення вектора стану

$$Q = -\frac{BX}{A}$$

Тоді, згідно з визначенням, чутливість розраховують як похідну **Q** за **X:**

$$\frac{dQ}{dX} = \frac{B}{A}.$$

Аналогічно чутливість вихідного вектора:

$$\frac{dY}{dX} = D - \frac{CB}{A}.$$

Якщо стан системи зберігається незалежно від зовнішніх збуджень, тоді це свідчить про її ***стійкість.*** Не можна ототожнювати поняття стійкості і рівноваги. Якщо стан рівноваги системи можна розглядати як деяку тотожність перевтілень, що відбуваються у ній, і визначати однаковий стан системи на будь-якому кроці її розвитку, тоді стан стійкості більш ємний.

Уявимо собі систему *S,* в якій характер перевтілень *T* такий, що перевтілення кожного операнда в образ відбувається з урахуванням збільшення чи зменшення його характеристики, а також постійної щодо кожного операнда величини Δk_i. При цьому кожен образ стає в наступному перевтіленні операндом. Наприклад, починаючи з величини k, T проходить траєкторію $k, k_1, k_2, k_{3,\dots,}$ де $k_1 - k = \Delta k$; $k_2 - k_1 = \Delta k_1$; $k_3 - k_2 = \Delta k_2$ т**а ін.**

Стійкість системи звичайно розглядають як позитивну її характеристику. Але інколи стійкість можна розглядати як небажану, що не допускає гнучкість в управлінні системи.

Трактування поняття стійкості системи дозволяє підійти до поняття ***інваріантості.*** **Інваріантість** у послідовності станів систем полягає в тому, що, незважаючи на зміни, які перетерплює система в цілому, деякі її властивості залишаються незмінними.

До понять стійкості і стану рівноваги системи близьке поняття *циклу* у перевтіленні системи.

Циклом називається така послідовність станів системи, при якій повторна зміна перевтілень змушує точку, що відображує, проходити повторно цю послідовність.

Розглянуті поняття стійкості і рівноваги у поведінці дуже ефективні при вивченні різних систем. Насамперед, стан системи вивчають з позицій можливої її рівноваги, тобто чи змінюється він, коли піддається, яким-небудь перевтіленням. Далі розглядають, чи є ця рівновага достатньо стійкою, і якщо так, тоді яким буде режим поведінки досліджуваної системи. Якщо є такий стан (чи безліч таких станів) і конкретні збудження, тоді аналізують, чи повернеться система після зміщення у свою вихідну область. І, якщо система безперервна, тоді розглядають, чи є вона стійкою проти всіх збуджень усередині деякої області значень?

Даний аналіз дозволяє відповісти на ряд запитань необхідних для вирішення задач управління поведінкою системи. Наприклад: «порушать якісь збудження рівноважний режим системи чи створять тільки періодичний режим?»; «чи мають сенс витрати з усунення збуджень або їхнього результату, якщо система досить стійка і буде повертатись у свою вихідну область?»; «які збудження мають бути обов'язково усунуті?»; «які зовнішні збудження можуть вивести систему з рівноваги?».

2.2. Принципи функціонування систем

Принцип - це правило, що виникло в результаті суб'єктивно осмисленого досвіду людей. Принципи функціонування систем сформульовані як узагальнення результатів експериментального і теоретичного дослідження динаміки станів технічних і біологічних систем.

Виділяють чотири групи принципів:

1) структурно-функціональні;
2) динаміки функціонування (роботи системи);
3) функціональної системи;
4) самоорганізації та адекватності.

Між структурою і функціями системи існує тісна єдність. У кожен момент часу функція формується на структурі. З іншого боку, необхідність виконувати якусь функцію протягом деякого часу неминуче приводить до формування нової структури.

Структура і функція системи пов'язані між собою принципом ***найпростішої конструкції.*** Цей принцип сформульований у 1943 році американським біофізиком Н. Рашевським. Відповідно з цим принципом *«...та конкретна структура чи конструкція, що ми дійсно знаходимо в природі, є найпростішою з можливих структур чи конструкцій, здатних виконувати дану функцію чи групу функцій».* У 1954 році американський учений Д. Кон підсилив формулювання цього принципу і запропонував принцип ***оптимальної конструкції,*** за яким органічна структура, що необхідна для виконання даної функції, повинна бути оптимальною відносно потрібної кількості матеріалу і необхідних витрат енергії. Подальше вивчення структурних особливостей біосистем привело Н. Рашевського у 1961 році до формулювання принципу ***адекватної конструкції організму:*** *«Конструкція має бути адекватною заданій функції при заданих умовах середовища, що змінюються».*

Узагальненням взаємин структури і функції є принцип *структурно-функціональної єдності.* Структура і функції представляють єдине ціле, причому функціональний ефект системи досягається за рахунок її внутрішнього структурування. Структура і функція взаємоадекватні.

Розглянуті принципи одержані як результат узагальнення закономірностей функціонування біологічних систем. Але через те, що в процесі створення технічних систем прототипами часто виступають біологічні, то дані принципи широко використовуються для побудови конструкцій технічних систем.

Принцип структурно-функціональної єдності справедливий для біосистем будь-якого ієрархічного рівня (рівень клітки, рівень органу, рівень організму, рівень суспільства). Наявність у біосистем властивості ієрархічності дозволило сформулювати принцип ***поверховості.*** У будь-якій складній системі поверхи обробки інформації взаємозалежні і впливають один на другий.

Одна з перших спроб формування принципів роботи системи належить П. де Мопертюі (XVIII ст.). Його принцип ***найменшої*** *дії* говорить: *коли в природі відбувається деяка зміна, кількість дії, необхідної для цієї зміни, є найменшою з можливих.* Надалі цей принцип був розвинутий І.М. Гельфандом і М.Л. Цетліним і названий принципом *найменшої взаємодії.* Відповідно до цього принципу, систему можна назвати доцільно працюючою в деякому зовнішньому середовищі, якщо система прагне мінімізувати взаємодію із середовищем. При цьому мірою взаємодії системи із середовищем може служити відхилення параметрів внутрішнього середовища системи від оптимальних значень.

Принцип Мопертюі, висловлений стосовно живої природи, отримав суворе трактування в додатку до механіки вагомих тіл і теорії синтезу технічних систем автоматичного управління в роботах Лагранжа, У. Гамільтона, К. Якобі, Г. Гельмгольца, Л.С. Понтрягіна.

Надалі даний принцип отримав розвиток у роботах К. Гауса, що сформулював принцип ***найменшого примушування.*** Відповідно з цим принципом вільний рух системи відбувається з найменшим можливим примушенням.

Аналіз дії властивостей біосистем, що зберігають систему, дозволив французькому хіміку А. Ле-Шательє у 1887 році сформулювати принцип ***самозбереження.*** Відповідно з цим принципом, якщо на систему, що знаходиться у стійкій рівновазі, подіяти ззовні, змінюючи яку-небудь з умов, що визначають положення рівноваги, тоді рівновага зміщується в тому напрямку, при якому ефект цього впливу зменшується.

У розвиток загальної теорії систем великий внесок зробив радянський біофізик Е. Бауер. Розглядаючи процеси обміну між біосистемою і середовищем речовиною та енергією, Е. Бауер сформулював деякі принципи, загальні для живих систем:

- *усім живим істотам властива насамперед мимовільна зміна свого стану, тобто зміна стану, що не викликана зовнішніми причинами, які лежать поза живим організмом;*
- *при зміні зовнішніх умов істота не просто протидіє зовнішній силі (інерції і тертю за принципом Даламбера), а в результаті протидії змінює стан середовища;*
- *робота живих систем при всякому навколишньому середовищі спрямована проти рівноваги, що мала б наступити в даному навколишньому середовищі при даному первісному стані системи.*

Узагальненням цих трьох принципів є сформульований Е. Бауером ***загальний закон біології:*** *всі і тільки живі системи ніколи не бувають у рівновазі і виконують за рахунок своєї вільної енергії постійно роботу проти рівноваги, що зумовлена законами фізики і хімії при існуючих зовнішніх умовах.*

Суттєве значення для роботи біологічних систем мають принцип ***функціонального гомеостазису,*** принцип ***одиничності*** і принцип ***причинності.*** Відповідно з першим принципом система має можливість підтримувати у визначених межах постійність своїх функціональних поведінок. Границями підтримки гомеостазису є область динамічної достатності. Суть принципу одиничності чи винятковості полягає в тому, що потреби біологічної системи, для задовольняння яких потрібні суперечливі (протилежні) дії, не можуть задовольнятись одночасно. їхнє задовольняння йде почергово.

Відповідно до принципу причинності, відгук системи на деякий вплив не може початись раніше самого впливу.

Принцип *функціональної системи* сформульований П.К. Анохіним у 1935 році. Задовго до розвитку кібернетики він додав закінчений вигляд ідеї рефлекторної дуги І.П. Павлова, характерної для систем зі зворотним зв'язком, висунув ідею про зворотну аферентацію, що вказує на роль результату функціонування як фактора, який утворює функціональну систему. За П.К. Анохіним результат завдяки зворотному зв'язку має можливість реорганізувати систему, створюючи таку форму взаємодії між її компонентами, що є найбільш сприятливою для отримання саме запрограмованого результату. Відповідно до принципу функціональної системи організм під даний результат мобілізує внутрішні структури, поєднує їх у функціональні системи. Структурні утворення, що складають функціональні системи, мають винятково рухливу змобілізованість. Це дає можливість системі бути пластичною, швидко змінювати свою архітектоніку в пошуках запрограмованого корисного результату.

Велику увагу дослідників привертає така фундаментальна властивість систем як здатність змінювати свою складність і організацію. У1958 році У.Р. Ешбі сформулював принцип *необхідної розмаїтості: «тільки розмаїтість може знищити розмаїтість».* Цей принцип означає, що система може функціонувати у середовищі, якщо число її станів дорівнює числу станів середовища.

Розвиток принципу необхідної розмаїтості передбачає встановлення рівноваги за числом станів не для всіх випадків взаємодії системи і середовища. Використовуючи міру складності, введену У.Р. Ешбі, можна записати функції числа станів для системи і середовища:

$$H_m^s = \log n_s; \quad H_m^e = \log n_e.$$

де H_m^s, H_m^e - максимальна ентропія системи і середовища відповідно;

n_s, n_e - число станів системи і середовища; індекс *s* відноситься до системи, індекс е – до середовища.

Неузгодженість між системою і середовищем за складністю в цьому випадку визначають так:

$$\varepsilon_n = H_m^e - k_H H_m^s = \log n_e - k_H \log n_s$$

де k_H - коефіцієнт пропорційності.

Величина $\varepsilon_п$ є керуючим сигналом, що діє на механізм зміни складності W_H за рахунок постійного припливу речовини *B* и енергії *E* (рис. 2.1). Складність системи змінюється за рахунок зміни розміри, числа елементів, числа зв'язків між елементами чи способом зміни параметрів елементів, динамічної конфігурації зв'язків між ними. Зміна функціональної складності можлива за рахунок використання раніше не задіяних структурних елементів, чи зміни порядку і розмаїтості реакцій системи.

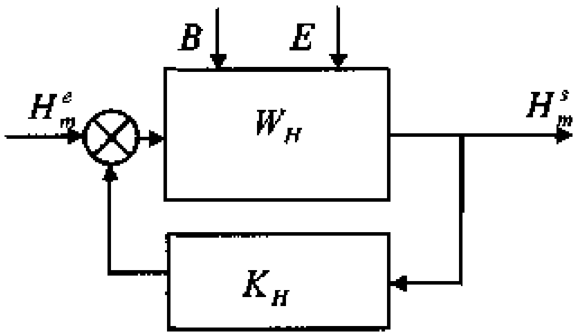

Рис. 2.1. Схема взаємодії за складністю

Зміна числа станів середовища може здійснюватись системою. При цьому можливі три випадки взаємодії системи із середовищем залежно від величини k_H:

1) $k_H = 1$- кожному стану середовища відповідає стан системи - пряме використання принципу необхідної розмаїтості Ешбі;

2) $k_H \le 1$ - кожному стану середовища може відповідати деяка безліч станів системи; такий випадок можливий, якщо система має можливість тонкого і структурного аналізу станів середовища;

3) $k_H \ge 1$ - визначеному набору числа станів середовища відповідає деякий менший набір числа станів системи. Цей випадок може відповідати процесу укрупнення показників середовища за тими чи іншими критеріями.

Знайдене тим чи іншим способом значення k_H може вказати на характер взаємодії системи із середовищем (аналіз, синтез, взаємооднозначна відповідність). Таким чином, у процесі взаємодії із середовищем система може перебудовувати власну структуру, мобілізовувати свої структурні і функціональні резерви. При цьому змінюється організація системи. Мимовільний перехід від неорганізованої системи до організованої чи перехід від погано організованої системи до добре організованої може розглядатись як ***процес самоорганізації.*** Згідно У.Р. Ешбі у ролі критерію самоорганізації може служити убування невизначеності, що виражається умовою.

$$\frac{dH}{dt} \le 0.$$

У 1960 р. Г. Ферстер ввів оцінку відносної організації і запропонував вважати систему, яка самоорганізується, якщо

$$\frac{dR}{dt} \geq 0\text{, чи } H\frac{dH_m}{dt} \geq H_m\frac{dH}{dt},$$

де *R* - відносна організація системи.

У1962 р. В.М. Глушковим був сформульований ***принцип самоорганізації:*** *система має здатність до самоорганізації в даному середовищі, якщо в результаті багаторазової взаємодії з нею невизначеність системи стосовно максимальної невизначеності, що проявляється системою в даному середовищі, убуває.*

Процес самоорганізації приводить до усунення неузгодженості між системою і середовищем за складністю. У результаті встановлюється стан адекватності (відповідності) системи середовищу. Для систем, що діють у середовищі постійної складності та організації, справедливий ***статичний принцип адекватності:*** *щоб успішно функціонувати в середовищі, складність і організація системи мають бути адекватними складності та організації середовища.* Умовно цей принцип можна виразити так:

$$H_m^s \approx H_m^e; \quad R_s \approx R_e.$$

Не менш важливим для функціонування систем є ***принцип динамічної адекватності:*** *при зміні складності та організації середовища система прагне досягти нового рівня адекватності за складністю і організацією із середовищем з мінімізацією часу, витрат речовини та енергії,* тобто

$$H_m^s(t) \approx H_m^e(t), \quad R_s(t) \approx R_e(t).$$

Різниця між показниками складності та організації системи і середовища служить для системи вхідним сигналом, що змушує її зменшувати початкову розбіжність (неузгодженість) між цими параметрами. Якщо позначити ці неузгодженості через

$$\varepsilon_H(t) = H_m^e(t) - k_m H_m^s(t), \quad \varepsilon_R(t) = R_e(t) - k_R R_s(t), \quad \varepsilon_Q(t) = Q_e(t) - k_Q Q_s(t),$$

тоді для виконання принципу динамічної адекватності необхідно, щоб

$$\frac{d\varepsilon_i(t)}{dt} \leq 0, \quad \frac{d\varepsilon_s(t)}{dt} \leq 0, \quad \frac{d\varepsilon_Q(t)}{dt} \leq 0.$$

Якщо в деякий момент часу між системою і середовищем установилась відповідність, і після цього починають змінюватись складність і організація середовища, тоді для виконання принципу адекватності досить пропорційності таких величин:

$$k_H \frac{dH_m^s}{dt} = \frac{dH_m^e}{dt}, \; k_R \frac{dR_s}{dt} = \frac{dR_e}{dt}, \; k_Q \frac{dQ_s}{dt} = \frac{dQ_e}{dt}.$$

Таким чином, у випадку статичної адекватності комплекс «біосистема - середовище» аналогічний системі стабілізації із зворотним зв'язком, а у випадку динамічної адекватності - системі, що стежить.

У результаті встановлення адекватності між біосистемою і середовищем взаємини біосистеми і середовища можуть відповідати трьом рівням: слабкі ймовірнісні, сильні ймовірнісні і детермінована взаємодія.

При слабкій імовірнісній взаємодії адекватність між біосистемою і середовищем установлюють за рівнем організації, тобто

$$R_s = R_e.$$

Взаємодія біосистеми із середовищем вважається сильною ймовірнісною, якщо в результаті взаємодії конкретному закону розподілу ймовірностей прийняття середовищем своїх станів відповідає єдиний закон, тобто

$$R_s = R_e, \qquad \varphi(p(n_s)) \approx \varphi(p(n_e)),$$

де φ - деякий закон розподілу.

У випадку взаємооднозначної відповідності між станами системи і станами середовища взаємодія називається детермінованою. Умови детермінованої взаємодії мають вигляд:

$$R_s = R_e, \qquad p_{si} = p_{ei}, \qquad i = 1,2,\ldots,n$$

$$p\left(\frac{n_{si}}{n_{ei}}\right) = 1$$

2.3. Розвиток систем

Слід розрізняти два типи динаміки системи: її функціонування і розвиток.

Під ***функціонуванням*** розуміють процеси, що відбуваються у системі (і навколишньому середовищі), яка стабільно реалізують фіксовану мету (наприклад, функціонують радіоприймачі, верстати, кінотеатри, школи, транспорт та ін.).

Розвитком називають те, що відбувається із системою при зміні її цілей. Характерною рисою розвитку є той факт, що існуюча структура перестає відповідати новій меті, і для забезпечення нової функції приходиться змінювати структуру, а іноді і склад системи, перебудовувати всю систему.

Як приклад, що ілюструє розвиток системи, розглянемо перевтілення масової системи.

Відповідно до визначення, масовою називається система, в якій безліч однотипних елементів з'єднані паралельно. Ці системи представляють надзвичайно розповсюджений клас систем:

1) тверде тіло, що складається з однакових молекул;

2) жива тканина, побудована з клітин одного виду;

3) група людей, що поєднує рівних індивідів та ін.

Останнє особливо важливо, тому що дозволяє розглянути процеси, які відбуваються в транспортних потоках на прикладі перевтілення масової системи.

Процес перевтілення масової системи описаний Г.А. Голідиним.

Рівняння зв'язку вектора вихідного сигналу з вектором вхідного для кожного елемента системи може бути представлене у вигляді:

$$\overline{Y} = f(X),$$

де *X* - вектор вхідного сигналу; $\overline{Y}$ - норма вектора вихідного сигналу; f - залежність між входом і виходом.

Норма вектора вихідного сигналу може розглядатись як точка притяжіння чи мета, до якої прагне система. Нормальні значення вихідних сигналів для різних елементів можуть бути різними, тобто $\overline{Y}_i \neq \overline{Y}_k$.

Дійсне значення вихідного сигналу У може відрізнятись від нормального. Відхилення $\Delta Y = Y - \overline{Y}$ дійсного значення вихідного сигналу від нормального можна назвати ***деформацією*** зв'язку. Деформацію і - го зв'язку під дією k-того ми будемо називати ***силою,*** що розвивається k-тим зв'язком і обмірюється за допомогою і-того.

Уведемо далі ***еталонний*** зв'язок і будемо вимірювати сили, що розвиваються різними зв'язками, за допомогою деформації одного і того ж еталонного зв'язку. Деформацію еталонного зв'язку, викликану дією і-того зв'язку, ми будемо називати силою і-того зв'язку, без указівки, за допомогою чого вона обмірюється, і позначати через h_i.

Інтеграл від сили за відхиленням будемо називати ***примушенням*** зв'язку чи потенціалом сили і позначати через U:

$$U = \int_{\overline{y}}^{y} h d\Delta y.$$

Тоді сили можна визначити як часткові похідні від потенціалу:

$$h_i = \frac{dU}{d\Delta y_i}$$

Похідну від сили за відхиленням будемо називати ***жорсткістю*** зв'язку (чи жорсткістю норми) і позначати через γ.

$$\gamma = \frac{dh}{d\Delta y}$$

Здатність реального зв'язку відхилятись від нормального стану під дією інших зв'язків будемо називати ***гнучкістю*** зв'язку. Усі реальні зв'язки є гнучкими. Гнучкість зв'язку можна визначити як величину, зворотню жорсткості:

$$\lambda = \frac{1}{\gamma}.$$

При послідовному з'єднанні зв'язків складаються їхні гнучкості:

$$\lambda_\Sigma = \lambda_1 + \lambda_2 + ... + \lambda_m.$$

При паралельному з'єднанні зв'язків складаються їхні жорсткості:

$$\gamma_\Sigma = \gamma_1 + \gamma_2 + ... + \gamma_m.$$

У випадку лінійності елементів системи жорсткість зв'язку постійна, усі динамічні величини знаходяться між собою в симетричних відносинах:

$$\gamma\lambda = 1,\ h = \gamma\Delta Y = \gamma\left(Y - \overline{Y}\right),\ \Delta Y = \lambda h,$$
$$U = \frac{1}{2}h\Delta Y = \frac{1}{2}\gamma\Delta Y^2 = \frac{1}{2}\lambda h^2$$

Основні властивості сили виходять як наслідок з її визначення:

- зв'язок, що знаходиться у нормальному стані, розвиває нульову силу;
- якщо на перемінну У накладено ***m*** зв'язків, тоді її дійсним станом буде точка, в якій сума всіх ***m*** сил дорівнює нулю.

Зазначені властивості сили дозволяють визначати норму стану масової системи з умови рівноваги сил:

$$\sum_{k=1}^{m}\gamma_k\left(\overline{Y}_\Sigma - \overline{Y}_k\right) = 0.$$

У лінійному випадку, коли жорсткість зв'язку постійна, сила пропорційна відхиленню $h_k = \gamma_k(\overline{Y}_\Sigma - \overline{Y}_k)$, то норма є середньозваженою за множиною $\overline{Y}_k$:

$$\overline{Y}_\Sigma = \frac{\sum_{k=1}^{m}\gamma_k\overline{Y}_k}{\sum_{k=1}^{m}\gamma_k}.$$

Жорсткості зв'язків γ_k відіграють роль вагових коефіцієнтів. Коли ці жорсткості однакові, нормою є просто середнє арифметичне:

$$\overline{Y}_{\Sigma} = \frac{1}{m}\sum_{k=1}^{m}\overline{Y}_k .$$

У системі, що розвивається, склад зв'язків постійно оновлюється. З кожним часом новий зв'язок нашаровується на вже існуючі і приєднує до них свій вплив. У той же час забувається і зникає із складу системи найбільш старий зв'язок. Норма такої системи $\overline{Y}_{\Sigma}(t)$ виходить як деяке середнє за безліччю минулих значень перемінної $Y(\tau_i)$ за часом T.

Обмежившись випадком, коли норму можна представити як середнє арифметичне:

$$\overline{Y}_{\Sigma}(t) = \frac{1}{2}\int_{t-T}^{t} Y(\tau)d\tau$$

Аналогічне вираження можна написати для дискретного часу:

$$\overline{Y}_{\Sigma}(t) = \frac{1}{M}\sum_{i=m-M+1}^{m} Y(\tau_k).$$

де M - дискретний аналог T (глибина пам'яті).

Якщо перемінна зберігає постійне значення, тоді норма як середнє збігається з цим значенням. Якщо перемінна стрибком міняється на величину ΔY, тоді норма як середнє буде мінятись у той же бік, поки не зіллється з новим значенням перемінної. Така зміна норми може розглядатись як *адаптація системи.*

Швидкість адаптації

$$\frac{d\overline{Y}(t)}{dt} = \frac{\Delta Y}{T},$$

чи для дискретного випадку

$$\Delta\overline{Y}_m = \frac{\Delta Y}{M}.$$

Звідси виходить, якщо зовнішній зв'язок досить довго утримує систему в ненормальному стані, тоді норма сама зрештою наближається до цього стану. У результаті ненормальний стан стає нормальним для системи.

Процеси додавання чи знищення зв'язків і елементів, зміни їхніх характеристик і вхідних величин, зміна норм стану системи, що перетворюються в остаточному підсумку в зміни вихідних сигналів, можна розглядати як розвиток системи. Розвиток систем підлеглий основному принципу - принципу

найменшої взаємодії. Якщо скористатись поняттям примушення, тоді даному принципу можна додати варіаційну форму: ***перемінна, на яку накладено кілька зв'язків, займає таке положення, в якому сумарне примушення усіх зв'язків мінімальне:***

$$U_{\Sigma} = \sum_{k=1}^{m} U_k(Y) = \min .$$

У системі слід розрізняти два види примушення: 1) внутрішнє, що викликане протиріччями між елементами і виникає при їхньому об'єднанні в систему; 2) зовнішнє, що обумовлене дією зовнішньої збуджуючої сили, тобто протиріччям між системою і зовнішнім світом. Примушення адитивне, і повне примушення k-того елемента дорівнює сумі внутрішнього і зовнішнього примушень $U_k = U_k^i + U_k^e$, а повне примушення системи дорівнює сумі примушень її елементів:

$$U_{\Sigma} = \sum_{k=1}^{m} U_k = \sum_{k=1}^{m} U_k^i + \sum_{k=1}^{m} U_k^e$$

Елементи поєднуються в систему, тому що це їм вигідно. Вигідно лише у випадку, коли примушення системи виявляється менше суми примушень розрізнених елементів, тобто

$$U_{\Sigma} < \sum_{k=1}^{m} U_k^i .$$

Дійсно, у випадку дії на системи зовнішньої сили h_e при об'єднанні елементів ця сила пропорційно «розподілиться» між ними. Коли ж елементи розрізнені, ця сила приходиться на якийсь один з них.

Сума примушень розрізнених елементів дорівнює

$$\sum_{k=1}^{m} U_k^i = U_k^i = \frac{h_e^2}{2\gamma}$$

Сума примушень елементів, об'єднаних у систему дорівнює

$$U_{\Sigma} = \frac{h_e^2}{2m\gamma} + \frac{m\gamma \overline{Y}^2}{2}$$

Перший доданок у правій частині цього рівняння характеризує зовнішнє примушення, другий доданок - внутрішнє.

Таким чином, умовою об'єднання елементів у систему є нерівність

$$\frac{m\gamma\overline{Y}^2}{2} < \frac{h_e^2(m-1)}{2m\gamma}$$

Дана нерівність показує, що об'єднання вигідне в таких випадках:

1) коли зовнішні умови несприятливі (чим більше h_e);
2) чим менше протиріччя між елементами (чим менше $\overline{У}$);
3) чим слабкіші елементи (чим менше жорсткість γ).

Таким чином, ***елементи поєднуються в систему при несприятливих зовнішніх умовах. Протиріччя між елементами не сприяють об'єднанню. Об'єднання вигідне слабким і невигідне сильним елементам.***

Остання нерівність дозволяє розглянути питання, до яких меж може збільшуватись число елементів у системі.

Критичне число елементів можна знайти з умови:

$$\frac{m_k\gamma\overline{Y}^2}{2} = \frac{h_e^2(m_k-1)}{2m_k\gamma}$$

Якщо m_k досить велике, можна зневажити одиницею в чисельнику. Тоді

$$\frac{m_k\gamma\overline{Y}^2}{2} \approx \frac{h_e^2}{2\gamma}$$

Звідси

$$m_k \approx \frac{h_e^2}{\gamma^2\overline{Y}^2}$$

Таким чином, *розміри системи можуть бути тим більшими, чим менше жорсткості її елементів* γ*, протиріччя між ними* $\overline{Y}$ *і чим сильніше зовнішнє збудження* h_e. *Подальше зростання числа елементів приводить до розпаду системи, її розподілу на кілька дочірніх систем.* Такий процес у транспортному потоці характерний для утворення так званих «груп автомобілів» у потоці.

У процесі розвитку системи через формування детермінованих зв'язків чи їхнього руйнування жорсткість її елементів може змінюватись. Для кожного сполучення внутрішніх і зовнішніх умов існує оптимальне значення жорсткості $\bar{\gamma}$, при якому примушення системи мінімальне. Оптимальне значення жорсткості можна знайти з умови:

$$\frac{dU_{\Sigma}}{d\gamma}=\frac{m\overline{Y}^{2}}{2}-\frac{h_{e}^{2}}{2m\overline{Y}^{2}}=0.$$

Звідси

$$\bar{\gamma}=\frac{h_{e}}{m\overline{Y}}.$$

З цього вираження виходить, що *жорсткість повинна зростати зі збільшенням зовнішніх протиріч і зменшуватись при збільшенні внутрішніх. Наявність зовнішньої погрози також виправдовує збільшення жорсткості внутрішніх зв'язків.* Але у будь-якому випадку збільшення жорсткості внутрішніх зв'язків вигідно для системи лише до визначеної величини $\bar{\gamma}$. Подальше збільшення жорсткості внутрішніх зв'язків приведе до старіння і загибелі системи.

Природний рубіж життя системи - це критичний стан, коли об'єднання елементів перестає бути вигідним унаслідок внутрішніх протиріч, тобто коли

$$\frac{m\gamma\overline{Y}^{2}}{2}=\frac{h_{e}^{2}}{2\gamma}.$$

Звідси критичне значення жорсткості

$$\gamma_{k}=\frac{h_{e}}{\overline{Y}\sqrt{m}}$$

Якщо γ_k взяти у відповідних одиницях, тоді її величина може характеризувати час життя системи. Відповідно з отриманою формулою для визначення γ_k:

1) *час життя тим більше, чим менше число елементів, що входять у систему (малі системи життєздатніші за великі);*

2) *час життя пропорційний зовнішньому збудженню* h_e *(зовнішнє збудження є стимулятором існування системи).*

Між процесами розділення і загибелі системи багато спільного, але існує найважливіше розходження. Розділення є наслідком зростання числа елементів системи і відбувається на фоні цього росту. Тому воно не веде до нескінченного здрібнення розмірів системи, а служить засобом стабілізації цих розмірів. Загибель системи, навпаки, є процес здрібнення і дроблення, аж до розкладання системи на окремі елементи.

2.4. Еволюція систем

Розглянуті правила перевтілення систем, які розвиваються, не дають відповіді на запитання, як проходить цей процес у часі. Відповідь на це запитання можуть дати закономірності еволюції систем. Під «еволюцією» розуміють процес історичного розвитку системи. Термін «еволюція» уперше був уведений у біології Ш. Бонні у 1762 році стосовно зміни поколінь організмів. Лише з появою дарвінівського еволюційного навчання у 1858 році термін «еволюція» у біології придбав сучасне тлумачення, хоча сам Ч. Дарвін його не використовував. Під еволюційним вченням стали розуміти систему знань про історичний розвиток живої природи, сукупність представлень про механізми і закономірності історичних змін в органічній природі. Історичний розвиток - це повільні якісні й кількісні зміни системи, що включають етапи функціонування і розвитку, періоди відкритого і закритого станів.

Надалі *під еволюцією стали розуміти історичний розвиток навколишнього світу: живої і неживої природи, суспільства, систем будь-якої природи.* Люди, не задумуючись, стали використовувати словосполучення: «еволюція ідей», «еволюція автомобільного транспорту», «еволюція науки» та ін. Термін «еволюція» став синонімом терміна «історичний розвиток», що відрізняється від терміна «розвиток». Розвиток відбувається в періоди відкритого стану системи і супроводжується її структурною і функціональною перебудовою. У процесі розвитку змінюються цілі існування системи. В закритому стані система функціонує, цілі її існування не змінюються. У свою чергу історичний розвиток включає як етапи розвитку, так і етапи функціонування системи.

Значний внесок у створення загальної теорії еволюції внесли С.С. Четвериков, М.І. Вавілов, В.Н. Сукачьов, І.І. Шмальгаузен, Н.Н. Дубинін, І.В. Тимофєєв-Ресовський, А. Ліма-де-Фаріа, Г. Паск, Л. Фогель, А. Оуенс, Л. Уолш, Х. Холленд, Р. Фокс, Р. Левонтін та ін. До кінця 50-х - початку 60-х років XX сторіччя в основному було закінчено створення основ сучасної синтетичної теорії еволюції, розвиток якої продовжується і теперішнім часом. Ця теорія приділяє основну увагу вивченню причинних механізмів і принципів еволюційних змін у системах різної природи.

Кажучи про історичний розвиток будь-якої системи чи еволюційну динаміку будь-якого явища, ми маємо на увазі, що нас цікавить зміна стану системи чи явища в часі. Ми прагнемо отримати відповіді на запитання, як, яким чином виникла система, які етапи проходила у своєму розвитку, які її історичні перспективи?

Незалежно від того, чи вивчаємо ми еволюцію суспільства, мов, видів, геологічних змін чи формацій зірок, можна виробити деяке формальне представлення, загальне для всіх цих сукупностей.

Нехай у деякий момент часу t система знаходиться у стані Q і нас цікавить її стан Q^1 у майбутньому чи минулому, що відстоїть від цього моменту на τ одиниць часу. Потім ми маємо побудувати закони перевтілення L, що дадуть нам можливість передбачити Q^1, виходячи з Q. Формально це можна записати

$$Q(t) \xrightarrow{L} Q^1(t+\tau).$$

Закони перевтілення не можуть бути довільними. По-перше, вони містять деякі параметри П, значення яких самі по собі не залежать від часу чи стану системи. По-друге, у загальному випадку вони будуть містити час τ, що минає. В окремому випадку рівноважних систем, в яких ніяких змін не відбувається, залежність від т відсутня.

Параметри П можуть мати чи не мати відношення до абсолютного часу t залежно від того, чи містить ця система у своєму нинішньому стані які-небудь зведення про своє минуле. Наприклад, якщо потрібно визначити, яка частина популяції залишиться в живих через рік, необхідно знати вік особи, тому що ймовірність виживання з віком міняється. Але для того, щоб оцінити, скільки чайних чашок із сервізу збережеться до кінця року, зовсім не потрібно знати, коли вони були виготовлені, тому що у першому наближенні середнє число розбитих за рік чашок не залежить від цього фактора. Нарешті, і це головне, закони перевтілення повинні враховувати стан системи Q у даний момент часу і відповідно визначити її новий стан Q^1. Ці закони представляють собою деякі механізми, необхідні для обробки поточної інформації про поточні стани системи Q, і, мабуть, одним з результатів їхньої діяльності буде новий стан Q^1. Це означає, що описування системи Q потрібно скласти таким чином, щоб на його підставі можна було дійсно створити закони перевтілення. Наприклад, не можна передбачити майбутнє положення автомобіля у приземному просторі, виходячи тільки з його положення в даний момент. Зокрема, не можна описати перевтілення $Q(t)$ у $Q^1(t+\tau)$, знаючи тільки положення Q автомобіля в тривимірному просторі. Необхідно, крім того, знати швидкість автомобіля тепер і його прискорення чи уповільнення у вигляді трьох проекцій на ортогональній осі. Можна сказати, що описування стану автомобіля, враховуючи усі дев'ять перемінних, буде *динамічно достатнім описуванням,* тому що на його підставі можна знайти закони $L(Q, \Pi, \tau, t)$:

$$Q(t) \xrightarrow{L(Q,\Pi,\tau,t)} Q^1(t+\tau).$$

Якщо розглядати не переміщення автомобіля у просторі, а його історичний розвиток, тоді закон переходу автомобіля із стану $Q(t)$ у стан $Q^1(t+\tau)$ буде містити перемінні параметри П, що залежать від вихідного стану автомобіля і часу, тобто

$$L\,[Q, \Pi(Q, t), \tau, t].$$

Тому динамічно достатнє описування еволюції автомобіля може бути представлене у вигляді

$$Q(t) \xrightarrow{L\,[Q,\,\Pi(Q,t),\,\tau,\,t]} Q^1(t+\tau).$$

Характерною рисою еволюції є той факт, що існуюча структура системи перестає відповідати новій меті, і для забезпечення нової функції приходиться змінювати структуру, а іноді й склад системи, перебудовувати всю систему. Мета знаходиться у середовищі. Тому еволюціонує не сама система, а сукупність «система-середовище».

Механізми еволюції давно привертали увагу дослідників складних систем, які цікавились, як у процесі еволюції виникають ті чи інші властивості біологічних систем високого ступеня складності. Так, відомим фахівцем-кібернетиком Г. Паском за допомогою моделі еволюції було досліджене запитання появи організованої структури в системі, що спочатку була безструктурною чи слабко організованою. Прикладом такої системи є популяція амеб. Амеби виділяють речовину акразин, що при наближенні до іншої амеби викликає її рух до джерела сигналу. Коли амеб збирається досить багато, вони утворюють організоване ціле, в якому різні клітини, виконують різні функції. При відсутності їжі новий кооперативний організм перебирається в інше місце, проходячи голодний район. Лише кооперативний організм успішно справляється з цим завданням, окремі амеби гинуть, перш ніж доберуться до їжі.

Модель такої системи представлена на рис. 2.2. Поведінка амеб моделюється поведінкою автоматів. Через мережу живлення на автомати надходить безперервний потік «їжі». Обмін сигналами між автоматами здійснюється через спеціальне середовище, що створює і підтримує канали передачі інформації за рахунок витрати «їжі». Таким чином, для кожного моменту активності системи виникає деяка структура, що визначається, створюється і підтримується витратою «їжі». Перед автоматами системи стоїть завдання вижити, тобто зберегти визначений рівень активності, що можливо лише в коаліції з іншими автоматами. При гарному живленні автомат збуджується, що виявляється в обміні сигналами з іншими автоматами. У результаті автомат вступає в коаліцію з іншими автоматами. Якщо кількість «їжі» у якому-небудь місці досягає критичної величини, то з декількох автоматів виникає новий автомат. Цей процес називають *нуклеацією*. Новий автомат відразу починає поглинати «їжу» і вступати в коаліцію з іншими автоматами. Якщо він буде як і раніше на тому ж місці, то, поглинувши всю «їжу», автомат загине. Щоб цього не відбулося, автомати наділені здатністю переміщатись і еволюціонувати. Переміщення відбуваються через мережу розподілу «їжі», що представлена у вигляді двомірних ґрат.

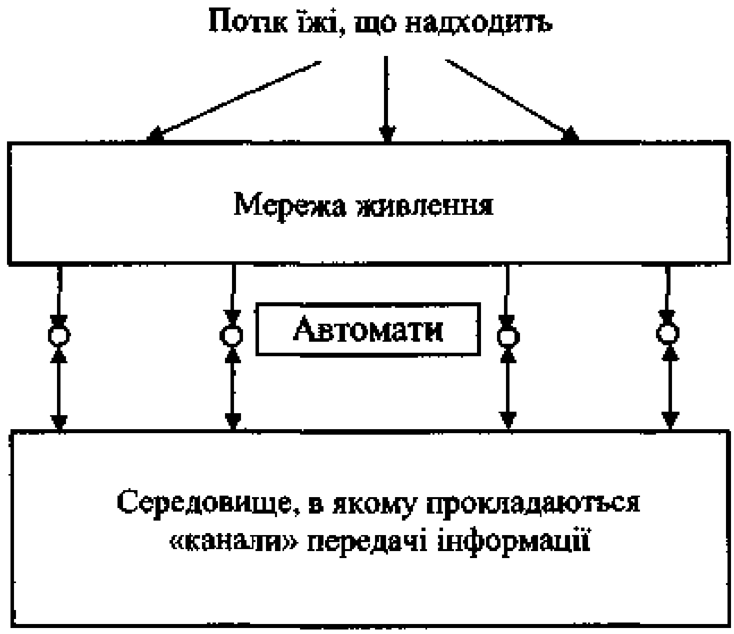

Рис. 2.2. Еволюційна схема Г. Паска

Г.Паск увів два класи примітивних автоматів. Одні автомати здатні переміщатись тільки в поздовжньому (вертикальному) напрямку, інші - тільки в поперечному (горизонтальному). Кожен з автоматів може робити лише три рухи уздовж свого напрямку: зрушитись на один крок вперед чи назад і залишитись на місці. Але, якщо в якомусь місці відбувається з'єднання (нуклеація) двох автоматів, що належать одному класу, то виходять автомати, здатні робити вже п'ять рухів. Додаються два рухи на два кроки вперед чи назад.

При подібному з'єднанні двох автоматів різних класів виходить новий вид автоматів з дев'ятьма рухами: на один і два кроки вперед та назад за кожним напрямком і можливість залишитись на місці. Таким чином, автомати можуть з'єднуватись до нескінченності (еволюціонувати), і кожен новий вид може виконувати всі рухи своїх попередників.

Щоб вижити, автомат рухається у бік більшої концентрації «їжі», залишаючи за собою смугу виснажених запасів. Прагнучи до коаліції, а виходить, і нуклеації, автомати збираються в районі великої концентрації «їжі».

Таким чином, у своїй моделі Г. Паск використав три еволюційні механізми: створення нових видів автоматів (аналог механізму рекомбінації), відбір, збереження і відтворення відібраних автоматів. Саме завдяки цим механізмам у колективі автоматів, як показано Г Паском, з'являється властивість стійкості і пристосовності до потоку «їжі», що змінюється.

Інший підхід реалізований Ю.І. Неймарком при вирішенні задачі оптимізації функції багатьох перемінних. Звичайно такі задачі зустрічаються в системах керування. За Ю.І. Неймарком роль аналога популяції виконує автоматний оптимізатор (АО). Цей оптимізатор представляється у вигляді деякого об'єму, що містить усередині себе автомати. Число автоматів більше числа каналів оптимізації і збігається з числом перемінної функції, що оптимізується. Тому в будь-який момент часу автомати поділяються на працюючі і ті, що імітують. Успішність роботи одних автоматів і імітація інших на відповідних каналах оптимізації відбивається у величинах компонент характеризуючого вектора, поставленого відповідно кожному автоматові. Число компонент збігається з числом каналів оптимізації. Кожна компонента кодує визначену ознаку. Крім того, Ю.І. Неймарк уводить механізм створення нових автоматів і відбору найбільш ефективних з них. Нові автомати виходять з пари автоматів шляхом «схрещування». «Схрещування» полягає в тому, що автоматові-нащадку передається одне з двох значень кожної компоненти ознакового вектора автоматів-батьків. Таку подію за аналогією з природним механізмом називають мутацією.

За допомогою автоматного оптимізатора вирішують два основні завдання оптимізації: регулювання (підтримка екстремального значення) і пошук екстремуму функції багатьох перемінних.

Аналоги механізмів випадкових мутацій і природного відбору успішно застосовують також у методі евристичної самоорганізації - методі групового урахування аргументів (МГУА). Метод МГУА ефективний при вирішенні завдань екстраполяції, ідентифікації і розпізнавання.

Авторами ідеї еволюційного моделювання Л. Фогелем, А. Оуенсом і Л. Уолшем запропоновано застосовувати штучну еволюцію при вирішенні завдань передбачення, управління, розпізнавання і проектування. У роботах цих авторів була продемонстрована ефективність еволюційної програми. Особливий інтерес викликає застосування еволюційного моделювання для проведення наукових досліджень. Наукове дослідження включає розробку моделей систем. Еволюційний аспект розробки моделей розглянутий Л. Фогелем.

На основі дослідних даних Л. Фогель будує первісну модель. За створеною моделлю виконується оцінка якогось невідомого аспекту даних. Подальші експерименти дозволяють перевірити справедливість первісної моделі і скорегувати її. Скорегована модель дозволяє отримати нову інформацію, що поряд з іншими емпіричними фактами використовується для створення іншої моделі, кращої стосовно конкретного середовища.

В отриманні нової інформації особливу цінність представляють моделі, правдивість оцінки яких підтверджується при незалежній перевірці. Кожна така модель несе додаткову і достовірну інформацію про середовище. Направивши свої зусилля на витяг цієї інформації із сукупності моделей, дослідник може створити модель тих моделей, що раніше виявились вірними. Ця модель, що погоджується з максимально можливим числом успішних у минулому моделей, знаходиться на іншому, більш високому рівні абстракції. Така модель буде мати велику ймовірність успіху через те, що вона охоплює деякі загальні властивості середовища, що можуть бути виявлені лише при розгляданні всіх дослідних даних у цілому.

Останнім часом еволюційне моделювання знайшло застосування у практиці прогнозування розвитку науки і техніки. Для цього використовується аналіз динамічних рядів зміни параметрів об'єкта прогнозування.

Одне з математичних описувань закону розвитку науки, що прискорюється, запропоноване Г. Вледуцем, В. Налімовим, Н. Стяжкиним і М. Карповим. Вихідною посилкою його служить те, що при відсутності обмежуючих факторів швидкість зростання наукових публікацій, як непрямих характеристик розвитку науки, визначається її існуючим рівнем, і кожна дійсно нова наукова праця породжує ряд наступних робіт. Механізм росту кількості публікацій *S* в такому разі описується диференціальним рівнянням

$$\frac{dS}{dt} = kS.$$

де dS/dt - швидкість зростання числа публікацій чи інших інформаційних характеристик розвитку науки; k - константа швидкості, властива деякій науці і конкретним умовам.

Вирішуючи це рівняння, одержуємо математичне вираження так званого експонентного закону зростання науки

$$S = S_0 e^{k(t-t_0)}, \quad (k > 0),$$

де S_0 - кількісна характеристика науки в початковий момент; e - основа натуральних логарифмів.

Вплив деяких обмежень на швидкість зростання можна врахувати, записавши диференціальне рівняння у вигляді

$$\frac{dS}{dt} = kS(N - S);$$

де *N* - число науковців $(0 < S < N; k > 0)$.

У цьому випадку існує межа зростання $(S < N)$, і відносна швидкість зростання є

$$\left(\frac{1}{S}\right)\left(\frac{dS}{dt}\right) = k(N - S)$$

лінійною функцією *S*. Вирішуючи це диференціальне рівняння, одержуємо рівняння логістичної кривої *(кривої сатурації)*

$$S = \frac{N}{[1 + S_0 \exp(-kNt)]}.$$

У роботах В.В. Налімова, З.М. Мульченко показано, що диференціальне рівняння логістичної кривої добре описує процес розвитку конкретних наукових напрямків протягом порівняно невеликих відрізків часу. Більш загальний випадок розвитку за тривалі інтервали часу краще апроксимується послідовністю експонент із значеннями коефіцієнта k, що змінюється на різних інтервалах. У цьому випадку можна записати

$$S_n = \sum_{i=0}^{n} S_{0i} \exp[k_i(t - t_i)], \; t - t_i > 0,$$

де *n* - число експонент; $t - t_i$ - зміщення початку відліку.

Як показали подальші дослідження, криві зростання науки можуть описуватись лінійними, логарифмічними і гіперболічними законами.

Для аналізу кривих зростання науки використовують аналоги біологічних процесів зростання, апарат математичної епідеміології та ін. Найважливіший висновок з цих досліджень полягає в тому, що розвиток науки як системи знань не може мати природних меж. Розвиток науки характеризується чергуванням періодів прискореного зростання з періодами уповільнених темпів в отриманні нових позитивних результатів. Так змінювались темпи відкриття нових хімічних елементів, нарощування потужностей прискорювачів елементарних часток, зростання коефіцієнта корисної дії засобів тяги залізничного транспорту, швидкостей руху автомобілів і літаків, вантажопідйомності і навантажень на задню вісь двохосьових вантажних автомобілів та ін. (рис. 2.3 - 2.5). У кожному випадку новий підйом вимагав переходу до нових методів дослідження, уточнення і відновлення фундаментальних поглядів, установлення нових можливостей використання пізнаних законів природи. Нові знання формуються у періоди уповільнених темпів у розвитку науки.

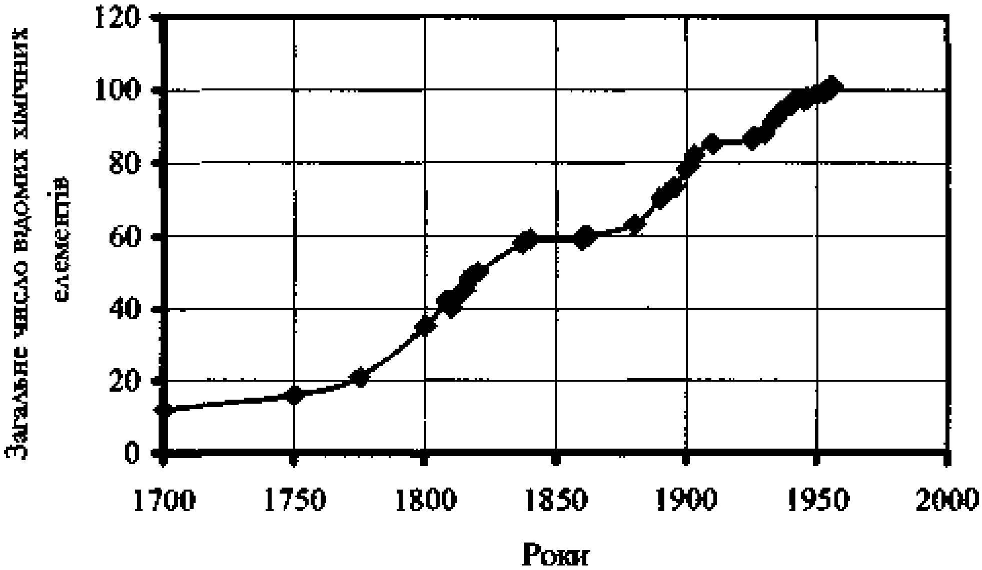

Рис. 2.3. Динаміка відкриття нових хімічних елементів

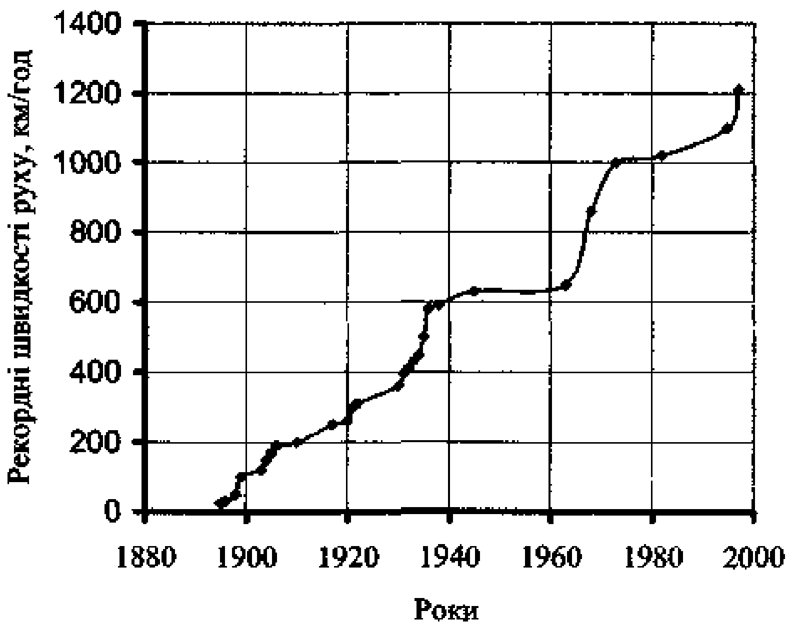

Рис. 2.4. Динаміка рекордів абсолютних швидкостей руху автомобілів

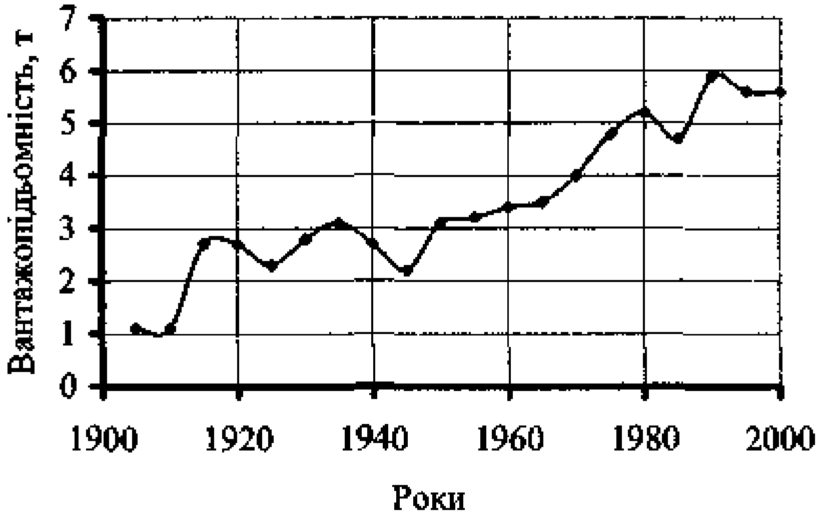

Рис. 2.5. Динаміка вантажопідйомності двохосьового вантажного автомобіля

У періоди прискореного зростання реалізуються потенційні можливості відкриттів, отриманих у періоди уповільнених темпів. Тривалість періодів уповільнених темпів у розвитку науки збільшується у два рази при переході від одного періоду до наступного. Це свідчить про те, що кожен крок вперед у науково-технічному прогресі пов'язаний із проникненням у більш глибокі закономірності і потребує докладання усе більших зусиль. У свою чергу, періоди прискорених темпів скорочуються в два рази при переході до наступного періоду. Останнє свідчить про те, що відновлення прикладних науково-технічних ідей сприяє скороченню часу промислової реалізації нових фундаментальних знань.

Розглянуті закономірності показують, що зміна темпів в отриманні нових корисних результатів характерна для історичного розвитку науки, тобто її еволюції. У періоди уповільнених темпів наука розвивається, у періоди прискорених темпів - функціонує.

Нові знання утворюються у результаті проведення наукових досліджень. У структурі наукового дослідження виділяють людину, засоби дослідження і частину природи, що підлягає дослідженню. У результаті наукового дослідження виходять знання, здатні задовольнити потреби в детермінованості й активності взаємодії людини з природою. Активною, пусковою частиною наукового дослідження є людина. Саме вона пускає в хід усю систему дослідження.

Схематичне зображення періодично повторюваного кругообігу задовольняння потреб людини в процесі пізнання в самому загальному вигляді можна представити в такій формі:

$$\ldots \to (r_2^-)(r_1^+) \to \overset{\overset{H}{\updownarrow}}{M}\ (m_0) \to N - \begin{bmatrix} r_n \\ \\ m_1 \end{bmatrix} \to (r_2^+)(r_1^-) \to M:(m_0,\ m_1) \to N:(r_n) \to (r_2^-)(r_1^+) \to$$

$$\to \overset{\overset{H}{\updownarrow}}{M}:(m_0,\ m_1) \to N - \begin{bmatrix} r_n \\ \\ m_2 \end{bmatrix} \to (r_2^+)(r_1^-) \to M:(m_0,\ m_1,\ m_2) \to N\ (r_n) \to (r_2^-)(r_1^+) \to \ldots$$

де r_k^- - незадовільнена потреба в детермінованості взаємодії з природою;

r_2^- - незадовільнена потреба в активності взаємодії з природою;

r_1^+, r_2^+ - задовільнені потреби в детермінованості й активності взаємодії з природою відповідно;

М - засоби дослідження; $m_1, m_2 ..., m_n$ - елементи засобів дослідження;

N - частина природи, що досліджується; r_n - отримані в результаті дослідження знання, здатні задовольнити потреби людини;

m_1 - знову створений засіб дослідження; *Н* - ноосфера; $\rightarrow$ - напрямок впливу.

При задовольнянні потреби r_2^- система наукового дослідження розімкнута (відкрита відносно до ноосфери) і тому в процесі взаємодії створюється новий засіб дослідження. При задовольнянні потреби r_1^- система наукового дослідження замкнута (закрита відносно до ноосфери) і реалізує лише потенційні можливості раніше створених засобів дослідження. Суперечливість потреб обумовлює ковзний (почерговий) режим їхнього задовольняння.

Створення нового засобу дослідження відкриває можливість переходу до удосконалення одного засобу за допомогою іншого. У результаті взаємодії двох засобів дослідження (старого і нового) породжується новий засіб дослідження, нове знаряддя праці людини. Якщо позначити результат впливу людини на засоби дослідження через R_a, а результат впливу засобів дослідження на природу через R_b, тоді, виходячи з вищевикладеного, можна зробити висновок, що початком отримання нових знань (нових понять і раціонального відображення дійсності взагалі) є встановлення в психіці людини взаємозв'язку: *«якщо* R_a, *тоді за допомогою М одержимо* R_b». Цю схему можна розглядати як вихідну *«клітинку»* логічного мислення людини.

Результат впливу людини на засоби дослідження можуть розглядатись як зовнішня дія на природу, а результат впливу засобів дослідження на природу - як реакцію природи (відповідь природи на зовнішню дію). Наслідком взаємодії R_a і R_b є зміни у стані природи Z. Процес взаємодії R_a і R_b підпорядкований ***принципу максимуму взаємної інформації.*** Цей принцип уперше був сформульований Г.О. Голіциним і В.М. Петровим. Формально сутність цього принципу представляється у вигляді:

$$I(R_a R_b) = \max_{R_a, R_b} .$$

Символи R_a і R_b під позначенням максимуму говорять про те, що отримати максимум інформації можна як за рахунок вибору R_b, так і за рахунок вибору R_a.

Справедливість цієї «троїчної» схеми R_a, R_b, Z є наслідком логічного мислення. Це просто перенесення на живу природу загальної схеми причинної взаємодії, в якій завжди є три агенти - умова, причина і наслідок.

Треба пам'ятати, що максимум інформації, який досягнуто, завжди умовний, тому що є обмеження за ресурсами *З*:

$$З_i(R_a, R_b) \leq const \; (i = 1, 2, ..., n) \; .$$

Це обмеження за енергією, матеріалами, часом, кількості операцій та ін. Такі ресурси, як енергія, час і кількість операцій пов'язані монотонною залежністю. Тому принцип максимізації інформації може бути представлений у вигляді максимізації Лагранжіана:

$$L(R_a, R_b) = I(R_a, R_b) - \sum_{i=1}^{n} \lambda_i З_i(R_a, R_b) = \max_{R_a, R_b} ,$$

де λ_i - невизначений множник Лагранжа.

Цю функцію можна розглядати, як цільову в процесі взаємодії людини з природою, або як функцію привабливості.

Функцію привабливості можна представити двома еквівалентними формами:

$$L(\bar{R}_a, \bar{R}_b) = H(R_a) - H\left(\frac{R_a}{R_b}\right) - \sum_{i=1}^{n} \lambda_i З_i, \quad L(R_a, R_b) = H(R_b) - H\left(\frac{R_b}{R_a}\right) - \sum_{i=1}^{n} \lambda_i З_i,$$

де *H(Ra), H(Rb)* - безумовні ентропії; *H(Ra/Rb),* Н *(Rb/Ra)* - умовні ентропії.

У першій формі акцент робиться на *Ra,* у другій - на *Rb.*

Розглянемо далі причинно-наслідкові відносини в системі «людина - засоби дослідження - досліджувана частина природи». Через те, що причина і наслідок є основним вимірником конкретного внутрісистемного складу процесів, то еволюція цих категоріальних параметрів відображує еволюцію відносин між частинами системи. Для формального описування цих відносин використовуємо мову теорії множин.

Використовуючи бінарні відношення *р* (причина чогось) і *с* (наслідок чогось) можна для першого періоду задовольняння потреби в активності взаємодії записати

1. $(r_2^-)(r_1^+)\, pd$, тобто $(r_2^-)(r_1^+)$ причина дії d,
2. $dp\,(r_2^+)(r_1^-)$, тобто d причина $(r_2^+)(r_1^-)$.

Для отримання композиції відношень можна застосувати теорему Трела: *«якщо R відносини між елементами а множини b множини В, aS - відносини між елементами b множини В і с множини С, тоді композиція RS відображується як відносини між елементами а множини А і с множини С таким чином: aRSc завжди, якщо для деякого b у множині В має місце aRb і bSs».*

Застосувавши теорему Трела, можна композицію відношень представити в такому вигляді

$$(r_2^+)(r_1^-)pp(r_2^-)(r_1^+),$$

тобто при наявності дії *d* величина $(r_2^+)\,(r_1^+)$ є наслідок $(r_2^+)\,(r_1^+)$.

Аналогічні композиції відношень можуть бути отримані для наступних періодів задовольняння потреби (r_2^-). Для цього введемо нові відношення - *G* (наслідок дії *d*), *S* (наслідок наслідку *c*), *R* (наслідок причини).

Відповідно з вищевикладеним у першому періоді задовольняння потреби формуються причинно-наслідкові відносини

1. ***pSc***, 2. ***dRp***, 3. ***cGd***.

У другому періоді задовольняння потреби через знову створені засоби досліджень формуються нові відносини, що відповідно до теореми Трела можуть бути представлені у вигляді

1. ***pGSd***, **тому що має місце *cGd* і *dRp*,**
2. ***dRSc***, **тому що має місце *dRp* і *cGd*,**
3. ***cGRp***, **тому що має місце *dRp* і *pSc*.**

У третьому періоді задовольняння потреби формуються відносини

1. ***pSSRGc***, **тому що має місце *pSGd* і *dRSc*,**
2. ***dRRGSp***, **тому що має місце *dRSc* і *cRGp*,**
3. ***cGGRSd***, **тому що має місце *cRGp* і *pGSd*.**

Таким чином, до третього періоду задовольняння потреби (r_2^-) перевтілення відносин у системі приводять до стану, в якому наступні причина, дія і наслідок знаходяться в значно більш складних відносинах одне з одним. Іншими словами, у результаті перевтілення відносин між частинами системи виходять нові причинно-наслідкові зв'язки. Останнє свідчить про те, що через створення нових засобів дослідження система переходить на якісно новий рівень.

Аналіз причинно-наслідкових відносин у системі наукового дослідження показує, що процес задовольняння потреб включає три етапи. У кожному етапі міститься три періоди замкнених і три періоди розімкнених станів системи. На третьому етапі композиції відносини у системі стають у 1500-2500 разів складніші, ніж на попередніх етапах. Після закінчення трьох етапів формується принципово новий засіб дослідження. Формування нових засобів дослідження заперечує існування попередніх, а отже, і всієї старої системи наукового дослідження. Тому момент формування композиції відносин третього етапу є вихідним для декомпозиції старих відносин у наступних періодах задовольняння потреби.

Дослідження М. Тода, Е.Х. Шуфорда показують, що між безліччю відносин, що беруть участь у композиції і декомпозиції, існує взаємно однозначна відповідність. Тому процес декомпозиції відносин у системі точно іде за порядком, зворотнім тому, в якому здійснюється композиція відносин.

Дослідження Л. Лефгрена і В.Г. Панова показали, що процес еволюції системи включає послідовну зміну циклічних причинно-наслідкових траєкторій. У самому спрощеному вигляді дані траєкторії формуються в результаті взаємодії між суб'єктом праці, знаряддям праці і досліджуваною частиною природи відповідно до схеми

$$\begin{tikzcd} & M \arrow[dr] & \\ S \arrow[ur] \arrow[rr, shift left] & & N \arrow[ll, shift left] \end{tikzcd} \qquad \begin{tikzcd} & d \arrow[dr] & \\ p \arrow[ur] \arrow[rr, shift left] & & c \arrow[ll, shift left] \end{tikzcd}$$

де S - суб'єкт праці.

Відповідно до розглянутих закономірностей формування відносин між компонентами системи наукових досліджень і результатами досліджень Л. Лефгрена квантування часу її еволюції може бути представлене в такому вигляді як на рис. 2.6.

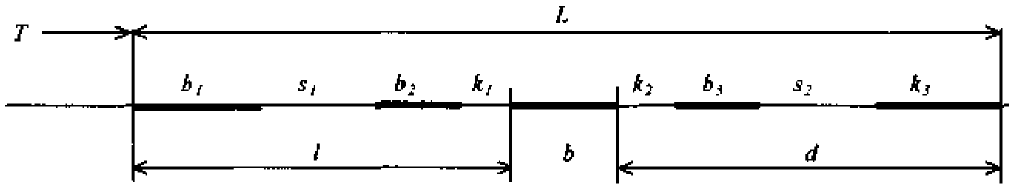

Рис. 2.6. Схема квантування часу еволюції системи наукового дослідження на останньому етапі його існування

Час існування системи включає інтервал передісторії T (стосовно до рис. 2.6 даний інтервал включає два попередніх етапи еволюції системи) та інтервал її еволюції L.

Інтервал еволюції

$$L = l + b + d,$$

де l - час формування системи; b - час стаціонарного стану системи; d - час руйнування системи.

Час формування системи включає

$$l = b_1 + s_1 + b_2 + k_1,$$

де b_1 - інтервал виділення із середовища споживання знаряддя праці (засобу дослідження); s_1 - інтервал асиміляції засобу дослідження; b_2 - інтервал виділення із середовища споживання знову створюваного засобу дослідження; k_1 - інтервал асиміляції нового засобу дослідження.

Інтервал руйнування старого детермінізму в системі включає

$$d = k_2 + b_3 + s_2 + k_3,$$

де k_2 - інтервал дисиміляції нового засобу дослідження; b_3 - інтервал вилучення зі структури системи нового засобу дослідження; s_2 - інтервал дисиміляції старого засобу дослідження; k_3 - інтервал вилучення зі структури системи старого засобу дослідження.

Асиміляція і дисиміляція старого і нового засобів дослідження не вимагає зміни організаційної норми системи. Тому в ці періоди максимальна ентропія системи не змінюється і система замкнена в організаційному відношенні.

Включення в систему нових компонентів чи вилучення з неї старих веде до зміни максимальної ентропії системи, що властиво розімкненому в організаційному відношенні стану.

Таким чином, відповідно до прийнятої схеми, *у процесі еволюції відбувається періодичне розмикання і замикання системи, перехід з хитливого у стійкий стани і назад.*

Відповідно з принципом «необхідної організації» Ю.Г. Антомонова адекватність між компонентами системи наукового дослідження в замкненому стані і навколишнім середовищем встановлюється при виконанні умови:

$$Q_i = Q_{сер.i}$$

де Q_i - абсолютна організація i-того компонента системи наукового дослідження; $Q_{сер.i}$ - абсолютна організація зовнішнього середовища для i-того компонента.

Під абсолютною організацією (згідно з Ю.Г. Антомоновим) розуміють реалізовану в системі невизначеність. В замкненому стані системи абсолютну організацію визначають за формулою

$$Q = H_m - H,$$

де H_m, H - максимальна і поточна ентропії системи.

Таке визначення називають *законом збереження ентропії.*

Враховуючи цю умову, рівняння статичної рівноваги i-того компонента

$$Q_{сер.i} - Q_i = 0.$$

При порушенні статичної рівноваги абсолютна організація i-того компонента системи змінюється на величину dQ_i, за елементарний інтервал часу dt. Тому умова динамічної рівноваги i-того компонента може бути представлена у вигляді:

$$\frac{dQ_i}{dt} = \Delta Q_{сер.i} - Q_i.$$

Якщо покласти, що кожен компонент системи може знаходитись лише в двох станах (фактичному і заданому), тоді

$$Q = 1 + P_i \log P_i + (1 - P_i)\log(1 - P_i).$$

Припустимо, що

$$\text{при } P_i \le 0{,}5 \qquad Q = 1 - 2P_i,$$
$$\text{при } P_i > 0{,}5 \qquad Q = 2P_i - 1.$$

Звідси

$$dQ_i = C_i dt.$$

де C_i - організаційна ємність і-того компонента.

Отже, рівняння динамічної рівноваги здобуває вигляд:

$$C_i \frac{dP_i}{dt} \Delta Q_{сер.i} - \Delta Q_i.$$

Середовищем для людини є старі, нові засоби дослідження і всі зовнішні фактори, що впливають на його стан. До зовнішнього можна віднести економічні, організаційні, погодно-кліматичні та інші фактори. Для старих засобів дослідження середовищем є людина, нові засоби дослідження і всі зовнішні фактори. Для нових засобів дослідження - людина, старі засоби дослідження і всі зовнішні фактори.

Оскільки кожен компонент системи прагне зрівноважитись із середовищем, то абсолютна організація цього середовища виступає в ролі своєрідної норми, до якої прагне компонент.

Норма абсолютної організації компонента системи наукових досліджень згідно Г.А. Голіцину може бути визначена як середньозважена з індивідуальних норм компонентів, що формують зовнішнє середовище

$$Q_{Hi} = \frac{\sum\limits_{j=1}^{m} \gamma_j^{(i)} Q_{ij}^{(i)}}{\sum\limits_{j=1}^{m} \gamma_j^{(i)}},$$

де Q_{Hi} - норма абсолютної організації і-того компонента системи; $Q_{Hj}^{(i)}$ - норма абсолютної організації j-того компонента середовища для і-того компонента системи; $\gamma_j^{(i)}$ - жорсткість норми $Q_{Hj}^{(i)}$.

Тут і далі під нормою розуміють оптимальну абсолютну організацію, що найбільшою мірою відповідає цілям і завданням функціонування системи.

Динаміка норми абсолютної організації описується рівнянням Г.А. Голіцина

$$Q_{Hi}(t) = \frac{1}{T_П} \int\limits_{t-T_П}^{t} Q_i(\tau) d\tau,$$

де Q_{Hi} - поточне значення абсолютної організації і-того компонента системи; $T_{П}$ - глибина пам'яті.

Тому норма і-того компонента поводиться подібно інерційній системі, що стежить: «відслідковує» дійсне значення перемінної. Якщо перехід зі стану $Q_i(0)$ у стан Q_{Hi} триває досить довго, тоді можна прийняти

$$Q_{Hi} = \frac{\sum_{J}^{3} \gamma_j^{(i)} Q_j^{(i)}}{\sum_{J}^{3} \gamma_j^{(i)}}.$$

Враховуючи вищевикладене, після розкладання в ряди Маклорена збільшень абсолютної організації і наступної лінеаризації, рівняння динамічної рівноваги компонентів системи наукового дослідження представляються у вигляді

$$C_ч \frac{d\Delta P_ч}{dt} - \left(\left(\frac{\partial Q_{сер\,ч}}{\partial P_а} \Delta P_а + \frac{\partial Q_{сер\,ч}}{\partial P_д} \right) - \frac{\partial Q_ч}{\partial P_ч} \Delta P_ч \right) = \frac{\partial Q_{сер\,ч}}{\partial P_с} \Delta P_с,$$

$$C_а \frac{d\Delta P_а}{dt} - \left(\left(\frac{\partial Q_{сер\,а}}{\partial P_ч} \Delta P_ч + \frac{\partial Q_{сер\,а}}{\partial P_д} \Delta P_д \right) - \frac{\partial Q_а}{\partial P_а} \Delta P_а \right) = \frac{\partial Q_{сер\,а}}{\partial P_с} \Delta P_с,$$

$$C_д \frac{d\Delta P_д}{dt} - \left(\left(\frac{\partial Q_{сер\,д}}{\partial P_ч} \Delta P_ч + \frac{\partial Q_{сер\,д}}{\partial P_а} \right) - \frac{\partial Q_д}{\partial P_д} \Delta P_д \right) = \frac{\partial Q_{сер\,д}}{\partial P_с} \Delta P_с,$$

де $\Delta P_ч$, $\Delta P_а$, $\Delta P_д$ - доданки імовірностей переходу людини, старого і нового засобів дослідження з фактичного в нормальний стан; $\Delta P_с$ - доданок імовірності переходу середовища з фактичного в нормальний стан.

Часткове рішення цієї системи рівнянь запишеться у вигляді

$$P_ч(\tau) = \frac{C_0}{\lambda_1} e^{\lambda_1 \tau} + \frac{C_1}{\alpha^2 + \beta^2} e^{\alpha\tau} (\beta \sin \beta\tau + \alpha \cos \beta\tau) +,$$

$$+ \frac{C_2}{\alpha^2 + \beta^2} e^{\alpha\tau} (\alpha \sin \beta\tau - \beta \cos \beta\tau) + 1;$$

$$P_д(\tau) = \frac{C_0 N_0}{\lambda_1} e^{\lambda_1 \tau} + \frac{C_1 N_1 + C_2 N_2}{\alpha^2 + \beta^2} e^{\alpha\tau} (\beta \sin \beta\tau + \alpha \cos \beta\tau) +,$$

$$+ \frac{C_2 N_1 - C_1 N_2}{\alpha^2 + \beta^2} e^{\alpha\tau} (\alpha \sin \beta\tau - \beta \cos \beta\tau) + 1;$$

$$P_а(\tau) = \frac{C_0 M_0}{\lambda_1} e^{\lambda_1 \tau} + \frac{C_1 M_1 - C_2 M_2}{\alpha^2 + \beta^2} e^{\alpha\tau} (\beta \sin \beta\tau + \alpha \cos \beta\tau) +,$$

$$+ \frac{C_1 M_2 + C_2 N_1}{\alpha^2 + \beta^2} e^{\alpha\tau} (\alpha \sin \beta\tau - \beta \cos \beta\tau) + 1,$$

де C_0, C_1, C_2 - довільні постійні, залежні від початкових умов; λ - корінь характеристичного рівняння; α - дійсна частина комплексно-сполучених коренів; β - кругова частота коливань; τ - безрозмірний вимір часу;

$$M_0 = \frac{1}{m_1 k_\partial^{(1)} - m_2 k_a^{(1)}}\left(-\frac{k_\partial^{(1)}\lambda_1^2}{r^2} - \frac{m_2}{r}\lambda + \left(m_2 - k_\partial^{(1)} m_0\right)\right);$$

$$M_1 = \frac{1}{m_1 k_\partial^{(1)} - m_2 k_a^{(1)}}\left(-\frac{k_\partial^{(1)}\left(\alpha^2 - \beta^2\right)}{r^2} - \frac{m_2}{r}\alpha + \left(m_2 - k_\partial^{(1)} m_0\right)\right);$$

$$M_2 = \frac{1}{m_1 k_\partial^{(1)} - m_2 k_a^{(1)}}\left(\frac{2\alpha\beta k_\partial^{(1)}}{r^2} - \frac{m_2}{r}\beta\right);$$

$$N_0 = \frac{1}{m_1 k_\partial^{(1)} - m_2 k_a^{(1)}}\left(-\left(k_a^{(1)} m_0 - m_1\right) + \frac{m_1}{r}\lambda + \frac{k_a^{(1)}\lambda_1^2}{r^2}\right);$$

$$N_1 = \frac{1}{k_a^{(1)} m_2 - k_\partial^{(1)} m_1}\left(\left(k_a^{(1)} m_0 - m_1\right) - \frac{m_1}{r}\alpha - \frac{k_a^{(1)}\left(\alpha^2 - \beta^2\right)}{r^2}\right),$$

$$N_2 = \frac{1}{\left(m_1 k_\partial^{(1)} - m_2 k_a^{(1)}\right)}\left(\frac{2\alpha\beta k_a^{(1)}}{r^2} + \frac{m_1}{r}\beta\right);$$

$$m_0 = \left(1 + k_a^{(1)} k_ч^{(2)} + k_\partial^{(1)} k_ч^{(3)}\right);$$

$$m_1 = \left(2k_a^{(1)} - k_\partial^{(1)} k_a^{(3)}\right);$$

$$m_2 = \left(2k_\partial^{(1)} - k_\partial^{(2)} k_a^{(1)}\right);$$

$$k_a^{(1)} + k_\partial^{(1)} + k_c^{(1)} = 1;$$

$$k_ч^{(2)} + k_\partial^{(2)} + k_c^{(2)} = 1;$$

$$k_ч^{(3)} + k_a^{(3)} + k_c^{(3)} = 1;$$

$$k_a^{(1)} = \frac{\gamma_a}{\gamma_a + \gamma_\partial + \gamma_c};$$

$$\tau = \frac{t}{r};$$

$$k_\partial^{(1)} = \frac{\gamma_\partial}{\gamma_\partial + \gamma_a + \gamma_c}; \qquad k_c^{(2)} = \frac{\gamma_c}{\gamma_ч + \gamma_\partial + \gamma_c};$$

$$k_c^{(1)} = \frac{\gamma_c}{\gamma_c + \gamma_a + \gamma_\partial}; \qquad k_ч^{(3)} = \frac{\gamma_ч}{\gamma_ч + \gamma_a + \gamma_c};$$

$$k_ч^{(2)} = \frac{\gamma_ч}{\gamma_ч + \gamma_\partial + \gamma_c}; \qquad k_a^{(3)} = \frac{\gamma_a}{\gamma_a + \gamma_\partial + \gamma_c};$$

$$k_\partial^{(2)} = \frac{\gamma_\partial}{\gamma_ч + \gamma_\partial + \gamma_c}; \qquad k_c^{(3)} = \frac{\gamma_c}{\gamma_ч + \gamma_\partial + \gamma_c};$$

$$r = \frac{1}{\sqrt[3]{\dfrac{n_0 a - n_1\left(m_0 k_\partial^{(1)} - m_2\right) - n_2\left(m_1 - k_a^{(1)} m_0\right)}{a}}};$$

$$a = m_1 k_\partial^{(1)} - m_2 k_a^{(1)}; \qquad n_1 = m_1 + m_0 k_a^{(1)} - m_2 k_a^{(3)};$$

$$n_0 = m_0 + m_1 k_ч^{(2)} + m_2 k_ч^{(3)}; \qquad n_2 = m_2 + m_0 k_\partial^{(1)} - m_1 k_\partial^{(2)};$$

r — **константа, яка є масштабом часу.**

У розімкненому стані системи йде формування нового засобу наукового дослідження. Останнє реалізується через збільшення числа елементів, що поступають у систему із середовища.

При зміні числа елементів і зв'язків між ними змінюється максимальна ентропія системи. Одночасно система прагне досягти нового рівня адекватності із середовищем.

Процес встановлення адекватності динамічний і для кожного моменту часу представляється у вигляді

$$H_m^s(t) \approx H_m^e(t),$$
$$H^s(t) \approx H^e(t),$$

де H_m^s, H_m^e - максимальні ентропії системи і середовища відповідно; H^s, H^e - поточні ентропії системи і середовища відповідно; *s*, *e* - індекси приналежності до системи і до середовища; *t* - час.

Відповідно з принципом Г. А. Голіцина про оптимальну швидкість задовольняння потреб, у рівноважному стані швидкості задовольняння потреб залишаються постійними і рівними оптимальній швидкості, тобто

$$\frac{dH_m^s}{dt} = a = \text{const}, \qquad \frac{dH^s}{dt} = b = \text{const}.$$

Якщо при $t = 0$, $H_m^s = H_{m0}^s$, $H^s = 0$, тоді

$$H_m^s = H_{m0}^s + at, \quad \text{чи} \quad H_m^s = H_{m0}^s + \frac{a}{b}H.$$

Якщо при $t = 0$, $H_m^s = H_{m0}^s$, $H^s = H_0^s$, тоді

$$H_m^s = H_{m0}^s + at, \quad H^s = H_0^s + bt,$$

Враховуючи закон збереження організації Ю.Г. Антомонова, можна записати

$$Q^s(t) = H_m^s(t) - H^s(t), \quad \text{чи} \quad Q^s(t) = H_{m0}^s - (b - a)t,$$

де Q^s - абсолютна організація системи.

Диференціюючи $Q^s(t)$ за часом, одержимо

$$\frac{dQ^s}{dt} = a - b.$$

Останнє рівняння дозволяє установити умову підвищення рівня організації системи наукового дослідження в розімкненому стані:

$$\frac{dQ^s}{dt} > 0, \qquad \text{якщо} \quad a > b.$$

Надходження речовини, енергії та інформації з зовнішнього середовища може відбуватись стрибкоподібно. У такому разі зміни максимальної ентропії системи будуть носити слабко коливальний характер із загасанням і прагненням процесу до деякого сталого значення адаптації з часом. Перехідні процеси такого типу описуються вирішенням неоднорідного диференціального рівняння другого порядку з комплексно-сполученими коренями характеристичного рівняння

$$\frac{d^2H_m}{dt^2} + p_1\frac{dH_m}{dt} + q_0H_m = R_3,$$

де R_3 - зовнішній вплив.

Виріїпення даного рівняння для ненульових початкових умов може бути представлене у вигляді

$$H_m = e^{-\alpha x}\left[\frac{H'_{m0} + \alpha\left(H_{m0} - H_{my}\right)}{\beta}\sin\beta t + \left(H_{m0} - H_{my}\right)\cos\beta t\right] + H_{my},$$

де H_{m0}, H'_{m0} - максимальна ентропія і її похідна при $t = 0$;

α - декремент загасання коливань,

$$\alpha = \frac{p_1}{2};$$

β - кругова частота,

$$\beta = \sqrt{q_0 - \frac{p_1^2}{4}};$$

H_{my} - значення максимальної ентропії, що встановилося (межа адаптаційних можливостей системи),

$$H_{my} = \frac{R_3}{q_1}.$$

При нульових початкових умовах, тобто при t = 0 $H_0 = \dot{H}_0 = 0$

$$\dot{H}_m = H_{my}\left[1 - e^{-\alpha t}\left(\frac{\alpha}{\beta}\sin\beta t + \cos\beta t\right)\right].$$

Аналогічна закономірність зміни поточної ентропії системи

$$H^s(t) = e^{-\alpha t}\left[\frac{H_0' + \alpha(H_0 - H_y)}{\beta}\sin\beta t + (H_0 - H_y)\cos\beta t\right] + H_y,$$

де H_y - значення поточної ентропії, що встановилося; H_0, $\dot{H}_0$- початкове значення поточної ентропії і її похідна.

На рис. 2.7 представлена типова схема адаптації. Коливальний характер кривої дозволяє говорити про перевагу тих чи інших механізмів у часі і розділити всю динаміку адаптації на етапи. На першому етапі від початку адаптації до першого максимального значення максимальної ентропії переважає мобілізація структурно-функціональних резервів, на другому, від першого максимуму до першого мінімуму, - накопичення ушкоджених структур. Третій (від першого мінімуму до другого максимуму) характеризується процесами відновлення функцій первісно перевтілених структур. Четвертий етап, до встановлення рівноваги, характеризується структурно-функціональним зрівноважуванням усіх процесів. П'ятий - рівновагою і значенням H_{my}.

Імовірності переходів з фактичного в заданий стан можуть використовуватись як вагові коефіцієнти цих станів. Тому в замкненому стані системи наукового дослідження її характеристики можуть бути передбачені на основі такої моделі:

$$X(t) = X_0 q(t) + X_3 p(t),$$

де $X(t)$ - поточна кількісна характеристика компонента системи; X_0 - кількісна характеристика компонента при $t = 0$; X_3 - задана характеристика компонента; $q(t)$ - імовірність того, що компонент системи не перейшов у заданий стан; $p(t)$ - імовірність переходу компонента у заданий стан.

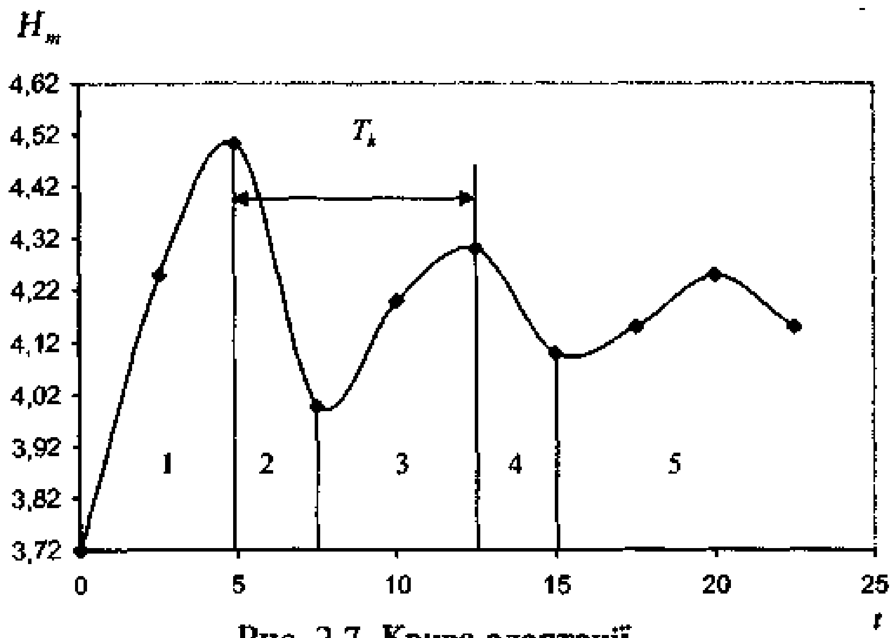

Рис. 2.7. Крива адаптації

У розімкненому стані можна прийняти, що збільшення кількісної характеристики стану системи визначається різницею між максимальною невизначеністю стану системи і її поточною абсолютною організацією (Ю.Г Антомонов, Л.І. Краснікова, О.Г. Чораян). З цього виходить емпірична формула

$$\Delta X^S(t) = \beta_i \left[H_m^S(t) - Q^S(t) \right] = \beta_i H^S(t),$$

де $\Delta X^S(t)$ - збільшення кількісної характеристики стану системи; H_m^S, H^S- максимальна і поточна ентропія системи; Q^S - абсолютна організація системи; β_i- коефіцієнт пропорційності і розмірності.

Отже, модель прогнозування характеристик стану розімкненої системи може бути представлена у вигляді

$$X^S(t) = X_0^S + \beta_i H^S(t),$$

де X_0^S - кількісна характеристика стану системи при $t = 0$.

Аналогічні моделі можуть бути отримані для кожного з компонентів системи.

Розглянуті закономірності еволюції системи наукового дослідження притаманні всьому класу систем «людина - техніка - середовище». Ці закономірності можуть використовуватись при вирішенні задач прогнозування властивостей компонентів систем і необхідних професійних якостей особистості людини для управління технічними системами.

3 МОДЕЛЬ І МОДЕЛЮВАННЯ В СИСТЕМНОМУ ДОСЛІДЖЕННІ

1 Модель об'єкта.
2 Критерій аналогії.
3 Теорія подоби та критерій подоби.
4 Класифікація моделей.
5 Класифікація різновидів зв'язків подоби і моделювання.
6 Методологія моделювання.
7 Вимоги до фактори, що враховуються при моделюванні.
8 Декомпозиція.
9 Модель-основа та вимоги до неї.
10 Компроміс між вимогою простоти і повноти аналізу.
11 Агрегування та основні агрегати.
12 Математична модель об'єкта, перший етап математичного моделювання.
13 Другий етап математичного моделювання.
14 Алгоритм установлення загальних характеристик системи, що моделюється.
15 Математичний апарат для побудови математичної моделі.
16 Третій етап математичного моделювання.
17 Попередній контроль математичної моделі об'єкта.
18 Імітаційна модель.
19 Схема алгоритму процесу імітації.

3.1 Класифікація моделей

У процесі діяльності в людини формуються представлення про ті чи інші властивості навколишніх об'єктів і їхніх взаємозв'язків. Вони формуються у вигляді описів цих об'єктів звичайною мовою, фіксуються на папері мовою рисунка, графіка, рівнянь і формул чи реалізуються у вигляді макетів та інших пристроїв. Подібні способи опису узагальнюються в єдиному понятті - ***модель,*** а побудова, вивчення і використання моделей об'єктів як спеціального засобу пізнання називають ***моделюванням.***

Під моделлю розуміють таку представлену подумки чи матеріально реалізовану систему, яка, відображаючи чи відтворюючи об'єкт дослідження, здатна заміщати його так, що її вивчення подає нову інформацію про цей об'єкт. Модель знаходиться у деякій відповідності з досліджуваним об'єктом. При побудові моделі об'єкт і його властивості подумки узагальнюються і спрощуються. Модель відображує лише суттєві властивості об'єкта, зайва деталізація утруднює дослідження, робить його громіздким. Тому модель виконують оптимальною за складністю виходячи з мети дослідження.

Моделювання спирається на строгу теоретичну базу, що включає ***теорії аналогії і подоби.*** Моделі, побудовані на базі аналогії, називають «моделями - аналогами». У випадку застосування теорії подоби, отримані моделі іменують «моделями - подобами».

Суть моделювання на основі теорії аналогії розглянемо на прикладі.

Тепловий потік q_T за законом Фур'є залежить від температурного перепаду:

$$q_T = -\lambda \frac{dt}{dx}$$

де λ - коефіцієнт теплопровідності.

Масоперенос чи перенос речовини q_B за законом Фіка визначають перепадом концентрацій речовини *c*:

$$q_B = -\mu \frac{dc}{dx}$$

де μ - коефіцієнт масопереносу.

Перенос електрики q_E за законом Ома обумовлений перепадом напруги:

$$q_E = -\frac{1}{\rho}\frac{du}{dx}$$

де ρ - коефіцієнт електропровідності.

Усі три явища відносяться до різних фізичних процесів, але описуються ідентичними математичними виразеннями. Це вказує на деяку аналогію розглянутих явищ і дозволяє використовувати один з розглянутих фізичних процесів для моделювання інших. Але ідентичність математичних виразів процесів не означає, що ці процеси абсолютно аналогічні. Для остаточної відповіді на це питання необхідне дотримання критерію аналогії.

Для отримання критерію аналогії необхідно математичні вираження процесів представити у безрозмірному вигляді. Для цього кожну змінну величину варто представити у вигляді добутку постійної розмірності на безрозмірну перемінну. Так, якщо розглядати процеси теплопереносу і переносу електрики, то перемінні, що входять у математичні вираження процесів, мають бути представлені таким чином:

$$q_T = q_{TП} q_{Tб};\ \lambda = \lambda_П \lambda_б;\ t = t_П t_б;\ x = x_П x_б;\ q_E = q_{EП} q_{Eб};\ \gamma = \gamma_П \gamma_б;\ u = u_П u_б .$$

Після підстановки перетворених перемінних у початкові рівняння одержимо ці ж рівняння у безрозмірному вигляді:

$$\left[\frac{q_{TП}x_П}{\lambda_П t_П}\right]q_{Tб} = \lambda_б \frac{dt_б}{dx_б}; \quad \left[\frac{q_{EП}x_П}{\gamma_П u_П}\right]q_{Eб} = \lambda_б \frac{du_б}{dx_б}.$$

Дані рівняння можна розглядати як ідентичні, якщо

$$\left[\frac{q_{TП}x_П}{\lambda_П t_П}\right] = \left[\frac{q_{EП}x_П}{\gamma_П u_П}\right]$$

Цю рівність називають ***критерієм аналоги.*** З його допомогою встановлюють параметри моделі по початковому рівнянню об'єкта.

Кількість критеріїв аналогії завжди на одиницю менше числа членів досліджуваного початкового вираження. Через те, що число невідомих більше числа рівнянь, то деякі параметри моделі задаються.

Теорія подоби - це навчання про подобу явищ. Основоположником теорії подоби є Г. Галілей, який вперше стосовно до механічного руху і механічних систем сформулював ідею про існування закономірних залежностей між різними параметрами, що визначають механічну систему. Г. Галілей показав, що подоба механічних систем не обмежується тільки їхньою геометричною подобою, а припускає визначені співвідношення, що зв'язують геометричні відношення з фізичними властивостями таких систем. Подальший розвиток цих ідей відбувався у роботах Ньютона, Бертрана, Букінгема, Кірпічьова.

Суть теорії подоби розглянемо на простому прикладі. Нехай є рад прямокутників, що відносяться до класу плоских фігур. Вони об'єднані загальними властивостями - мають по чотири сторони і по чотири прямих кута. З цього класу можна виділити одиничну фігуру, що має конкретне значення сторін l_1 і l_2. Якщо сторони l_1 і l_2 помножити на K_1 якому можна додати будь-яке значення, то одержимо серію подібних плоских фігур із відношеннями

$$\frac{l_1'}{l_2'} = \frac{l_1''}{l_2''} = \frac{l_1'''}{l_2'''} = \ldots = K_1$$

де K_1 - означає «однаково для всіх розглянутих об'єктів».

Величину K_1 називають критерієм подоби.

Такий спосіб визначення критерію подоби застосовується не тільки для геометричних фігур, але й для різних фізичних процесів. У широкому смислі критерій подоби - це комплексне вираження, що складається з констант, яке зв'язує їх за деяким законом і зберігає постійність у подібних систем.

Критерії подоби формують усередині класу явищ деяку групу подібних. У цій групі явища відрізняються тільки масштабами.

Теорія подоби базується на трьох теоремах.

Теорема 1. *Два фізичних явища подібні, якщо вони описуються однією і тією ж системою диференціальних рівнянь і мають подібні (граничні) умови однозначності, та їхні визначальні критерії подоби чисельно рівні.*

Теорема 2. *Якщо фізичні процеси подібні, то критерії подоби цих процесів рівні між собою.*

***Теорема** 3. Рівняння, що описують фізичні процеси, можуть бути виражені диференціальним зв'язком між критеріями подоби.*

Критерії подоби будь-якого явища можуть перетворюватись в критерії іншої форми за допомогою операцій множення чи ділення раніше знайдених критеріїв один на другий.

Критерії подоби є математичним формулюванням тих умов, при яких модель може вважатись закономірно відображаючою оригінал. У табл. 3.1-3.4 приведені критерії подоби деяких найбільш характерних процесів.

Таблиця 3.1. **Критерії електричної подоби**

Критерії	**Формальне вираження**
Додаткові умови подоби систем з розподіленими параметрами	$R_0 G_0 l^2 = idem$
Приблизного електромагнітної та електродинамічної подоби	$T_1 = T_2 = idem$
Намагнічування феромагнітних тіл	$l\sqrt{\gamma\mu'\omega} = idem$
Електромагнітної подоби середовища, що рухається	$\lambda\mu v / t^{-2} = idem$

Таблиця 3.2. **Критерії теплової подоби**

Критерій	**Формальне вираження**
Фурьє	$at/l^2 = idem = [Fo]$ (тут $a = \lambda / c_p \xi$ — коефіцієнт температуропровідності; λ — коефіцієнт теплопровідності; c_p — коефіцієнт теплоємкості; ξ — питома вага).
Пеклє (для рідини, що рухається з заданим тепловим станом)	$Vl/a = idem = [Pe]$
Нусельта	$al/\lambda = idem = [Nu]$
Кірпічьова	$Rl/\lambda = idem = [Ki]$ (тут R — коефіцієнт теплопередачі)
Прандтля	$\gamma_0 / a = idem[Pr]$

Таблиця 3.3. **Критерії механічної і гідродинамічної подоби**

Критерій	Формальне вираження
Ньютона	$Ft^2 / Ml = [Ne]$
Фруда	$v^2 / ql = [Fr]$
Ейлера	$p / \rho v^2 = [Eu]$
Рейнольдса	$v / \mu_0 = [Re]$
Архімеда	$\frac{gl^3}{\gamma_0^2} \frac{\rho - \rho}{\rho} = [Ar]$

Примітки: t — час; M — маса; l — геометричний розмір; v — швидкість; g — прискорення сили ваги; p — сила, тиск; ρ — щільність рідини; μ_0 — в'язкість; γ_0— коефіцієнт кінематичної в'язкості.

Таблиця 3.4. **Критерії електромагітної подоби**

Загальні критерії	Формальне вираження
Подоба електромагнітних явищ	$\pi_1 = \frac{\mu\gamma l^2}{t} = idem$
Гомохронності	$[Ho] = \omega t = idem$
Подоба ланцюгів	$\pi_* L_a = L_a / R_a t = idem$ $\pi_* c_a = c_a / Gt = idem$
Електродинамічна подоба	$T_{*j} = \frac{T_j}{t} = idem$

Примітки: μ — коефіцієнт магнітної проникності; γ — провідність; l — геометричний розмір; t — час; ε — діелектрична постійна; ω — кутова швидкість; $\omega = 2\pi f$, f — частота; T — постійна часу; L — індуктивність ланцюга; R — омічний опір; c — ємність; G — провідність на одиницю довжини електричної лінії.

До теорем подоби є ряд додаткових положень, якими необхідно керуватись при моделюванні:

1. Подоба складних систем, які складаються з декількох підсистем, відповідно подібних окремо, забезпечується подобою всіх схожих елементів, що є загальними для підсистем. Як наслідок цього, подібні складні системи залишаються подібними після будь-яких спрощень, якщо ці спрощення були проведені в системах відповідно однаково.

2. Усі теореми й умови подоби справедливі для систем різної складності. Вони можуть бути поширені на нелінійні системи, чи системи з перемінними параметрами, якщо виконуються умови співпадіння відносних характеристик, схожих параметрів, що є нелінійними чи перемінними.

3. Умови подоби, справедливі для ізотропних систем, які характеризуються однаковістю фізичних властивостей (електропровідність, теплопровідність, пружність та ін.) по всіх координатах усередині даної системи, можуть бути поширені і на анізотропні системи, що мають неоднакові властивості на різних напрямках.

При цьому відносні анізотропії в системах, що порівнюються, мають бути відповідно однакові.

4. У системах, геометрично не подібних, але таких, що мають нелінійну подобу простору, процеси можуть бути фізично подібні, маючи в схожих точках простору подібні зміни параметрів процесу.

5. Всі умови подоби, що відносяться до детерміновано заданих систем, справедливі для стохастично визначених систем за умовою співпадіння в цих системах щільностей імовірностей схожих параметрів, представлених у вигляді відносних характеристик. При цьому дисперсії і математичні очікування всіх параметрів, враховуючи масштаби, мають бути у подібних систем однаковими. Додатковою умовою подоби є виконання вимоги фізичної реалізованості схожої кореляції між стохастично заданими параметрами, що входять в умову однозначності.

Слід зазначити, що теорія подоби обмежується встановленням співвідношення між якісно однорідними явищами, що відносяться до однієї й тієї ж форми руху матерії.

Усі моделі можна класифікувати за двома ознаками:

1) *за способом побудови;*

2) *за якісною специфікою процесу чи об'єкта, що моделюють.* Залежно від *способу побудови* моделі поділяють на два класи: матеріальні (реальні, речовинні); ідеальні (уявні, умоглядні, гадані) (рис. 3.1).

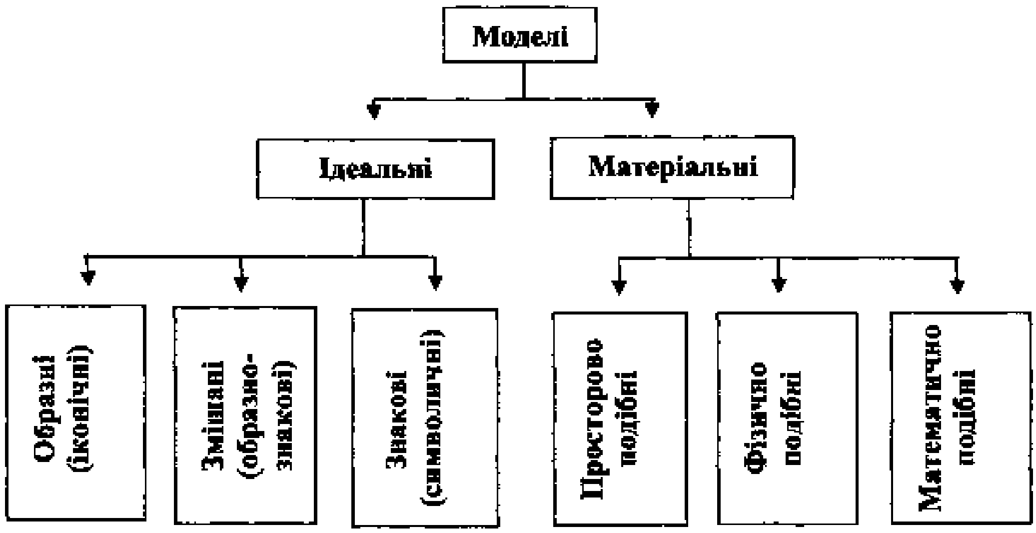

Рис. 3.1. Класифікація моделей

Матеріальні моделі поділяють на три підкласи: просторово подібні; фізично подібні; математично подібні.

У свою чергу, ідеальні моделі також поділяють на три підкласи: образні (іконічні); змішані (образно-знакові); знакові (символічні).

Просторово подібні моделі відтворюють просторові властивості чи відношення реального об'єкта. Відношення цих моделей до об'єкта характеризується геометричною подобою як обов'язковою умовою. До цих моделей відносять різні макети і муляжі.

Фізично подібні моделі створюються з метою відтворити не стільки просторові властивості, скільки динаміку досліджуваних процесів, різного роду залежності і закономірності. Фізична подоба припускає однаковість чи подібність фізичної природи моделі й об'єкта і тотожність законів їхнього руху.

Математично подібні моделі не вимагають виконання умов фізичної і геометричної подоби. Тут відносини між моделлю і реальним об'єктом зводяться до аналогії. Ця аналогія може бути структурною чи функціональною (ізоморфізм чи ізофункціоналізм), що знаходить відображення у наявності однакового математичного формалізму. До цих моделей відносять всілякі аналогові моделі (наприклад, електричні моделі механічних, теплових, акустичних явищ та ін.), структурні, цифрові і кібернетичні функціональні моделі.

Матеріальні моделі нерозривно пов'язані з ідеальними, тому що людина, перш ніж побудувати модель з яких-небудь матеріалів, подумки уявляє собі майбутню матеріальну модель.

Ідеальні моделі залишаються уявними навіть тоді, коли вони втілені в якій-небудь матеріальній формі, у вигляді рисунка, креслення, схеми чи просто системи знаків (математичних формул). Усі перетворення в них, усі переходи в інший стан здійснюються подумки, тобто у свідомості людини. Без уявних перетворень рисунки, креслення та інше позбавляються смислу.

Іконічні моделі конструюються у свідомості людини з образів якихось реальних явищ чи об'єктів (рідини, гази, пари, потоки та ін.).

У змішаних моделях іконічна модель доповнюється рисунком, кресленням, схемою, графіком. Це додає іконічній моделі властивість наочності.

У знакових моделях властивість наочності відсутня, тому що поняття знака виключає подібність між знаком і предметом чи явищем, що він позначає. У силу знаковості така модель за своєю фізичною природою не має вже нічого спільного з характером елементів об'єктів, що моделюються.

За ***якісною специфікою*** процесу чи об'єкта, що моделюється, розрізняють такі види моделей: динамічні і статичні; безперервні і дискретні; детерміновані, квазідетерміновані та імовірнісні; цілісні і часткові; концептуальні; структурні; функціональні; субстанціональні; процедурні та ін.

Відзначимо, що як і реальні системи, їхні моделі можуть класифікуватись *за типом системи, що моделюється, мовою опису функціонування системи (логічні, інформаційні), областю досліджень системи (економічні, психологічні, фізіологічні та ін.), призначенням (дескриптивні, нормативні).*

Дескриптивні моделі призначені для пояснення факторів, що спостерігаються, чи прогнозу поведінки об'єкта. Вони відповідають на питання «як це відбувається, як буде розвиватись?». *Нормативні* - призначені для визначення найкращого чи допустимого з погляду дослідника стану об'єкта. Вони відповідають на запитання «як має бути?».

Усі **різновиди зв'язків подоби і моделювання** можуть бути класифіковані таким чином:

Повне моделювання. Відбиває подобу руху матерії в основних формах її існування, тобто в часі й просторі. Процеси, що відповідають досліджуваним явищам, подібно змінюються і в часі й у просторі. Повна подоба і відповідно повне моделювання математично характеризують таким співвідношенням параметрів моделі *x* й оригіналу *y*.

$$x_j = m_j y_j$$

де m_j - масштабний коефіцієнт; y_j - параметри системи чи режиму її функціонування,

$$y_j = \varphi\left(y_1,\ldots,y_{j-1},\ldots,y_{j+1},\ldots,y_k,l_x,l_y,l_z,t\right)$$

де l_x, l_y, l_z - геометричні розміри; t - час.

Неповне (часткове, локальне) ***моделювання.*** Протікання в моделі всіх основних процесів, що характеризують досліджуване явище, подібно частково або тільки в часі, або тільки в просторі.

У першому випадку

$$y_j = \varphi\left(y_1,\ldots,y_{j-1},\ldots,y_{j+1},..,t\right).$$

У другому випадку

$$y_j = \varphi_1\left(y_1,\ldots,y_{j-1},\ldots,y_{j+1},..,l_x,l_y,l_z\right).$$

Наближене моделювання. Деякі фактори, що свідомо впливають, але не здійснюють вирішальної дії на досліджуваний процес, моделюються приблизно чи зовсім не моделюються. У зв'язку з цим між деякими параметрами моделі не існує співвідношень подоби і

$$x_j \neq m_j y_j \quad \text{чи} \quad x_j \approx m_j y_j$$

Це свідомо обумовлює погрішність, яку можна тим чи іншим способом оцінити кількісно.

Усі види моделювання можуть бути або детермінованими, що відображують процес з однозначно визначеними причинами і їх наслідками, або стохастичними, що відображують імовірнісні події. Моделювання будь-якого виду може проводитись як у реальному, так і в зміненому часі.

При системному моделюванні на перший план висувається можливість деяких узагальнень чи виявлення інтегральних властивостей. Для цього в умови подоби вводять узагальнені величини, що називають *інтегроформами,* за допомогою яких виключають перемінні коефіцієнти диференціальних рівнянь. Виявляється, можливо вивести такі інтегральні варіанти тотожних перетворень, що характеризують процес у цілому, відображаючи так звані *метаякісні* явища даного класу. У результаті об'єкти відображують не змінними величинами, а інтегральними співвідношеннями, і, таким чином, у дію вводиться як би «вторинна інформація», що заміняє систему первинної інформації, одержаної на основі критерію подоби.

У розробці моделей поряд з теоріями аналогії і подоби істотне значення грає аналіз розмірностей. Застосування аналізу розмірностей дозволяє згрупувати перемінні в безрозмірні комбінації, число яких менше вихідних перемінних.

Розмірними називають величини, чисельні значення яких залежать від прийнятих масштабів, тобто системи одиниць вимірювання. Це, наприклад, довжина, сила, час, енергія, температура та ін.

Безрозмірними називають величини, чисельні значення яких не залежать від прийнятих масштабів і застосовуваної системи одиниць вимірювання. Це кут у радіанах (відношення дуги до радіуса), відношення двох довжин, відношення енергії до моменту сили та ін.

Якщо частину фізичних величин прийняти за основні й установити їхні одиниці вимірювання, тоді одиниці вимірювання інших величин можна виразити через одиниці вимірювання інших. Одиниці вимірювання основних величин називають ***основними,*** всі інші - ***похідними.***

Розмірність фізичної величини - це вираження, що відбиває зв'язок даної фізичної величини з основними величинами системи одиниць, у якому коефіцієнт пропорційності прийнятий рівним одиниці. Це вираження являє собою добуток основних величин, зведених у відповідні ступені. Основними механічними одиницями Міжнародної системи одиниць (СІ) є: грам, секунда, метр, градус Кельвіна.

Символ одиниці довжини позначають L, маси M, температури K, часу T і т. ін. (табл. 3.5) Розмірність похідної одиниці знаходять за допомогою визначеного рівняння, що виражає вимірювану величину в математичній формі. Наприклад, одиницю вимірювання швидкості визначають з рівняння

$$[\mathrm{V}] = \frac{[\mathrm{l}]}{[\mathrm{t}]} = \frac{\mathrm{L}}{\mathrm{T}} = \mathrm{LT}^{-1}$$

де l і t - відповідно довжина і час.

Означальним рівнянням сили є другий закон Ньютона

$$\boldsymbol{F = ma}.$$

Одиницю вимірювання сили визначають з рівняння

$$[F] = [m] \cdot [a] = MLT^{-2}.$$

Розмірність одиниці роботи визначають з рівняння

$$[A] = [F] \cdot [l] = ML^2T^{-2}.$$

Таблиця 3.5. Значення розмірностей механічних величин

Найменування величини	Позначення	Розмірність
Довжина	L	L
Об'єм	V	L^3
Кривизна	G	L^{-1}
Швидкість	V	LT^{-1}
Прискорення	G	LT^{-2}
Кутова швидкість	ω	T^{-1}
Щільність	ρ	ML^{-3}
Кількість руху	mV	MLT^{-1}
Момент кількості руху	mVR	ML^2T^{-1}
Сила	F	MLT^{-2}
Робота и енергія	A	ML^2T^{-2}
Потужність	N	ML^2T^{-3}
Кінематична в'язкість	ν	L^2T^{-1}
Поверхневий натяг	σ	MT^{-2}
Тиск	P	$ML^{-1}T^{-2}$
В'язкість	μ	$ML^{-1}T^{-1}$

Примітка: M — маса; L — довжина; T — час.

Якщо на об'єкт дослідження діє кілька факторів, тоді з них можуть бути створені безрозмірні комбінації - співвідношення параметрів об'єкта дослідження, складеш так, що одиниці вимірювання вхідних у неї величин скорочуються. Для одержання безрозмірних комбінацій використовують теорему Букінгема: *1) якщо яке-небудь рівняння однорідне відносно розмірностей, тоді його можна перетворити до співвідношення, що містить набір безрозмірних комбінацій величин; 2) якщо існує однозначне співвідношення*

$$\varphi(y, x_1, x_2, ..., x_{n-1}) = 0$$

між n фізичними величинами, для описування яких використовують k основних одиниць, тоді існує також співвідношення

$$\varphi'(\pi_1, \pi_2, ..., \pi_{n-k}) = 0$$

між (n - k) безрозмірними комбінаціями, складеними з цих фізичних величин.

Однорідним відносно розмірностей називають рівняння, форма якого не залежить від вибору основних одиниць вимірювання. Наприклад, запис другого закону Ньютона не залежить від одиниць вимірювання маси тіла, що рухається.

Процес одержання безрозмірних комбінацій розглянемо на прикладі аналізу роботи пневматичного вимірювального пристрою для контролю товщини плоских деталей (рис. 3.2). У даному пристрої вихідним для проведення вимірювальних процедур є тиск повітря p_0. Повітря через дросель із прохідним отвором d_2 надходить у вимірювальну камеру, з якої виходить через кільцевий зазор, утворений вихідним дроселем діаметром d_1, а також зазором між ним і вимірюваною деталлю h. При роботі приладу тиск p_0 постійний, а тиск p_1 залежить від зазору. Величину p_1 вимірюють.

Для аналізу розмірностей потрібно знати характер і число перемінних об'єкта дослідження. Необхідно побудувати модель при таких допущеннях: повітря нестискаєме, тертя повітряних потоків у приладі зневажливо мале, рух повітря сталий, інерційні сили відсутні. Перемінні і їхні розмірності наведені у табл. 3.6.

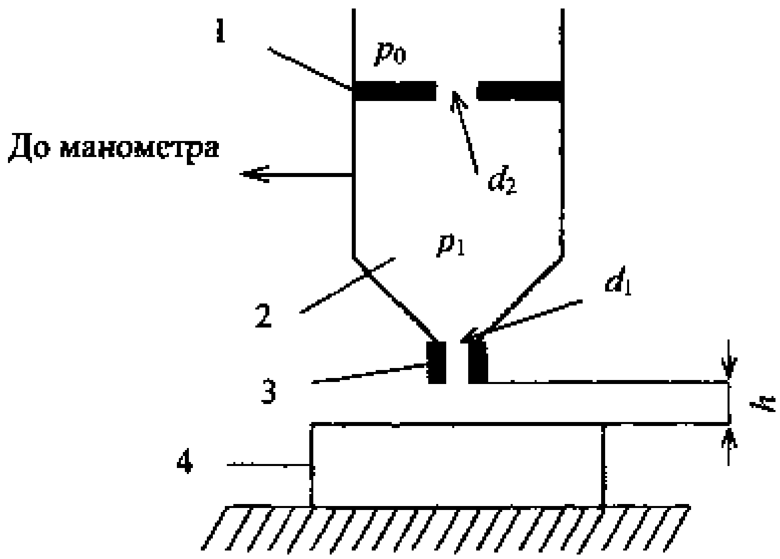

Рис. 3.2. Пневматичний товщиномір:
1 — вхідний дросель; 2 — вихідне сопло;
3 — вимірювальна камера; 4 — деталь

Таблиця 3.6. **Перемінні експерименту і їхні розмірності**

Найменування величини	Позначення	Розмірність
Тиск на манометрі	p_1	$ML^{-1}T^{-2}$
Постійний тиск	p_0	$ML^{-1}T^{-2}$
Діаметр вимірювального дроселя	d_1	L
Діаметр вхідного дроселя	d_2	L
Відстань між соплом і деталлю	h	L

Таким чином, відповідно до побудованої експериментальної моделі отримане вираження для визначення значення p_1

$$p_1 = \varphi(p_0, h, d_1, d_2)$$

У найбільш загальному випадку між перемінними існує співвідношення

$$p_1 = \varphi\left(p_0^a, h^b, d_1^c, d_2^c\right)$$

Розмірність перемінних може бути виражена за допомогою трьох основних одиниць: маси M, довжини L і часу T. Підстановка їх у розглянуте співвідношення дає

$$ML^{-1}T^{-2} = \varphi\left[\left(ML^{-1}T^{-2}\right)^a, L^b, L^c, L^d\right]$$

Застосовуючи до останнього вираження першу частину теореми Букінгема, тобто вимагаючи, щоб вираження було однорідним відносно розмірностей, зрівнюємо показники ступеня в правій і лівій частинах:

$$\begin{array}{ll} \textbf{для } \boldsymbol{M} & \left\{ \begin{array}{l} 1 = a; \\ -1 = -a + b + c + d; \\ -2 = -2a. \end{array} \right. \\ \textbf{для } \boldsymbol{L} & \\ \textbf{для } \boldsymbol{T} & \end{array}$$

Таким чином, ми одержимо три рівняння з чотирма невідомими. Виключаючи із отриманої системи рівнянь величини *a* і с, маємо

$$1 = a; \quad c = -b - d; \quad 1 = a;$$

У результаті

$$p_1 = \varphi\left(p_0^1, h^b, d_1^{-b-d}, d_2^d\right)$$

Об'єднавши члени з однаковими показниками ступеня, одержимо такі безрозмірні комбінації величин:

$$\frac{p_1}{p_0} = \varphi\left(\frac{h}{d_1}, \frac{d_2}{d_1}\right)$$

Виходить, замість первісних п'яти перемінних в експерименті можуть брати участь тільки три.

3.2. Методологія моделювання

Розробка моделей процесів і об'єктів починається з установлення цілей, тому що система - є засіб досягнення мети. Моделі того самого фізичного процесу можуть мати мало спільного, якщо вони розроблялись з різною метою.

Цілі, що ставить перед собою людина, рідко досяжні тільки за рахунок її власних можливостей чи зовнішніх засобів, що є у неї в даний момент. Такий збіг обставин називається *проблемною ситуацією.*

Проблема - це сукупність складних теоретичних і практичних завдань, потреба у вирішенні яких назріла в суспільстві. Із соціально-психологічних позицій проблема - це відображення протиріччя між суспільною потребою в знанні і відомими шляхами його одержання, протиріччя між знанням і незнанням. Проблема виникає, коли людська практика натрапляє на труднощі чи навіть наштовхується на «неможливість» у досягненні мети. Проблема, як деяке негативне, небажане явище, пов'язана з іншими системами. Тому її вирішення позначається на стані цих систем. Отже, у будь-якій реальній проблемі знаходить відображення «клубок» взаємозалежних проблем. Даний «клубок проблем» називають ***проблематикою.***

Системне дослідження всякої проблеми починається з її розширення до проблематики, тобто знаходження системи проблем, суттєво пов'язаних із досліджуваною, без урахування якої вона не може бути вирішена.

Проблемність існуючого положення усвідомлюється в декілька «стадій»: від неясного відчуття, що «щось не так», до усвідомлення *потреби,* а потім до виявлення *проблеми* і, нарешті, до формулювання *мети.* ***Мета*** представляється людині як суб'єктивний образ (абстрактна модель) неіснуючого, але бажаного стану середовища, що вирішило б виниклу проблему.

В інженерній і науковій практиці момент постановки мети (формулювання технічного завдання) - один із найскладніших і найважливіших етапів створення систем. У процесі розробки моделей мета неодноразово уточнюється з багаторазовими змінами і доповненнями. Мета відповідає на запитання «навіщо потрібна система?». Формулювання мети - це одночасно і формулювання напрямку вирішення проблемної ситуації. Напрямків багато, а вибрати потрібно тільки один, дійсно правильний, а не удаваний таким.

При формулюванні мети слід уникати підміни мети засобами її досягнення. Не менш важливим є і виключення змішання цілей. Змішання цілей добре ілюструється прикладом, приведеним А. Вольштетером з роботи дорожньої поліції: «Батьки міста хотіли б зменшити число порушень правил дорожнього руху. Вони хотіли б також карати штрафом чи в'язницею якнайбільше порушників. Існує два добре відомих альтернативних способи досягнення цієї мети - поліцейська засідка чи відкрите патрулювання. Перший спосіб збільшує ймовірність піймання порушника. Другий - відбиває бажання до порушень. Якщо наша мета - максимізувати число покараних порушників чи збільшити міський бюджет за рахунок штрафів, то, очевидно, для цього краще підійде засідка. Якщо ж наша мета зменшити кількість дорожніх подій або числа спроб порушити правила (навіть, якщо самі такі спроби стануть більш успішними, тому що порушник буде діяти, точно знаючи про присутність поліцейського), то цілком може виявитись, що більш часта і відкрита присутність поліцейських, здатних негайно покарати порушника, спонукає водіїв до обережності і тим самим мета буде досягатись якнайкраще».

Після встановлення мети визначають параметри моделі і критерії моделювання: область дії чи границі досліджуваного об'єкта або процесу; глибину деталізації, прийняті допущення, фізичні обмеження й обмеження, що накладаються умовами безпеки; необхідну точність; змінні величини стану і відомі перемінні управління; збудження та інші некеровані перемінні.

Критерій - це подоба мети, її апроксимація, модель. Визначення значення критерію для даної альтернативи є непрямим виміром ступеня її придатності як засобу досягнення мети.

При вивченні реальних об'єктів створюють моделі, в яких відбиті тільки головні, істотні зв'язки оригіналу, використовують ті чи інші припущення. Припущення приймають для відсівання факторів, якими можна зневажити без суттєвого перекручування умов завдання. Багато припущень приймають неусвідомлено. Не існує якого-небудь набору правил, що дозволяють швидко приймати припущення при вирішенні проблемних завдань. Дослідник намагається подумки побудувати деяку модель реального фізичного явища, що може дати результати, які мають смисл, і в той же час досить просту. Яке саме припущення слід приймати для цієї мети, значною мірою залежить від обмежень, обумовлених методом вирішення завдання.

Якщо потрібно знати лише порядок величин, то підійде одна група припущень. Якщо ж потрібно отримати точність 1%, то правильним буде вибір іншої групи припущень. Тут слід керуватись здоровим глуздом і накопиченим досвідом. Помилки, обумовлені припущеннями, можуть згодом спливати і переслідувати дослідника на всьому протязі подальшої роботи. Уникнути цієї пастки можна лише шляхом ретельного і повного урахування усіх факторів при побудові моделі і формулюванні припущень. Потрібно глибоко зрозуміти досліджуване явище, усвідомити його суттєві риси. Нарешті, при побудові моделі повинні використовуватись найбільш загальні принципи і закони. Це дозволяє врахувати всі припущення, прийняті при одержанні формул і виражень, і точно визначити область їхнього застосування.

Часто припущення приймають на основі наближених розрахунків, що мають назву оцінки порядку величини. У математиці звичайно зневажають членами так званих високих порядків.

Суттєвим у розробці моделей є відбір факторів, які визначають об'єкт дослідження та його стан. Якщо до початку дослідження відсутні які-небудь теоретичні проробки, тоді об'єкт дослідження розглядається як «чорна скринька». У цьому випадку при відборі факторів дотримуються деяких правил.

Фактори мають бути визначені операційно. Це еквівалентно вказівці способу встановлення будь-якого можливого значення фактора за допомогою ручки настройки відповідного регулятора. Кожен фактор повинен мати область визначення. Останню встановлюють принциповими і технічними обмеженнями.

Прикладом принципового обмеження може бути абсолютний нуль температури в звичайних термодинамічних системах. Прикладом технічного обмеження може бути температура плавлення матеріалу, з якого виготовляється модель об'єкта.

До точності фіксування значень факторів не пред'являють особливих вимог. Важливо тільки, щоб вона була відома. Але, чим вище точність, тим швидше будуть отримані результати з заданою вірогідністю. Слід вказати на дві найважливіші вимоги, що пред'являють до сукупності факторів. Це вимога відсутності кореляції між будь-якими двома факторами і вимога сумісності факторів. Вимога некорельованості не означає, що між факторами не має бути ніякого зв'язку. Досить, щоб цей зв'язок не був лінійним. Несумісність факторів виникає тоді, коли деякі комбінації їхніх значень не можуть бути здійснені. Може, наприклад, виявитись, що деяка концентрація речовин приводить при їхньому зміщенні до вибуху.

Таким чином, фактори мають бути *визначені операційно, для них вказують область визначення, точність вимірювання і підтримки. Крім того вони мають бути некорельованими і сумісними.*

Якщо виникає необхідність вивчити в статичній формі динамічні особливості об'єкта дослідження, використовують складні фактори - критерії подоби.

При складанні списку факторів слід врахувати всі можливі фактори, як би велике не було їхнє число. Краще включити кілька зайвих несуттєвих факторів, ніж пропустити один суттєвий. Подальший відбір факторів здійснюють експертним методом або за результатами попереднього пошукового експерименту.

Процедура відбору факторів експертним методом зводиться до наступного: широкому колу фахівців, що працюють в області дослідження, пропонують розташувати фактори в порядку убування ступеня їхнього впливу на вихідні характеристики об'єкта, проран-жувати фактори. Кожен експерт може включати додаткові фактори, якщо список, на його думку, неповний. Результати опитування представляють у вигляді матриці і діаграми рангів. Матрицю рангів використовують для обчислення коефіцієнта конкордації, що характеризує ступінь узгодженості думок експертів. Коефіцієнт конкордації розраховують за формулою

$$W = \frac{12S}{m^2\left(n^3 - n\right)}$$

де W - коефіцієнт конкордації; S - сума квадратів відхилень суми рангів по кожному фактору від середньої суми; m - число експертів; n - число факторів, які оцінюють.

Часто експерти не можуть вказати порядок проходження для двох чи декількох факторів, що стоять поруч. У такому разі їм приписують той самий номер, а при обчисленнях вводять так звані дробові ранги. Нехай, наприклад, експерт не може віддати перевагу другому і третьому факторам. Тоді кожному з них надають ранг 2,5. Коефіцієнт конкордації обчислюють за формулою

$$W = \frac{S}{\frac{1}{12}m^2\left(n^3 - n\right) - m\sum_i T_i}$$

де $T_i = \frac{1}{12}\sum_j \left(t_j^3 - t_j\right)$; t_j - j-е число однакових рангів у i-тому ранжуванні.

Коефіцієнт конкордації змінюється від 0 до 1. На діаграмі рангів фактори розміщають в послідовності, що погоджується із сумою рангів, отриманих фактором (рис. 3.3).

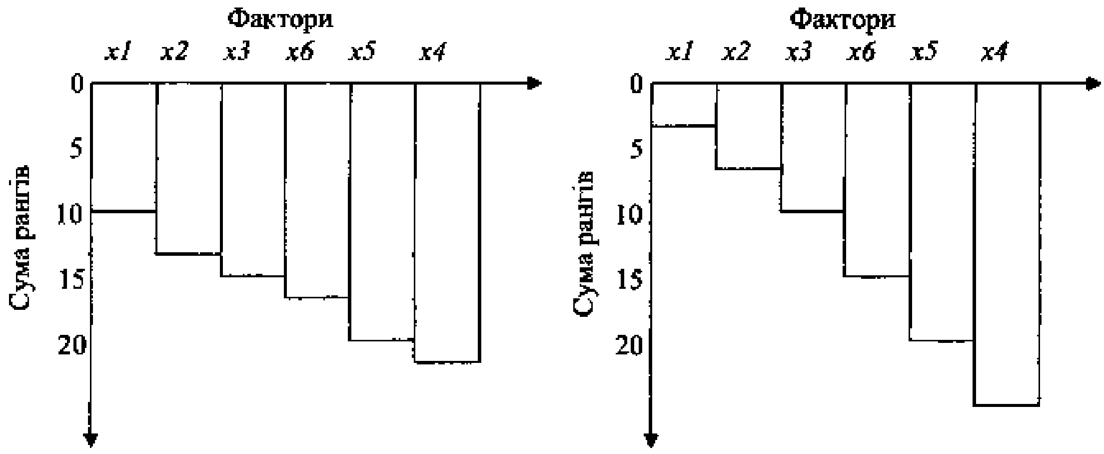

Рис. 3.3. Діаграми рангів

На осі ординат діаграми виносять суму рангів, на осі абсцис - фактори. На перше місце ставлять фактор, що отримав найменшу суму рангів, на останнє - фактор з найбільшою сумою.

З вигляду діаграми рангів приймають рішення про відсівання факторів. Якщо суми рангів факторів на діаграмі розподілені рівномірно, тоді усі фактори мають бути включені в модель і подальший фізичний експеримент. Якщо розподіл нерівномірний, тоді виникає питання: чи є падіння ступеня впливу ефектом швидким або повільним, плавним? При швидкому експонентному падінні ступеня впливу відсівають фактори за точкою перегину експоненти. При плавному лінійному падінні ступеня впливу відсівають фактори, сума рангів яких різко відхиляється від вирівнюючої лінії. При великому числі факторів, що залишилися, їхнє подальше відсівання здійснюють експериментально (методом випадкового балансу, дисперсійного аналізу, факторного аналізу та ін.).

Базою для розробки моделей фізичних процесів є закони фізики і хімії. Трактування основних законів фізики і хімії повинні виходити з єдиного підходу, у ролі якого служить системний аналіз.

Основною операцією аналізу є поділ цілого на частини. Задача розпадається на підзадачи, система - на підсистеми, цілі - на підцілі. При необхідності цей процес повторюється, що приводить до ієрархічних деревоподібних структур.

На початковому етапі поділ досліджуваного об'єкта на частини виконує експерт. Як основу для розподілу експерт використовує деяку модель. Операція поділу (декомпозиції) представляється як зіставлення об'єкта аналізу з моделлю і виділення в ньому того, що відповідає елементам узятої моделі. Використану експертом модель називають ***моделлю-основою.***

У ролі моделей-основ на початковому етапі використовують моделі формальних типів: моделі «чорної скриньки», складу, структури, конструкції та ін. Прикладом формальної моделі є загальна схема діяльності людини, у якій відображені всі можливі компоненти і всі можливі зв'язки між ними (рис. 3.4). Іншим прикладом формальної моделі є модель організаційної системи (рис. 3.5). Звичайно формальна модель носить глобальний, загальний характер і відноситься до проблемовирішуючої ситуації. Формальна модель-основа - це укрупнене представлення досліджуваного об'єкта і його зв'язків з іншими системами і середовищем.

Формальна модель має бути повною, тому повинна враховувати якнайбільше зв'язків і відношень. У дослідженнях із штучного інтелекту формальні моделі називають *фреймами.*

Вибір формальної моделі лише підказує, якого типу має бути модель-основа. Надалі формальна модель повинна бути наповнена змістом, щоб стати основою для декомпозиції. Так, наприклад, у фреймовій моделі входів оргсистеми (рис. 3.5) потрібно визначити конкретно, що саме розуміють під «суттєвим середовищем», тобто взаємодія з якими реальними системами зовнішнього середовища має ввійти в основу.

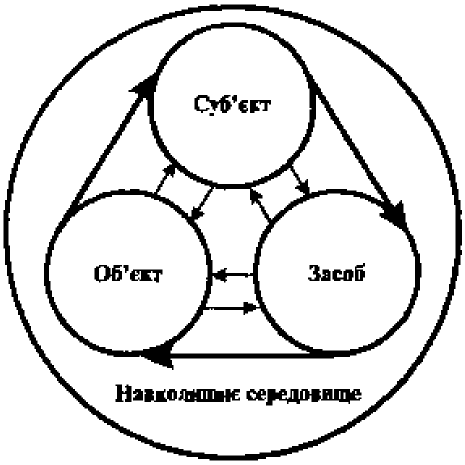

Рис. 3.4. **Загальна схема діяльності**

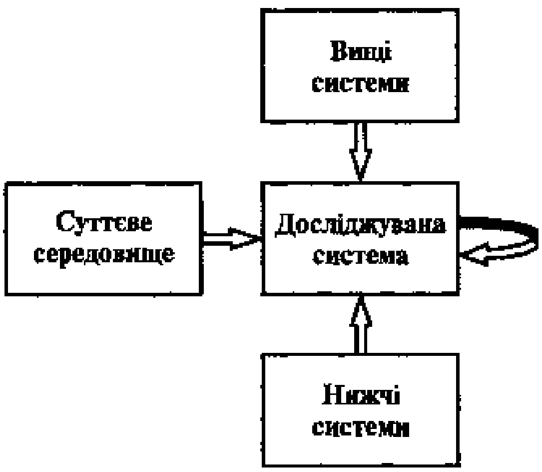

Рис. 3.5. **Схема входів організаційної системи**

Щоб зберегти повноту і можливість розширення змістовної моделі здійснюють логічне замикання переліку елементів моделі компонентою «все інше». Ця компонента, як правило, буде вважатись «той, що мовчить», тому що до неї віднесено усе, що здається несуттєвим, але її присутність буде постійно нагадувати експерту, що, можливо, він не врахував щось важливе.

Змістовна модель-основа має відповідати двом суперечливим вимогам. Вона повинна бути одночасно *простою* і *повною.* Вимога простоти змушує використовувати найбільш компактні моделі-основи. Навпаки, вимога повноти змушує брати найбільш розвинуті, докладні моделі.

Компроміс досягається за рахунок того, що в модель включають тільки компоненти і зв'язки, суттєві стосовно мети аналізу. Такі компоненти і зв'язки називають *релевантними.*

Вирішення питання про суттєвість покладається на експерта. Аналогічні вимоги пред'являються і до глибини деталізації моделі. *Декомпозиція* моделі (розклад на елементарні частини) закінчується, коли вона приводить до результату, що не вимагає подальшого розкладання, тобто результату простого, зрозумілого, реалізованого, забезпеченого, свідомо здійсненного. Такий результат називають елементарним. Поняття елементарності в деяких математичних задачах може бути конкретизовано до формальної ознаки. В інших задачах перевірка фрагментів на елементарність доручається експертам.

Неелементарний фрагмент підлягає подальшій декомпозиції за іншою моделлю-основою (що не використовувалася раніше). Нові моделі-основи одержують способом введення у початкові інші суттєві елементи через розщеплення вже наявних. Через те, що неелементарний фрагмент відноситься до деякого компонента моделі-основи, то розщепленню піддається саме цей компонент. На цій же стадії використовують і нові елементи з початкової компоненти «все інше».

Таким чином, компроміс між вимогою простоти і повноти змістовної моделі досягається через *суттєвість, елементарність, поступову деталізацію та ітеративність процесу декомпозиції* (рис. 3.6).

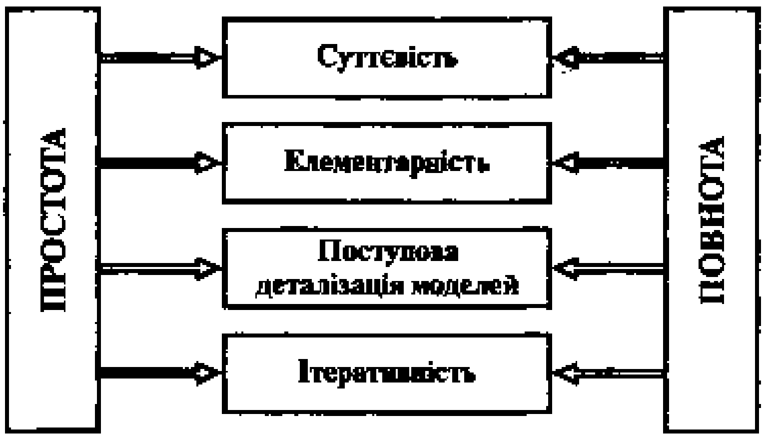

Рис. 3.6. Схема компромісів між вимогами простоти і повноти аналізу

Ітеративність процесу декомпозиції додає йому варіабельність, можливість користатись моделями різної детальності на різних галузях, поглиблювати деталізацію скільки завгодно (якщо це буде потрібно). Але, незважаючи на наявність ітеративності, може наступити момент, коли експерт визнає, що його компетентності недостатньо для подальшого аналізу. Тоді треба залучити іншого експерта. У висновку можна сказати, що процес декомпозиції не дає нових знань, а лише «витягає» знання з експертів, структурує й організує їх. Укрупнена схема алгоритму процесу декомпозиції представлена на рис. 3.7.

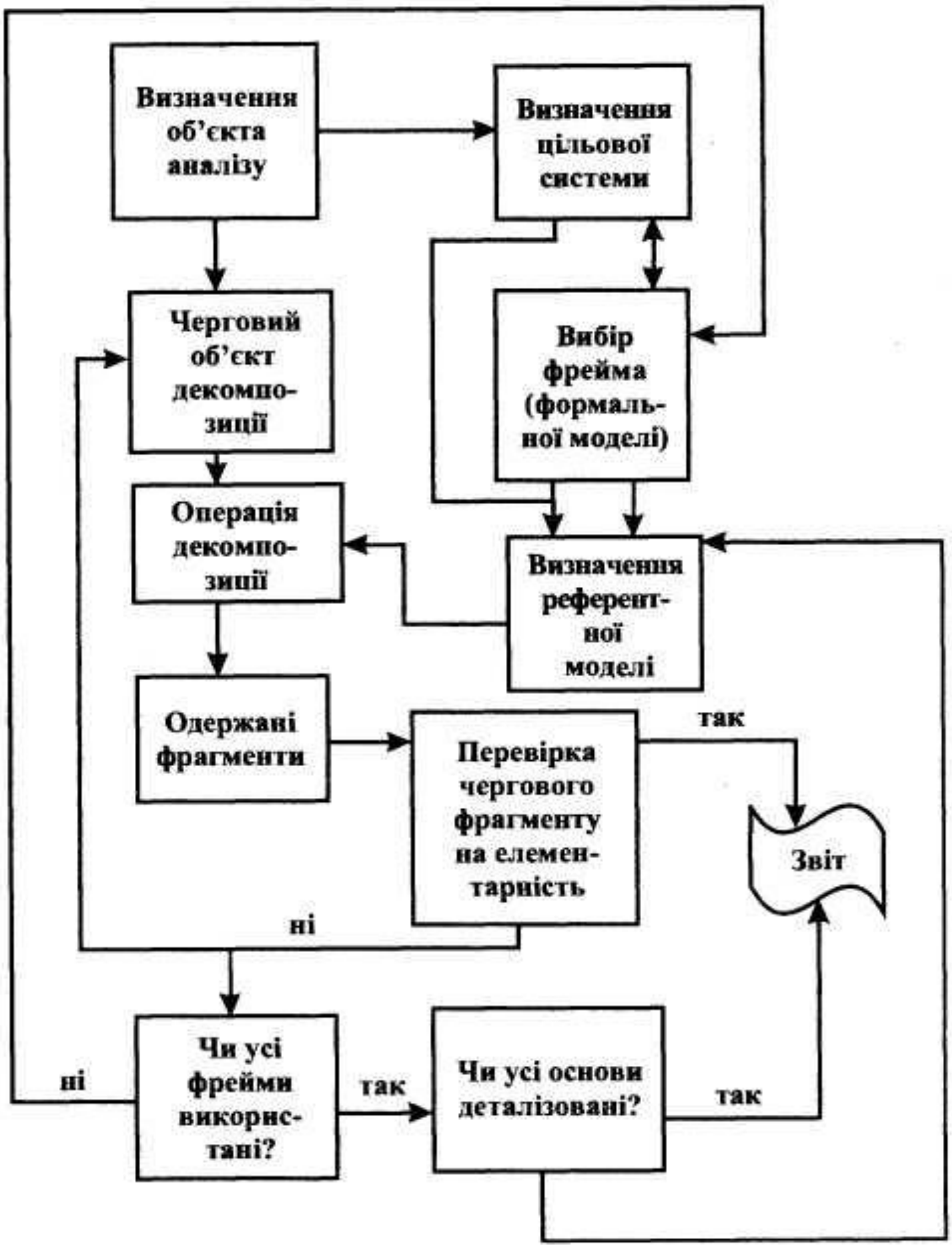

Рис. 3.7. **Укрупнена схема алгоритму декомпозиції**

У процесі декомпозиції систему розчленовують на усе менш складні (і кінець кінцем, прості) частини.

При цьому розкривають лише структуру системи, одержують знання про те, як система працює, але залишається відкритим питання чому і навіщо вона це робить. Ці знання одержують на другому етапі моделювання, що називають *агрегуванням* (об'єднанням частин у ціле).

Техніка агрегування ґрунтується на використанні моделей частин системи, отриманих на етапі декомпозиції. У процесі агрегування встановлюють відносини між цими моделями. Залежно від того які відносини встановлюються (тобто виявляються чи встановлюються) виходить досить велика кількісно і різноманітна якісно безліч задач агрегування.

Основними агрегатами, типовими для системного аналізу є: *конфігуратор, агрегати-оператори, агрегати-структури.*

Конфігуратор - це сукупність мов описування системи. Усякий процес вимагає різнобічного, багатопланового описування, розгляду з різних точок зору. Тільки спільне (агреговане) описування у термінах декількох якісно різних мов дозволяє охарактеризувати явище з достатньою повнотою. Наприклад, автомобільна катастрофа повинна розглядатись не тільки як фізичне явище, викликане технічним станом автомобіля чи дорожнього покриття, але і як явище медичного, соціального, економічного, юридичного характеру. Разом з тим, для описування системи використовують не нескінченне число мов, а мінімально необхідне для досягнення заданої мети. Наприклад, у радіотехніці для описування одного і того ж приладу використовується конфігуратор, що включає блок-схему, принципову схему і монтажну схему.

Назвавши мови, на яких будемо говорити про систему, ми тим самим визначаємо тип системи, фіксуємо наше розуміння природи системи.

Агрегати-оператори - конкретизація відносини, зокрема класифікація, упорядкування, числові функції, пошук закономірностей та ін. Даний агрегат поєднує частини у щось ціле, єдине, окреме. Найпростіший спосіб агрегування складається у встановленні відносини еквівалентності між агрегованими елементами, тобто утворення класів. У результаті є відповідь на запитання, до якого класу відноситься конкретний елемент.

Інший тип агрегату - оператора виникає, якщо агреговані ознаки фіксуються у числових шкалах. Тоді з'являється можливість задати відношення на безлічі ознак у вигляді числової функції багатьох перемінних, яка є агрегатом. Прикладом може бути перехід від багатокритеріальної оптимізаційної задачі до однокритеріальної за допомогою агрегування декількох критеріїв в один суперкритерій. Побудова супер-критеріальної функції, власне кажучи, є побудовою моделі системи.

Іншим прикладом агрегатів-операторів є *статистики,* тобто функції вибіркових значень. Статистичним агрегуванням є факторний аналіз, у якому декілька перемінних зводяться в один фактор.

Агрегати-структури - це описування зв'язків на всіх мовах конфігуратора. При синтезі ми створюємо, визначаємо, нав'язуємо структуру проектованій системі. Якщо це не абстрактна, а реальна система, то в ній цілком реально виникнуть, установляться і почнуть «працювати» не тільки ті зв'язки, що ми спроектували, але і безліч інших, не передбачених нами, що випливають із самої природи, зведених в одну систему елементів. Тому при проектуванні системи важливо задати її структури у всіх суттєвих відносинах, тому що в інших відносинах структури складуться самі, стихійним чином. Сукупність всіх суттєвих відносин визначається конфігуратором системи, і звідси виходить, що проект будь-якої системи має містити розробку стількох структур, скільки мов включено в її конфігуратор.

3.3 Математичне моделювання

Математична модель - це система математичних співвідношень - формул, функцій, рівнянь, систем рівнянь, що описують ті чи інші сторони досліджуваного об'єкта, явища, процесу.

Першим етапом математичного моделювання є *постановка задачі,* визначення об'єкта і мети дослідження, завдання критеріїв (ознак) вивчення об'єктів і управління ними. Неправильна чи неповна постановка задачі може звести нанівець результати всі наступних дій.

Дуже важливим на цьому етапі є встановлення границь області впливу досліджуваного об'єкта. Границі області впливу об'єкта визначають областю значимої взаємодії з зовнішніми об'єктами. Дана область може бути визначена на основі таких ознак: *границі області охоплюють ті елементи, вплив яких на досліджуваний об'єкт не дорівнює нулю; за границями області дія досліджуваного об'єкта на зовнішні об'єкти прагне до нуля.*

Врахування області впливу об'єкта при математичному моделюванні дозволяє включити в цю модель всі суттєві фактори і розглядати систему, що моделюють, як замкнуту, тобто з відомим ступенем наближення, незалежну від зовнішнього середовища. Останнє значно спрощує математичне дослідження. Загальна схема алгоритму першого етапу представлена на рис. 3.8.

Другим етапом моделювання є *вибір типу (класу)* математичної моделі. Для орієнтовного вибору типу математичної моделі системи можна використовувати класифікаційну діаграму (рис. 1.10). Це вимагає попередньої оцінки ступеня складності і ступеня детермінованості системи.

Ступінь складності системи може бути оцінений за величиною її максимальної невизначеності. Останню розраховують за формулою

$$H_m = \log_2 n$$

де H_m - максимальна невизначеність системи; *n* - число станів системи.

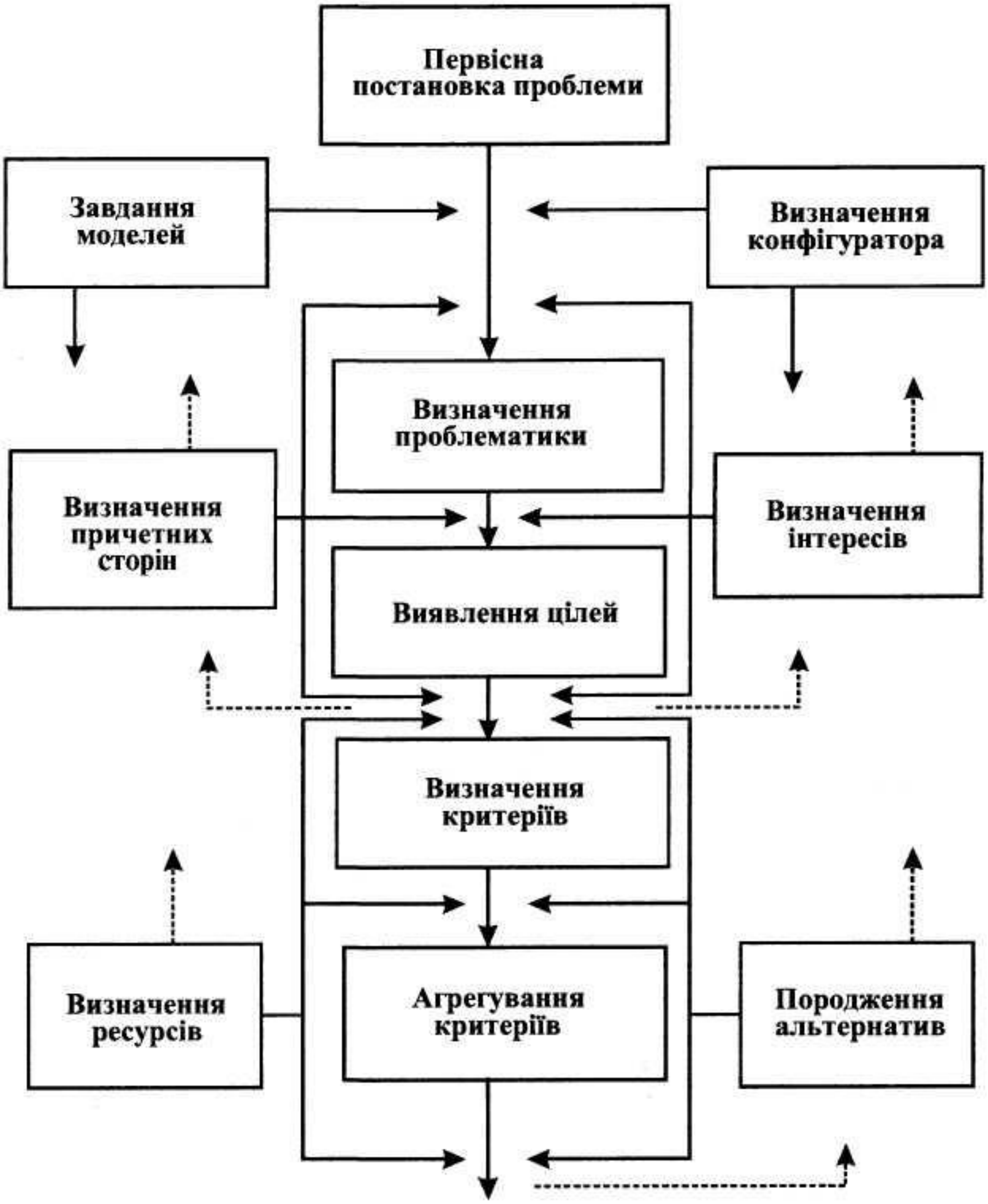

Рис. 3.8. **Загальна схема алгоритму першого етапу моделювання**

Якщо хочуть отримати оцінку максимальної невизначеності системи за деяким безперервним показником, тоді H_m визначають за формулою

$$H_m = \log_2 \frac{y_{max} - y_{min}}{\Delta y}$$

де y_{max}, y_{min} - максимальне і мінімальне значення параметра; Δy - точність вимірювання параметра.

Фактичну невизначеність системи розраховують за формулою К. Зенона

$$H = -\sum_{i=1}^{n} P_i \log_2 P_i$$

де P_i - частота появи і-того стану системи.

Якщо оцінюють стан системи за безперервним показником, фактичну невизначеність системи обчислюють за формулою

$$H = -\sum_{i=1}^{n} \frac{m_i}{N} \log_2 \frac{m_i}{N}$$

де m_i - число появ стану системи з і-тим показником за час спостереження; *N* - повне число спостережень (число значень показника, що спостерігається).

Ступінь детермінованості системи оцінюють за величиною відносної організації системи, що розраховують за формулою Г. Ферстера

$$R = 1 - \frac{H}{H_m}$$

Визначення типу системи за класифікаційною діаграмою дозволяє підібрати математичний апарат, який щонайкраще описує поведінку системи. Так, для описування простих і складних детермінованих систем (області 11, 21) цілком можливий апарат диференціальних рівнянь і теорії автоматичного управління. Для областей 12 і 22 діаграми можна скористатись нелінійними диференціальними рівняннями, параметричними рівняннями і рівняннями, коефіцієнти яких підкоряються деяким законам розподілу. Для областей 13, 23 застосовують теорію автоматів, теорію марковських і випадкових процесів, алгебраїчні і диференціальні рівняння для описування ймовірностей і параметрів законів розподілу. Для дуже складних імовірнісних систем можна використовувати ентропійне (інформаційне) моделювання.

Не менш важливим при встановленні типу моделі є визначення її лінійності чи нелінійності, динамічності чи статичності, стаціонарності чи нестаціонарності.

Лінійність моделі встановлюють за характером статичної характеристики досліджуваного об'єкта. Якщо статична характеристика об'єкта виявляється лінійною, то моделювання цього об'єкта здійснюють з використанням лінійних функцій. Нелінійність статичної характеристики і наявність запізнювання в реагуванні об'єкта на зовнішній вплив є чіткою ознакою нелінійності об'єкта. У цьому випадку для його моделювання має бути прийнята нелінійна математична модель.

Застосування лінійної математичної моделі значно спрощує її подальший аналіз, тому що така модель дозволяє скористатись принципом суперпозиції.

Установлення динамічності чи статичності здійснюють із поводження досліджуваних показників об'єкта в часі. Стосовно до детермінованих систем можна говорити про їх статичність чи динамічність з характеру вихідної характеристики системи. Якщо середнє арифметичне значення вихідного сигналу на різних відрізках часу не виходить за допустимі межі, обумовлені точністю методики виміру досліджуваного показника, то це свідчить про статичність об'єкта.

Стосовно до імовірнісних систем, їхню статичність установлюють за зміною рівня їх відносної організації. Якщо мінливість цього рівня не перевищує допустимі межі, тоді систему визначають як статичну.

Дуже важливим є вибір відрізків часу, на яких встановлюють статичність чи динамічність об'єкта. Якщо об'єкт на малих відрізках часу виявився статичним, тоді при збільшенні цих відрізків результат не зміниться. Якщо ж статичність установлена для великих відрізків часу, тоді при їхньому зменшенні результат може змінитись і статичність об'єкта може перейти в динамічність.

При виборі типу моделі імовірнісного об'єкта важливе встановлення його стаціонарності. Звичайно про стаціонарність чи нестаціонарність імовірнісних об'єктів судять за зміною в часі параметрів законів розподілу випадкових величин. Найчастіше для цього використовують середнє арифметичне випадкової величини $M(\tau_i)$ і середнє квадратичне відхилення випадкових величин σ_i від середнього арифметичного. При цьому досліджують зміну середнього арифметичного і середнього квадратичного відхилення в часі. З ряду середніх арифметичних $M(\tau_1)$, $M(\tau_2)$, ..., $M(\tau_i)$ вибирають мінімальне значення $M(\tau_{min})$ і будують інтервали з границями

$$M(\tau_{min}) + \Delta y, \qquad M(\tau_{min}) - \Delta y,$$

де Δy - точність методики вимірювання досліджуваного показника.

Якщо значення $M(\tau_i)$ укладається в цей інтервал, тоді об'єкт визначають як стаціонарний за середнім арифметичним $M(\tau_i)$. Аналогічно визначають стаціонарність за середнім квадратичним відхиленням. Граничні значення при встановленні стаціонарності визначають за формулами

$$\sigma_1 = \left(\frac{D_1}{n}\right)^{\frac{1}{2}}, \quad \sigma_2 = \left(\frac{D_2}{n}\right)^{\frac{1}{2}},$$

$$D_1 = \left\{\sum_{i=1}^{n}\left[y_i - \left(M(\tau_{min}) + \Delta y\right)\right]^2\right\}\frac{1}{n-1};$$

$$D_1 = \left\{ \sum_{i=1}^{n} \left[y_i - \left(M(\tau_{min}) - \Delta y \right) \right]^2 \right\} \frac{1}{n-1};$$

n - число спостережень.

Якщо всі значення σ укладаються в інтервал $\sigma_1 + \sigma_2$, тоді об'єкт вважають стаціонарним. В іншому випадку об'єкт визначають як нестаціонарний, навіть якщо величина середнього арифметичного *M* не змінюється в часі.

Ознакою нестаціонарносгі об'єкта є також зміна в часі рівня відносної організації R. Але при цьому важко вибрати інтервал стаціонарності для R, який би враховував точність вимірювання показника.

Процес установлення загальних характеристик системи, що моделюють, може бути представлений у вигляді такого алгоритму:

1. Ідентифікація детермінізму.

- ранжування усього початкового інформаційного масиву (вимірюваних показників);
- розбивання ранжувального ряду на розряди з точністю вимірювання показника;
- визначення числа станів за формулою

$$n = E\left(\frac{y_{max} - y_{min}}{\Delta y} \right) + 1$$

де n - число станів; *E* - ціла частина дробу. (При знаходженні числа станів значення показника, що потрапили на границю розряду, відносять до наступного розряду.)

- оцінка складності системи;
- визначення частот появи вимірюваних показників у кожному розрядів;
- оцінка фактичної ентропії системи;
- оцінка відносної організації системи;
- ідентифікація детермінованості системи за класифікаційною діаграмою.

2. Ідентифікація динамічності детермінованих систем.

- обчислення середнього арифметичного для різних розрізів за часом;
- визначення границь інтервалів зміни середнього арифметичного;
- ідентифікація властивостей статичності чи динамічності.

3. Ідентифікація нестаціонарності системи.

- обчислення середнього арифметичного для різних розрізів за часом;
- визначення дисперсії і середнього квадратичного відхилення для тих же моментів часу;
- розрахунок границь середнього квадратичного відхилення;
- ідентифікація властивості стаціонарності чи не стаціонарності системи.

Установлення загальних характеристик системи дозволяє вибрати ***математичний апарат***, на базі якого будують математичну модель. Вибір математичного апарата може бути здійснений відповідно за схемою, представленою на рис. 3.9.

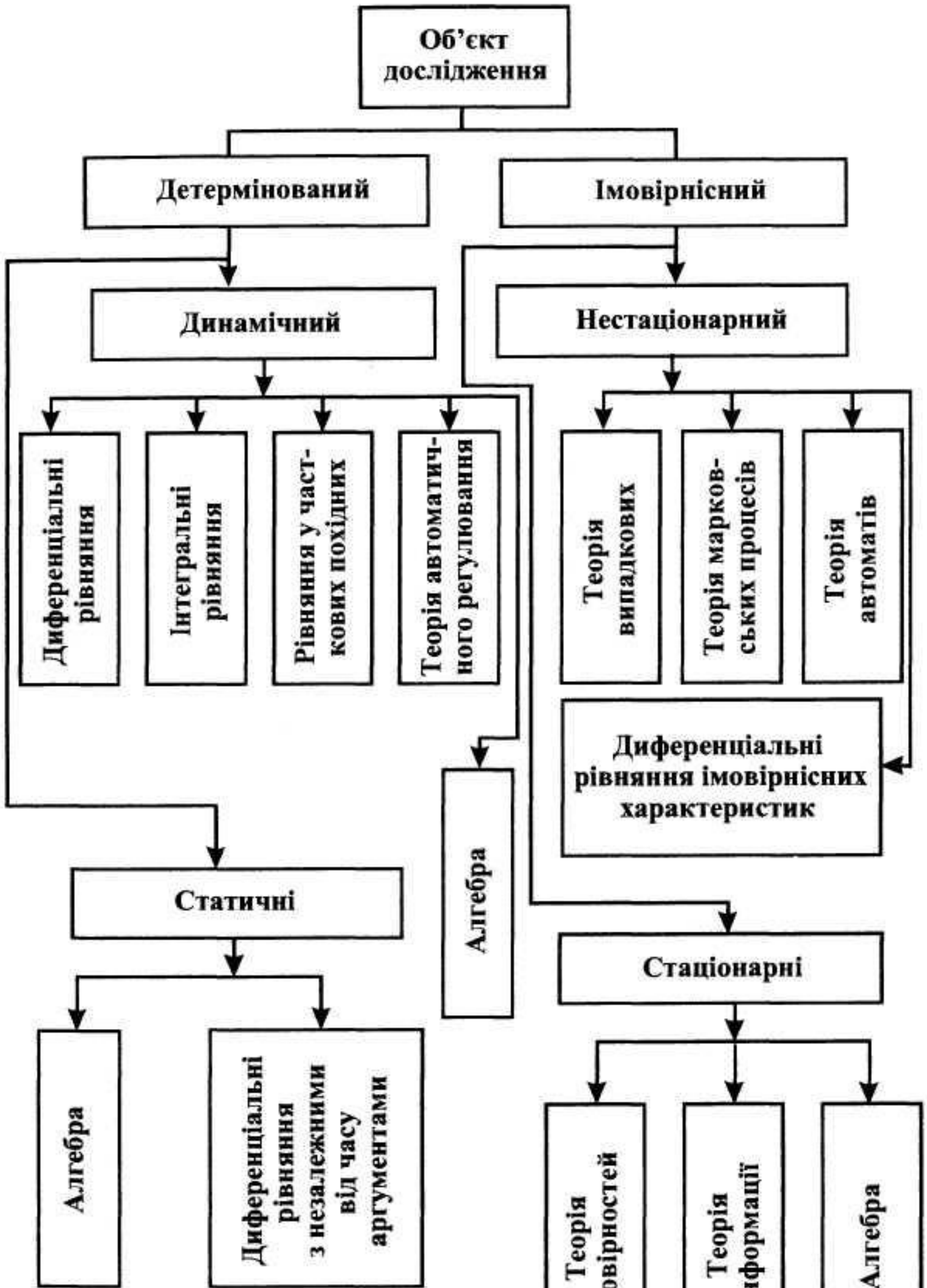

Рис. 3.9. **Математичний апарат для побудови математичної моделі**

Як видно з даної схеми, вибір математичного апарата не є однозначним і жорстким. Так, для детермінованих об'єктів може використовуватись апарат лінійної і нелінійної алгебри, теорії диференціальних і інтегральних рівнянь, теорії автоматичного регулювання. Адекватним математичним апаратом для моделювання імовірнісних об'єктів є теорія детермінованих і випадкових автоматів з детермінованими і випадковими середовищами, теорія випадкових процесів, теорія марковських процесів, евристичне програмування, методи теорії інформації, методи теорії управління та оптимізації.

При описуванні квазідетермінованих (імовірнісно-детермінованих) систем може використовуватись теорія диференціальних рівнянь з коефіцієнтами, що підкоряються деяким законам.

Мета і задачі, що ставлять при математичному моделюванні, відіграють важливу роль при виборі типу моделі. Практичні задачі вимагають простого математичного апарата, а фундаментальні - більш складного, допускають проходження ієрархії математичних моделей, починаючи від чисто функціональних і кінчаючи моделями, що використовують твердо встановлені закономірності і структурні параметри.

Не менше значення для вибору типу моделі має аналіз інформаційного масиву, отриманий як результат аналітичного огляду результатів досліджень інших авторів чи пошукового експерименту.

Розподіл масиву на залежні і незалежні фактори, на вхідні і вихідні перемінні, попередній пошук взаємозв'язку між різними даними вибірки дозволяє визначити адекватний математичний апарат. Аналіз інформаційного масиву дозволяє встановити безперервність чи дискретність досліджуваного показника і системи в цілому. У безперервних системах усі сигнали є безперервними функціями часу. У дискретних системах усі сигнали квантуються за часом і амплітудою. Якщо сигнали квантуються тільки за часом, тобто представляються у вигляді імпульсів з рівною амплітудою, тоді такі системи називають дискретно-безперервними.

Встановлення безперервності систем дозволяє використовувати для моделювання диференціальні рівняння. У свою чергу, дискретність системи визначає використання для математичного моделювання апарата теорії автоматів.

На встановлення типу (класу) моделі може вплинути необхідність відображення гіпотези. Врахування мети, задач математичного моделювання, характеру гіпотези та аналіз інформаційного масиву дозволяють конкретизувати модель, тобто в обраному типі (класі) моделей визначити їхній вид.

Третім етапом математичного моделювання *є вибір виду математичної моделі* в даному їхньому класі. Даний етап пов'язаний із завданням областей визначення досліджуваних параметрів об'єкта, тобто значень, що є допустимими, і встановленням залежностей між ними. Для кількісних (числових) параметрів залежності задають у вигляді систем рівнянь (алгебраїчних чи диференціальних рівнянь), для якісних - використовують табличні способи завдань функцій.

Якщо параметри описують суперечливими залежностями, тоді визначають їхні вагові коефіцієнти, виражені в частках одиниці, балах. Тим самим суперечливі залежності переводять у імовірнісні.

Для описування складних об'єктів з великою кількістю параметрів можна розбивати об'єкт на елементи (підсистеми), встановлювати ієрархії елементів і описувати зв'язки між ними на різних рівнях ієрархії.

Особливе місце на етапі вибору виду математичної моделі займає описування перетворення вхідних сигналів у вихідні характеристики об'єкта. Якщо на попередньому етапі встановлено, що об'єкт є статичним, тоді побудову функціональної моделі здійснюють за допомогою алгебраїчних рівнянь. При цьому крім найпростіших алгебраїчних залежностей використовують регресійні моделі і системи алгебраїчних рівнянь. Якщо характер зміни досліджуваного показника відомий заздалегідь, тоді число можливих структур алгебраїчних моделей різко скорочують і перевагу віддають тій структурі, що виражає найбільш загальну закономірність чи загальновідомий закон.

Якщо характер зміни досліджуваного показника заздалегідь невідомий, тоді ставлять пошуковий експеримент. Перевагу віддають тій математичній формулі, що дає найкращий збіг з даними пошукового експерименту. Результати пошукового експерименту та апріорний інформаційний масив дозволяють установити схему взаємодії об'єкта із зовнішнім середовищем за співвідношенням вхідних і вихідних величин (рис. 1.3).

При одномірно-одномірній взаємодії статичного стаціонарного детермінованого об'єкта із зовнішнім середовищем постійний вхідний вплив позв'язується з постійним вихідним сигналом через постійний коефіцієнт. Якщо ж цей об'єкт є нестаціонарним, тоді зазначений зв'язок описують різними функціями $y=f(x)$. Найчастіше дана функція описується поліномом.

У випадку багатомірно-одномірної схеми статичний стаціонарний детермінований об'єкт описують такою моделлю:

- при рівнозначності зовнішніх впливів

$$y = a\sum_{i=1}^{m} x_i$$

- при нерівнозначності зовнішніх впливів

$$y = \sum_{i=1}^{m} a_i x_i$$

де a_i - коефіцієнт; m - число зовнішніх впливів (факторів).

Для статичного нестаціонарного об'єкта (при тій же схемі взаємодії) часто використовують модель у вигляді повного ступінного полінома

$$y = a_0 + \sum_{i=1}^{m} a_i x_i + \sum_{i=1}^{m_1}\sum_{j=1}^{m_1} a_{ij} x_i x_j + \sum_{i=1}^{m_2}\sum_{j=1}^{m_2}\sum_{\gamma=1}^{m_2} a_{ij\gamma} x_i x_j x_\gamma + \ldots;$$

де m_1, m_2 - число парних і потрійних сполучень факторів $\left(m_1 = C_m^2\right)$, $\left(m_2 = C_m^3\right)$.

При одномірно-багатомірній схемі статичний стаціонарний і нестаціонарний об'єкт описують аналогічно одномірно-одномірній схемі взаємодії статичного стаціонарного об'єкта із зовнішнім середовищем. При цьому визначають окремо математичні моделі вхідного впливу з кожним вихідним сигналом. Вихідні сигнали вважаються незалежними.

Багатомірно-багатомірні взаємодії зводять до багатомірно-одно-мірних і математичну модель об'єкта приймають аналогічною викладеній вище. Для нестаціонарної одномірно-одномірної (багатомірної) взаємодії алгебраїчні функції можуть бути рішеннями диференціальних рівнянь. При цьому необхідно розглядати похідні математичного очікування по перемінному фактору.

Наприклад, експонентна залежність може бути рішенням такого диференціального рівняння

$$\frac{d\overline{y}}{dx} + a\overline{y} = a\overline{y}^m, \qquad \left(y = 0 \quad \text{при} \qquad x = 0\right)$$

де $\overline{y}$ - максимальне значення математичного очікування.

Вибір виду моделі динамічного об'єкта зводиться до складання диференціальних рівнянь. Щоправда, моделі динамічного об'єкта можуть бути побудовані й у класі алгебраїчних функцій. Але такий підхід є обмеженим, тому що не дозволяє в математичному описуванні врахувати вплив вхідних дій на динаміку виходу без суттєвої перебудови самих алгебраїчних функцій (структури і коефіцієнтів). Тому за повнотою моделі віддається перевага математичним моделям, побудованим у класі диференціальних рівнянь. Слід зазначити, що диференціальні рівняння особливо зручні для описування систем із плавно мінливими перемінними, але їх не можна застосовувати безпосередньо для дослідження систем з суттєво дискретними станами.

Якщо перемінні, що цікавлять дослідника, є функціями тільки часу, тоді для моделювання використовують звичайні диференціальні рівняння. Якщо ж ці перемінні є функціями не тільки часу, але і просторових координат, тоді для описування таких об'єктів недостатньо звичайних диференціальних рівнянь і слід використовувати більш складні диференціальні рівняння в часткових похідних.

Методологія моделювання динамічних систем у класі диференціальних рівнянь суттєво залежить від схеми взаємодії об'єкта із середовищем і ступеня знання входу і виходу об'єкта.

Розглянемо випадок, коли вхід і вихід об'єкта відомі. При одномірно-одномірній і одномірно-багатомірній взаємодії детермінованого об'єкта із середовищем структура диференціального рівняння визначається із виду вихідної характеристики об'єкта для типового вхідного впливу. Якщо вхідний вплив є ступінчастим, то однією з найбільш простих вихідних характеристик може бути лінійна (рис. 3.10, *а)*. Така зміна виходу системи визначається рішенням диференціального рівняння

$$\frac{dy}{dt} = kx, \qquad y(0) = 0.$$

де k>0 - коефіцієнт розмірності і пропорційності; y(0) - початкове значення вихідного сигналу; t - час.

Якщо $y(0) \neq 0$, тоді вихідна характеристика об'єкта відповідає схемі, представленої на рис. 3.10, *б.* Але диференціальне рівняння залишається незмінним.

Більш складний вид реакції об'єкта на ступінчастий вхідний вплив (рис. 3.10, *в*) може бути описаний повним неоднорідним диференціальним рівнянням першого порядку:

$$\frac{dy}{dt} + a_0 y = kx, \qquad y(0) = y_0,$$

де a_0 - коефіцієнт.

Реакція системи, що відповідає рис. 3.10, г, дозволяє використовувати диференціальне рівняння другого порядку як математичну модель об'єкта

$$\frac{d^2y}{dt^2} + a_1 \frac{dy}{dt} + a_0 y = kx.$$

Розглянуті види математичних моделей відповідають постійному вхідному впливу *(x = const).*

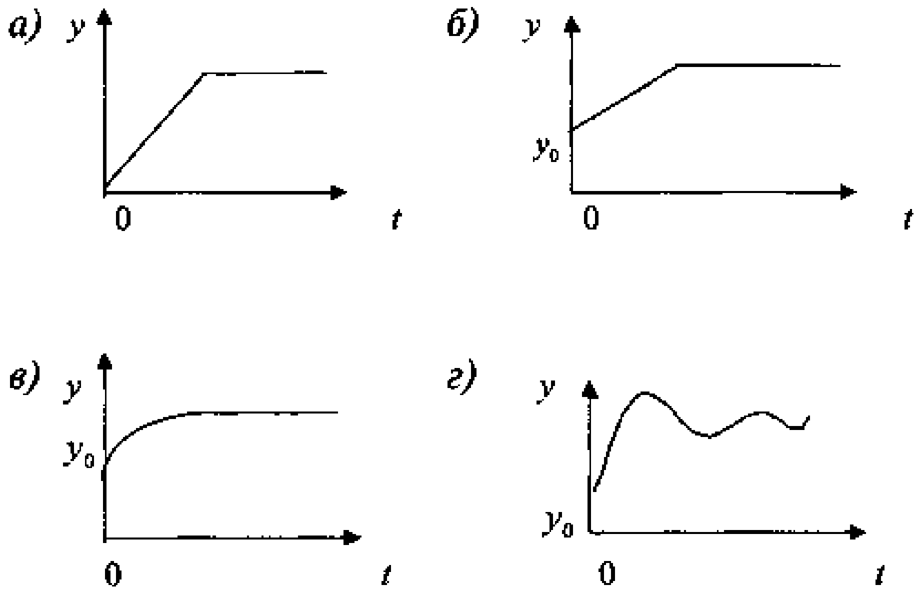

Рис. 3.10. Вихідні характеристики детермінованого об'єкта при ступінчастому зовнішньому впливі

Якщо вхідні впливи є деякими функціями часу, зокрема алгебраїчними, тоді в приведених диференціальних рівняннях змінюються праві частини $x=f(t)$.

При багатомірно-одномірній взаємодії у випадку детермінованого об'єкта динамічні моделі також знаходять у класі диференціальних рівнянь. При цьому допускається, що вхідні фактори є незалежними за їхньою дією на об'єкт. Усі фактори приводять до суми коефіцієнтами чутливості в правій частині диференціального рівняння. Диференціальне рівняння підбирають за видом вихідної характеристики об'єкта.

Якщо припущення про незалежність дії факторів неправомірне, тоді попередньо встановлюють вплив кожного з факторів на вихідний показник об'єкта і за видом вихідних характеристик підбирають відповідні диференціальні рівняння. При одночасній дії факторів покладають, що вихідна характеристика об'єкта - це сума рішень незалежних диференціальних рівнянь, які відповідають кожному фактору.

Для багатомірно-багатомірній взаємодії побудова динамічних моделей може бути зведена до багатомірно-одномірної схеми взаємодії.

При відсутності апріорної інформації про вхід і вихід об'єкта диференціальні рівняння, що моделюють динаміку об'єкта, складають на основі припущень чи знань про властивості, структуру та ін. об'єкта.

Універсального методу складання диференціальних рівнянь не має, можна лише використовувати деякі загальні підходи до складання рівнянь першого порядку.

Геометричні чи фізичні задачі звичайно приводять до одного з трьох видів рівнянь:

a) *диференціальні рівняння в диференціалах;*

b) *диференціальні рівняння в похідних;*

c) *найпростіші інтегральні рівняння з наступним перетворенням їх у диференціальні рівняння.*

Розглянемо, як складають рівняння кожного з приведених видів окремо.

1. Рівняння в диференціалах. При складанні диференціальних рівнянь першого порядку зручно застосовувати «метод диференціалів». Сутність його полягає в тому, що з умови задачі складають наближене співвідношення між диференціалами. Для цього малі збільшення величин заміняють їхніми диференціалами; фізичні процеси, що нерівномірно протікають протягом малого проміжку часу Л, розглядають як рівномірні. Ці припущення і заміни не відбиваються на остаточних результатах, через те, що заміна збільшень диференціалами зводиться до відкидання нескінченно малих величин вищих порядків малості. Через те, що відношення диференціалів функції й аргументу є границею відношення їхніх збільшень, то, як збільшення прагнуть до нуля, і прийняті припущення виконуються з більшою точністю. Диференціальні рівняння, що отримують при цьому, виявляються точними, якщо вони однорідні і лінійні щодо диференціалів.

2. Рівняння в похідних. Для складання диференціальних рівнянь використовують видозмінений метод диференціалів, який називають методом похідних.

Сутність методу похідних полягає в тому, що з умови задачі складають наближені співвідношення між швидкостями зміни функції й аргументу. При цьому часто використовують геометричну інтерпретацію швидкості - кутовий коефіцієнт дотичної. Відсутність нескінченно малих у методі похідних удаване, оскільки швидкість зміни досліджуваної величини сама виявляється із розглядання нескінченно малих елементів.

3.Найпростіші інтегральні рівняння. Роботу сил, об'єми тіл, площу криволінійних поверхонь та ін. можна описувати за допомогою визначеного інтеграла чи інтегральних формул. Якщо при такому описуванні невідомі функції підпадають під знак інтеграла, тоді одержаний формальний запис називають інтегральним рівнянням. Наступне диференціювання інтегрального рівняння перетворює його в диференціальне.

При складанні диференціальних рівнянь регульованих об'єктів необхідно насамперед визначити умови одержання рівноважного режиму роботи об'єкта, тобто рівняння статичної рівноваги. У багатьох випадках рівняння статичної рівноваги виявляється загальним для різних об'єктів дослідження. Наприклад, при поступальному русі досліджуваний об'єкт буде знаходитись у стані статичної рівноваги (рух буде рівномірним) тільки в тому випадку, коли рушійні сили $F_д$ дорівнюють силам опору F_c. Рівняння статичної рівноваги прийме вигляд

$$F_д - F_c = 0$$

У тому випадку, коли об'єкт дослідження (колінчатий вал двигуна, ротори електродвигуна, генератора, турбіни та ін.) робить обертальний рух, то умовою статичної рівноваги є рівність моменту, що крутить, $M_д$ і моменту опору M_c, тобто

$$M_д - M_c = 0$$

Для регульованого резервуара, в якому необхідно підтримувати постійний рівень, умовою статичної рівноваги є рівняння

$$V_{пр} - V_{вит} = 0$$

де $V_{пр}$ - прихід рідини в резервуар; $V_{вит}$ - витрата рідини.

Аналогічно виглядає рівняння статичної рівноваги регулятора ресивера з газом чи парою

$$G_{пр} - G_{вит} = 0$$

де $G_{пр}$, $G_{вит}$ маси газу, що приходять чи виходять з ресивера.

Якщо регульований об'єкт є об'єм, у якому повинна підтримуватись постійна температура, тоді умова статичної рівноваги прийме вигляд

$$Q_{пр} - Q_{вит} = 0$$

де $Q_{пр}$ - кількість теплоти, що надходить в об'єм в одиницю часу;

$Q_{вит}$ - кількість теплоти, що виходить з об'єму в одиницю часу.

Перехідний процес у досліджуваному об'єкті може з'явитись тільки тоді, коли буде порушена статична рівновага. Поява збільшення одного з членів рівняння статичної рівноваги неминуче спричинить появу збільшення і другого члена рівняння, причому такі збільшення, як правило, нерівні між собою. У результаті умова статичної рівноваги порушується і тоді при поступальному русі

$$F_{д} + \Delta F_{д} \neq F_{c} + \Delta F_{c};$$

при обертальному русі

$$M_{д} + \Delta M_{д} \neq M_{c} + \Delta M_{c};$$

при заповненні резервуара рідиною

$$V_{пр} + \Delta V_{пр} \neq V_{вит} + \Delta V_{вит};$$

при заповненні ресивера газом чи парою

$$G_{пр} + \Delta G_{пр} \neq G_{вит} + \Delta G_{вит};$$

при порушенні теплового режиму

$$Q_{пр} + \Delta Q_{пр} \neq Q_{вит} + \Delta Q_{вит};$$

Отримані нерівності можуть бути спрощені, якщо врахувати умови статичної рівноваги:

$$\Delta F_{д} \neq \Delta F_{c}; \quad \Delta M_{д} \neq \Delta M_{c};$$
$$\Delta V_{пр} \neq \Delta V_{вит}; \quad \Delta G_{пр} \neq \Delta G_{вит}; \quad \Delta Q_{пр} \neq \Delta Q_{вит}.$$

Таким чином, при порушенні умов статичної рівноваги в досліджуваній системі виникає надлишок чи недолік рушійних сил (чи моментів), надходження рідини, газу, надлишок чи недолік тепла. Цей надлишок чи недолік викликає зміну характеру роботи системи.

Подальше перетворення отриманих нерівностей здійснюють способом залучення загальновідомих залежностей, принципів, законів чи аналізу вихідних характеристик системи, отриманих у пошуковому експерименті. При цьому можна використовувати феноменологічні закони (такі, як закони Гука, Фур'є та ін.), напівемпіричні співвідношення (наприклад, додаткове співвідношення Ньютона в теорії удару) і чисто емпіричні співвідношення.

Порушення статичної рівноваги досліджуваної системи, що має поступальний рух, приведе до прискореного чи уповільненого руху. Об'єкт, що рухається, має масу «m», а прискорення його «a», тоді згідно з принципом Даламбера рівняння динамічної рівноваги (рівняння руху) може бути записане у вигляді

$$ma = \Delta F_{д} - \Delta F_{c} \text{ чи, тому що } a = \frac{dv}{dt},$$

де v - швидкість руху; t - час,
тоді

$$m\frac{dv}{dt} - \Delta F_{д} - \Delta F_{c}.$$

Для об'єкта, що має обертальний рух, рівняння динамічної рівноваги записують також згідно з принципом Даламбера. Якщо J - момент інерції деталей двигуна, що обертаються чи поступально рухаються, приведений до осі вала двигуна, ω - кутова швидкість обертання вала, тоді рівняння руху буде

$$J\frac{d\omega}{dt} = \Delta M_{д} - \Delta M_{c}.$$

При порушенні статичної рівноваги роботи резервуара в ньому відбувається накопичення (в алгебраїчному смислі) рідини і, як наслідок, зміна її рівня.

Якщо dV - зміна об'єму рідини в резервуарі за час dt, тоді

$$dV = (\Delta V_{пр} - \Delta V_{вит})dt,$$

або $$\frac{dV}{dt} = \Delta V_{пр} - \Delta V_{вит}.$$

При порушенні рівноваги у роботі ресивера кількість газу в ньому змінюється на величину dG за елементарний інтервал часу dt, тому

$$dG = (\Delta G_{пр} - \Delta G_{вит})\, dt, \text{ і } \frac{dG}{dt} = \Delta G_{пр} - \Delta G_{вит}.$$

Порушення теплового режиму деякого об'єму приводить до зміни кількості тепла, зосередженого в обраному об'ємі, на dQ за інтервал часу dt.

Тому

$$dQ = (\Delta Q_{пр} - \Delta Q_{вит})\, dt, \text{ чи } \frac{dQ}{dt} = \Delta Q_{пр} - \Delta Q_{вит}.$$

Порівняння отриманих диференціальних рівнянь для різних регульованих об'єктів показує їхню ідентичність. Подальше перетворення цих рівнянь можливо на основі знання фізики загальних властивостей газів, теплопередачі та ін. Наприклад, відомо, що стан ідеальних газів описується рівнянням Клайперона

$$pV = GTR,$$

де p - тиск газу в ресивері; V - об'єм ресивера; G - кількість газу, що знаходиться в ресивері; T - абсолютна температура газу; R - газова постійна.

З рівняння Клайперона виходить

$$G = \frac{pV}{RT}.$$

Тоді рівняння динамічної рівноваги ресивера буде

$$\frac{V}{RT}\frac{dp}{dt} = \Delta G_{пр} - \Delta G_{вит}.$$

Описування збільшень у правій частині розглянутих рівнянь також вимагає попереднього знання їхнього зв'язку з властивостями об'єкта.

Введення в рівняння динамічної рівноваги залежностей, що описують збільшення, часто приводить до підвищення порядку диференціальних рівнянь. Але при деяких спрощеннях порядок диференціальних рівнянь удається знизити. Таким спрощенням у більшості випадків є зневага інерційності об'єкта чи лінеаризація збільшень. Для лінеаризації останніх часто використовують розкладання функції в ряд Маклорена.

Нехай, наприклад, крутячий момент об'єкта (вала двигуна), що має обертальний рух, залежить від кутової швидкості обертання ω і положення h органа управління паливного насоса двигуна, тобто

$$M_д = f(\omega, h).$$

Для визначення збільшення крутячого моменту двигуна залежно від збільшень ω і h отриману функцію слід розкласти в ряд Маклорена

$$M_д + \Delta M_д = M_д + \frac{\partial M_д}{\partial \omega} d\omega + \frac{\partial^2 M_д}{\partial \omega^2}\frac{d\omega^2}{2!} + \ldots + \frac{\partial M_д}{\partial h} dh + \frac{\partial^2 M_д}{\partial h^2}\frac{dh^2}{2!} + \ldots$$

Якщо в розкладанні замінити нескінченно малі величини $d\omega$ і dh величинами $\Delta\omega$ і Δh кінцевими, але досить малими, тоді рад буде мати вигляд

$$M_д + \Delta M_д = M_д + \frac{\partial M_д}{\partial \omega}\Delta\omega + \frac{\partial^2 M_д}{\partial \omega^2}\frac{\Delta\omega^2}{2!} + \ldots + \frac{\partial M_д}{\partial h}\Delta h + \frac{\partial^2 M_д}{\partial h^2}\frac{\Delta h^2}{2!} + \ldots$$

Похідні в цьому розкладанні повинні підраховуватись у положенні рівноваги режиму роботи двигуна, при якому виконується умова статичної рівноваги.

При малих кінцевих збільшеннях $\Delta\omega$ і Δh і нерозривності функції $M_d = f(\omega, h)$ можна відкинути всі члени ряду з $\Delta\omega$ і Δh в ступенях вище першої без внесення суттєвої помилки, тобто практично зробити заміну дійсної функції $M_d = f(\omega, h)$ її дотичною в точці рівноважного режиму. Таку заміну називають лінеаризацією. Після лінеаризації збільшення крутячого моменту приймає вигляд

$$\Delta M_д = \frac{\partial M_д}{\partial \omega}\Delta\omega + \frac{\partial M_д}{\partial h}\Delta h.$$

Нехай момент опору є тільки функцією кутової швидкості

$$M_c = f(\omega).$$

Тоді, після розкладання цієї функції в рад Маклорена і лінеаризації, одержимо

$$\Delta M_c = \frac{dM_c}{d\omega}\Delta\omega.$$

Після підстановки перетворених збільшень у рівняння динамічної рівноваги об'єкта, отримаємо:

$$J\frac{d\omega}{dt} + \left(\frac{\partial M_c}{\partial \omega} - \frac{\partial M_д}{\partial \omega}\right)\Delta\omega = \frac{\partial M_д}{\partial h}\Delta h;$$

Будь-які диференціальні рівняння - це модель цілого класу явищ, тобто сукупність явищ, що характеризуються однаковими процесами. При інтегруванні рівнянь одержують велику кількість вирішень, що задовольняють початковому диференціальному рівнянню. Щоб отримати з безлічі можливих вирішень одне, що задовольняє розглянутому процесу, необхідно задати додаткові умови до диференціального рівняння. Вони повинні чітко виділяти досліджуване явище з усього класу явищ. Умови, що характеризують усі особливості явища, називаються ***умовами однозначності.*** Вони мають такі ознаки: ***геометрію системи*** (форму і розміри тіла); ***фізичні властивості тіла*** (теплопровідність, вологопровідність, пружність, в'язкість та ін.); ***початкові умови,*** тобто стан системи в початковий момент; ***граничні умови,*** тобто умови взаємодії системи на границях з навколишнім середовищем. Початкові і граничні умови називають ***крайовими.***

Розглянуті приклади вибору виду моделі об'єкта мають відношення лише до таких об'єктів, що за класифікаційною діаграмою визначаються як детерміновані.

Розглянемо далі способи вибору виду математичної моделі для імовірнісних об'єктів. Як і раніше, для статичних об'єктів під стаціонарністю входу будемо розуміти постійне його значення. Якщо вхідний сигнал приймає кілька значень, тоді його будемо вважати нестаціонарним.

Нехай є одномірно-одномірна схема взаємодії об'єкта з зовнішнім середовищем. Якщо на вхід об'єкта подається постійна в часі дія, тоді в якості математичної моделі статичного імовірнісного об'єкта може бути деякий закон розподілу вихідної величини. Якщо вхідна дія приймає різні значення і кожному значенню відповідає рад значень вихідної величини об'єкта, тоді в якості моделі імовірнісного об'єкта приймають набір законів розподілу вихідної величини для всіх значень вхідної дії.

При моделюванні імовірнісних об'єктів, крім законів розподілу вхідних і вихідних величин, суттєвим є зв'язок між ними. Тому до складу моделі включають коефіцієнти взаємної кореляції і функції

$$H_m = f(x);\ R = f(x);\ y_{cp} = f(x);\ \sigma = f(x),$$

де x - вхідний вплив; y_{cp} - середньозважене значення вихідної характеристики; H_m - максимальна ентропія системи; R - відносна організація системи; σ - середньоквадратичне відхилення.

При багатомірно-одномірній схемі взаємодії статичного імовірнісного об'єкта з зовнішнім середовищем задача математичного моделювання зводиться до одномірно-одномірної схеми для кожного сполучення постійних вхідних впливів.

Нестаціонарний випадок відрізняється тим, що кожний вхідний вплив може приймати кілька значень. При цьому для кожного конкретного сполучення значень входів задача аналізу зв'язку між входами і виходом може вирішуватись аналогічно задачі для багатомірно-одномірної стаціонарної схеми.

Оцінка ступеня зв'язку виходу з входами проводиться способом зіставлення статистичних параметрів і обчислення коефіцієнтів взаємної кореляції.

Моделювання об'єкта при одномірно-багатомірній і багатомірно-багатомірній схемах взаємодії виконується аналогічно вищевикла-деному

Розглянемо далі моделювання динамічних режимів імовірнісних об'єктів. На відміну від детермінованих систем і процесів, де стаціонарність розглядалась стосовно самої координати, для імовірнісних систем під стаціонарністю розуміють незмінність у часі деяких статистичних параметрів входу (математичного очікування, дисперсії, моментів більш високого порядку).

При одномірно-одномірній схемі взаємодії об'єкта з зовнішнім середовищем і багаторазовим надходженням на його вхід однієї і тієї ж функції часу x(t) можливі два випадки: на виході об'єкта спостерігається стаціонарний випадковий процес; на виході об'єкта спостерігається нестаціонарний випадковий процес.

У першому випадку за математичну модель вихідної величини об'єкта приймають закон розподілу значень вихідної величини об'єкта, що має ті самі параметри для всіх зрізів за часом. Зрізи за часом назначають з кроком Δt.

Модель доповнюють залежностями

$$H_m\,(i,\,\Delta t) = f(x);\; R\,(i,\,\Delta t) = f(x);\; i = 1,\,2,\,3,\,\ldots\,.$$

Ці залежності можуть бути представлені алгебраїчною функцією чи диференціальними рівняннями.

У другому випадку (нестаціонарний вихід об'єкта) за математичну модель об'єкта приймають функціональні зв'язки

$$M_y\,(i,\,\Delta t) = f(x);\; \sigma_y\,(i,\,\Delta t) = f(x);\; i = 1,\,2,\,3,\,\ldots\,.$$

де $M_y(i, \Delta t)$ - математичне очікування розподілу вихідних величин *y* по дискретних відрізках часу з кроком Δt, $\sigma(i, \Delta t)$ - середньоквадратичне відхилення.

Ці зв'язки також описують за допомогою алгебраїчних чи диференціальних рівнянь.

Якщо на вхід об'єкта багаторазово надходять різні функції часу $x_i(t)$, тоді їх розглядають як випадкові реалізації випадкового процесу. Для кожного перерізу за часом $(i, \Delta t)$ визначають інформаційні і імовірнісні характеристики $H_m^{x_i}$, R_{x_i}, M_{x_i}, σ_{x_i}. Аналогічно для кожної сукупності виходу для тих же перерізів за часом визначаються інформаційні та імовірнісні характеристики $H_m^{y_i}$, R_{y_i}, M_{y_i}, σ_{y_i}. За математичну модель об'єкта приймають функціональні зв'язки

$$H_m^{y_i}(i, \Delta t) = f_{1i}\left[H_m^{x}(i, \Delta t)\right];$$

$$R_{y_i}(i, \Delta t) = f_{2i}\left[R_x(i, \Delta t)\right];$$

$$M_{y_i}(i, \Delta t) = f_{3i}\left[M_x(i, \Delta t)\right];$$

$$\sigma_{y_i}(i, \Delta t) = f_{4i}\left[\sigma_x(i, \Delta t)\right], (i = 1, 2, \ldots, k)\ldots$$

Іноді всі масиви реалізацій вихідної величини розглядають як єдиний масив. Тоді за математичну модель об'єкта приймають функціональні зв'язки між узагальненими параметрами виходу і параметрами входу.

При багатомірно-одномірній схемі взаємодії імовірнісного об'єкта із середовищем у нестаціонарному режимі в якості його математичної моделі приймають функціональні залежності:

$$H_m^y = f_{1l}\left(H_m^{x_1}\right); \qquad M_y = f_{3l}\left(M_{xl}\right);$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$H_m^y = f_{1m}\left(H_m^{x_1}\right); \qquad M_y = f_{3m}\left(M_{xn}\right);$$

$$R_y = f_{2l}\left(R_{x1}\right); \qquad \sigma_y = f_{1l}\left(\sigma_{xl}\right);$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$R_y = f_{2m}\left(R_{xm}\right); \qquad \sigma_y = f_{2m}\left(\sigma_{xn}\right).$$

При одномірно-багатомірній схемі взаємодії об'єкта в нестаціонарному режимі за математичну модель об'єкта приймають функціональні залежності:

$$H_m^{y_i} = f_{1i}\left(H_m^x\right);\ R_y^i = f_{2i}\left(R_x\right);$$
$$M_{y_i} = f_{3i}\left(M_x\right);\ \sigma_{y_i} = f_{4i}\left(\sigma_x\right),\ (i = 1, 2, ..., n)$$

При багатомірно-багатомірній схемі взаємодії у нестаціонарному режимі встановлюють функціональні залежності інформаційних та імовірнісних характеристик для кожного входу і виходу При цьому можна вважати, що будь-який вихід залежить від усіх входів, і відповідно вибрати вид математичних моделей.

При моделюванні квазідетермінованих систем часто їх відносять до класу детермінованих, тому що математичний апарат описування детермінованих систем дає великі можливості конкретного аналізу.

Суттєвим моментом математичного моделювання є ***вибір структури*** моделі. Так, у класі алгебраїчних функцій під структурою моделі розуміють вид і ступінь полінома. У класі диференціальних рівнянь - вид і порядок диференціального рівняння.

Вибрати структуру математичної моделі в заданому класі допомагає залучення різних додаткових критеріїв, що відбивають фізичні закономірності чи специфіку розв'язуваної задачі.

Для вибору структури математичної моделі в класі алгебраїчних функцій звичайно використовують два методи: повний перебір рівнянь різної структури (метод А.Г. Івахненко); синтез структури математичної моделі, заснований на дисперсійному аналізі.

У методі А.Г. Івахненко перебір починається з найпростішої лінійної моделі від одного аргументу виду

$$y = a_0 + a_1 x_i\ ,\ 1 \le i \le m.$$

Потім модель відшукують в класі лінійних рівнянь для кожних двох факторів

$$y = a_0 + a_1 x_i + a_2 x_j,\ 1 \le i \le m,\ 1 \le j \le m.$$

Далі вводять коваріації факторів (попарні добутки)

$$y = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j,\ 1 \le i \le m,\ 1 \le j \le m,\ i \ne j.$$

Наступним етапом ускладнення моделі є введення до неї квадратів, кубів факторів і відповідних коваріацій. Під складністю математичної моделі в даному випадку розуміють число невідомих коефіцієнтів $a_0, a_1, ...$ На кожному кроці ускладнення з набору моделей рівної складності вибирають найкращу за критерієм мінімуму середньої квадратичної помилки між теоретичними й експериментальними значеннями, а також мінімуму середньої квадратичної помилки стосовно до інформаційного масиву перевірочної послідовності. Для визначення коефіцієнтів і характеристик моделі на кожному кроці використовують частину експериментального масиву, що утворює «навчальну» послідовність. Інші експериментальні значення входять у «перевірочну» послідовність.

Метод синтезу структури алгебраїчної моделі, заснований на дисперсійному аналізі, дозволяє вибрати найкращу з декількох можливих моделей. Звичайно вибір здійснюють з обмеженого числа моделей. У процесі вибору використовують такі критерії (чи їхні комбінації): найменше число коефіцієнтів моделі, сумісне з різною помилкою; найпростішу математичну форму моделі, сумісну з розумною помилкою; розумні передумови (структура моделі виходить з деякого закону); мінімальну суму квадратів відхилень між передбаченими характеристиками за допомогою моделі та їхніх емпіричних значень; мінімальну дисперсію, що включає дисперсію експерименту і дисперсію математичної моделі; математичну модель мінімальної складності, що задовольняє критерію Фішера.

При виборі структури моделі у класі диференціальних рівнянь записують систему n однорідних лінійних диференціальних рівнянь першого порядку, у якій враховують всі зв'язки між усіма перемінними:

$$\frac{dy_1}{dt} = a_{i1}y_1 + a_{i2}y_2 + \ldots + a_{in}y_n;$$
$$\frac{dy_n}{dt} = a_{n1}y_1 + a_{n2}y_2 + \ldots + a_{nn}y_n,$$

де y_i - досліджувана перемінна системи; a_{ij} - невідомі параметри математичної моделі.

Структуру моделі визначають значеннями параметрів: частина параметрів може дорівнювати нулю, а частина параметрів мати малі значення порівняно з іншими. Параметри, що мають нульові і малі значення, відкидають, інші, що характеризують суттєві зв'язки між перемінними, залишають. У результаті виходить система диференціальних рівнянь, що відбиває структуру суттєвих зв'язків між перемінними.

Пошук структури моделі системи в класі лінійних диференціальних рівнянь є першим етапом. На другому етапі праві частини системи рівнянь записують у вигляді нелінійних алгебраїчних поліномів для суттєвих перемінних, що залишились. Ускладнення ступеня і виду полінома проводять поступово і цей процес йде до задовольняння деякого критерію узгодженості математичної моделі й експериментальних даних. Визначення параметрів відбувається способом вирішення системи лінійних алгебраїчних рівнянь.

Слід помітити, що в класі диференціальних рівнянь можливі два типи моделей:

1. Система із зосередженими параметрами. Це система, в якій невідомі перемінні є функціями тільки однієї незалежної перемінної. Поведінку такої системи описують звичайними диференціальними рівняннями.

2. Лінійні і нелінійні системи. Якщо усі функції, що стоять у правій частині рівняння

$$\frac{dx}{dt} = Q(x(t), t);$$

лінійні щодо своїх аргументів, тоді система рівнянь лінійна. В іншому випадку їх називають нелінійними.

Розглянуті вище методи визначення структури моделей базуються на формальних критеріях. В окремих випадках задачі моделювання можуть вимагати використання неформальних критеріїв. Так, для функціонального моделювання критеріями визначення структури можуть виступати: критерій форми, що дозволяє зв'язати динаміку вихідного показника з диференціальними рівняннями, вирішення яких можуть мати аналогічний вигляд; критерій мінімальної математичної складності, що дозволяє вибрати мінімальний порядок диференціального рівняння.

При структурно-функціональному моделюванні ці критерії використовують для вибору математичних моделей підсистем, а зв'язки між підсистемами визначають на основі додаткових гіпотез.

Процес вибору класу і структури математичної моделі об'єкта закінчують її попереднім ***контролем***. При цьому здійснюють контроль розмірностей, контроль порядків, контроль характеру залежностей, контроль екстремальних ситуацій, контроль граничних умов, контроль математичної замкнутості, контроль фізичного смислу, контроль стійкості моделі.

Контроль розмірностей зводиться до перевірки виконання правила, відповідно з яким прирівнюватись і складатись можуть тільки величини однакової розмірності.

Контроль порядків спрямований на спрощення моделі. При цьому визначають порядки величин, які складають, і явно малозначні доданки відкидають.

Контроль характеру залежностей зводиться до перевірки напрямку і швидкості зміни одних величин при зміні інших. Напрямки і швидкість, що виходять з математичної моделі, мають відповідати фізичному смислу задачі.

Контроль екстремальних ситуацій зводиться до перевірки наочного смислу вирішення при наближенні параметрів моделі до нуля чи нескінченності.

Контроль граничних умов полягає в тому що перевіряють відповідність математичної моделі граничним умовам, які виходять із смислу задачі. При цьому перевіряють, чи дійсно граничні умови поставлені і враховані при побудові функції, яку знаходять, і що ця функція насправді задовольняє таким умовам.

Контроль математичної замкнутості зводиться до перевірки того, що математична модель дає однозначне вирішення.

Контроль фізичного смислу - це перевірка фізичного смислу проміжних співвідношень, використаних при побудові математичної моделі.

Контроль стійкості моделі складається в перевірці того, що варіювання вихідних даних у рамках наявних даних про реальний об'єкт не приведе до суттєвої зміни вирішення.

3.4. Імітаційне моделювання

Загальноприйнятого визначення імітаційної моделі немає. В.Н. Бусленко визначає імітаційну модель як модель, у якій тим чи іншим способом імітують випадкові процеси. К. Шенон під імітайним моделюванням розуміє процес конструювання моделі реальної системи і постановки експериментів на цій моделі з метою зрозуміти поводження системи або оцінити різні стратегії, що забезпечують функціонування даної системи.

Найчастіше ***під імітаційною моделлю розуміють програму, що у процесі реалізації її на ЕОМ дозволяє імітувати поводження реальної системи в різних умовах.***

Кожна імітаційна модель включає комбінацію компонентів, параметрів, перемінних, функціональних залежностей, обмежень і цільових функцій. У загальному вигляді імітаційну модель описують так:

$$E = f(x_i, y_j)$$

де E - результат дії системи; f - функціональна залежність; x_i - перемінні, якими можна керувати; y_i - перемінні, якими керувати не можна.

Перемінні моделі поділяються на два види: ***екзогенні*** й ***ендогенні.*** Екзогенні - вхідні перемінні, породжені поза системою. Ендогенні - перемінні стану, коли вони характеризують стан чи умови, реалізовані в системі, або вихідні перемінні, коли вони відносяться до виходів системи.

Поведінку перемінних і параметрів описують за допомогою функціональних залежностей, що по природі можуть бути детермінованими чи стохастичними. Детерміновані і стохастичні співвідношення звичайно виражають у формі математичного рівняння, що установлює взаємозв'язок між ендогенними й екзогенними перемінними.

Границі зміни значень перемінних чи інтервал можливих значень випадкової величини, встановлюють за допомогою обмежень. Обмеження можуть виходити або з фізичної сутності самої системи, або призначатись дослідником.

Параметри моделі вибирають до початку імітаційного моделювання і не змінюють в процесі дослідження поведінки моделі.

Компоненти моделі - це всі її складові частини, що відповідають частинам реальної системи.

Цільова функція (функція критерію) служить для відображення цілей чи задач системи і необхідних правил оцінки їхнього виконання. У якості цільових часто використовують функції двох типів - ***збереження і придбання.***

Функції збереження пов'язані зі збереженням яких-небудь станів (наприклад, безпеки, комфорту) чи підтримки яких-небудь ресурсів (наприклад, енергетичних, матеріальних, фінансових).

Функції придбання пов'язані з придбанням нових ресурсів чи досягненням деяких станів, до яких прагне система.

Застосування імітаційного моделювання є доцільним при виконанні кожної з умов:

- експерименти з реальною системою є неможливими в результаті порушення нормальної роботи системи чи при неможливості дослідження альтернативних варіантів її поведінки;
- важко забезпечити підтримку тих самих умов при кожному повторенні експерименту;
- для отримання статистично значимих результатів необхідні великі витрати часу і засобів;
- не існує закінченої математичної постановки даної задачі і розроблених аналітичних методів її вирішення або математичні методи настільки трудоємні, що імітаційна модель дає більш простий спосіб рішення задачі.

Конструювання імітаційної моделі починають з побудови статичної моделі об'єкта. Таку модель представляють у вигляді логічної блок-схеми, органограми, технологічної карти або технологічної діаграми.

Процес переходу від блок-схеми до імітаційної моделі системи включає два етапи. Спочатку модель системи представляють у вигляді діаграми причинно-наслідкових зв'язків і орієнтованого графа, у вершинах якого вказують параметри моделі, а напрямки дуг показують, що зміна одного параметра викликає зміну іншого параметра. Далі дуги графа навантажують функціональними залежностями чи вхідними діями, під впливом яких здійснюється перехід з однієї вершини графа до іншої. Ця процедура перетворює орієнтований граф у графоїд кінцевого автомата. Тепер кожен блок імітаційної моделі і всю модель можна розглядати як кінцевий детермінований чи стохастичний автомат.

Задаючи правила переходів і взаємодії автоматів, можна цілком описати поведінку імітаційної моделі.

У процесі перетворення блок-схеми в імітаційну модель необхідно визначити призначення моделі, які компоненти мають бути включені в її склад, параметри і перемінні, функціональні співвідношення та ін. До складу імітаційної моделі включають усі компоненти системи, що впливають як позитивно, так і негативно на ефективність її роботи. Призначення компонентів системи полягає в тому, щоб перетворювати вхідні сигнали у вихідні. Прийнято розрізняти три основних види компонентів, що складають функціональні блоки складних систем:

1) елементи перетворення, в яких один чи кілька вхідних сигналів *x,* що оброблені деяким наперед заданим чином, перетворяться в один чи кілька вихідних сигналів *y,* причому $x \geq y$.

2) елементи класифікації, в яких один чи кілька вхідних сигналів *x* розподіляються або сортуються за двома чи декількома різними виходами *y,* причому $x \leq y$.

3) елементи зворотного зв'язку, в яких вхідний сигнал *x* деяким чином змінюється залежно від вихідного сигналу *y,* тобто $x = f(y)$, причому зворотний зв'язок може бути як позитивний, так і негативний. Вирішення питання, які

компоненти необхідно включати, а які виключати, визначають прийнятим рівнем деталізації чи кількістю перемінних і параметрів, що входять у модель.

Далі визначають функціональні зв'язки між компонентами і перемінними і задають значення параметрів. При визначенні функціональних зв'язків виникають прямі і зворотні задачі. Якщо знають рівняння, які описують поведінку динамічної системи Q, тоді вирішенням прямої задачі можна знайти вихідний сигнал системи на заданий вхідний сигнал:

$$\alpha:(x,Q)\rightarrow y$$

Зворотну задачу тоді представляють так: за заданим математичним описуванням системи чи компонента Q і відомому відгуку потрібно знайти вхідний сигнал x, що викликав цей відгук:

$$\beta:(y,Q)\rightarrow x$$

Якщо задано сукупності вхідних *x* і відповідних вихідних сигналів *y,* а необхідно знайти математичне описування самого компонента чи всієї системи Q, то це теж зворотна задача, що представляється у вигляді

$$\gamma:(x,y)\rightarrow Q$$

Ця задача відома як задача ідентифікації, або структурного синтезу.

Отримані описування функціональних зв'язків, поведінки компонентів і системи в цілому переводять на будь-яку універсальну алгоритмічну мову програмування, що потім використовують при проведенні машинного експерименту на ЕОМ.

В імітаційних моделях можуть відтворюватись тимчасові і просторові закономірності функціонування систем.

Алгоритм процесу імітації може бути представлений у вигляді ланцюжка елементарних кроків чи дій:

1.***Визначення системи.*** Установлюють границі системи, визначають обмеження на параметри, визначають перемінні, встановлюють критерії оцінки.

2.***Формування моделі.*** У результаті процесів спрощення й абстрагування реальний об'єкт чи система заміняються деякою логічною схемою.

3.***Підготовка даних.*** Відбирають і представляють у відповідній формі ті вихідні дані, що є найбільш суттєвими для реалізації моделі.

4. ***Трансляція моделі.*** На відповідній мові програмування складають програму, що дозволяє імітувати динамічні процеси системи, що моделюється.

5.***Оцінка адекватності.*** За результатами, отриманими після реалізації програми на ЕОМ, робиться підсумок про коректність висновків про реальну систему, одержаних на основі моделі.

6.***Стратегічне планування.*** Складання загального плану проведення експериментів з моделлю для отримання необхідної інформації про систему.

7.***Тактичне планування.*** Відповідно зі стратегічним планом визначають способи проведення кожної серії експериментів.

8.***Експеримент.*** Прогін моделі на ЕОМ з метою одержання відповідей на поставлену задачу.

9. ***Інтерпретація.*** Заданими, отриманими на моделі, формують підсумки і висновки про діяльність реальної системи.

10.***Реалізація.*** Практичне використання моделі і результатів моделювання.

11.***Документування.*** Реєстрація ходу здійснення проекту та його результатів, а також документування процесу створення і використання моделі.

Схема алгоритму процесу імітації представлена на рис. 3.11.

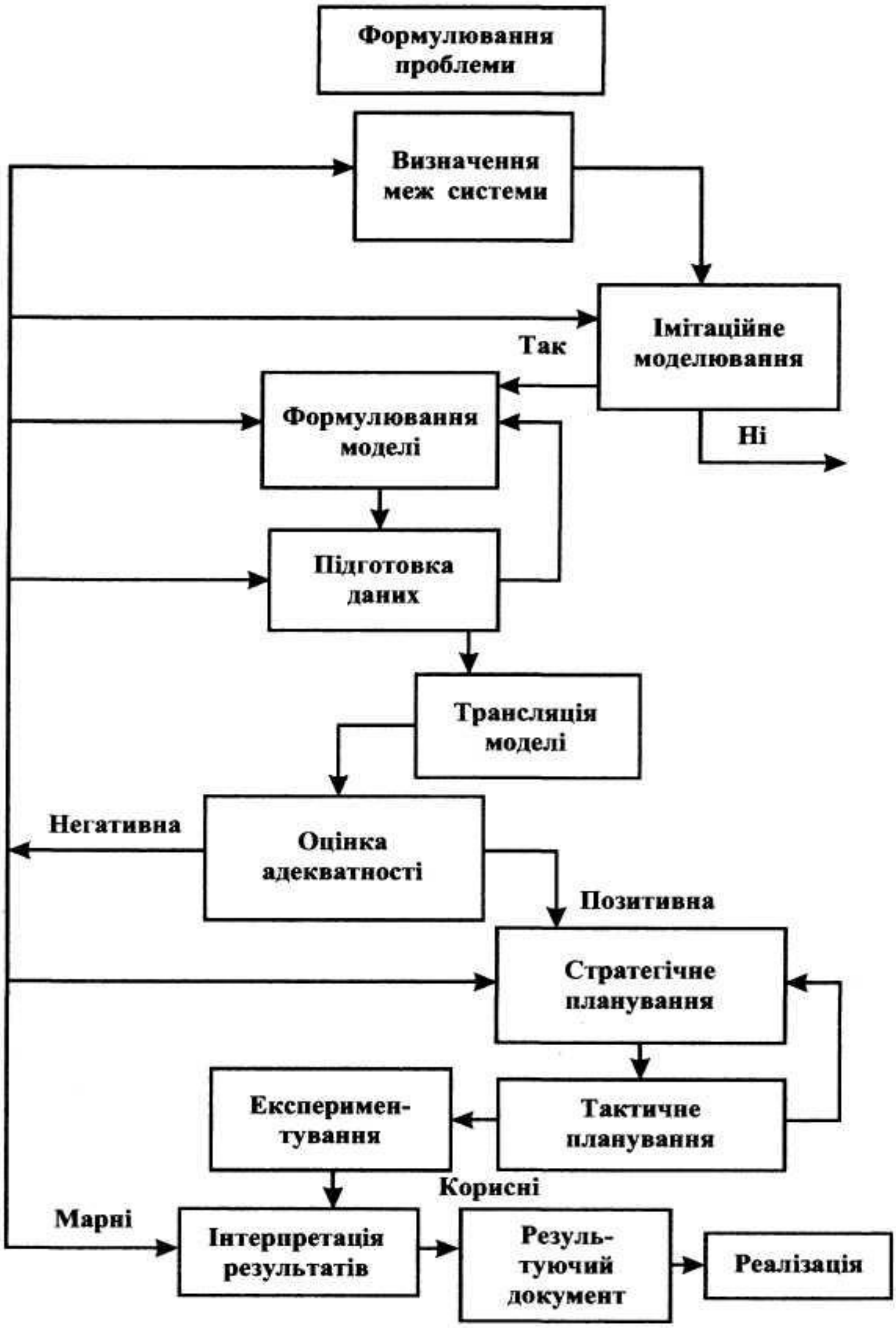

Рис. 3.11. **Схема алгоритму процесу імітації**

4 МЕТОДИ ДОСЛІДЖЕННЯ СИСТЕМИ

1 Зовнішня і внутрішня правдоподібність дослідження
2 Фазова траєкторія. Фазовий та структурний портрет системи
4 Передатна функція системи
5 Врахування зв'язків в системі при визначенні її передатної функції
6 Частотна характеристика системи
7 Компартмент і використання компартментального аналізу
8 Основні інформаційні характеристики систем
9 Оцінка ефективності і надійності систем
10 Основні етапи математичного дослідження
11 Критерій і оптимізація
14 Сутність оптимізації в рамках лінійного і нелінійного програмування
15 Динамічне програмування
16 Оптимізація за обмежуючими критеріями
17 Корекція системи критеріїв
18 Застосування теорії ігор для оптимізації
20 Експеримент
21 Класифікаційні ознаки експериментів
22 Структура експерименту
23 Методика
24 Класифікація вимірювань
25 Методи вимірювань

4.1. Вибір методу дослідження математичної моделі

Математична модель дозволяє: досліджувати об'єкт із відключенням несуттєвих чи суттєвих зв'язків; реалізувати ті умови, що недосяжні на практиці; прогнозувати стан об'єкта; перевіряти гіпотези, одержувати їхні наслідки та ін.

Вибір методу дослідження математичної моделі безпосередньо пов'язаний з такими поняттями як ***зовнішня*** і ***внутрішня*** правдоподібність дослідження.

Під ***зовнішньою правдоподібністю*** дослідження розуміють очікуваний ступінь адекватності математичної моделі реальному об'єкту за властивостями, цікавими для дослідника.

Під ***внутрішньою правдоподібністю*** дослідження розуміють очікуваний ступінь точності вирішення отриманих рівнянь, що прийняті за математичну модель об'єкта.

Якщо тип моделі вже обрано, тоді зовнішню правдоподібність моделі вважають фіксованою і вибір методу дослідження буде цілком визначатись необхідним ступенем внутрішньої правдоподібності.

У більшості випадків при виборі методу дослідження керуються ***принципом відповідності зовнішньої і внутрішньої правдоподібності (рівної правдоподібності).*** Цей принцип аналогічний відомому правилу наближених обчислень: *ступінь точності обчислень має відповідати ступеню точності вихідних даних.* Але залежно від умов і завдань дослідження можливі відхилення від цього принципу. Наприклад:

1) якщо мова йде про розробку нового методу досліджень, який передбачається застосовувати до широкого заздалегідь не фіксованого класу моделей, тоді потрібно прагнути до максимальної внутрішньої правдоподібності дослідження незалежно від рівня зовнішньої правдоподібності;

2) якщо здійснюється перевірка зовнішньої правдоподібності моделі, тоді внутрішня правдоподібність обраного методу перевірки має бути максимальною;

3) якщо модель настільки проста, що для неї легко отримати точне вирішення, тоді штучно знижувати строгість вирішення безглуздо.

В інших випадках перевага віддається «принципу рівної правдоподібності».

Вибір методу дослідження тим ефективніше, чим чіткіше поставлена мета і є більше зведень про кінцеве вирішення задачі. Такі зведення можуть бути отримані способом попередніх досліджень моделі чи її елементів. У процесі попередніх досліджень здійснюють порівняння величин окремих членів рівнянь у досліджуваному діапазоні зміни перемінних і параметрів моделі. Відносно малі доданки відкидають, нелінійні залежності заміняють на лінійні. Деякі з компонентів моделі апроксимують грубими рівняннями. Усе це дозволяє швидко отримати грубе вирішення задачі. Знання, хоча б саме грубе, якісних і кількісних характеристик вирішення, що відшукують, допомагає при виборі точності методу дослідження. Іноді навіть грубе вирішення виявляється достатнім. Як приклад можна вказати на задачу про пошук екстремального значення функції. Якщо точка екстремуму є стаціонарною, тоді навіть груба помилка в її відшукуванні мало позначиться на підрахунку цього значення. Тому застосування високоточних методів пошуку такого екстремуму нераціональне. Громіздкі точні обчислення в цьому випадку створюють лише ілюзію точності. Якщо ж вибирають грубу математичну модель, не потрібно застосовувати громіздкі обчислювальні процедури.

Вибір методу дослідження математичної моделі багато в чому визначається її видом. Так, статичні системи, представлені за допомогою алгебраїчних рівнянь, досліджують за допомогою визначників, методів Крамера і Гауса. У випадку утруднень з аналітичними вирішеннями використовують наближені методи: графічний метод, метод хорд, метод дотичних, метод ітерацій. У методі ітерацій доцільно використовувати ЕОМ. Застосування ЕОМ вимагає контролю точності (числа значущих цифр) залежно від грубості обчислювального методу.

Дослідження динамічних режимів функціонування об'єкта, представлених у класі диференціальних рівнянь, також визначається

класом, до якого відноситься розв'язуване рівняння. Якщо в результаті вирішення алгебраїчних рівнянь виходять числа, то при вирішенні диференціальних рівнянь виходять функції.

Для рішення диференціальних рівнянь широко використовують метод розподілу перемінних, метод підстановки, метод інтегруючого множника, метод якісного аналізу та ін. Для одержання наближених вирішень використовують метод послідовних наближень, метод функціональних рядів, метод Рунге-Кута, чисельні методи інтегрування та ін.

Для докладного вивчення моделей динамічних систем, побудованих у класі диференціальних рівнянь (особливо нелінійних), використовують якісну теорію диференціальних рівнянь. Остання дозволяє вивчити всі можливі вирішення - регулярні й особливі. В основі якісної теорії лежить поняття ***фазового портрета,*** що представляє собою сімейство фазових траєкторій на фазовій площині. ***Фазова траєкторія*** - зображення одного з можливих вирішень диференціального рівняння.

У деяких випадках системи диференціальних рівнянь містять параметри, зміна яких, як правило, приводить до зміни фазового портрета системи. Тому для упорядкування фазових портретів використовують *структурний портрет* системи.

Структурний портрет - це лінії, що поділяють площину суттєвих параметрів системи на області, усередині яких зберігається тип системи. Основними лініями в структурному портреті є лінія нейтральності і лінія кратних коренів. ***Лінія нейтральності*** - це геометричне місце точок, де система має чисто мнимі корені. ***Лінія кратних коренів*** - це границя дійсних і комплексних коренів.

Багато задач досліджують за допомогою варіаційного числення. Щоб сформулювати задачу варіаційного числення вводять поняття функціонала. Нехай, наприклад, є плоска крива $y = f(x)$ з областю визначення $x_0 \le x \le x_1$. Неважко побачити, що довжина кривої S_1, площа P криволінійної трапеції, об'єм тіла обертання V залежать від вигляду заданої кривої.

$$S_1 = \int_{x_0}^{x_1} \sqrt{1 + [y'(x)]^2}; \qquad P = \int_{x_0}^{x_1} y(x)dx; \qquad V = \pi \int_{x_0}^{x_1} [y'(x)]^2 dx.$$

Таким чином, функція $y = f(x)$ однозначно визначає величини S_1, P, V тобто відіграє роль своєрідного «аргументу». У цьому випадку величини S_1, P, V називаються функціоналами щодо функції $y = f(x)$.

Суть задачі варіаційного числення полягає в тому, якщо задано функціонал $F(y')$ в області $x_0 \le x \le x_i$, тоді потрібно знайти таку функцію $y = f(x)$ у заданій області визначення функціонала $F(y')$, при якій цей функціонал приймає мінімальне чи максимальне значення.

При дослідженні процесів методами варіаційного числення знаходять такі закономірності, при яких їхній розвиток енергетично найбільш економний. Дуже часто вони описуються експонентними функціями, що задовольняють принципам варіаційного числення.

При рішенні задач теорії пружності, аеродинаміки та ін. широко використовують теорію функцій комплексної перемінної. В основі цієї теорії лежить положення про конформне перетворення. Відповідно до цього положення дві пересічні криві z_1z_2 і z_1z_3 з області Z завжди можна перенести в область W з відповідними кривими ϖ_1, ϖ_2 і ϖ_1, ϖ_3, що зберігають рівність кутів між ними в кожній парі. Це дозволяє змінити координати таким чином, щоб спростити громіздкі математичні перетворення.

Розглянуті аналітичні методи, як правило, дозволяють успішно вирішувати лише відносно прості задачі. В той же час, часто виникає необхідність використання складних диференціальних рівнянь чи їхніх систем зі складними початковими і граничними умовами. їхнє вирішення дуже складне або невідоме. У цих випадках вдаються до тих чи інших наближених методів обчислень за допомогою методу кінцевих різниць або сіток.

Ідея методу кінцевих різниць чи сіток полягає в наступному:

1) у плоскій області G, де шукають вирішення, будують сіткову область G_h, що складається з однакових комірок і наближається до області G;

2) задане диференціальне рівняння заміняють у вузлах побудованої сітки відповідними кінцево-різницевими рівняннями;

3) на підставі граничних умов установлюють значення вирішення, яке відшукують, в граничних вузлах області G_h;

4) вирішивши отриману систему кінцево-різницевих рівнянь, знаходять значення функції у вузлах сітки.

Вибір сіткової області виконують залежно від конкретної задачі, але у всіх випадках контур сіткової області G_h слід вибирати так, щоб він можливо краще апроксимував контур заданої області G. Сіткова область може складатись з квадратних, прямокутних, трикутних та інших кліток.

При вирішенні задач, пов'язаних із системами управління, широко застосовують методи перетворення вихідних рівнянь за допомогою логарифмування, перетворень Лапласа, Фур'є та ін.

Логарифмування є найпростішим способом перетворень. Для приклада розглянемо вирішення найпростішого рівняння

$$y = a^{0,2}.$$

Звести число «а» у ступінь 0,2 прямим методом важко. Уникнути цього можна способом перетворення початкового рівняння за допомогою логарифмування:

$$\log y = 0{,}2 \log a$$

При цьому початкове рівняння можна назвати *оригіналом функції,* а перетворене - її *зображенням.* Процес логарифмування переводить вихідну функцію з *простору оригіналів* у *простір зображень.*

Після логарифмування операція зведення в ступінь приводить до простого множення числа 0,2 на log a, що не зустрічає ніяких утруднень. Потім за допомогою антилогарифмування отриманий результат переводять з простору зображень у простір оригіналів.

Аналогічно ведуть обчислення і за допомогою перетворень Лапласа чи Фур'є. Наприклад, у перетвореннях Лапласа вихідна функція часу $f(t)$ переводиться з простору оригіналів у простір зображень за допомогою інтеграла

$$F(p) = \int_0^{\infty} f(t) e^{-pt} dt,$$

де p - оператор Лапласа, $p = \frac{d}{dt}$.

Перевід функції з простору зображень у простір оригіналів здійснюють за допомогою інтеграла

$$f(t) = \frac{1}{2\pi} \int_{c-i_{\infty}}^{c+i_{\infty}} F(p) e^{pt} dp$$

Величину «с» вибирають так, щоб забезпечити збіжність інтеграла.

Перетворення Лапласа широко використовують при вирішенні диференціальних і інтегральних рівнянь. У процесі вирішення цих рівнянь користуються таблицями прямих і зворотних перетворень функцій так, як це виконують у випадку логарифмування.

Ґрунтуючись на методі перетворення функцій, вирішують задачі аналізу перехідних процесів у системах управління. У процесі вирішення оперують передатними функціями.

Під ***передатною функцією*** розуміють відношення вихідної характеристики лінійної системи до функції вхідного сигналу при нульових початкових умовах. У просторі зображень:

$$W(p) = \frac{y(p)}{x(p)},$$

де W(p) - передатна функція; x(p) - перетворення Лапласа функції вхідного сигналу; y (p) - перетворення Лапласа вихідної характеристики системи.

Вид передатної функції може бути отриманий з диференціального рівняння системи управління способом заміни операції диференціювання за часом оператором Лапласа p, а інтегрування за часом - заміною на 1/p.

Наприклад, якщо система управління описується диференціальним рівнянням

$$\frac{dy(t)}{dt} + ay(t) = kx(t),$$

тоді, зробивши заміну d/dt на p і переходячи до зображень, одержуємо

$$py(p)+ay(p)=kx(p),$$

Таке представлення диференціального рівняння називається операційним.

Для перебудування передатної функції цієї системи досить зробити нескладні алгебраїчні перетворення. У результаті одержимо

$$W(p)=\frac{y(p)}{x(p)}=\frac{k}{p+a}.$$

Для складних структур систем управління з *n* ланок існують деякі правила отримання передатних функцій:

1) передатна функція послідовного з'єднання ряду ланок системи дорівнює добутку передатних функцій окремих ланок,

$$W_c(p)=\prod_{i=1}^{n} W_i(p);$$

2) передатна функція паралельно з'єднаних ланок дорівнює сумі передатних функцій окремих ланок,

$$W_c(p)=\sum_{i=1}^{n} W_i(p);$$

3) передатну функцію з'єднання двох ланок зі зворотним зв'язком визначають за формулою

$$W_c(p)=\frac{W_n(p)}{1\pm W_{0c}(p)W_n(p)};$$

де $W_n(p)$ - передатна функція ланки прямого зв'язку; $W_{0c}(p)$ - передатна функція ланки зворотного зв'язку.

Знак «+» відповідає негативному зворотному зв'язку, знак «-» - позитивному.

Крім методу передатних функцій для аналізу систем управління широко використовують метод частотних характеристик, що складає теоретичну базу узагальненого гармонійного аналізу.

Частотні характеристики систем управління використовують при аналізі стійкості систем, якості перехідних процесів і динамічної точності, синтезу коригувальних пристроїв.

Під *частотною характеристикою* системи розуміють відношення комплексних зображень вихідної і вхідної амплітуд у режимі гармонійних коливань, що встановився:

$$W(j\omega) = \frac{A_2(\omega)}{A_1} e^{j\phi(\omega)},$$

де A_1 - амплітуда синусоїдального вхідного сигналу; ω - кругова частота; $A_2(\omega)$ - амплітуда коливань вихідного сигналу; $\phi(\omega)$ - зсув за фазою синусоїдальних коливань вихідного сигналу

Аналітично $W(j\omega)$ можна отримати з передатної функції заміною параметра перетворення Лапласа *p* на $j\omega$ з наступним виділенням модуля комплексного числа і фазового зсуву.

Крім названих методів при вирішенні задач управління широко використовують метод простору станів, метод компартментального аналізу і різні інформаційні методи.

Метод простору станів використовують для встановлення режимів поведінки системи, оцінки її динамічних властивостей, чутливості й прогнозування вихідних характеристик (див. п. 2.1).

При описуванні систем часто виявляється, що в них можна виділити деякі відносно незалежні кількості чи об'єми речовини, що беруть участь у кінетичних процесах як деяке ціле. У такому випадку при математичному описуванні покладають, що ця кількість речовини рівномірно розподілена в деякій області простору, тобто її концентрація у відповідній області постійна. Зручним способом вивчення таких систем є *компартментальтй аналіз.*

Компартментом звичайно називають деяку кількість речовини, що виступає як самостійна одиниця. Як правило, компартмент описують масою речовини, його концентрацією чи об'ємом, який ця речовина займає в просторі. Моделі систем, в яких використовують представлення про компартменти, називають компартментальними.

При компартментальному аналізі систем основну увагу приділяють процесам обміну, перетворення та утилізації компонентів.

Аналіз інформаційних характеристик систем включає оцінку кількості інформації, переданої через систему, швидкість передачі, пропускну здатність, перешкодостійкість, ефективність і надійність системи.

Кількість переданої через систему інформації визначають за формулою

$$I(X, Y) = H(x) - H\left(\frac{X}{Y}\right),$$

де H(X)- ентропія (невизначеність) сигналів на вході системи; H(X/У) - середня невизначеність того, який елемент був переданий системою, тобто результат дії шуму.

Швидкість передачі інформації розраховують за формулою

$$V_h = \frac{n}{T}\left[H(X) - H\left(\frac{X}{Y}\right)\right],$$

де V_h - швидкість передачі інформації; *n* - число елементів у інформації, яку передано через систему; *T* - час передачі інформації.

Пропускна здатність - це максимально можлива швидкість передачі інформації системою:

$$C = \max\left[H(X) - H\left(\frac{X}{Y}\right)\right]\frac{n}{T},$$

де *C* - пропускна здатність системи.

Реальні системи включають послідовно і паралельно з'єднані між собою елементи. Пропускна здатність послідовно з'єднаних елементів дорівнює мінімальному значенню з усіх значень пропускних здатностей окремих елементів:

$$C = \min C_i.$$

Пропускна здатність паралельно з'єднаних елементів дорівнює сумі пропускних здатностей окремих елементів:

$$C = \sum_i C_i.$$

Перешкодостійкість - це здатність системи протистояти дії перешкод. Перешкодостійкість оцінюють за показниками вірності прийому (W), середньої ймовірності помилки $(p_п)$, середньої невизначеності прийнятої інформації щодо переданого (H_ω):

$$W = \log\frac{1}{P_п} = \log\frac{1}{1 - P_{пр}};$$

$$P_п = \sum_{i=1}^{m} p(x_i)[1 - p(x_i, y_i)];$$

$$H_\omega = -(1 - p_п)\log(1 - p_п) - p_п \log p_п,$$

де $p_п$ - імовірність помилки; $p_{пр}$ - імовірність правильного прийому; $p(x_i)$ -

імовірність прийому x_i-того елемента; $p(x_i, y_i)$ - імовірність того, що буде прийнятий саме той елемент, що був переданий.

У практичних розрахунках за показник перешкодостійкості часто приймають відношення потужності сигналу P_c до потужності шуму $P_ш$:

$$\rho = \frac{P_c}{P_ш}.$$

Ефективність системи характеризують коефіцієнтом

$$\eta = \frac{V_h}{C}.$$

Надійність системи оцінюють імовірністю безвідмовної роботи, інтенсивністю відмов, наробітком на відмову та ін.

Імовірність безвідмовної роботи системи чи її елемента обчислюють за формулою

$$p(t) = e^{-\int_0^t \lambda(\tau) d\tau}$$

де λ - інтенсивність відмов; τ - період функціонування системи. Під інтенсивністю відмов розуміють число появи відмов в одиницю часу.

Якщо $\lambda = \text{const}$, тоді

$$p(t) = e^{-\lambda t}$$

Для складних інформаційних систем проблема надійності є принциповою проблемою, тому що при великому числі навіть якісних елементів частота відмов зростає. При послідовному з'єднанні елементів імовірність безвідмовної роботи системи дорівнює добутку імовірностей безвідмовної роботи елементів, що входять у систему:

$$p_s(t) = \prod_{i=1}^{n} p_i(t).$$

де $p_s(t)$ - імовірність безвідмовної роботи системи; $p_i(t)$ - імовірність безвідмовної роботи і-того елемента системи; *n* - число елементів у системі.

При паралельному з'єднанні елементів у системі імовірність її безвідмовної роботи визначають:

$$p_s(t) = \prod_{i=1}^{n} p_i(t) + \{p_i(t)[1 - p_{i+1}(t)] + p_{i+1}(t)[1 - p_i(t)]\}.$$

Під наробітком на відмову розуміють інтервал часу між двома послідовними відмовами.

Узагальнюючим показником якості системи є її *добротність*

$$Q = W\eta$$

де W- перешкодостійкість; η - коефіцієнт ефективності.

Елементи системи можуть бути залежними і незалежними. Імовірність стану системи у випадку незалежності станів її елементів, відповідно до теореми множення імовірностей визначають так:

$$p(x, y) = p(x)p(y).$$

Прологарифмуємо це рівняння:

$$\log p(x, y) = \log p(x) + \log p(y).$$

Ентропія системи може бути представлена у вигляді математичного очікування

$$H(X) = M[-\log P(X)].$$

Підстановка $\log p(x, y)$ в останню формулу дає

$$H(X, Y) = M[-\log P(x) - \log P(y)].$$

Таким чином,

$$H(X, Y) = H(X) + H(Y).$$

Дана рівність являє собою теорему складання ентропій, яка говорить, що при об'єднанні незалежних систем їхня ентропія складається.

Для довільного числа незалежних систем справедливе таке співвідношення:

$$H(X_1, X_2, \ldots, X_N) = \sum_{k=1}^{N} H(X_k),$$

що представляє собою загальний запис адитивності ентропії.

У випадку об'єднання залежних систем її ентропія дорівнює ентропії однієї з складових частин, що сумується з умовною ентропією другої частини щодо першої:

$$H(X, Y) = H(X) + H\left(\frac{Y}{X}\right) = H(Y) + H\left(\frac{X}{Y}\right).$$

У процесі досліджень завжди виникає питання, якою мірою суттєво впливає той чи інший фактор або їхня комбінація на досліджуваний процес? Методи встановлення основних факторів і їхній вплив на досліджуваний процес розглядають у спеціальному розділі теорії імовірностей і математичної статистики - дисперсійному аналізі (одно- і багатофакторному).

Суть однофакторного дисперсійного аналізу розглянемо на прикладі.

Нехай необхідно перевірити ступінь точності групи нівелірів (m приладів) і встановити, чи є їхні систематичні помилки однаковими, тобто вивчити вплив одного фактора на погрішність вимірювання. Кожним приладом виконано n вимірювань об'єкта. Усього виконано nm вимірювань. Позначимо окремі вимірювання через x_{ij}, де i - номер нівеліру; j - номер виконаного на цьому приладі вимірювання. Значення i змінюються від 1 до m, а j - від 1 до n.

Дисперсійний аналіз допускає, що відхилення підкоряються нормальному закону розподілу. Тому порядок аналізу зводиться до наступного.

Обчислюють для кожної серії вимірювання середньоарифметичне значення і середнє з показань першого нівеліра і т. п. для кожного з n_i вимірювань і m_i нівелірів. У результаті таких розрахунків установлюють величину суми квадратів відхилень між вимірюваннями серій (Q_1). Вона показує ступінь розбіжності в систематичних погрішностях усіх приладів, тобто характеризує розсіювання досліджуваного фактора між приладами. Дана сума розраховується за формулою

$$Q_1 = n\sum_{i=1}^{m}\left(\bar{x}_i - \bar{x}\right)^2,$$

де $\bar{x}_i$ - середньоарифметичне для *n* вимірювань; $\bar{x}$ - середньоарифметичне для всіх серій вимірювань, тобто загальне середнє значення.

Далі визначають суму квадратів відхилень усередині серії (Q_2), що характеризує залишкове розсіювання випадкових погрішностей одного нівеліра:

$$Q_2 = \sum_{i=1}^{m}\sum_{j=1}^{n}\left(x_{ij} - \bar{x}_i\right)^2,$$

де x_{ij} - окреме i-те вимірювання на j -тому нівелірі.

Метод дисперсійного аналізу допускає гіпотезу: центри нормальних розподілів випадкових величин рівні (чи рівні з деяким ступенем точності). Отже, усі mn вимірювань можна розглядати як вибірку з однієї й тієї ж нормальної сукупності. Щоб переконатися в можливості такого припущення, обчислюють показник

$$J = \frac{\frac{Q_1}{(m-1)}}{\frac{Q_2}{m(n-1)}}.$$

Можна побачити, що чисельник і знаменник цього показника - це дисперсії σ^2 для *m* і *nm* спостережень. Залежно від чисельних значень $K_1 = m-1$ і $K_2 = m(n-1)$ числа ступенів свободи та ймовірності p (наприклад, 0,95, 0,99 та ін.) складені табличні значення даного показника J_T. Якщо $J \le J_T$, тоді гіпотеза стверджується, тобто в даному прикладі нівеліри мають однакові (допустимі) систематичні помилки. При $J > J_T$ гіпотеза не стверджується.

Дисперсійний аналіз називають багатофакторним, якщо він має два фактори і більше. Суть його не відрізняється від однофакторного, але ускладнюється викладення і суттєво збільшується кількість розрахунків.

Для дослідження складних процесів імовірнісного характеру можна використовувати *метод Монте-Карло.* За допомогою цього методу вирішують широке коло задач, у яких ставлять мету відшукати найкраще вирішення з безлічі розглянутих варіантів: відшукати найкращий варіант розміщення баз, складів, підприємств; визначити оптимальну кількість автомобілів, що обслуговують екскаватор чи змішувач; установити найкращі параметри продукції, що випускається; уточнити пропускну здатність транспортних шляхів та ш.

Метод Монте-Карло часто називають методом статистичного моделювання чи статистичних випробувань. Він заснований на використанні випадкових чисел, що моделюють імовірнісні процеси, і являє собою чисельний метод вирішення складних задач. Результати вирішення дозволяють установити емпіричні залежності досліджуваних процесів. Математичною основою методу є закон великих чисел, що згідно П.Л. Чебишеву формулюється так: *при великому числі статистичних випробувань імовірність того, що середньоарифметичне значення випадкової величини прагне до її математичного очікування, дорівнює 1,0, тобто*

$$\lim p\left\{\left|\frac{\sum x_i}{n} - m(x)\right| < \varepsilon\right\} \to 1, \qquad n \to \infty$$

де ε - будь-яке мале позитивне число

З даної формули видно, що в міру збільшення числа випробувань середньоарифметичне необмежено (асимптотично) наближається до математичного очікування.

Послідовність вирішення задач методом Монте-Карло така: збір, обробка та аналіз статистичних спостережень досліджуваного процесу; відбір головних і відкидання другорядних факторів і складання адекватної математичної моделі (рівнянь, графіків, циклограм та ін.); складання алгоритмів і вирішення задачі на ЕОМ.

Для вирішення задачі методом Монте-Карло необхідно мати статистичний ряд, знати закон його розподілу, середнє значення *x* і математичне очікування $m(\bar{x})$, а також середньоквадратичне відхилення.

При нормальному законі розподілу випадкових величин точність результатів, отриманих за методом Монте-Карло, можна оцінити за формулою

$$p|\bar{x} - m(x)| < \frac{3\sigma}{\sqrt{n}}.$$

За допомогою даної формули можна вирішити задачу забезпечення заданої точності вирішення, тобто $\bar{x} \to m(\bar{x})$.

Нехай, наприклад, за умовою задачі задана допустима помилка ε_d. Якщо при наявному числі ряду n_i і σ_1 помилка ε_{d1} виявиться більше, ніж ε_d, то збільшують число випробувань до n_2 і обчислюють нове значення помилки доки не буде дотримана умова $\varepsilon_{d1} \le \varepsilon_d$ (і - число випробувань).

Вирішення задач методом Монте-Карло ефективне лише у випадку використання швидкодіючих ЕОМ.

На практиці вирішення задач транспортного будівництва, вантажних і пасажирських перевезень широке поширення отримала *теорія масового обслуговування (ТМО).* Методами цієї теорії вирішують задачі відшукання умов, що забезпечують високу ефективність роботи системи «вимога-обслуговування». Під обслуговуванням розуміють задоволення потреби у який-небудь заявці. Наприклад, навантаження щебеню у кар'єрі в автомобілі-самоскиди. Тут у якості вимоги виступає необхідність подачі під навантаження автомобілів за заявками, а в якості обслуговування - навантаження щебеню навантажувальними засобами (екскаватором, автонавантажувачем, транспортером та ін.). Таким чином, система обслуговування включає потік вимог, обслуговуючий пристрій і вихідний потік. Залежно від умов функціонування системи число вимог складає чергу на обслуговування. Так, при надлишку автомобілів неминуче виникають простої під навантаженням.

Основними характеристиками системи масового обслуговування є інтенсивність надходження вимог чи заявок на обслуговування λ, інтенсивність обслуговування (пропускна здатність системи) μ, коефіцієнт використання системи $\varphi = \lambda / \mu$, час очікування в черзі до обслуговування t_o, час безпосереднього обслуговування $t_{об}$, загальна тривалість обслуговування $t_1 = t_0 + t_{об}$, число вимог у черзі n, математичне очікування числа вимог у системі n_c. Дані характеристики знаходяться в такому взаємозв'язку: $t_{об} = 1/\mu$; $n = n_c \varphi$; $t_0 = n/\lambda$. У практичних розрахунках приймають середні з випадкових значень $\lambda, t_0, t_1, t_{об}, n$. Розподіл часу обслуговування за тривалістю найчастіше виражають показовим законом. Звичайно інтенсивність обслуговування вище інтенсивності вимог. Але черги на обслуговування усе-таки виникають, оскільки інтервал між обслуговуваннями неритмічний.

Задачею теорії масового обслуговування є встановлення найбільш достовірних залежностей між інтенсивністю потоку вимог і пропускною здатністю системи, кількістю вимог і ефективністю роботи системи.

При звичайному розрахунку необхідну кількість самоскидів N_a визначають за формулою

$$N_a = \frac{t_2}{t_3}.$$

де t_2 - повний час одного циклу роботи самоскида; t_3- час навантажування самоскида.

Ця кількість самоскидів виявляється достатньою тільки при строгому дотриманні графіка роботи, циклічності подачі автомобілів, високій надійності роботи навантажувача і самоскидів. Але на практиці такої роботи не спостерігається. Час t_2 неоднаковий і змінюється на 100-200% у менший і більший бік від середнього значення. Час навантажування t_3 також неоднаковий і носить імовірнісний характер. Зі збільшенням відсотка використання навантажувача зростає потреба в автомобілях самоскидах, а варіабельність часу навантажування приводить до формування черг. Використання звичайних методів розрахунку необхідної кількості автомобілів приводить до того, що навантажувач використовується не цілком(максимум на 75%), а простої автомобілів під навантаженням досягають 10%. Тому аналіз роботи системи слід виконувати імовірнісними методами, зокрема методами теорії масового обслуговування.

4.2. Проведення математичного дослідження

Математичне дослідження є не що інше, як експеримент над математичною моделлю. Задачею дослідження є прогноз функціонування чи розвитку досліджуваного об'єкта. Для її вирішення задають зовнішні впливи на об'єкт протягом розглянутого часу і за допомогою побудованої математичної моделі визначають динаміку перемінних чи розподіл імовірнісних значень параметрів, що характеризують систему.

Математичне дослідження передбачає: виявлення причинно-наслідкових відносин у досліджуваному об'єкті, ідентифікацію параметрів математичної моделі, визначення особливостей поведінки функцій, відшукування точок екстремуму, інтервалів монотонності, опуклості функцій, установлення керованих параметрів і параметрів, які управляють, оптимізацію рішень за заданими критеріями, установлення коректності поставленої задачі, визначення наявності вирішення задачі та одиничності цього вирішення.

Для одержання прогнозу функціонування чи розвитку об'єкта виконують аналіз результатів дослідження на відповідність заданим цілям і критеріям, розробляють пропозиції з поліпшення виду математичної моделі. Цей процес повторюють доти, поки не буде отримано задовільного результату.

Ідентифікація параметрів математичної моделі - це процес визначення їхніх кількісних значень. Для відшукування значень параметрів використовують експериментальні дані і методи, що дозволяють мінімізувати погрішність чи інтегральну помилку.

Звичайно відшукування параметрів моделі є громіздкою процедурою і вимагає використання обчислювальної техніки. Одночасно використовують і формальні методи, розроблені в математиці.

Але надалі кожен результат, отриманий після використання формальних методів, піддають перевірці на фізичну, біологічну та ін. коректність.

У складі формальних методів часто використовують методи відшукування екстремуму функцій чи функціоналів: метод Ньютона, градієнтний, найшвидкішого спуску в просторі параметрів моделі, випадкового пошуку та інші.

Якщо модель об'єкта чи процесу представлена алгебраїчним або тригонометричним поліномами, тоді оцінку параметрів здійснюють за інтерполяційними формулами Лагранжа, Ньютона, за методом найменших квадратів. Для ступінних поліномів часто використовують метод Крамера.

Нехай, наприклад, маємо таблицю експериментальних даних. Ступінь інтерполяційного полінома визначається кількістю експериментальних точок в таблиці, тобто кількістю її вузлів. Взагалі ступінь полінома на одиницю менше числа вузлів таблиці.

Припустимо далі, що модель процесу описується поліномом ступеня *n,*

$$P(x) = a_0 + a_1 x + \ldots + a_n x^n,$$

для якого виконуються умови:

$$P(x_0) = y_0, \quad P(x_1) = y_1, \ldots, \quad P(x_n) = y_n.$$

Задача ідентифікації буде вирішена, якщо будуть визначені значення невідомих коефіцієнтів $a_0, a_1, \ldots, a_n$.

Для визначення коефіцієнтів моделі процесу використовуємо крайові умови і запишемо систему з $(n+1)$ лінійних алгебраїчних рівнянь з $(n+1)$ невідомими коефіцієнтами $a_0, a_1, \ldots, a_n$

$$
\begin{aligned}
&a_0 + a_1 x_0^1 + \ldots + a_n x_0^n = y_0;\\
&a_0 + a_1 x_1^1 + \ldots + a_n x_1^n = y_1;\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_0 + a_1 x_n^1 + \ldots + a_n x_n^n = y_n;
\end{aligned}
$$

Дана система рівнянь має єдине вирішення, якщо її визначник

$$
A = \begin{vmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^n \\ 1 & x_1 & x_1^2 & \ldots & x_1^n \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x_n & x_n^2 & \ldots & x_n^n \end{vmatrix} \neq 0.
$$

Відомо, що це визначник Вандермонда:

$$
\begin{aligned}
A = &\left[(x_n - x_0)(x_n - x_1)\ldots(x_n - x_{n-1})\right]\left[(x_{n-1} - x_0)(x_{n-1} - x_1)\ldots(x_{n-1} - x_{n-2})\right]\ldots\\
&\ldots\left[(x_2 - x_0)(x_2 - x_1)\right](x_1 - x_0);
\end{aligned}
$$

причому $A \neq 0$, якщо всі табличні значення аргументу відмінні один від другого, тобто жодна з дужок у добутках не дорівнює нулю.

Таким чином, вирішуючи отриману систему алгебраїчних рівнянь методом Крамера, одержимо

$$
a_k = \frac{A_k}{A},
$$

де A_k - визначник A, у якому замість k-того стовпця стоїть стовпець вільних членів $y_0, y_1, \ldots, y_n$.

Крім розглянутого методу для оцінки параметрів моделі може використовуватись інтерполяційна формула Лагранжа

$$P(x)=\frac{(x-x_1)(x-x_2)\dots(x-x_n)}{(x_0-x_1)(x_0-x_2)\dots(x_n-x_{n-1})}y_0+$$
$$+\frac{(x-x_0)(x-x_2)\dots(x-x_n)}{(x_1-x_0)(x_1-x_2)\dots(x_1-x_n)}y_1+\dots$$
$$\dots+\frac{(x-x_0)(x-x_1)\dots(x-x_{n-1})}{(x_n-x_0)(x_n-x_2)\dots(x_n-x_{n-1})}y_n.$$

Якщо розглянуті значення аргументу відрізняються один від другого на одну і ту ж величину, тобто таблиця задана з постійним кроком, то ідентифікація параметрів значно спрощується. Для цього обчислюють різниці значень функції $y=f(x)$ у суміжних вузлах:

$$y_1-y_0=f(x_0+h)-f(x_0)=\Delta f(x)=\Delta y_0;$$
$$y_2-y_1=f(x_0+2h)-f(x_0+h)=\Delta f(x_0+h)=\Delta y_1;$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$
$$y_n-y_{n-1}=f(x_0+nh)-f(x_0+(n-1)h)=\Delta f(x_0+(n-1)h)=\Delta y_{n-1};$$

Ці різниці називають різницями першого порядку. Можна побудувати різниці другого порядку, використовуючи різниці першого порядку:

$$\Delta^2 y_0=\Delta y_1-\Delta y_0;\ \Delta^2 y_1=\Delta y_2-\Delta y_1;\ \Delta^2 y_{n-2}=\Delta y_{n-1}-\Delta y_{n-2};\ \dots;$$
$$\Delta^2 y_m=\Delta y_{m+1}-\Delta y_m.$$

Різниці k-того порядку при використанні різниць $(k-1)$-го порядку складають:

$$\Delta^k y_m=\Delta^{k-1}y_{m+1}-\Delta^{k-1}y_m.$$

Використовуючи зазначені різниці обчислюють коефіцієнти полінома

$$a_n=\frac{\Delta^n y_0}{n!h^n},$$

де h - крок між вузлами.

Застосування методу найменших квадратів для ідентифікації параметрів може бути проілюстровано наступним прикладом.

Нехай у результаті експериментального дослідження залежності величини y від величини x отримано ряд значень величини y при різних значеннях x:

$$\begin{array}{c||cccc} x & x_1 & x_2 & \dots & x_N \\ y & y_1 & y_2 & \dots & y_N \end{array}.$$

Між величинами x і y обрана функціональна залежність

ё

Якщо вимірювання не містять випадкових помилок, тоді для визначення невідомих параметрів $a_0, a_1, \dots, a_n$, досить було б провести стільки досліджень, скільки невідомих містить обрана функціональна залежність, і отримати при цьому $n + 1$ пару значень x, y:

$$\begin{array}{c||cccc} x & x_1 & x_2 & \dots & x_{n+1} \\ y & y_1 & y_2 & \dots & y_{n+1} \end{array}.$$

Підставивши ці пари в обрану функціональну залежність, отримують систему з $n + 1$ рівнянь з n невідомими

$$y = f(x_k, a_0, \dots, a_n),\ k = 1, 2, \dots, n + 1 \dots$$

Вирішення даної системи дозволяє визначити параметри, що відшукують.

На практиці величини y_k містять у собі випадкові помилки. Лежандр і Гаус у цьому випадку запропонували для визначення $a_0, a_1, \dots, a_n$ використовувати умову мінімуму суми квадратів нев'язань. Якщо при цьому усі вимірювання $y_1, y_2, \dots, y_n$ виконані з однаковою точністю, то формують суму

$$S = \sum_{k-1}^{n} \left[f\left(x_k, a_0, a_1, \dots, a_n\right) - y_k \right]^2.$$

Для знаходження невідомих $a_0, a_1, \dots, a_n$, що мінімізують S, необхідно скласти рівняння

$$\frac{\partial S}{\partial a_0} = 0;\ \frac{\partial S}{\partial a_1} = 0;\ \dots,\ \frac{\partial S}{\partial a_n} = 0,$$

які називаються нормальними.

Система нормальних рівнянь містить стільки невідомих, скільки рівнянь. Тому в загальному випадку ця система завжди визначена і вирішується алгебраїчними методами.

Якщо вимірювання y_k нерівноточні, тобто виконані з різними дисперсіями $\sigma_1, \sigma_2, \ldots, \sigma_n$ тоді в цьому випадку мінімізують суму

$$S_p = \sum_{k=1}^{n} \left[f(x_k, a_0, a_1, \ldots, a_n) - y_k \right]^2 p_k ,$$

де p_k - ваги вимірювань, зворотно пропорційні дисперсіям, що обчислюють за формулою

$$p_k = \frac{\sigma_0^2}{\sigma_k^2}$$

тут σ_0 - довільне позитивне число, найкраще, коли його значення близьке до середнього з σ_k.

Аналогічно ідентифікують параметри функцій двох і більше перемінних. Так, для функції Z від двох перемінних x, *y* оцінки невідомих параметрів $a_0, a_1, \ldots, a_n$ визначають з умови мінімуму суми

$$S = \sum_{k=1}^{n} \left[z_k - f(x_k, y_k, a_0, a_1, \ldots, a_n) - y_k \right]^2 p_k .$$

Числові значення виходять при вирішенні системи нормальних рівнянь:

$$\frac{\partial S}{\partial a_0} = 0; \quad \frac{\partial S}{\partial a_1} = 0; \ldots, \quad \frac{\partial S}{\partial a_n} = 0.$$

Якщо в обрану математичну модель параметри $a_0, a_1, \ldots, a_n$ входять лінійно, тоді розглянута нормальна система рівнянь також буде лінійна щодо невідомих параметрів.

Обчислені у такий спосіб параметри лінійно залежать від випадкових чисел y_k.

У випадку незалежних помилок, отриманих при вимірюванні y_k суттєво спрощується обчислення дисперсій параметрів, що відшукують. От чому при використанні методу найменших квадратів необхідно домагатися лінійності умовних рівнянь. Якщо ж параметри в обрану емпіричну формулу входять не лінійно, то проводять перетворення, що зводить задачу до випадку лінійної залежності.

Динамічні моделі об'єкта звичайно представляють у вигляді системи диференціальних рівнянь виду:

$$\frac{dy_1}{dt} = f_1(y_1,\dots,y_n, x_1,\dots,x_m, a_{11},\dots,a_{1l});$$
$$\frac{dy_2}{dt} = f_2(y_1,\dots,y_n, x_1,\dots,x_m, a_{21},\dots,a_{2l});$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$
$$\frac{dy_n}{dt} = f_n(y_1,\dots,y_n, x_1,\dots,x_m, a_{n1},\dots,a_{nl}).$$

Задача ідентифікації параметрів $a_{ij}(i=1,\dots,n;\ j=1,\dots,l)$ даної моделі складається у визначенні їхніх чисельних значень, що відповідають мінімальному відхиленню передбачених математичною моделлю динамічних залежностей $y_i(t)$, $(i=1,\dots,n)$, від наявних експериментальних даних $y_i^e(t)$.

Вирішення системи диференціальних рівнянь мають бути отримані для функцій $y_i(t_0)=y_{i0}(i=1,\dots,n)$, відповідних зовнішнім впливам і початковому стану об'єкта в експерименті.

Як міру відхилень модельних вирішень від експериментальних залежностей приймають один з функціоналів:

$$J[y_1(t),\dots,y_n(t)] = \int_{t_1}^{t_n} \sum_{i=1}^{n} \left[y_i(t) - y_i^e(t)\right]^2 dt;$$
$$J[y_1(t),\dots,y_n(t)] = \int_{t_1}^{t_n} \sum_{i=1}^{n} \left|y_i(t) - y_i^e(t)\right| dt;$$

де t_1, t_n - границі інтервалу спостережень.

Якщо при обраних зовнішніх впливах $x_k(t)$ виконуються умови існування й одиничності вирішення системи диференціальних рівнянь, розглянуті функціонали стають функціями кінцевого числа перемінних a_{ij} ,

$$\bar{J}[y_1(t),\dots,y_n(t)] = \bar{J}[a_{1l},\dots,a_{nl}];$$
$$\bar{J}[y_1(t),\dots,y_n(t)] = \bar{J}[a_{1l},\dots,a_{nl}],$$

а задача ідентифікації параметрів формулюється таким чином:

$$\bar{J}(\bar{a}_{1l}, \ldots, \bar{a}_{nl}) = \min_{a_{1l}, \ldots, a_{nl}} \bar{J}(a_{1l}, \ldots, a_{nl}):$$

$$\bar{J}(\bar{a}_{1l}, \ldots, \bar{a}_{nl}) = \min_{a_{1l}, \ldots, a_{nl}} \bar{J}(a_{1l}, \ldots, a_{nl}),$$

де $\bar{a}_{ij}$ - шукані значення параметрів моделі, що відповідають мінімальному значенню функціонала.

Якщо спостереження в експерименті виконують через рівні проміжки часу і для обчислення інтегралів застосовують формулу прямокутників, задача ідентифікації параметрів перетвориться до виду:

$$J[y_1(t), \ldots, y_n(t)] = \sum_{S=1}^{N-1} \sum_{i=1}^{n} \left[y_i(t_s) - y_i^э(t_s)\right]^2 \Delta t;$$

$$J[y_1(t), \ldots, y_n(t)] = \sum_{S=1}^{N-1} \sum_{i=1}^{n} \left|y_i(t_s) - y_i^э(t_s)\right| \Delta t,$$

де $t_S(S = 1, \ldots, N)$ - момент спостережень; Δt - проміжок часу між сусідніми спостереженнями.

Через те, що величина Δt не залежить від вибору значень постійних $a_{1l}, \ldots, a_{nl}$, задача ідентифікації невідомих параметрів зводиться до мінімізації функцій:

$$\hat{J}(a_{1l}, \ldots, a_{nl}) = \sum_{S=1}^{N} \sum_{i=1}^{n} \left[y_i(t_s) - y_i^э(t_s)\right]^2;$$

$$\hat{J}(a_{1l}, \ldots, a_{nl}) = \sum_{S=1}^{N} \sum_{i=1}^{n} \left|y_i(t_s) - y_i^э(t_s)\right|,$$

причому підсумовування доцільне проводити за всіма спостереженнями (S змінюється від 1 до N).

При багаторазовому повторенні експериментів над досліджуваним об'єктом у тих самих умовах функції $y_i^e(t)$ мають бути замінені їх середніми значеннями $\bar{y}_i^e(t)$ за серією експериментів.

Якщо експериментальні дані отримані при різних зовнішніх впливах $x_k(t)$ і для різних початкових станів об'єкта, тоді як міру неузгодженості теоретичних і експериментальних даних можна прийняти узагальнений функціонал, що представляє собою зважену суму функціоналів

$$L = \alpha \int_{t_1}^{t_N} \sum_{i=1}^{n} \left[y_i(t) - y_i^э(t)\right]^2 dt + \beta \int_{t_1}^{t_N} \sum_{i=1}^{n} \left|y_i(t) - y_i^э(t)\right| dt,$$

де α, β - коефіцієнти ваги.

Якщо система диференціальних рівнянь, що описують динамічну модель об'єкта, може бути вирішена в елементарних функціях, тоді для вирішення задачі ідентифікації можуть бути використані результати класичного математичного аналізу. Для цього у функціонал, що використовується як міра відхилень, підставляють вирішення диференціальних рівнянь:

$$\bar{J}(a_{1l}, \ldots, a_{nl}) = \int_{t_1}^{t_N} \sum_{i=1}^{n} \left[y_i(t, a_{1l}, \ldots, a_{nl}) - y_{ij}^{э}(t) \right]^2 dt,$$

Якщо інтеграл у даному функціоналі обчислюють в елементарних функціях, тоді величина $\bar{J}(a_{1l}, \ldots, a_{nl})$ представляється у вигляді

$$\bar{J}(a_{1l}, \ldots, a_{nl}) = \Phi(a_{1l}, \ldots, a_{nl}).$$

Параметри моделі визначають способом вирішення системи рівнянь

$$\frac{\partial \Phi(a_{1l}, \ldots, a_{nl})}{\partial a_{1l}} = 0;$$
$$\frac{\partial \Phi(a_{1l}, \ldots, a_{nl})}{\partial a_{2l}} = 0;$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\frac{\partial \Phi(a_{1l}, \ldots, a_{nl})}{\partial a_{nl}} = 0.$$

Дана система рівнянь нелінійна, тому її вирішення не обов'язково однозначне. Якщо вирішення задачі ідентифікації існує, воно має належати безлічі рішень розглянутої системи рівнянь. Це системи nl рівнянь відносно nl невідомих параметрів $(a_{1l}, \ldots, a_{nl})$.

Для спрощення позначень перенумеруємо невідомі параметри, привласнивши їм одне-цифрові індекси і ввівши для них нові позначення:

$$a_{11}, a_{12}, \ldots, a_{nl} \to b_1, b_2, \ldots, b_r, \ (r \to nl).$$

Тоді розглянута система рівнянь перепишеться так:

$$\frac{\partial \Phi(b_1, ..., b_r)}{\partial b_1} = 0;$$
$$\frac{\partial \Phi(b_1, ..., b_r)}{\partial b_2} = 0;$$
$$..........................$$
$$\frac{\partial \Phi(b_1, ..., b_r)}{\partial b_r} = 0.$$

Позначимо вирішення цієї системи рівнянь через $b_1^{\alpha},...,b_r^{\alpha}$.

Через те, що в стаціонарних точках функція $\Phi(b_1,...,b_r)$ може приймати як мінімальні, так і максимальні значення, необхідно досліджувати квадратичну форму

$$d^2\Phi(b_1^{\alpha}, ..., b_r^{\alpha}) = \sum \frac{\partial^2 \Phi(b_1^{\alpha}, ..., b_r^{\alpha})}{\partial b_i \partial b_j} \Delta b_i \Delta b_j ,$$

де $\Delta b_i, \Delta b_j$ - збільшення параметрів b_i, b_j.

Стаціонарні точки, в яких квадратична форма позитивно визначена $\left(d^2\Phi\left(b_1^2,...,b_r^2\right) > 0\right)$ для будь-яких $\Delta b_i, \Delta b_j$ є точками локального мінімуму функції $\Phi(b_1,...,b_r)$.

Права частина рівності квадратичної форми симетрична. Це дозволяє вважати квадратичну форму позитивно визначеною, якщо позитивний кожен з визначників:

$$\frac{\partial^2\Phi}{\partial b_1^2}, \begin{vmatrix} \frac{\partial^2\Phi}{\partial b_1^2} & \frac{\partial^2\Phi}{\partial b_1 \partial b_2} \\ \frac{\partial^2\Phi}{\partial b_2 \partial b_1} & \frac{\partial^2\Phi}{\partial b_2^2} \end{vmatrix}, ..., \begin{vmatrix} \frac{\partial^2\Phi}{\partial b_1^2} & \frac{\partial^2\Phi}{\partial b_1 \partial b_2} & \cdots & \frac{\partial^2\Phi}{\partial b_1 \partial b_r} \\ \frac{\partial^2\Phi}{\partial b_2 \partial b_1} & \frac{\partial^2\Phi}{\partial b_2^2} & \cdots & \frac{\partial^2\Phi}{\partial b_2 \partial b_r} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2\Phi}{\partial b_r \partial b_1} & \frac{\partial^2\Phi}{\partial b_r \partial b_2} & \cdots & \frac{\partial^2\Phi}{\partial b_r^2} \end{vmatrix}.$$

Вирішення задачі ідентифікації належить безлічі знайдених таким чином точок локального мінімуму. Якщо ця безліч містить більше однієї точки, необхідно обчислити значення функції $\Phi(b_1,...,b_r)$ у всіх точках локального мінімуму. Вирішенням задачі ідентифікації слід вважати ті значення параметрів $b_1^{\alpha},...,b_r^{\alpha}$ при яких величина $\Phi(b_1^{\alpha},...,b_r^{\alpha})$ мінімальна.

Якщо рівняння математичної моделі не інтегруються в елементарних функціях, тоді для їхнього вирішення можуть використовуватись чисельні методи. У цьому випадку при вирішенні задачі ідентифікації задаються деяким початковим наближенням для параметрів, які відшукують:

$$a_{11} = a_{11}^0,\ a_{12} = a_{12}^0,\ ...,\ a_{nl} = a_{nl}^0,$$

які підставляють у рівняння моделі:

$$\frac{dy_1}{dt} = f_1(y_1, ..., y_n, x_1, ..., x_m, a_{11}, ..., a_{1l});$$

$$\frac{dy_2}{dt} = f_2(y_1, ..., y_n, x_1, ..., x_m, a_{21}, ..., a_{2l});$$

....

$$\frac{dy_n}{dt} = f_n(y_1, ..., y_n, x_1, ..., x_m, a_{n1}, ..., a_{nl}).$$

Дані рівняння при $x_k = x_k(t)$, $(k = 1,...,m)$ і початкових умовах $y_i(t) = y_{i0}, (i = 1,...,n)$ відповідних зовнішнім впливам і початковому стану об'єкта в експерименті, інтегрують одним з методів чисельного інтегрування. Отримані вирішення підставляють у функцію ідентифікації й обчислюють значення функціонала $\widehat{J}_0 = \widehat{J}_0\left(a_{11}^0,...,a_{nl}^0\right)$. Якщо $\widehat{J}_0 = 0$ тоді значення перемінних можуть бути поліпшені за допомогою прямих методів мінімізації функції $\widehat{J}(a_{11},...,a_{nl})$ тому що необхідні умови екстремуму не можуть бути записані в явному виді взагалі. У цьому випадку задаються довільним набором значень параметрів, що відшукують

$$b_1 = b_1^0, \quad b_2 = b_2^0,...,b_r = b_r^0, \quad b_1 = b_1^0,$$

і обчислюють величину Φ при цих значеннях $\Phi_0 = \Phi(b_1^0,...,b_r^0)$. Потім, використовуючи інформацію про властивості функції $\Phi(b_1,...,b_r)$, визначають збільшення Δb_i так, щоб значення функції Φ, обчислене при нових значеннях параметрів $b_1^1 = b_1^0 + \Delta b_1$, $b_2^1 = b_2^0 + \Delta b_2,...,b_r^1$, виявилось менше Φ^0 тобто

$$\Phi_1 = \Phi(b_1^1, b_2^1, ..., b_r^1) < \Phi^0.$$

Якщо $\Phi_1 \neq 0$, описаний процес повторюють доти, поки або Φ звернеться в нуль, або будуть знайдені такі значення параметрів $b_i = b_i^\alpha$, $(i = 1,...,r)$, для яких не існує збільшень Δb_i, що дозволяють зменшити отримане значення $\Phi^\alpha = \Phi\left(b_1^\alpha, b_2^\alpha,...,b_r^\alpha\right)$. Ці значення параметрів відповідають точці локального мінімуму функції $\Phi(b_1,...,b_r)$.

На практиці часто обмежуються наближеним визначенням параметрів

$\bar{b}_i^{\alpha}$ вважаючи, що локальний мінімум знайдений, якщо величина Ф близька до нуля

$$\Phi(\bar{b}_1^{\alpha}, \bar{b}_2^{\alpha}, \ldots, \bar{b}_r^{\alpha}) \leq \delta.$$

Або навіть якщо збільшення функції Ф при зміні значень кожного з аргументів, віднесені до збільшень Δb_i малі:

$$\frac{\left|\Phi\left(\bar{b}_1^{\alpha} + \Delta b_1, \bar{b}_2^{\alpha}, \ldots, \bar{b}_i^{\alpha}\right) - \Phi\left(\bar{b}_1^{\alpha}, \bar{b}_2^{\alpha}, \ldots, \bar{b}_i^{\alpha}\right)\right|}{\Delta b_1} \leq \xi;$$

$$\frac{\left|\Phi\left(\bar{b}_1^{\alpha}, \bar{b}_2^{\alpha} + \Delta b_2, \ldots, \bar{b}_i^{\alpha}\right) - \Phi\left(\bar{b}_1^{\alpha}, \bar{b}_2^{\alpha}, \ldots, \bar{b}_i^{\alpha}\right)\right|}{\Delta b_2} \leq \xi;$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\frac{\left|\Phi\left(\bar{b}_1^{\alpha}, \bar{b}_2^{\alpha}, \ldots, \bar{b}_i^{\alpha} + \Delta b_1\right) - \Phi\left(\bar{b}_1^{\alpha}, \bar{b}_2^{\alpha}, \ldots, \bar{b}_i^{\alpha}\right)\right|}{\Delta b_i} \leq \xi.$$

Значення δ і ξ фіксують заздалегідь.

Існує безліч способів обчислення «поліпшуючих» збільшень, що спрямовані на зменшення трудомісткості задачі. Дуже часто задачу вибору величин Δb_i розбивають на дві: визначають напрямок «руху» у r-мірному просторі параметрів, вибирають довжину «кроку» у цьому напрямку. Вирішення задачі здійснюють методами відшукування екстремуму: градієнтний метод; метод покоординатного спуску та ін.

Суть градієнтного методу пояснимо на наступному прикладі. Припустимо, що необхідно знайти екстремум функції $\Phi\left(\vec{b}_1^{\alpha}, \vec{b}_2^{\alpha}\right)$, що описує деяку поверхню (рис. 4.1). Знаходження екстремуму починають з будь-якої точки поверхні $A_0\left(\vec{b}_{01}^{\alpha}, \vec{b}_{02}^{\alpha}\right)$. З цієї точки визначають напрямок підйому чи спуску, що є найбільш крутим; його називають градієнтом і позначають $\vec{q}$. Звідси ж починають рух у напрямку за градієнтом до оптимуму з кроком $\vec{c}\vec{q}$, де $\vec{c}$ - постійна величина, що залежить від точності вимірювання чи оцінки ξ. У результаті одержують нову точку $A_1\left(\vec{b}_{11}^{\alpha}, \vec{b}_{22}^{\alpha}\right)$ в якій знову повторюють описану процедуру доти, поки не визначать екстремум.

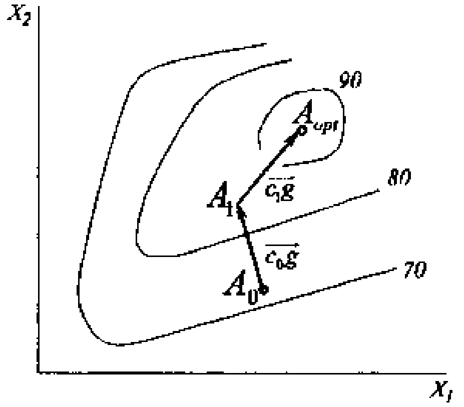

Рис. 4.1. Схема руху до оптимуму за градієнтом (круте сходження)

Для зменшення трудомісткості задачі використовують менш точні методи ідентифікації, але такі, що дають гарні результати. Розглянемо один з найбільш простих прийомів. Вихідну систему диференціальних рівнянь перетворять до вигляду:

$$\frac{dy_1}{dt} - f_1(y_1, ..., y_n, x_1, ..., x_m, a_{11}, ..., a_{1l}) = 0;$$
$$\frac{dy_2}{dt} - f_2(y_1, ..., y_n, x_1, ..., x_m, a_{21}, ..., a_{2l}) = 0;$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\frac{dy_n}{dt} - f_n(y_1, ..., y_n, x_1, ..., x_m, a_{n1}, ..., a_{nl}) = 0.$$

Замість $y_i(t)$ у дану систему рівнянь підставляють залежності $y_i^e(t)$, отримані в експерименті. При цьому для довільних значень параметрів $a_{1l},...,a_{nl}$ ліві частини диференціальних рівнянь не обов'язково звертаються в нуль:

$$\frac{dy_1^э}{dt} - f_1(y_1^э, ..., y_n^э, x_1, ..., x_m, a_{11}, ..., a_{1l}) = \xi_1(t);$$
$$\frac{dy_2^э}{dt} - f_2(y_1^э, ..., y_n^э, x_1, ..., x_m, a_{21}, ..., a_{2l}) = \xi_2(t);$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\frac{dy_n^э}{dt} - f_n(y_1^э, ..., y_n^э, x_1, ..., x_m, a_{n1}, ..., a_{nl}) = \xi_n(t).$$

Як міра неузгодженості теоретичних і експериментальних даних на інтервалі спостереження (t_1, t_N) може бути прийнятий функціонал

$$\varphi[\xi_1(t),...,\xi_n(t)] = \max,$$

Величини $y_i(t)$ залежать від параметрів моделі $a_{1l},...,a_{nl}$. Тому даний функціонал є функцією nl перемінних:

$$\varphi(a_{11},...,a_{nl}) = \max \sum_{i=1}^{N} \xi_i^2 \left(\underset{t_1 \le t \le t_N}{t}, a_{i1}, a_{i2}, ..., a_{il} \right).$$

У загальному випадку сума квадратів нев'язань у правій частині функціонала може бути замінена зваженою сумою

$$\overline{\varphi}(a_{11},...,a_{nl}) = \max \sum_{i=1}^{N} \xi_i^2 \left(\underset{t_1 \le t \le t_N}{t}, a_{i1}, a_{i2}, ..., a_{il} \right).$$

де $p_i > 0$ - вагові коефіцієнти, що характеризують відносну значимість вимірювань величин y_i^e у ході експерименту.

У такому разі задача ідентифікації параметрів моделі може бути сформульована як задача відшукування значень $a_{1l},...,a_{nl}$, що доставляють мінімум функції φ чи $\overline{\varphi}$:

$$\varphi[\xi_1(t),...,\xi_n(t)] = \max, \qquad t_1 \le t \le t_N.$$

Часто замість функціонала $\varphi[\xi_1(t),...,\xi_n(t)]$ для оцінки міри неузгодженості теоретичних і експериментальних даних використовують інтегральний функціонал

$$\chi[\xi_1(t), \xi_2(t), ..., \xi_n(t)] = \int_{t_1}^{t_N} \sum_{i=1}^{N} [\xi_i(t)]^2 \, dt,$$

що представляє собою, як і φ, функцію параметрів моделі, які відшукують.

Задача ідентифікації в цьому випадку формулюється у вигляді:

$$\chi(\overline{a}_{11}, ..., \overline{a}_{nl}) = \min \int_{t_1}^{t_N} \sum_{i=1}^{n} [\xi_i(t, a_{1_{i1}}, ... a_{il})]^2 dt,$$

Інтеграл у правій частині даної рівності має обчислюватись одним з методів чисельного інтегрування.

Застосування інтегрального функціонала для ідентифікації параметрів моделі знижує жорсткість обмежень, що накладаються на функцію $\xi_1(t)$. В результаті надійність обчислень параметрів моделі знижується. Таким чином, даний метод ідентифікації зручний для одержання більш-менш локальних параметрів моделі і для локальної перевірки правильності закладених у модель принципів.

4.3. Оптимізація результатів за заданими критеріями

На етапі аналізу й інтерпретації отриманого математичного результату часто здійснюють оптимізацію процесів, об'єктів чи їхніх окремих характеристик. Оптимізація означає вибір з нескінченної безлічі можливостей. Цей вибір здійснюють відповідно з прийнятими критеріями.

Критерій - це ознака, що дозволяє класифікувати процеси, характеристики та інше на гарні й погані з точки зору поставленої мети. Кількісне вираження ознаки називають *показником.*

Оптимізація дій може розглядатись як уроджена схильність людини Кароль Адамецький (1866 - 1933), читаючи лекції у Варшавському політехнічному інституті, привів приклад уродженої оптимізації. Він розглядав дороги, що з'єднують рибацьке селище з причалом човнів для лову риби (рис 4 2)

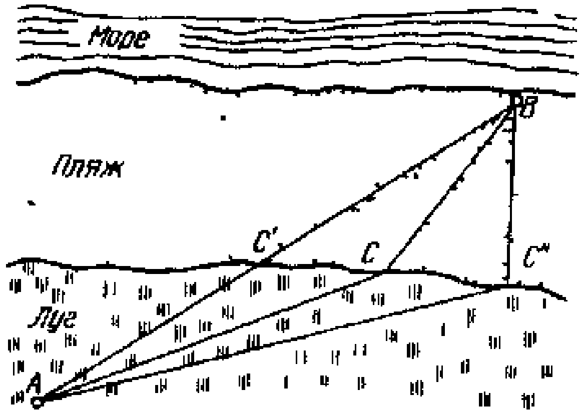

Рис 4 2 Дороги на місцевості з різною проходимістю

Точка А на даному рисунку вказує на початок шляху із селища, а точка В - на місце розташування причалу Дорога до причалу відповідала не прямій АВ, що проходить через точку С, а ламаній, що проходить через точку С Цей шлях - є результат інстинктивного прагнення до мінімізації роботи з подолання шляху з точки А в точку В Ця мінімізація враховує різні опори рухові на лузі і на піщаному пляжі

В аналітичній формі задача оптимізації вибору шляху руху рибалок представляється у вигляді

$$E = \alpha a + \beta b,$$

$$\left(\frac{\beta}{\alpha} \to 1\right) \Rightarrow \left(C \to C' \Rightarrow E_{min}\right),$$

$$\left(\frac{\beta}{\alpha} \to \infty\right) \Rightarrow \left(C \to C'' \Rightarrow E_{min}\right),$$

де *a* - довжина ділянки дороги на луці, b - довжина ділянки дороги на пляжі, α - коефіцієнт витрат енергії при русі лугом, β - коефіцієнт витрат енергії при русі пляжем, *E*- витрати енергії на подолання шляху ACB

Вимога мінімізації витрат енергії показує, що при $\alpha = \beta$ шлях руху з A у B виглядав би у вигляді прямої лінії AC'B При $\alpha > \beta$ шлях, що обирається, відповідає ламаній лінії AC"B.

У розглянутій задачі витрати енергії рибалок можуть бути обміряні Але у більшості практичних випадків існують лише обмежуючі критерії Тоді мірою служить границя, яку не слід переступати.

Коли використовують лише один критерій, то виконують безпосередню (пряму) оптимізацію Задачу безпосередньої оптимізації можна записати у вигляді (рис 4 3, а)

$$y = f(x) \underset{x}{\rightarrow} \min$$

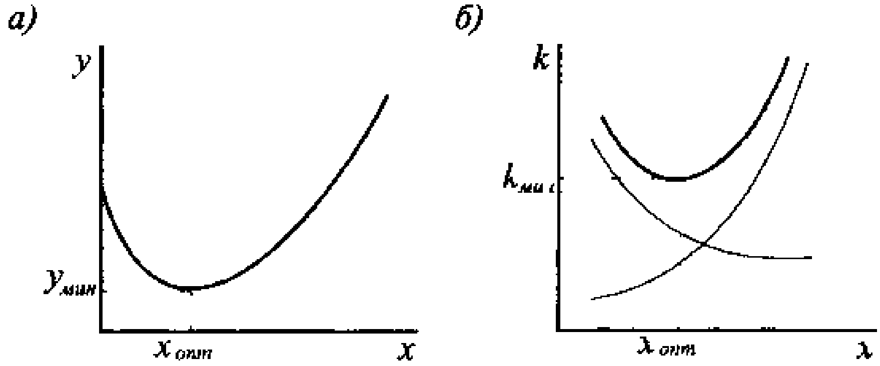

Рис 4 3 Приклад безпосередньої (*а*) оптимізації та оптимізації з протилежними критеріальними залежностями (*б*).

Таку функцію називають цільовою

Інший приклад безпосередньої оптимізації представлений на рис 4 3, б Цей випадок називають оптимізацією з протилежними критеріальними залежностями Модель оптимізації представляється у вигляді

$$K = K_1 + K_2 \underset{x}{\rightarrow} \min$$
$$K_1 = f(x), K_2 = f(x).$$

На практиці установити абсолютний оптимум не вдається, тому що неможливо розглянути незліченну безліч варіантів вирішення

Тому оптимізацію заміняють раціоналізацією, тобто знаходженням найкращого вирішення з практично розглянутих.

Вирішення задачі оптимізації може існувати й у багатомірному просторі $(x_1, x_2, ..., x_n)$. Якщо цільова функція і всі обмежувальні функції лінійні щодо перемінних $x_j\left(j=1,n\right)$, тоді такі задачі вирішуються методами лінійного програмування. Якщо до того ж поставлена умова: «усі (чи деякі) x_j - цілі числа», виникає задача цілочисельного лінійного програмування. У всіх інших випадках говорять про задачі нелінійного програмування. Запис задачі оптимізації у рамках лінійного програмування може бути представлений у вигляді

$$K=\omega_1 x_1+\omega_2 x_2+\ldots+\omega_n x_n \underset{x_j}{\rightarrow} \min(\max),$$

$$f_1=a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n \geq r_1,$$
$$f_2=a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n \geq r_2,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$f_m=a_{n1}x_1+a_{n2}x_2+\ldots+a_{mn}x_n \geq r_m.$$

Теперішнім часом задачі лінійного програмування вивчені досить повно, причому для їхнього вирішення потрібні порівняно нескладні обчислення. Для багатьох з них є стандартні програми для ЕОМ. Найбільше поширення отримали симплекс-метод і метод Гомори.

Вирішення задач нелінійного програмування базуються на методі невизначених коефіцієнтів Лагранжа. Представляючи собою великий самостійний інтерес, цей метод дозволив одержати ряд узагальнень, що привели до розробки алгоритмів вирішення задач з некласичними умовами (у першу чергу - задач опуклого квадратичного програмування).

Застосування методу Лагранжа можна розібрати на прикладі. Розглянемо систему, що складається з планувального органа (центра) і *n* підприємств-виробників. Завдання центра - призначити план кожному підприємству за умовою, що сумарний випуск продукції дорівнював би заданій кількості *R,* а сумарні витрати на виробництво продукції були мінімальними. Позначимо x_i- план, z_i- витрати і-того підприємства на випуск продукції в кількості x_i Приймемо далі, що z_i знаходиться в нелінійному зв'язку з x_i:

$$z_1(x_i)=\frac{1}{2r_i}x_i^2$$

де r_i - параметр (коефіцієнт) ефективності виробництва.

Враховуючи викладене, задача оптимізації представляється у вигляді

$$\sum_{i=1}^{n}\frac{1}{2r_i}x_i^2 \underset{x_i}{\rightarrow} \min, \qquad \sum_{i=1}^{n}x_i=R$$

Вирішення даної задачі починають із складання функції Лагранжа (лагранжіана):

$$L(x,\lambda)=\sum_{i=1}^{n}\left(\frac{1}{2r_i}x_i^2\right)-\lambda\left(\sum_{i=1}^{n}x_i-R\right)$$

де λ - невизначений коефіцієнт Лагранжа.

Умови існування оптимуму *x* представляються у вигляді системи рівнянь

$$\begin{cases}\dfrac{dL(x,\lambda)}{dx_i}=0\\ \dfrac{dL(x,\lambda)}{d\lambda}=0\end{cases}$$

Диференціюючи $L(x,\lambda)$ по x_i одержимо $x_i=\lambda r_i$. У свою чергу, диференціюючи $L(x,\lambda)$ по λ одержимо

$$\sum_{i=1}^{n}x_i=R$$

Підстановка знайденого значення x_i в останнє вираження дозволяє визначити значення λ:

$$\lambda=\frac{R}{\sum_{i=1}^{n}r_i}$$

Нарешті, зворотна підстановка знайденого значення λ вираження $x_i=\lambda r_i$ дає оптимальний план для і-го підприємства

$$x_i=\frac{r_i}{H}R,\qquad \text{де}\quad H=\sum_{i=1}^{n}r_i$$

Вирішення ряду задач управління динамічними процесами здійснюють за допомогою методу динамічного програмування.

Динамічне програмування (динамічне планування) - це математичний метод оптимізації вирішень, спеціально пристосований до багатокрокових (чи багатоетапних) операцій. Припустимо, що досліджувана операція - це процес, що розвивається в часі і розпадається на ряд «кроків» чи «етапів». Деякі операції розчленовують на кроки природно. Наприклад, при плануванні господарської діяльності групи підприємств природним кроком є господарський рік. В інших операціях поділ на кроки приходиться вводити штучно. Наприклад, процес тепловоложистої обробки залізобетонних виробів можна умовно розбити на етапи (підйом температури, ізотермічна витримка, остигання), кожен з яких займає деякий відрізок часу. Цей процес є керований. На кожному кроці приймають якесь вирішення, від якого залежить успіх даного кроку й операції в цілому.

В основу задач динамічного програмування покладені принципи оптимальності. Звичайно цільова функція виражається сумою

$$\omega = \sum_{k=0}^{N-1} f_0\left[x(k), u(k)\right] \to \max(\min)$$

де N - загальне число інтервалів (кроків); $u(k)$ - керуючі впливи; $x(k)$ - значення координати в дискретні моменти часу k.

При оптимальному управлінні даний функціонал має бути мінімізований (чи максимізований). Оптимальний процес стане відомий, якщо будуть знайдені значення керуючого впливу $u_0, u_1, \ldots, u_{N-1}$ і в усі дискретні моменти часу $k = 0, 1, \ldots, N-1$.

Щоб вирішити задачу динамічного програмування, необхідно відшукати максимум (чи мінімум) складної дискретної функції з великою кількістю перемінних. Метод динамічного програмування зводить цю задачу до простої - мінімізуються прості функції в зворотному порядку - від кінця до початку процесу.

Для оптимізації процесу методами нелінійного чи динамічного програмування немає стандартних вирішень. У кожному конкретному випадку застосовують свій підхід (свій метод) до вирішення задачі.

Іноді можливі випадки оптимізації тільки з обмежуючими критеріями без додаткових умов вибору екстремальних значень. Тут можливі два варіанти критеріїв: обмежуючі критерії, що сходяться і частково розбіжні обмежуючі критерії.

Прикладом обмежуючих критеріїв, що сходяться, можуть бути: критерій фізичних можливостей *A;* критерій технічних можливостей *B;* економічний критерій С(рис. 4.4).

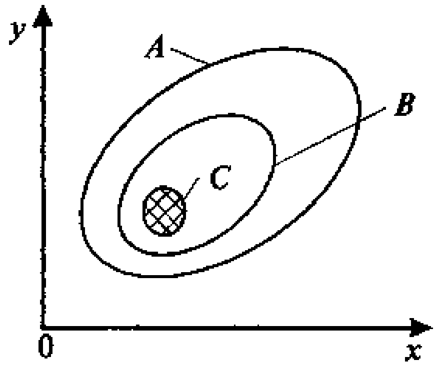

Рис. 4.4. Приклад критеріїв, що сходяться

Критерії, що сходяться, поступово виключають з можливої безлічі елементи, що знаходяться поза областю допустимих елементів. Це означає, що критерій більш низького рівня (критерій *A*) охоплює елементи безлічі, що вважаються допустимими з позицій критерію більш високого рівня (критерію *B*), а також і критерію С, тому що критерій *B* охоплює елементи, допустимі за критерієм С. Такі критерії являють собою як би сита з послідовно зменшуваними комірками. Приклад частково розбіжних критеріїв представлений на рис. 4.5. Тут оптимальне вирішення можливе лише у випадку, якщо

$$A \cap B \cap C \equiv X, \quad X \neq 0.$$

де $\cap$ - позначає приналежність (пересічення) однієї безлічі іншою.

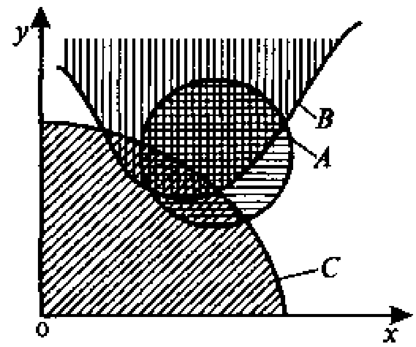

Рис. 4.5. Приклад частково розбіжних критеріїв

Іншими словами, для існування оптимального вирішення необхідна наявність пересічення всіх розглянутих безлічей.

Відсутність пересічення двох з розглянутих безлічей говорить про відсутність оптимального вирішення і наявність цілком розбіжних критеріїв (рис. 4.6). Такий випадок відповідає умові

$$A \cap C \equiv 0,$$

тобто пересічення безлічей А і С дорівнює нулю.

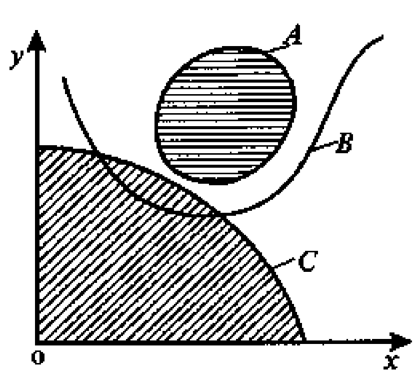

Рис. 4.6. Приклад цілком розбіжних критеріїв

Прикладом цілком розбіжних критеріїв є критерії якості управління дорожнім рухом. Так, вимога попередження забруднення навколишнього середовища і підвищення рівня безпеки руху показує, що оптимальним є повна заборона всякого дорожнього руху. Але таке вирішення вступає в конфлікт із вимогою найбільш ефективного використання доріг.

У деяких випадках цілком розбіжних критеріїв можна застосувати компромісний критерій. Найчастіше такий критерій одержують способом корекції основних критеріїв. Нехай, наприклад, конструктивні вирішення автомобіля приймають виходячи з трьох критеріїв: швидкість руху V, маса автомобіля m, надійність руху P. Задачу конструювання за цими критеріями формулюють так: необхідно забезпечити швидкість руху, перевищуючу мінімально допустиму, тобто $V \geq V_{min}$; маса автомобіля має бути менше максимально допустимої, тобто $m \leq m_{max}$; надійність руху має бути більше мінімально допустимої, тобто $P \geq P_{min}$. Нехай далі виконуються умови

$$\Delta = (V \geq V_{min}) \cap (m \leq m_{max}) \neq 0,$$
$$\Theta = (m \leq m_{max}) \cap (P \geq P_{min}) \neq 0,$$
$$\Gamma = (V \geq V_{min}) \cap (P \geq P_{min}) \neq 0.$$

Крім того $\Delta \cap \Theta \cap \Gamma = 0$, що свідчить про те, що ми маємо справу з цілком розбіжними критеріями. Отже, у даній постановці задачі оптимального вирішення немає. Вихід із ситуації, що створилася, може бути знайдений за допомогою корекції системи критеріїв.

Приймаючи до уваги, що вимога до надійності руху зростає, критерій надійності залишаємо незмінним і віддаємо йому пріоритет. Коректуванню піддаємо критерії швидкості і маси автомобіля (рис. 4.7).

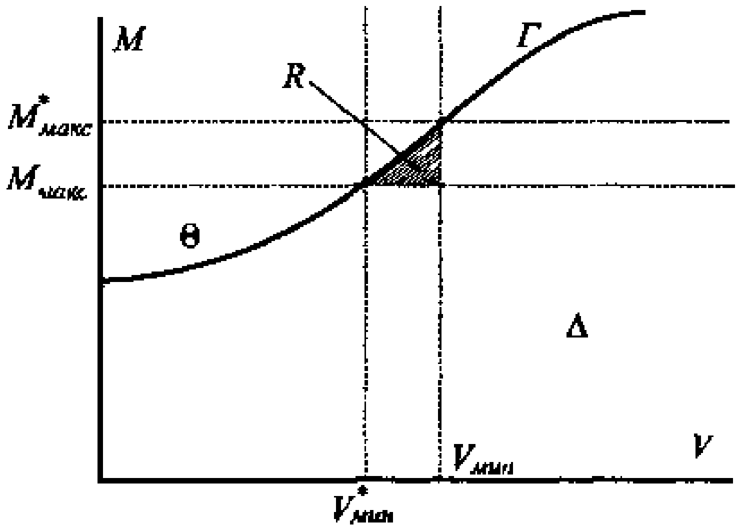

Рис. 4.7. Ефект зміни критеріальної швидкості ($v_{min} \rightarrow v^*_{min}$) і критеріальної маси ($M_{max} \rightarrow M^*_{max}$)

У результаті коректування може бути отримане нове (квазіоптимальне) вирішення, що відповідає зміненим вимогам

$$\left(V \geq V_{min}^{1}\right) \cap \left(m \geq m_{max}^{1}\right) \cap \left(P \geq P_{min}\right) = R.$$

При цьому величини маси і швидкості автомобіля являють собою безлічі, що відрізняються від нуля

$$\{V, m\} = R \neq 0.$$

На практиці досить часто проявляється прагнення до однокритеріальної оптимізації. Це пояснюється впливом економічних шкіл, що намагаються звести оцінки до грошового виміру. При цьому передбачається, ніби-то усе можна звести до такого вимірювання і визначити усілякі витрати деяким способом, що дозволяє одноманітно оцінити їх. Економічний підхід виходить з вигоди для людського суспільства (вигоди для себе). Це приводить часто до вирішень, що виявляються згубними для природи і самих людей.

Більш широкий підхід до проблеми оптимізації виявляється при використанні критеріїв соціальної адекватності (екологічних, ергономічних, естетичних та ін.) з обмеженнями по грошовим витратам.

Зведення багатокритеріальної задачі до однокритеріальної здійснюють способом формування *суперкритерію,* тобто скалярної функції векторного аргументу

$$q_0(x) = q_0\left[q_1(x), q_2(x), \ldots, q_p(x)\right].$$

Суперкритерій дозволяє упорядкувати альтернативи за величиною q_0, виділивши тим самим найкращу (у смислові цього критерію). Вигляд функції q_0 визначається тим, як ми уявляємо собі внесок кожного критерію в суперкритерій. Звичайно для формування суперкритерію використовують адитивні чи мультиплікативні функції

$$q_0 = \sum_{i=1}^{p} \frac{\alpha_i q_i}{s_i}; \qquad 1 - q_0 = \prod_{i-1}^{p} \left(1 - \frac{\alpha_i q_i}{s_i}\right);$$

де α_i - коефіцієнти, що відбивають відносний внесок часткових критеріїв у суперкритерій;

$$s_i = \sum_{i=1}^{p} \alpha_i.$$

Коефіцієнти α_i визначають за допомогою методу експертних оцінок, методу аналітичного синтезу узагальнених координат стану (тобто аналітичне визначення вагових коефіцієнтів), методу зважування координат стану чи системи локальних критеріїв на вагах механізму самозбереження системи.

Сутність методу експертних оцінок зводиться до статистичного усереднення постулатів окремих фахівців, що беруть участь в експертизі. Метод експертних оцінок свавільний, суб'єктивний і має інші недоліки, органічно властиві постулуванню.

У методі аналітичного синтезу покладають, що всі локальні критерії підвищуються, тобто з ростом q_i зростає і q_0. Суперкритерій представляють у вигляді

$$q_0 = \sum_{i=1}^{p} \nu_i q_i; \qquad \text{де} \quad \nu_i = \frac{\alpha_i}{\sum_{i=1}^{\Gamma} \alpha_i}.$$

Кожен локальний критерій визначений у вирішенні системи

$$\frac{dq_i}{dt} = \Phi(\cdot).$$

де $\Phi(\cdot)$ - вектор-функкція узагальненої сили.

Дане диференціальне рівняння визначене в області простору локальних критеріїв і вектора управління u

$$Z(q,u) \geq 0.$$

Для кожного локального критерію q_i відомо його граничне значення $q_{in}, (q_i \leq q_{in})$.

Крім локальних критеріїв, що входять у суперкритерій, можуть бути обмежені й інші локальні критерії, що не ввійшли в суперкритерій. Тому обмеження, що накладається на локальні критерії, доцільно представити у вигляді

$$\varphi_k(\cdot) \leq A_k, \qquad k \in [1,s], \qquad s \geq n.$$

де $\varphi_k(\cdot)$ - функціонал системи.

Функціонал $\varphi_k(\cdot)$ називають втратами, що супроводжують процес функціонування системи, що оцінюється. Тоді відносні втрати

$$\varphi_k(\cdot) = \frac{1}{A_k}\varphi(\cdot), \qquad \varphi_k(\cdot) \geq 0, \qquad k \in [1,s].$$

Для оцінки вагових коефіцієнтів ν_i локальних критеріїв, що входять у суперкритерій q_0, використовують два принципи:

1) принцип чебишевської рівномірної оптимізації, при якій процес функціонування вважається гарним, якщо при його виконанні досягається найбільш рівномірна зміна рівня усіх відносних втрат;
2) принцип інтегральної оптимізації, при якій динамічний процес оцінюється як гарний, якщо сума відносних втрат виявляється мінімальною.

У випадку застосування принципу інтегральної оптимізації задаються довільними значеннями вагових коефіцієнтів ν_i з деякої області їхніх допустимих значень Γ_ν. Потім методами варіаційного числення виконують оптимізацію процесу $dq/dt = \Phi(\cdot)$ із введеними обмеженнями. У результаті одержують аналітичне вираження для безлічі псевдооптимальних процесів

$$\{q(t,\nu), u(t,\nu)\}, \qquad t_0 \le t \le t_1, \qquad \nu \in \Gamma_\nu .$$

Підставляючи функцію відносних втрат у це вираження, одержують

$$\varphi_{0k} = \varphi_{0k}(\nu), \qquad k \in [1, s],$$

з якого виходить, що відносні втрати представляються тепер як функції вагових коефіцієнтів.

Надалі слід враховувати, що для виключення тривіальних вирішень на коефіцієнти звичайно накладається додаткова умова

$$\sum_{i=1}^{n} \nu_i = 1.$$

У випадку застосування принципу чебишевської рівномірної оптимізації вводиться система рівнянь

$$\nu_1 = 1, \quad \varphi_{0j}(\nu) = \varphi_{0,\,j=s}(\nu), \quad j \in [1, s-1]$$

і неправильна міра її існування у вигляді

$$W(\nu) = \sum \left(\varphi_{0,k} - \varphi_{0,k+1}\right)^2.$$

Дана міра має природний мінімум $W(\nu) = 0$. Тому, досліджуючи цю міру на мінімум, одержують систему спільних рівнянь у вигляді

$$\nu = 1, \quad \frac{dW(\nu)}{d\nu_j} = 0, \quad j \in [2, n].$$

Вирішення останньої системи рівнянь дає ті значення коефіцієнтів v, при яких процес характеризується близькістю відносних втрат. Отримані вирішення перевіряють на оптимальність за Парето.

Більш простим і менш трудомістким методом відшукування значень вагових коефіцієнтів є метод зважування локальних критеріїв на вагах механізму самозбереження досліджуваної системи. Покладають, що найбільшу вагу в структурі суперкритерію має той локальний критерій, що має більше значення для забезпечення самозбереження системи. У цьому випадку коефіцієнт v, визначають за формулою

$$\nu_i = \frac{q_{i\text{н}}}{q_{in}},$$

де $q_{i\text{н}}$, q_{in} - нормальне і граничне значення локального критерію.

Під нормальним розуміють оптимальне значення локального критерію у випадку оптимізації суперкритерію лише за цим локальним критерієм.

Так само, як і в попередньому випадку, на коефіцієнти накладають додаткову умову

$$\sum_{i=1}^{n} \nu_i = 1.$$

Таким чином, у випадку зведення багатокритеріальної задачі до однокритеріальної оптимізація зводиться до максимізації суперкритерію

$$x^* = \arg\max q_0 [q_1(x), \ldots, q_p(x)], \qquad x \in X.$$

Застосування для оптимізації суперкритерію супроводжується деякими труднощами і недоліками. Так, упорядкування точок у багатомірному просторі в принципі не може бути однозначним і цілком визначатись видом функції, яка упорядковує. Суперкритерій відіграє роль цієї функції, що упорядковує, і його навіть невелика зміна може привести до того, що оптимальна в новому смислові альтернатива виявиться такою, що дуже відрізнятиметься від старої. Крім того лінійна агрегація локальних критеріїв справедлива лише для стаціонарних процесів і систем.

Один із способів багатокритеріальної оптимізації складається у відмовленні від виділення єдиної «найкращої» альтернативи й угоді про те, що перевагу одній альтернативі перед другою можна віддавати тільки, якщо перша за всіма критеріями краще другої. Якщо ж перевага хоча б за одним критерієм розходиться з перевагою за іншим, то такі альтернативи визнають непорівнянними.

У результаті попарного порівняння альтернатив усі гірші за всіма критеріями альтернативи відкидають, а всі останні непорівнянні між собою (недомінуючі) приймають. Якщо всі максимально досяжні значення часткових критеріїв не відносяться до однієї й тієї ж альтернативи, тоді прийняті альтернативи утворюють множину Парето і вибір на цьому закінчується. При необхідності вибору єдиної альтернативи слід залучати додаткові міркування: вводити нові додаткові критерії й обмеження, або кидати жереб, або удаватись до послуг експертів.

Для оптимізації процесів часто використовують методи теорії ігор, що розглядає розвиток процесів як зміну випадкових ситуацій. ***Теорія ігор*** - це математична теорія конфліктів. Конфлікт полягає в тому, що інтереси двох сторін не співпадають (виникає боротьба інтересів) або сторони переслідують протилежні цілі.

Прикладом конфліктної ситуації є спортивна гра. Гравець вибирає таку сукупність правил поведінки (стратегію), що забезпечує йому бажаний виграш. Як правило, теорія ігор розглядає конфліктні ситуації, при яких приходиться приймати вирішення з частковою чи повною відсутністю даних про обстановку. Тому можуть бути і випадкові ходи, ефект яких можна оцінити в середньому математичним очікуванням. Результат гри оцінюють кількісними показниками чи умовними числами: виграш - +1, нічия - 0, програш - -1.

Методи теорії ігор застосовують не тільки для дослідження в буквальному розумінні конфліктних ситуацій, але і для вирішення задач, в яких, наприклад, як «супротивник» виступає природа. Такі задачі виникають при будівництві різних споруд, організації виконання робіт, організації транспортного процесу та ін.

За допомогою теорії ігор можна оцінити найбільш сприятливі чи несприятливі ситуації і на основі отриманих даних прийняти оптимальне для даних умов вирішення.

У теорії ***ігор*** важливе значення має поняття стратегії, під яким розуміють правила поведінки кожної сторони у відповідь на дію іншої сторони.

Частіше застосовують найбільш розроблену теорію парної гри, коли досліджується задача з двома протилежними сторонами *А* и *В.* Якщо допустити, що кожна із сторін дотримується оптимальної стратегії, тоді вони можуть розраховувати на рівноважний середній виграш, названий ціною гри. Вирішити гру, значить знайти пари оптимальних стратегій для *А* и *В* і ціну гри. Звичайно розрахунки ведуть за принципом «обережності», тобто знаходять таку стратегію, коли сторона *А* одержує найкращий результат при найгірших діях сторони *В.* Цей принцип є основним у теорії ігор, він забезпечує деякий запас в інженерних розрахунках.

4.4. Експериментальні дослідження

Експеримент (лат. *experimentum)* - це науково поставлене дослідження із точно врахованими і керованими умовами. Основною метою експерименту є виявлення властивостей досліджуваних об'єктів чи явищ, перевірка справедливості гіпотез, а також більш широке і глибоке вивчення теми дослідження.

Експериментальні дослідження класифікують за такими ознаками: ***галуззю науки*** (фізичні, хімічні, біологічні, психологічні, соціальні та ін.); ***способом формування експериментальних умов*** (природні, штучні); цілями дослідження (перетворюючі, констатуючі, контролюючі, пошукові, вирішальні); ***організацією проведення експерименту*** (лабораторні, натурні, відкриті, закриті); ***структурою досліджуваних об'єктів і явищ*** (прості, складні, дуже складні); ***характером зовнішніх впливів на об'єкт дослідження*** (речовинні, енергетичні, інформаційні); ***характером взаємодії засобів експериментального дослідження з досліджуваним об'єктом*** (звичайні, модельні); ***типом моделей, досліджуваних в експерименті*** (матеріальні, уявні); ***числом варіюючих факторів*** (однофакторні, багатофакторні); ***характером досліджуваних об'єктів чи явищ*** (технологічні, соціометричні); ***ступенем впливу на експериментальні умови*** (пасивні, активні).

Приведена класифікація експериментальних досліджень не є повною й остаточною, тому що з розширенням наукового знання розширюється й область застосування експерименту. Залежно від завдань експерименту різні його типи можуть поєднуватись, утворюючи комплексний чи комбінований експеримент.

Структура експерименту включає такі елементи: цілеспрямовану дослідницьку діяльність людини; зразок для експерименту; комплекс взаємодіючих дослідницьких пристроїв і наукових приладів. У модельному експерименті замість зразка використовують матеріальну модель досліджуваного об'єкта. Дана модель не тільки заміщає об'єкт дослідження, але й умови, у яких вивчається об'єкт.

Звичайний експеримент включає таку послідовність робіт: висування гіпотези; планування експерименту; конструювання експериментальної установки; проведення експерименту; інтерпретація й узагальнення отриманих даних; оцінка адекватності гіпотези; виведення наслідків; обговорення результатів експерименту.

Особливе місце в експерименті займає методика. ***Методика*** - це сукупність уявних і фізичних операцій, розміщених у визначеній послідовності, відповідно до якої досягається мета дослідження. При розробці методик проведення експерименту необхідно передбачати проведення попереднього цілеспрямованого спостереження над досліджуваним об'єктом чи явищем з метою визначення вихідних даних (гіпотез, відбору факторів, що варіюють); створення умов, у яких можливе експериментування; підбор об'єктів для експериментального впливу; усунення впливу випадкових факторів; визначення меж вимірювань; систематичне спостереження за ходом розвитку досліджуваного явища і точне описування факторів; проведення систематичної реєстрації результатів вимірювань і оцінок факторів різними засобами і способами; створення повторюваних ситуацій; зміна характеру умов і перехресних впливів; створення ускладнених ситуацій з метою підтвердження чи спростування раніше отриманих даних; перехід від емпіричного вивчення до логічних узагальнень, до аналізу і теоретичної обробки отриманого фактичного матеріалу; оцінка практичної застосовності, відповідності сучасному рівню науки і техніки.

Методика проведення досліджень має забезпечувати їхню відтвореність. ***Відтворене дослідження*** - це експеримент, при проведенні якого внесені зміни настільки незначні, що не піддаються виявленню.

Для відтворених експериментальних досліджень можна застосувати два варіанти послідовності: ***послідовний план,*** відповідно з яким необхідно обрати верхнє чи нижнє значення основного досліджуваного фактора, змінювати його через визначений інтервал до отримання іншого граничного значення; ***рандомізоватй (випадковий) план*** при реалізації якого значення основного фактора можна чергувати випадково.

Послідовний план проведення експерименту доцільно застосовувати при дослідженні матеріалів, течії рідин та ін.

Для більшості експериментальних досліджень, особливо у виробничих умовах, доцільний рандомізований план, що дозволяє зменшувати вплив невідомих систематичних факторів на результати. Це зменшення досягається через те, що систематично діючі фактори в результаті випадкової послідовності зміни значень досліджуваного основного фактора переводяться в розряд випадкових.

Обсяг експериментальних даних визначають виходячи з таких вимог: забезпечення точності результатів; зниження трудомісткості експериментальних робіт.

Кожен з факторів, що беруть участь в експерименті, змінюється у визначених межах. Межі зміни факторів задають цілями дослідження, конструктивними чи технологічними обмеженнями експериментальних установок і приладів. Межі зміни факторів обмежують область досліджуваних величин. Останню поділяють на рад інтервалів, в яких визначають значення показників об'єкта дослідження. Вибір інтервалів виконують залежно від характеру очікуваної функціональної залежності або з умови забезпечення однакової точності дослідів.

Число дослідів у кожній серії визначають виходячи з очікуваної функціональної залежності. Але у реальних умовах проведення експериментальних досліджень, для зменшення випадкових помилок дослід необхідно повторити з метою підвищення точності й вірогідності результатів. Число повторень залежить від розкиду значень паралельних дослідів.

Часто у процесі планування перед дослідниками виникає питання, які експерименти виконувати. Найбільш розповсюдженою стратегією є виконання тих експериментів, що максимізують чисту очікувану корисність.

Особливим випадком планування експерименту є вибір моменту його зупинки при проведенні повторних дослідів. Це завдання вперше було вирішено Вальдом у Колумбійському університеті. Він показав, що якщо існують фіксовані вартості помилок в остаточному вирішенні і якщо вартість проведення експерименту чи спостереження лінійно пов'язана з числом виконаних спостережень, тоді існує оптимальна стратегія дій: продовжити спостереження доти, поки апостеріорна ймовірність одного із станів середовища не сіане досить високою, і потім вибрати дію, обумовлену цим станом. Значення критерію, при яких приймається остаточне вирішення, залежить від винагороди за правильні дії, вартості помилок, вартості спостереження і діагностичності чи інформативності спостережень.

Придатною статистичною характеристикою для використання як основи вибору моменту зупинки дослідів є байєсовська ймовірність. Збір даних і обчислень цієї ймовірності варто продовжувати доти, поки байєсовська ймовірність не досягне заздалегідь визначеного значення. Для послідовної вибірки оптимальним є критичне значення апостеріорної ймовірності.

Традиційним способом випробувань є однопараметричний експеримент, при якому усі фактори, крім досліджуваного, фіксуються

на постійних рівнях. Фактори, варіюючи почергово, отримують залежності, що не відбивають характер взаємодії факторів. Для усунення цього недоліку застосовують факторне планування експерименту: повний факторний експеримент; дробовий факторний експеримент; експеримент, заснований на ортогональному центральному композиційному плануванні; Д *А* и *Е* - оптимальні плани; симплексно-ґратчасті плани; критеріальні плани та ін.

Результати дослідів реєструють, фіксують за допомогою вимірювань, тобто зображення результатів досліду у вигляді символів, номерів або чисел. За характером вимірювання можуть бути ***кількісні*** **і** ***якісні, прямі і непрямі, абсолютні*** **і** ***відносні, сукупні*** **і** ***спільні.***

При *прямих* вимірюваннях величину, що відшукують, встановлюють безпосередньо з досліду, при *непрямих* - функціонально від інших величин, визначених прямими вимірюваннями:

$$b = f(a)$$

де b - величина, знайдена за допомогою непрямих вимірювань; а - те ж, але за допомогою прямих вимірювань.

Абсолютні - це прямі вимірювання в одиницях вимірюваної величини.

Відносні вимірювання представляють у вигляді відношення вимірюваної величини до однойменної величини, що відіграє роль одиниці вимірювання цієї величини стосовно однойменної, прийнятої за початкову.

При ***сукупних*** вимірюваннях проводять одночасно вимірювання декількох *однойменних* величин, а величину, що відшукують, при цьому знаходять способом вирішення системи рівнянь. У випадку *спільних* вимірювань одночасно проводять вимірювання не однойменних величин для знаходження залежності між ними.

У процесі вимірювань використовують такі методи: безпосередньої оцінки;

порівняння з мірою; протиставлення; диференціальний; нульовий; заміщення; збігів.

Метод безпосередньої оцінки зводиться до безпосереднього визначення значення вимірюваної величини на відліковому пристрої вимірювального приладу прямої дії.

Метод порівняння з мірою - вимірювану величину порівнюють з величиною, що відтворюється мірою, наприклад, вимірювання маси на важільних вагах із зрівноважуванням гирями.

Метод протиставлення - вимірювана величина і величина, що відтворюється мірою одночасно впливають на прилад, за допомогою якого встановлюється співвідношення між цими величинами.

Диференціальний метод - порівняння з мірою, при якому на вимірювальному приладі реєструється різниця вимірюваної і відомої величини, що відтворюється мірою.

Нульовий метод - результуючий ефект впливу величини на приладі доводять до нуля.

Метод заміщення - вимірювану величину заміщають відомою величиною, що відтворюється мірою.

Метод збігу - різниця між вимірюваною величиною і величиною, що відтворюється мірою, вимірюють, використовуючи збіг позначок шкал чи періодичних сигналів.

МОДУЛЬ 2
5 МЕТОДИ ПРИЙНЯТТЯ ВИРІШЕНЬ

Що собою представляє вибір
Назвіть етапи прийняття вирішення
Назвіть шкали корисності для оцінки наслідків прийняття вирішення
Які мови використовують для описування вибору
Що розуміють під вибором за допомогою критеріальної мови
Як представляється сімейство вирішення
Які підходи використовують для формування результату вирішення
Назвіть класичні критерії прийняття вирішення
Назвіть похідні критерії вибору
Як формуються складені критерії
Що враховують гнучкі критерії
Як визначають величину ризику
Що таке суперкритерій
Які положення лежать в основі мови бінарних відношень
Назвіть способи завдання бінарних відношень
За допомогою яких відношень задають перевагу однієї альтернативи над іншою
Що таке функція вибору
У чому полягає суть вибору в умовах статистичної невизначеності
Що таке вирішальна функція
Що таке функція правдоподібності
Що така розпливчаста множина
Що виражає функція приналежності
У чому полягає ідея прийняття вирішення у розпливчастій ситуації
Що таке груповий метод вибору
У чому полягає смисл теореми про неможливість (парадокса Ерроу)
У чому полягає суть методу агрегування переваг

5.1. Загальна структура процесу прийняття вирішення

Прийняття вирішення - це дія над множиною альтернатив, у результаті якої виходить підмножина обраних альтернатив. ***Процедура прийняття вирішення*** включає такі етапи чи фази: ініціативу, описування проблеми, аналіз ситуації, постановку задачі, аналіз наявної інформації, дискретизацію і комбінування зовнішніх умов, розробку альтернатив, перевірку результатів, оформлення вирішення.

Ініціатива може бути зовнішньою (замовленою) чи власною, коли на підставі власних спостережень приходять до переконання, що потрібно шукати раціональний спосіб досягнення поставленої мети. При цьому мета може бути споконвічно відомою чи може характеризуватись великим числом варіантів. Часто немає необхідності докладно описувати саму задачу, тому що її структура досить ясна і спосіб вирішення певним чином випливає з життєвого досвіду. Такі рутинні вирішення звичайно протікають за схемою: ініціатива (замовлення), ознайомлення з задачею, порівняння з аналогічними чи схожими вирішеннями, визначення раціональних варіантів. Для складних чи нових задач з однозначними параметрами необхідна точна і докладна постановка задачі. У цьому випадку необхідно мати значний обсяг інформації, що стосується і мети задачі. Безпосередньо з постановки задачі випливає безліч раціональних варіантів її вирішення.

Дискретизація варіантів вирішення може бути природною чи штучною. В останньому випадку її вибирають приймаючим вирішення. При цьому немає загального підходу. Надійним є ітераційний метод. Спочатку проводять грубу дискретизацію і розраховують вирішення в першому наближенні. Потім біля цього наближеного вирішення формують ряд більш детально дискредитованих альтернатив і з більшою точністю наближаються до оптимуму.

При дискретизації параметрів зовнішніх умов вибирають їхні представницькі значення так, щоб розв'язувана задача описувалась досить точно, але щоб число значень параметра в інтересах спрощення розрахунків було якнайменше. Звичайно використовують середні і крайні значення параметрів. Якщо можливі стани параметра представляють багатьма значеннями, тоді доцільно виходити з рівномірного поділу діапазону зміни параметра, а як представницькі значення вибирають середини інтервалів. Бажане формулювання задачі виходить з комбінації обраних представницьких значень, що мають адекватно характеризувати аналізований параметр зовнішніх умов.

Через те, що число комбінацій впливає на витрати, пов'язані з їхньою обробкою, то з безлічі всіх комбінацій слід вибирати такі сполучення, які б щонайкраще характеризували стан вихідних даних і зміну стану. Відбір комбінацій має задовольняти таким вимогам: 1) число значень параметра має відповідати його впливові на результат; 2) сполучення значень параметра, що вчиняють однаковий чи подібний вплив на результат, не повинні бути представлені у числі всіх комбінацій багаторазово; 3) відбір має бути в максимальному ступені незалежний від суб'єктивного відношення дослідника; 4) ті з обраних сполучень, що за характером розв'язуваної задачі малоймовірні, мають бути виключені.

Л.С. Бєляєв запропонував три формалізовані методи відбору комбінацій параметрів, що, незважаючи на трудомісткість, дозволяють у значній мірі уникнути впливу суб'єктивного фактора на процес відбору.

У першому методі для безлічі значень параметрів розглядають паралелепіпед чи іншу n-мірну область. У цю область вносять *N* сфер однакового і максимально можливого діаметра. Центри сфер визначають, наприклад, на границі паралелепіпеда. Вони підлягають вибору реалізації параметрів і одночасно сполученню вихідних даних. Цей метод застосовують для невеликого числа невизначених параметрів, приблизно для 10. Крім того, для цього методу ще не існує загального вирішення.

За другим методом усередині обраної області розташовують рівномірну сітку. На цій сітці вибирають задане число вузлових точок таким чином, щоб відстань їх одна від другої була максимальною. Для відбору застосовують теорію лінійних кодів.

Третій метод формального відбору заснований на використанні методу Монте-Карло і методу класифікацій. За заданими яким-небудь чином розподілами параметра визначають статистично велике число точок в області невизначеності параметра. Отриману безліч точок розділяють на *N* груп. Групові «центри» вибирають при цьому так, щоб середньоквадратична відстань між точками в групі була мінімальною, а відстані між центрами - максимальними.

Для того щоб зробити розумний вибір між різними варіантами вирішення, необхідно оцінити його наслідки. Причому дана оцінка має бути зроблена за однозначними правилами.

Наслідки вирішення можуть оцінюватись за допомогою різних **шкал корисності:** *номінальної шкали, шкали упорядкованості, інтервальної, масштабної.*

За допомогою ***номінальної шкали*** безліч наслідків поділяють на підмножини, такі, як коло, овал чи прямокутник, області з гладкою чи нерівною границею та ін. Такі шкали застосовують для найпростіших тимчасових вирішень, коли не ставиться мета досягти оптимального вирішення, а потрібно знайти лише прийнятне. Ця шкала часто складається тільки з двох градацій.

Іноді номінальну шкалу називають шкалою найменувань чи класифікаційною.

Шкали упорядкованості встановлюють між підмножинами, на які розбивають безліч результатів вирішення, визначені жорсткі співвідношення. Ці співвідношення формулюють у формі *аксіом лінійності чи повної упорядкованості, транзитивності і рефлективності.* Відповідно до аксіоми лінійності про два будь-які результати можна зробити такі висновки: 1) e_1 не гірше, ніж e_2; 2) e_2 не гірше, ніж e_1; 3) e_1 і e_2 рівноцінні. Аксіоми транзитивності і рефлективності розглянуто нижче.

Шкали упорядкованості достатні для прийняття вирішення у задачах з однозначними параметрами.

У структурі шкал упорядкованості виділяють шкали простого порядку, слабкого порядку, часткового порядку.

Інтервальні шкали встановлюють, чи є різниця у корисностях e_1-e_2 однаковою, більшою чи меншою, ніж різниця e_2-e_3. Співвідношення між різницями корисностей зберігаються, коли різницю множать на будь-яку константу чи складають з нею Назва «шкала інтервалів» підкреслює, шо в ній тільки інтервали мають сенс дійсних чисел і тільки над інтервалами слід виконувати арифметичні операції. Якщо зробити арифметичні операції над самими відліками на шкалі, забувши про їхню відносність, тоді є ризик одержати безглузді результати.

Якщо потрібно порівнювати відносини корисностей, тоді останні потрібно вимірювати у ***масштабній шкалі***. Ці шкали говорять про рівність або розходження сум чи добутків розглянутих величин. Шкали довжини, маси та інші є масштабними.

Універсальною і єдиною у сенсі можливостей перетворень показань на цій шкалі є ***абсолютна шкала***. Вона має й абсолютний нуль, й абсолютну одиницю Важливою особливістю абсолютної шкали є абстрагованість (безрозмірність) і абсолютність п одиниці. Саме такі якості є й у числової осі Числова вісь дозволяє робити над своїми показниками такі операції, що недопустимі для показників інших шкал, - уживати ці показники як показник ступеня чи аргументу логарифма.

5 2 Мови для описування вибору

Для описування вибору використовують три основні мови: ***критеріальну, бінарних відношень, функцій вибору.***

При використанні критеріальної мови вибір розглядають як пошук варіанта з його безлічі, що максимізує деякий критерій. Нехай, наприклад, кожен варіант $E_i \in E$ однозначно визначають деяким результатом e_i. Цей результат допускає кількісну оцінку. Ми відшукуємо варіант із найбільшим значенням результату, тобто метою нашого вибору є max e_i. При цьому ми припускаємо, що результат e_i характеризує такі величини, як, наприклад, виграш, корисність чи надійність. Протилежну ситуацію з оцінкою витрат чи втрат можна досліджувати точно так само способом мінімізації оцінки чи, як це робиться частіше, за допомогою розгляду негативних величин корисності. Таким чином, вибір оптимального варіанта виконують за допомогою критерію

$$E_o = \left\{ E_{io} \middle| E_{io} \in E \Lambda e_{io} = \max_i e_i \right\}.$$

Це правило вибору читається так: множина E_0 оптимальних варіантів складається з тих варіантів E_{i0}, що належать множині Е усіх варіантів, і оцінка e_{i0} яких максимальна серед всіх оцінок e_i. Логічний знак $\wedge$ читається як «і» й вимагає, щоб обидва твердження, які зв'язуються ним, були щирі.

У більш складних випадках кожному допустимому варіанту вирішення E_i можуть відповідати різні зовнішні умови (стани) F_j і результати e_{ij} вирішення. Під результатом e_{ij} розуміють оцінку, що відповідає варіанту E_i й умовам F_j і ефективність, що характеризує, корисність чи надійність досліджуваного об'єкта Такий результат часто називають *корисністю вирішення.*

Сімейство вирішень при різних зовнішніх умовах представляють у формі матриці вирішень (табл.. 5.1)

Таблиця 5 1 **Матриця вирішень,** $\|e_{ij}\|$

	F_1	F_2	F_3	.....	F_j	....	F_n
E_1	e_{11}	e_{12}	e_{13}	..	e_{1j}	. ..	e_{1n}
E_2	e_{21}	e_{22}	e_{23}	.....	e_{2j}	.. ..	e_{2n}
E_3	e_{31}	e_{32}	e_{33}	.	e_{3j}	.....	e_{3n}
E_i	e_{i1}	e_{i2}	e_{i3}	.....	e_{ij}	.....	e_{in}
E_m	e_{m1}	e_{m2}	e_{m3}	.....	e_{mj}	.....	e_{mn}

Щоб прийти до однозначного варіанта вирішення необхідно ввести придатні оцінні (цільові) функції При цьому матриця вирішення $\|e_{ij}\|$ зводиться до одного стовпця. Кожному варіанту E_i приписується деякий результат e_{ir} , що характеризує всі наслідки цього вирішення.

Якщо, наприклад, наслідки кожного з альтернативних вирішень характеризуються комбінацією з його найбільшого і найменшого результатів, тоді можна прийняти

$$e_{ir} = \min_j e_{ij} + \max_j e_{ij}.$$

Тоді найкращий у цьому сенсі результат має вигляд

$$\max_i e_{ir} = \max_i \left(\min_j e_{ij} + \max_j e_{ij} \right).$$

Формуючи таким чином результат ми виходимо з компромісу між оптимістичним і песимістичним підходами

Оптимістична позиція може бути представлена у вигляді

$$\max_i e_{ir} = \max_i \left(\max_j e_{ij} \right).$$

Позиція нейтралітету:

$$\max e_{ir} = \max_i \left(\frac{1}{n} \sum_{j=1}^{n} e_{ij} \right).$$

Песимістична позиція:

$$\max_i e_{ir} = \max_i \left(\min_j e_{ij} \right).$$

Позиція відносного песимізму:

$$\max_i e_{ir} = \min_i \max_j \left(\max_j e_{ij} - e_{ij}\right).$$

Класичними критеріями прийняття вирішення є: ***мінімакстй критерій (ММ-критерій); критерій Байєса - Лапласа (BL-критерій); критерій Севіджа (S-критерій).***

У ***мінімаксному критерії (ММ)*** використовують песимістичну оцінну функцію

$$Z_{MM} = \max_i e_{ir} \quad і \quad e_{ir} = \min_j e_{ij}.$$

Безліч оптимальних варіантів вирішення будується відповідно до співвідношення

$$E_o = \left\{ E_{io} \middle| E_{io} \in E \Lambda e_{io} = \max_i \min_j e_{ij} \right\}$$

Правило вибору вирішення відповідно з ***ММ-критерієм*** формулюють так матриця вирішень $\|e_{ij}\|$ доповнюється ще одним стовпцем з найменших результатів кожного рядка. Вибрати слід ті варіанти E_{i0} у рядках яких стоять найбільші значення e_{ir} цього стовпця.

Застосування ***ММ-критерію*** виправдано у ситуаціях, коли про можливість появи зовнішніх умов F_j нічого не відомо; коли вирішення реалізується лише один раз; коли необхідно виключити який би то ні було ризик, тобто ні при яких умовах не допускається отримувати результат менший, ніж max e_{ir}.

Тут

$$e_{ir} = \min_j e_{ij}.$$

Критерій Байєса - Лапласа (BL) враховує ймовірність q_j появи зовнішньої умови F_j. Тому для ***BL-критерію***

$$Z_{BL} = \max_j e_{ir},$$

$$e_{ir} = \sum_{j=1}^{n} e_{ij} q_j,$$

$$E_o = \left\{ E_{io} \middle| E_{io} \in E \Lambda e_{io} = \max_i \sum_{j=1}^{n} e_{ij} q_j, \sum_{j=1}^{n} q_j = 1 \right\}.$$

Тут Z_{BL} - оцінна функція ***BL-критерію.***

Правило вибору формулюють так: матриця вирішень $\|e_{ij}\|$ доповнюється ще одним стовпцем, який містить математичне очікування значень кожного з рядків. Вибирають ті варіанти E_{i0} у рядках яких стоїть найбільше значення e_{ir} цього стовпця.

Припускають, що ситуація прийняття вирішення характеризується такими обставинами

- імовірності появи умов F_j відомі і не залежать від часу;
- вирішення реалізуються (теоретично) нескінченно багато разів;
- для малого числа реалізацій вирішень допускають деякий ризик.

Критерій Севіджа (S) використовує позицію відносного песимізму За допомогою позначень

$$a_{ij} = \max_i e_{ij} - e_{ij} \text{ і } e_{ir} = \max_j a_{ij} = \max_j \left(\max_i e_{ij} - e_{ij}\right)$$

формують оцінну функцію Севіджа

$$Z = \min e = \min\left[\max\left(\max_i e_{ij} - e_{ij}\right)\right]$$

і будують безліч оптимальних варіантів вирішення

$$E_o = \left\{E_{,}|E \in E \Lambda e = \min e_{ir}\right\}$$

Правило вибору відповідно до критерію Севіджа формулюють так: кожен елемент матриці вирішень $\|e_{ij}\|$ віднімається з найбільшого результату $\max\limits_i e_{ij}$ відповідного стовпця.

Різниці a_{ij} утворюють матрицю залишків $\|a_{ij}\|$. Ця матриця поповнюється стовпцем найбільших різностей e_{ir}. Вибирають ті варіанти E_{i0} у рядках яких стоїть найменше для цього стовпця значення.

Звичайно у процесі прийняття вирішень класичні критерії застосовують почергово Після цього серед декількох відібраних оптимальних варіантів вольовим чином виділяють остаточне вирішення.

Прагнення одержати критерії, які б краще пристосовувались до наявної ситуації, привело до побудови ***похідних, складених і гнучких критеріїв.*** Так, намагаючись зайняти найбільш урівноважену позицію, Гурвиц запропонував ***критерій (HW),*** оцінна функція якого знаходиться десь між точками зору граничного оптимізму і крайнього песимізму

$$Z_{HW} = \max_i e_{ir}, \; e_{ir} = c \min_j e_{ij} + (1-c)\max_j e_{ij}$$

Тоді

$$E_o = \left\{E_{io} \middle| E_{io} \in E \Lambda e_o = \max_i \left[c \min_j e_{ij} + (1-c)\max_j e_{ij}\right] \Lambda 0 \le c \le 1\right\}.$$

де c - ваговий множник.

Правило вибору згідно ***HW-критерію*** формулюють так матриця вирішення $\|e_{ij}\|$ доповнюється стовпцем, що містить середні зважені найменшого і найбільшого результатів для кожного рядка. Вибирають ті варіанти E_{i0}, у стовпцях яких стоять найбільші елементи e_{ir} цього стовпця.

HW-критерій пред'являє до ситуації, коли приймають вирішення, такі вимоги:

- про імовірності появи умов F_j нічого не відомо;
- з появою умов F_j необхідно рахуватись;
- реалізується лише мала кількість вирішень;
- допускається деякий ризик.

Критерій Ходжа - Лемана (HL) спирається одночасно на ***ММ*** і ***BL-критерії.*** За допомогою параметра *v* виражають ступінь довіри до використаного розподілу ймовірностей. Якщо ця довіра велика, то акцентують ***BL-критерій,*** у противному випадку перевагу віддають ***ММ-критерію.***

Оцінну функцію визначають рівністю

$$Z_{HL} = \max_i e_{ir},$$

$$e_{ir} = v\sum_{j=1}^{n} e_{ij} q_j + (1-v)\min_j e_{ij},\ 0 \le v \le 1,$$

а безліч ***HL-оптимоіільних*** вирішень записують у вигляді

$$E_{HL} = \left\{ E_{i0} \middle| E_{i0} \in E \wedge e_{i0} = \max_i \left[v\sum_{j=1}^{n} e_{ij} q_j + (1-v)\min_j e_{ij} \right] \wedge 0 \le v \le 1 \right\}$$

Правило вибору, що відповідає ***HL-критерію,*** формулюють так матриця вирішень $\|e_{ij}\|$ доповнюється стовпцем, складеним із середніх зважених (з постійними вагами) математичного очікування і найменшого результату кожного рядка. Відбирають ті варіанти вирішень E_{i0} у рядках яких стоїть найбільше значення цього стовпця.

Розглянутий критерій застосовують у випадках:

- імовірності появи умови F_j невідомі, але деякі припущення про розподіли ймовірностей можливі;
- прийняте вирішення теоретично допускає нескінченно багато реалізацій;
- при малих числах реалізацій допускається деякий ризик.

Для відшукування придатних до компромісу вирішень в області поліоптимізації - тобто всіх вирішень, що не вважаються свідомо гіршими ніж інші, запропонований ***критерій Гермейера*** **(G),** що має в деякому відношенні визначену еластичність. Оцінну функцію ***критерію*** **(G)** представляють у вигляді

$$Z_G = \max_i e_{ir},$$

$$e_{ir} = \min_j e_{ij} q_j .$$

Безліч ***G-оптимальних*** вирішень записують так:

$$E_o = \left\{ E_{i0} \middle| E_{i0} \in E \wedge e_{i0} = \max_i \min_j e_{ij} q_j \wedge e_{ij} 0 \right\}.$$

Правило вибору відповідно до ***критерію Гермейера (G)*** формулюють так: матриця вирішень $\|e_{ij}\|$ доповнюється ще одним стовпцем, що містить у кожному рядку найменший добуток наявного в ній результату на ймовірність відповідного стану F_j. Вибирають ті варіанти E_{i0} у рядках яких знаходиться найбільше значення цього стовпця.

Умови застосування ***критерію (G)*** такі:

- імовірності появи зовнішніх умов F_j відомі;

- з появою тих чи інших зовнішніх умов, окремо або в комплексі, необхідно рахуватись;

- вирішення може реалізуватись один чи багато разів;

- допускається деякий ризик.

Критерії Гурвица, Ходжа - Лемана і Гермейера називають ***похідними критеріями.***

Їхня комбінація дає **складені критерії**: ***BL (MM)-критерій; BL (S)-критерій* та ін.**

Останнім часом у практиці прийняття вирішення знайшла застосування П-операція, яку в теорії нечітких множин використовують для фільтрації інформації. На її основі побудовано ***критерій добутків (P)***.

Оцінна функція цього критерію

$$Z_P = \max_i e_{ir},$$

$$e_{ir} = \prod_j e_{ij}.$$

Тоді

$$E_o = \left\{ E_{i0} \middle| E_{i0} \in E \wedge e_{i0} = \max_i \prod_j e_{ij} \wedge e_{ij} > 0 \right\}.$$

Правило вибору формулюють так: матриця вирішень і $\|e_{ij}\|$ доповнюється новим стовпцем, що містить добутки всіх результатів кожного рядка. Вибирають ті варіанти E_{i0} у рядках яких знаходяться найбільші значення цього стовпця.

Умови застосування ***(P) -критерію*** такі:

- імовірності появи зовнішніх умов F_j невідомі;

- з появою кожної з умов F_j окремо необхідно рахуватись;

- критерій пристосований у першу чергу для випадків, коли усі e_{ij} позитивні;

- критерій застосовується і при малому числі реалізацій вирішення;

- деякий ризик допускається.

Гнучкі критерії вибору вирішення враховують якісні характеристики початкової інформації і число майбутніх реалізацій вирішення. Крім того ці критерії дозволяють врахувати можливість ризику ухвалення помилкового вирішення. Ризик обмежують допустимою величиною.

Величину ризику ***R*** визначають через добуток величини події ***A*** на міру можливості його настання. Мірою можливості настання події служить імовірність ***q*** її настання. Звідси виходить:

$$R = Aq.$$

Часто під ризиком розуміють просто ймовірність настання деякого сполучення несприятливих подій.

У джерелах ризику розбираються способом систематичного аналізу. Допоміжний засіб для цього - дерево помилок, що будують аналогічно дереву вирішень. Наслідки задають стосовно до конкретної проблеми. Аналіз інформації проводять так само, як і при кількісній оцінці ситуації прийняття вирішення, і визначають імовірність настання небажаних подій.

Кількісно описування ризику спирається на теоретико-імовірнісний підхід. Безліч несприятливих подій S представляється у вигляді

$$S = \{s_1, s_2, \ldots, s_n\}.$$

У деякій ситуації можуть одночасно наступати кілька з цих подій. Кожне сполучення таких подій може бути позначене через ***К***. Безліч усіх можливих сполучень ***К*** в математиці називається *булеаном S* (безліччю всіх підмножин). До сполучень ***К*** зараховують саму множину ***S*** і пусту множину $\varnothing$ (пуста множина відповідає відсутності несприятливих подій). Таким чином, сполучення ***К*** є підмножиною несприятливих подій множини ***S***

$$K = \{s_{k1}, s_{k2}, \ldots, s_{kl}\}; \qquad s_{kj} \in S. \qquad j = 1, \ldots, l.$$

У безлічі всіх сполучень можна виконувати звичайні операції алгебри множин об'єднання, пересічення, різниця, доповнення.

Нехай з деяким ризикованим варіантом вирішення пов'язані елементарні сполучення несприятливих подій K_{i1}, $K_{i2}, \ldots$, K_{ik_i}. Формула «елементарні сполучення несприятливих подій» означає, що ніяка власна підмножина сполучення K_{ij} не може сама зустрічатись як сполучення несприятливих подій. Якщо ще позначити через N_i гарантовану відсутність несприятливих подій для ризикованого варіанта вирішення E_i, тоді

$$\overline{K}_i := \{K_{i1}, K_{i2}, \ldots, K_{ik_i}, N_i\}$$

утворить повну, пов'язану з вирішенням E_i, систему подій.

Покладемо далі, що кожному сполученню несприятливих подій $K_{ij}(j = 1, \ldots, K_i)$, що може реалізуватись у результаті ухвалення вирішення

$E_i \subset E$, а також події N_i можна приписати ймовірності $p_i(K_{ij})$ і, відповідно, $p_i(N_i)$:

$$0 \le p_i(K_{ij}) \le 1, \quad \sum_{j=1}^{k_i} p_i(K_{ij}) + p_i(N_i) = 1$$

Якщо кожному сполученню K_{ij} може бути поставлене у відповідність кількісно описуваний наслідок A_{ij}, тоді величину ризику R_i, що є супутнім вирішенню E_i, визначають за формулою

$$R_i = \sum_{j=1}^{k_i} A_{ij} p_j(K_{ij}).$$

Якщо всі ймовірності реалізації сполучення несприятливих подій $K_{ij} \in \overline{K}_i$ однакові, тобто $p_i(K_{ij})$, тоді

$$R_i = p_j \sum_{j=1}^{k_i} A_{ij}.$$

Варіанти вирішення $E_i \in E$ без урахування можливості несприятливій наслідків будуть мати корисність e_i. Тоді відповідному варіанту вирішення E_i величину $G_i = e_i - R_i$, називають *сумарним ефектом вирішення*.

Безліч раціональних варіантів вирішення позначають

$$\overline{E}^+ := \{E_i \in E : G_i > 0\}.$$

Варіант вирішення E_i^* називають оптимальним у випадку

$$G_i^* = \max G_i. \quad E_i \in E.$$

При цьому в рамках конкретної практичної задачі безліч допустимих варіантів вирішення може бути додатково обмежена межами ризику.

Складність відшукування найкращої альтернативи зростає, коли її приходиться оцінювати за декількома критеріями, що якісно розрізняються між собою. Наприклад, при виборі конструкції автомобіля приходиться враховувати технічні, технологічні, економічні, соціальні, екологічні, ергономічні та інші критерії. Часто в цьому випадку багатокритеріальну задачу зводять до однокритеріальної. Це означає введення ***суперкритерію,*** тобто скалярної функції векторного аргументу:

$$q_0(x) = q_0(q_1(x), q_2(x), \ldots, q_p(x)).$$

Суперкритерій дозволяє упорядкувати альтернативи за величиНою q_0, виділивши тим самим найкращу. Вид функції q_0 докладно розглянутий у п. 4.3.

Більш загальною порівняно з критеріальною є ***мова бінарних відношень.*** Основні припущення цієї мови зводяться до такого:

1) окрема альтернатива не оцінюється, тобто критеріальна функція не вводиться;

2) для кожної пари альтернатив (x, у) можна встановити, що одна з них переважніше іншої або вони рівноцінні чи непорівнянні (найчастіше останні два поняття ототожнюються);

3) відношення переваги усередині будь-якої пари альтернатив не залежить від інших альтернатив, пред'явлених до вибору.

Бінарне відношення між *x* і *y* позначають у вигляді *xRy.*

Основними способами завдання бінарних відношень є: безпосереднє перерахування пар альтернатив, між якими існує бінарне від ношення; матричний спосіб завдання переваг; завдання графа переваг; завдання відношень перерізами.

Завдання відношень перерізами використовують на нескінченних множинах. Так, множину $R^+(x) = \{y \in X | (x, y) \in R\}$ називають *верхнім перерізом* відношення *R,* а множину $R^-(x) = \{y \in X | (x, y) \in R\}$ - *нижнім перерізом.*

Верхній переріз - це множина усіх $y \in X$, які знаходяться у відношенні yRx з заданим елементом, $x \in X$, а нижній переріз - множина усіх $y \in X$, з якими заданий елемент x знаходиться у відношенні R. Відношення однозначно визначаться одним із своїх перерізів.

Перевагу однієї альтернативи над іншою задають за допомогою відношень *еквівалентності, порядку і домінування.*

Відношенням еквівалентності називається відношення *R* на множині *X,* якщо воно *рефлексивно, симетрично і транзитивно.*

Бінарне відношення *R* на множині *X* є рефлексивним, якщо *xRx* для кожного $x \in X$; симетричним, якщо xRy => $yRx \forall x$, $\nu \in X$; транзитивним, якщо *для всіх x, y* $z \in X(xRy, yRz) \Rightarrow xRz$.

Відношення порядку поділяють на *суворе* і *несуворе.* Відношенням несуворого порядку називають рефлексивне, антисиметричне і транзитивне відношення. Відношенням суворого порядку називають антирефлексивне, асиметричне і транзитивне відношення.

Бінарне відношення *R* на множині *X* називають антирефлексивним, якщо $x\overline{R}x \forall x \in X$ (тобто *R* може виконуватись тільки для незбіжних елементів). Відношення *R* на множині *X* асиметричне, якщо для всіх *x* $y \in X(xRy, yRx) \Rightarrow x \Rightarrow y$.

Відношенням домінування називають відношення антирефлексивне й асиметричне.

Найважливішим положенням у мові бінарних відношень є незалежність упорядкування двох альтернатив від будь-якої третьої.

Для здійснення вибору мовою бінарних відношень часто використовують орієнтований граф переваг. Стрілки цього графа направляють убік альтернативи, що має меншу перевагу. Вершини графа, з яких стрілки тільки виходять, є найкращими альтернативами.

Мова функцій вибору описує вибір як операцію над довільною безліччю альтернатив *X*, що ставить цііі безлічі у відповідність деяку її підмножину $C(X): C(X) \subseteq X$. *Функція вибору - це відображення сукупності множин у сукупність множин без поелементного відображення однієї множини на іншу і без відображення множин на числову вісь.* Розмаїтість функцій вибору набагато перевершує розмаїтість графів переваг. Розходження між класами правил вибору можна відобразити через різні обмеження, яким підкоряється той чи інший тип функцій вибору. Обмеження формулюють у вигляді деяких вимог, пропонованих до вибору. Вимоги представляють у вигляді аксіом:

Аксіома спадкування (H):

$$X^1 \subseteq X \Rightarrow C\left(X^1\right) \subseteq C(X) \cap X^1.$$

Смисл цієї аксіоми зводиться до вимоги, щоб у вибір на підмножині X^1 увійшли всі ті альтернативи з X, що входили у вибір X (рис. 5.1, а).

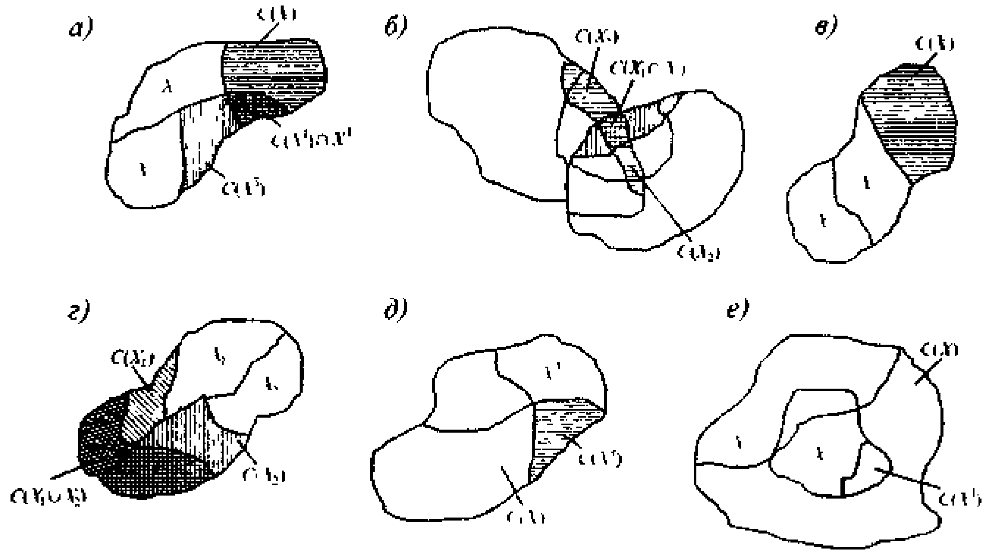

Рис. 5.1. Ілюстрація різних аксіом, що накладаються на функції вибору одержуючи при цьому різні типи вибору

Аксіома згоди (C):

$$C(X_i) \subseteq C(\underset{i}{\cup} X_i).$$

Відповідно з даною аксіомою у вибір з об'єднання множин обов'язково мають входити альтернативи, загальні для виборів із усіх множин (рис. 5.1, б).

Аксіома відкидання (В):

$$C(X) \subseteq X^1 \subseteq X \Rightarrow C(X^1) = C(X).$$

Дана аксіома вказує, якщо відкинути будь-яку частину відкинутих при виборі альтернатив, тоді вибір на множині, що залишилась, не зміниться (рис. 5.1, в). Аксіому відкидання часто називають умовою незалежності від відкинутих альтернатив.

Аксіома Плотта (КС):

$$C(X_1 \cup X_2) = C(C(X_1) \cup C(X_2)).$$

Дана аксіома відображує вимоги, що накладаються при багатоступінчастих виборах. Вона вказує на умову незалежності від способу ухвалення вирішення (рис. 5.1, г). Функції вибору, що задовольняють їй, називають квазісуматорними.

Аксіома переваги (П):

$$X^1 \subseteq X \Rightarrow C(X) \cap X^1 = C(X^1).$$

Аксіома вимагає, щоб при звуженні безлічі альтернатив у виборі залишались тільки ті альтернативи, що входили до нього раніше (рис. 5.1, д).

Спільне підпорядкування функції вибору аксіомам *Н* і *С* дає вибір, який описують мовою бінарних відношень.

Спільне накладання на вибір аксіом *Н,* Сі *О* приводить до вибору паретовської множини.

Вимога *КС* еквівалентна спільному виконанню *Н* і *О.* Отже, з'єднання *КС* і С теж приводить до паретовського вибору (рис. 5.1, е).

Розглянуті аксіоми можна послабляти чи підсилювати (наприклад, П є посилення Н). Аксіому Плотта можна підсилити до ***аксіоми суматорності***

$$C(X_1 \cup X_2) = C(X_1) \cup C(X_2);$$

можна накладати нові, незалежні вимоги. Наприклад, ***аксіома мультиплікаторності***

$$C(X_1 \cap X_2) = C(X_1) \cap C(X_2),$$

аксіома монотонності

$$X_1 \subseteq X_2 \Rightarrow C(X_1) \subseteq C(X_2).$$

5.3 Вибір в умовах статистичної і розпливчастої невизначеностей

Для задач прийняття вирішення в умовах статистичної невизначеності загальним є необхідність вибору на підставі непрямих чи прямих, але обов'язково «зашумлених» даних Ці дані носять статистичний характер Зв'язок між істинною, але невідомою альтернативою *О є в,* яку відшукують, і даними x_h x_2, , x_N, що спостерігаються, адекватно описують розподілом імовірностей Загальна схема ухвалення вирішення в таких умовах приведена на рис 5 2

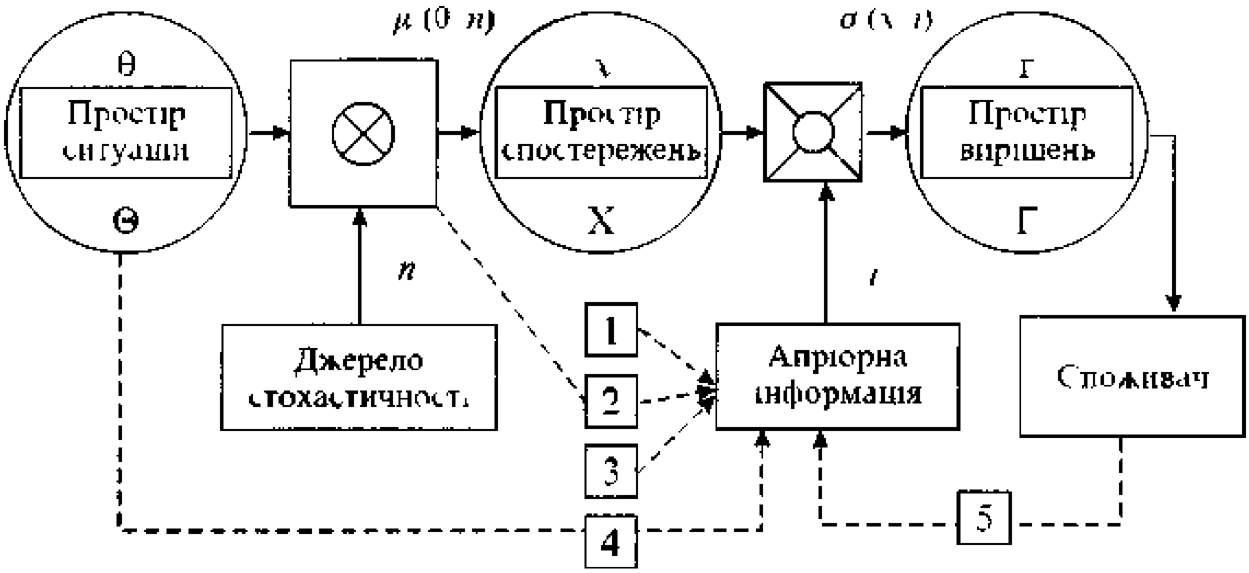

Рис 5 2 Загальна схема прийняття статистичних вирішення

Процедура вибору зображена як дія деякого оператора *а* над вибіркою *х* Даний оператор називають *вирішальною функцією,* оскільки він кожній вибірці х ставить у відповідність вирішення 7= *а{х, і)* Аргумент *і* показує, що ту саму вибірку можна обробляти по-різному

Проблема побудови вирішальної функції тісно пов'язана з роллю апріорної інформації

Апріорна інформація — це зведення, які вже є до того, як сформована вирішальна функція Вона характеризує простір ситуацій 0, природу випадкових факторів *п,* оператор /І, що визначає характер взаємодії *О in;* простір спостережень *X,* вимоги споживача до якості вирішення.

Залежно від повноти апріорної інформації по-різному ставлять і вирішують статистичні задачі вибору. Найбільш повне завдання апріорної інформації включає: розподіл *Р(&),* Єє 0; умовний розподіл вибіркових значень *F(x/&), хе X, в є &* функцію втрат *I (у, в),* що виражає відношення споживача вирішення до розбіжності між *у* і дійсним станом *в.* Середнє значення втрат, яке пов'язане з конкретним алгоритмом обробки спостережень, називають *байєсовим ризиком R* і приймають як міру якості цього алгоритму. Оптимальну в цьому смислові процедуру *f* вважають найкращим вирішенням задачі:

$$\gamma^o(x) = \arg\min_{\gamma(x)} R(\gamma(x)) = \arg\min_{\gamma(\cdot)} M_{X\Theta}\, l(\gamma(x), \theta) =$$

$$= \arg\min_{\gamma(\cdot)} \int_{\Upsilon}\int_{\Theta} l(\gamma(x), \theta)\, dF\left(\frac{x}{\theta}\right) dP(\theta).$$

Менш повне завдання апріорної інформації виключає *Р(в)*. У синтезі алгоритмів бере участь тільки інформація про сімейство функцій *F(x | в)*. Для оцінки розглянутого параметра ***в*** використовують ***метод максимальної правдоподібності,*** відповідно з яким

$$\hat{\theta}_{МП} = \arg\max_{\theta} L(\theta | x_1, \ldots, x_N).$$

Функція *L* може бути отримана способом підстановки у функцію щільності *f(x | в)* вибіркових значень *X*,..., x_N. Якщо розглядати дану функцію **в** залежності від *в,* тоді

$$f(x_1, \ldots, x_N | \theta) = L(\theta | x_1, \ldots, x_N).$$

Цю функцію називають ***функцією правдоподібності.***

Коли за вибіркою хь ..., x_N слід прийняти вирішення на користь однієї з конкуруючих гіпотез $H_O I H_b$ обчислюють ***відношення правдоподібності*** /(ль ..., x_N *\Hi)* */f*(x_b ..., x_N ІД,). Якщо це відношення перевищує заданий поріг, тоді вибирають гіпотезу Д, якщо ні, тоді вибирають гіпотезу Д.

При відсутності апріорної інформації про розподіл *F(x)* прийняття вирішення здійснюють у рамках *непараметричної статистики.*

Для цього використовують методи теорії шваріантостей й евристики Теорія шваріантостей пропонує методи, шо використовують властивості симетрії деяких задач, а також метод інтерпретації статистик як функціоналів від оцінок розподілів Останній метод заснований на двох ідеях По-перше, усяка статистична задача може бути зведена до вибору значення деякого функціонала *J(F)* від невідомого розподілу *F(x)* По-друге, оцінку цікавлячого нас функціонала (тобто вибір альтернативи) передбачають одержувати, підставляючи в нього замість невідомих розподілів їхні непараметричні оцінки

Останнім часом на практиці прийняття вирішень отримала поширення *робасна статистика* її основна ідея полягає в тому, щоб трохи поступитись оптимальшстю на опорному розподілі, що задається точно, але зате забезпечити деякий гарантований рівень якості вирішення на всіх інших розподілах, що лежать біля опорного вирішення Наприклад, для отримання робасних ощшок середнього рекомендуься обчислювати середнє арифметичне, відкинувши з вибірки найбільше і найменше спостереження

Як описування альтернатив, так і описування правил їхнього порівняння дають в термінах тієї чи іншої вимірювальної шкали Іноді виникають ситуації, описати які можна лише в розмитих шкалах У результаті ми приходимо до задач вибору в умовах розпливчастої невизначеності Для описування розпливчастої невизначеності використовують понятійний апарат теорц розпливчастих множин, запропонованої Л Заде

Розпливчаста множина *А* складається з невизначеного числа елементів *х*

Ознаки, за якими елементи включають у розпливчасту множину, не дозволяють однозначно відокремити всі елементи, шо входять до неї, від елементів, ги не приналежних

Важливим є поняття *функціїприналежності/$и_A(х)$* Ця функція виражає ступінь приналежності елемента А ДО розпливчастої множини *А* і змшюється у межах 0 < *$ц_А$ (х)* < 1 Існує кілька підходів до визначення функцц *$ц_А$* (А)

1) евристичний підхід, коли суб'єкт сам визначає, як вш розуміє ступінь приналежності,

2) статистичний підхід, при якому *$ц_А(х)$* визначають усередненням функцій, що задаються різними експертами,

3) часткове завдання uj(x) пояснюючими прикладами і наступне визначення усієї функції придатним методом,

4) штервальне визначення типу завдання песимістичної та оптимістичної границь для */$и_Л$* (А) функції,

5) кратна розпливчастість, тобто завдання *$ц_А(х)$,* як розмитої множини за допомогою функції приналежності другого порядку *$ц_А$Л$ц_А$(х))*

Ідея прийняття вирішення у розпливчастій ситуації полягає в тому, щоб і мети і обмеження представляти як розмиті множини на безлічі альтернатив У випадку однієї мети й одного обмеження це відповідає завданню множини *G= {х, ц,,(х)} \С={х, Цс(х)}* Наступний важливий крок є у визначенні розмитого вирішення *D* як перетинання розмитої мети *G* і розмитого обмеження С, тобто

$$\mu_D(x) = \min\left[\mu_G(x), \mu_C(x)\right]$$

Якщо з розмитої множини *D* потрібно виділити якусь одну альтернативу, тоді можна поступити по-різному (аж до рандомізащі вибору), але можливий варіант складається у максимізаші *$ц_о(х)$*

$$x^* = \arg\max_{x \in X} \mu_D(x)$$

При такому підході функція приналежності перетворюється в суперкритерш якості

У практиці застосування розпливчастих варіантів критеріальних задач вибору зустрічаються задачі з використанням відстаней між точками в просторі альтернатив При розпливчастому описуванні альтернатив пропонується «відстань» визначати через модулі різностей функцій приналежності, наприклад

$$d(x_i, x_j) = \left(\frac{1}{m}\sum_{r=1}^{m}\left|\mu_r(x_i) - \mu_r(x_j)\right|^p\right)^{\frac{1}{p}}$$

де *$іл_г(х)$* — функція приналежності до цікавлячої нас множини за г-тою ознакою

Розпливчасті варіанти вибору можливо описувати і мовою бінарних відношень *Розмите відношення R слабкого порядку* визначають як задовольняюче розмитим умовам зв'язності і транзитивності

$$x_i \neq x_j \Rightarrow \mu_R(x_i, x_j) > 0 \text{ чи } \mu_R(x_i, x_j) > 0 \text{ (зв'язність)},$$
$$\mu_R(x_i, x_k) = \max\{\min[\mu_R(x_i, x_j), \mu_R(x_i, x_j)]\} \text{ (транзитивність)}$$

Для розмитого відношення сильного порядку справедлива **умова**

$$\mu_R(x_i, x_j) > 0 \Rightarrow \mu_R(x_j, x_i) = 0,$$

тобто асиметрія

Л Заде показав, що будь-яке розпливчасте відношення *R* допускає розкладання за а у вигляді об'єднання нерозмитих безлічей R_a з функціями приналежності

$$\mu_{R\alpha}(x_i, x_j) = \begin{cases} \alpha \text{ при } (x_i, x_j) \in R_\alpha, \\ 0 \text{ при } (x_i, x_j) \notin R_\alpha, \end{cases}$$

де 0 < *a* < 1 (що можна представити як розшарування обсягу під $ц_к$ на горизонтальні шари рівнів *a*

Це дозволяє перейти від розпливчастого описування колективного упорядкування альтернатив до нерозпливчастої множини альтернатив, відібраних «зі ступенем згоди на рівні а»

5 4 Груповий метод вибору

Вибір за принципом формування «загальної думки» є розповсюдженим методом прийняття вирішення Звичайно загальна думка формується з урахуванням індивідуальних переваг учасників групового вибору Нехай, наприклад, на безлічі альтернатив *X* задано *n* бінарних відношень *Яі, R₂, , Rₙ* Ставиться задача *про вироблення деякого нового відношення R, що узгоджує тдивідусиіьт вибори, виражає в якомусь смислові «загсиіьну думку» і приймається за груповий вибір*

$$R = \Phi(R_1, R_2, \ \ , R_i)$$

Один з популярних способів колективного вибору — голосування Вирішення приймають за *правилом більшості.* Прийнятою усіма вважають альтернативу, шо одержала найбільше число голосів

Ця широко застосовувана й v багатьох випадках успішна процедура з очевидними перевагами має деякі скриті особливості і недоліки Правило більшості лише узагальнює індивідуальні переваги, але його результат не є критерієм істини Це правило не враховує сили перевага кожного з голосуючих Для усунення цього недоліку формуються різні прийоми Наприклад, у криміналістичній практиці експертам пропонують в одному ряду з двома варіантами *a* \ *b* упорядкувати за перевагою ще кілька альтернатив, скажемо *c, die* Нехай перший експерт дав упорядкування *(c d, a, Ь, є),* а другий — *(b, c, d, e, a)* Тоді можна зробити висновок, що ступінь переваги *b* порівняно з *a* у другого експерта більше, ніж у першого, і прийняти вирішення на користь *b* Але і в цьому випадку прийняте вирішення сумнівне, тому що передбачається однакова компетентність експертів, які вимагають перевірки у відповідальних випадках

Другий прийом — це можливість відмовлення від вибору через недосягнення необхідної більшості Наприклад, три експерти більшістю голосів вирішують питання, яка з альтернатив більш краща При такій постановці питання вони не можуть не зробити вибір Але у цьому випадку правило голосування виявляється нетранзитивним Причина нетранзитивності групового вибору в цьому випадку складається в циклічності сукупності індивідуальних переваг Але це лише окремий випадок більш загального явища, що отримало *назву* ***парадокса*** *Ерроу (чи теореми про неможливість).* Смисл даної теореми полягає в наступному

Вибір можна вважати узгодженим при виконанні таких умов

1 *Умова монотонності* Якщо в результаті групового вибору перевага була віддана альтернативі л, тоді це вирішення не повинно мінятись, якщо хто-небудь, що раніше відкидав х, змінив свою перевагу на його користь Крім того, *п > 2,* число альтернатив більше або дорівнює 31 /'визначена для будь-яких *{R,}*

2 *Умова незаіїежності не зв язаних альтернатив* Якщо змши індивідуальних переваг торкнулися визначених альтернатив, тоді в новому груповому упорядкуванні порядок цих альтернатив не повинен мінятись

3 *Умова суверенності Для* будь-якої пари альтернатив *х* і *у* існує такий набір індивідуальних переваг, для якого $\Phi(K_b \quad, R_n) = (x > y)$.

4 *Умова відсутності диктаторства* Не повинно бути такого щдищдуума, для якого з його переваги *х> у* (при будь-яких л і *у)* виходить, що *0{R\,   , R_r) = (х >* у) не залежно від переваг інших індивідуумів Парадокс Ерроу полягає в тому, що перші три умови суперечать четвертій Не існує правила вибору, що задовольняє всім умовам

Основною причиною цього є можливість циклічних безлічей ранжирувань

Задачі групового вибору часто все-таки МОЖУТЬ бути дозволені По-перше, у ряді випадків циклічні ранжирування можуть бути відсутніми, або вони не охоплюють «найбільш важливі» альтернативи, або приймаються заходи для їхнього виявлення й усунення По-друге, у багатьох випадках «диктаторський» принцип узгодження не є прийнятним По-третє, перехід (коли це можливо) до використання єдиної числової, а не індивідуальних порядкових шкал переваг, може взагалі анулювати проблему нетранзитивності По-четверте, v реальних ситуаціях мажоритарні правила застосовуються в комбінації з іншими правилами, так що, утворивши, наприклад, коаліцію, групи суб'єктів можуть блокувати дію голосування

Привабливий метод агрегування переваг запропонований Кон-дорсе Ідея цього методу зводиться до пошуку порядку колективного добробуту (ПКД), що якнайменше відрізняється від (зваженого) турніру бінарних порівнянь Кондорсе припускав, що існує деяке об'єктивне ранжирування альтернатив, яке може бути витягнуте з профілів індивідуальних переваг Він запропонував просту біноміальну модель помилок вибірників (чи експертів) Для будь-якого бінарного порівняння для кожного вибірника задана деяка ймовірність того, що альтернативи будуть упорядковані правильно Передбачають, що вибірники мають рівні можливості і не існує кореляції між порівняннями в різних парах У випадку бінарного вибору останній за найбільшою правдоподібністю є, мабуть, правилом більшості Звідси, зпдно з Кондорсе, виходить головне обґрунтування принципу більшостіЯкщо різні вибірники (експерти) мають різну компетентність у порівнянні почапав, тоді правильний вибір за найбільшою правдоподібністю не настільки очевидний Виявляється, шо погрібно визначити зважену більшість думок агентів, при якому ваги експертиз змінюються як log *[p, / {*1 - *p,)*], якщо *p,* — імовірність правильної відповіді для /-того експерта

Якщо вибірники мають однакову компетентність, об'єднання різних думок вибірників здійснюється так проводиться мажоритарний турнір, що виходить при всіх парних голосуваннях з фіксацією поданих голосів Далі альтернативи ранжують відповідно до найбільш ймовірної комбінації думок У термшолопі сучасної статистики — це критерій максимальної правдоподібності

У методиках організації групового вибору спеціальну увагу приділяють створенню сприятливих умов і нейтралізаци факторів, що несприятливо впливають на роботу експертів Дуже важливу роль відіграють фактори психологічного характеру особистісш якості експертів (конформізм, особиста зацікавленість, скла характеру, самолюбність та ш), міжособистісш відносини з іншими експертами, критика оцінок, висловлених експертом, відповідальність за використання результатів експертизи Через те, що взаємодія між експертами може, як стимулювати, так і придушувати їхню діяльність, у різних випадках використовують методики експертиз, які мають різні ступені і характер взаємного впливу експертів один на другого анонімні і відкриті опитування, анкетування, наради, ділові ігри, мозковий штурм та ш

Компенсація впливу взаємної критики оцінок експертів реалізується у методі «Делфі»

За методом «Делфі» оцінку альтернатив здійснюють у кілька турів Після кожного туру групі експертів повідомляють повний його підсумок При цьому зберігається анонімність оцінок альтернатив кожним з експертів Це приводить до того, шо кожен з експертів починає критикувати загальний підсумок і прислухатись до критики, що відноситься до нього особисто Усунення психологічних труднощів, пов'язаних з персональною критикою, додає самій критиці велику об'єктивність, вона легше сприймається

Оцінку альтернатив за методом «Делфі» здійснюють у чотири етапи

1) роздавання анкет, збір оцінок, їхнє узагальнене представлення з указівкою розкиду думок,

2) повідомлення підсумків і запит пояснень причин індивідуального відхилення від середньої чи медіанної оцінки першої ітерації,

3) повідомлення всіх пояснень і запит контраргументів на них,

4) повідомлення заперечень і запит нових оцінок альтернатив, якщо експерт побажає їх змінити, знаходження остаточного підсумку Залежно від методики експертизи для ухвалення вирішення використовують різні оцінні функції, що виражають загальну думку ,л

Якщо, наприклад, експерти оцінюють альтернативи в числовий шкалах, тоді для вираження загальної думки використовують ви-t біркове середнє

$$R = \frac{1}{n}\sum_{j=1}^{n} q_j(x_i),$$

де $\partial_j(x_i)$ — оцінка /-тої альтернативи/-тим експертом, n — число експертів

Коли альтернативи не можна оцінити одним числом при потребі урахування декількох характеристик системи (економічних, естетичних, ергономічних, функціональних та ш), експертам пропонують дати оцінки окремо за кожною характеристикою, а також ступінь п важливості У цьому випадку для вираження загальної думки використовують оцінну функцію у формі

$$R = \frac{1}{n}\sum_{j}\sum_{k}\lambda_{jk} q_{jk}(x_i),$$

де <7д(х,) — оцінка ;-тоі альтернативи/-тим експертом за А.-тою ознакою, $X_{lк}$ — оцінка важливості А.-тш ознаки /-тим експертом

Якщо група експертів неоднорідна за кваліфікацією, тоді враховують компетентність експертів за допомогою коефіцієнта *а,* = ^*(щ* / X «/*) В цьому випадку оцінна функція представляється у вигляді

$$R = \sum_{j}\sum_{k}\alpha_j \lambda_k q_{jk}(x_i),$$

де *(Xij* — коефіцієнт, що враховує ошнкл компетентності /-того експерта /-тим, (0<а,,< 1)

Якщо експерти лише упорядковують альтернативи, тоді оцінку виконують за відносною частотою переваг даної альтернативи або за результатами ранжирування

6 УПРАВЛІННЯ СИСТЕМОЮ

Мета управління системою
Етапи, що включає процес управління
Завдання регулювання
Класифікація системи управління за ознакою керованості і спостережності
Умови керованості і спостережності для системи з діагональною матрицею
Умови повної керованості і спостережності
Показник якості управління
Інтегральні показники якості, що використовують для оцінки процесу управління
Сутність оптимального управління
Надійність системи управління
Показники надійності
Адекватність системи, що управляє, об'єкту управління за складністю

6.1. Загальні положення

Під ***управлінням*** *розуміють здійснення сукупності дій, спрямованих на підтримку і поліпшення функціонування об'єкта, яким управляють, відповідно з метою і програмою управління.* Управління здійснюють способом реалізації комплексу заходів, що включають політичні, соціальні, адміністративні, юридичні, економічні та ін.

Теорію управління почали розробляти в рамках кібернетики. ***Кібернетика*** (від грецького kibernetike - майстерність управління) - наука про загальні закони отримання, збереження, переробки і перетворення інформації в системах будь-якої природи.

Управління системою вимагає забезпечити її цілеспрямоване поводження при умовах роботи, що змінюються. Це досягається належною організацією системи, під якою розуміють її структуру і спосіб функціонування.

Якщо організація системи однозначно визначена при її створенні, то управління нею зводиться до забезпечення розрахункових значень її перемінних при відхиленнях зовнішніх умов і параметрів системи від розрахункових. В інших випадках, коли компоненти системи і способи їхнього об'єднання вибираються залежно від класу розв'язуваних завдань (наприклад, агрегатний принцип побудови обчислювальних систем), то вибір структури і способу її функціонування є завданнями управління.

При формуванні системи її елементи, якими управляють, поєднуються в частину, якою управляють - об'єкт управління (ОУ). Сукупність елементів, що управляють, утворюють систему, яка управляє (УС). Обидві частини взаємодіють за допомогою кінцевого числа інформаційних зв'язків.

Необхідне поводження системи досягається способом управління її входами *X* чи незалежними від входів координатами стану *Q* - параметрами системи, або спільно тими й іншими.

Процес управління звичайно складається з двох тісно пов'язаних етапів.

Перший етап - розробка програми (планування), що визначає необхідне поводження об'єкта управління

Другий етап - реалізація програми Другий етап часто називають регулюванням, керівництвом, оперативним управлінням

Чітке розмежування функцій планування і регулювання не завжди можливе Чим менше об'єкт управління і динамічніше ситуація, тим тісніше переплітаються ці функції.

Управління динамічною системою, що піддається різноманітному і часто мінливому впливу зовнішнього середовища, сполучено з необхідністю залучення величезного обсягу інформації. Тому структура системи, що управляє (УС), будується за ієрархічним принципом.

У раціонально організованій ієрархічній системі, що управляє:

- кожен її рівень m здійснює управління ступінню (m - 1) и одночасно управляється рівнем (m + 1) Усі рівні інформаційно пов'язані між собою;

- інформація, що надходить від об'єкта управління, рухається в протилежному напрямку - від нижніх рівнів до верхнього і при цьому послідовно «стискується». Найчастіше нижчий рівень управління постає перед вищим як «чорна скринька», що інформує його лише про результати своєї діяльності, але не про внутрішні процеси, пов'язані з її реалізацією;

- чим самостійніше функціонує кожний рівень управління, тим більшу частину інформації, що надходить, він «поглинає» і відносно меншу частину передає на наступний рівень. Самостійність кожного рівня управління в рамках компетенції і послідовне «стиснення» інформації - головні умови ефективності багаторівневого управління;

- функціонування системи управління як єдиного цілого досягається узгодженням цілей управління кожним її елементом і їхніми сукупностями з цілями, що стоять перед системою. Це значить, що ієрархія системи, що управляє, ставиться у відповідність ієрархи цілей.

Розробка програми управління, з яким би ступенем деталізації вона не виконувалась, охоплює лише основні фактори, що впливають на поводження об'єкта. Вона відображує деякі ідеалізовані умови й обмеження, пов'язані з реалізацією програми. Тому програма завжди базується на математичному трактуванні зв'язків між показниками. Без цього їхній попередній розрахунок узагалі неможливий.

У регулюванні приходиться враховувати нескінченну безліч факторів і зв'язків між ними. Іноді неможливо заздалегідь оцінити кожний з них і математичне трактування зв'язків між ними. Тому вирішальне значення в регулюванні здобуває *принцип розробки впливу, що управляє, по відхиленню фактичного значення величини, якою управляють, від її необхідного (розрахункового) значення незалежно від причин, що викликали зазначене* *відхилення.*

Практична реалізація даного принципу здійснюється за допомогою зворотного зв'язку (як правило, негативного).

Розрізняють три типи основних завдань регулювання ***стабілізація, програмне регулювання* і *спостереження.***

Мета ***стабілізації*** - підтримування заданого постійного значення вихідної величини об'єкта регулювання. Стабілізація здійснюється за допомогою регулятора (наприклад, регулятор Уатта).

Програмне регулювання забезпечує зміну вихідної перемінної об'єкта управління відповідно заданій програмі Зміна вихідної перемінної може бути задана у вигляді функції часу або іншого аргументу, наприклад, інтенсивності входу об'єкта. Завданням системи регулювання є в даному випадку реалізація цієї програми при наявності тих чи інших перешкод. Характерно, що стабілізація є частковим випадком програмного регулювання

Спостереження відрізняється тим, що програма не розраховується заздалегідь, а визначається поводженням об'єкта, який спостерігають. Прикладом може служити автоматичне підстроювання радіоприймача, що стежить за обраною хвилею.

Наявності зворотного зв'язку не завжди досить для забезпечення стійкості управління. Запізнювання, інерція системи, приховані нелінійності та ін. не можуть бути належним чином враховані при виборі параметрів зворотного зв'язку. Недостатність апріорної інформації про їхні фактичні значення приводить до того, що ефективний. У деяких умовах зворотний зв'язок не може бути реалізований і стійкість не може бути забезпечена. Тоді найефективнішим стає регулювання, при якому регулятору надаються властивості ***пристосування (адаптації)*** до характеристик середовища, що змінюються, і самого об'єкта управління. Програма для такого регулятора формується у вигляді мети управління й обмежень на перемінні і параметри, а її уточнення і корегування здійснюється самим регулятором за допомогою широко розвинутих зворотних зв'язків.

Адаптація поклала початок створенню штучних систем, які мають властивості пристосування до середовища. Таке пристосування досягається через навчання. *Під **навчанням** розуміють накопичення інформації про хід процесу регулювання в минулому і її використання для удосконалювання цього процесу на основі деякого набору правил і стимулювання поводження системи.*

Власне кажучи, при адаптації змінюється співвідношення програмного (що завдає) блоку і регулятора в системах управління. Частина роботи з формування програми переміщується в регулятор, а адаптація самого блоку, що завдає, складається у зміні цілей і обмежень на перемінні і параметри регулятора.

Іноді системи, що мають і програмний блок і регулятор, називають ***системами, що самоорганізуються.*** У їхньому складі виділяють системи, що самонавчаються, що самоорганізуються і самоудосконалюються. Різниця цих систем полягає у способі пристосування до середовища.

Система регулювання залежно від її організації і використаних технічних засобів функціонує як «чисто людська» чи в умовах автоматичної системи управління (АСУ) як «людино-машинна». В іншому випадку за людиною залишаються лише функції прийняття рішень.

Управління системою завжди спрямоване на обмеження числа ступенів її свободи чи діапазонів зміни її перемінних, а найчастіше того й іншого. Так досягається зменшення розмаїтості в поводженні системи - саме до цього зводиться завдання управління.

Для вирішення завдань управління необхідно, щоб інформаційна потужність системи, що управляє, (чи її власна інформаційна розмаїтість) були не менше розмаїтості об'єкта управління (тобто розв'язуваного завдання управління). Дана умова ілюструє ***принцип необхідної розмаїтості*** стосовно до завдання управління необхідно, щоб розмаїтість системи, що управляє (УС), була не менше розмаїтості об'єкта управління (ОУ). У реальних системах управління розмаїтість ОУ і середовища настільки велика, що зазначена умова не виконується. Тому УС формує гомоморфну модель ОУ, використовуючи агрегування, лінеаризацію зв'язків, заміну стохастичних залежностей детермінованими та ін. Часто вплив не врахованих факторів враховують за допомогою *«зовнішнього доповнення»* до моделі. Таким доповненням є деяка «чорна скринька», наприклад, рандомізатор (датчик випадкових чисел). Зовнішнє доповнення на основі додаткових відомостей, знань і досвіду вносить поправки в модельні розрахунки.

6.2. Керованість і спостережність

Поняття керованості і спостережності специфічні для методу *простору станів.* При класичному описанні динамічних систем у термінах *вхід* - *вихід* проблема керованості і спостережності не виникає.

При використанні методу *простору станів* не втрачається цілісна картина об'єкта. При записі рівняння стану передбачається, що в об'єкті можуть відбуватись інші процеси й існувати перемінні, не доступні для спостереження чи ті, що не піддаються управлінню.

Розглянемо проблему керованості і спостережності на якісному прикладі, запропонованому Дж. Медич.

Нехай динамічна система описується вектором стану ***Q,*** вектором входів ***X*** і вектором виходів ***Y***. Схема системи представлена на рис. 6 1, де ***Y*** - вектор, компонентами якого є перші ***k*** компоненти вектора стану, $q_1, q_2, \ldots, q_k$. Зі структури системи ясно, що значення інших компонентів вектора стану $(q_{k+1}, q_{k+2}, \ldots, q_m)$ не можна визначити на основі наявних відомостей про вихідний вектор ***Y,*** тому що ці перемінні не впливають на $q_1, q_2, \ldots, q_k$ і не включені до складу вектора, ***Y,*** який спостерігають. Отже, система не є тією, що спостерігається. Але, якщо ***X*** впливає на всі перемінні стани ***Q,*** система є керованою.

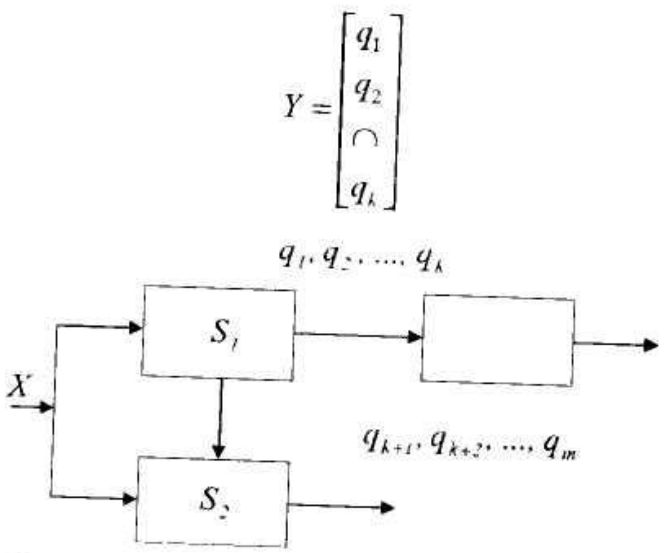

Рис. 6.1. **Схема системи, що не спостерігається, але керована**

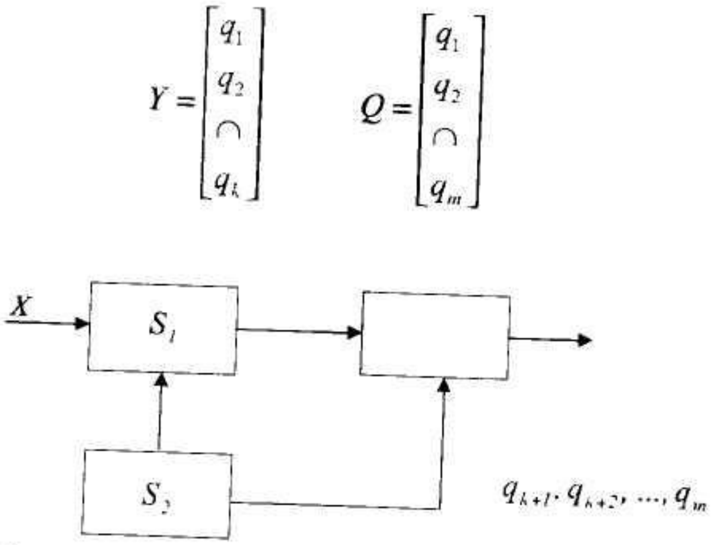

Рис. 6.2. **Схема системи, що спостерігається, але некерована**

Аналогічно система, показана на рис. 6.2, буде тією, що спостерігається, але не керованою, тому що сигнал ***X*** впливає тільки на перемінні $q_1, q_2, ..., q_k$, а на перемінні $\left(q_{k+1}, q_{k+2}, ..., q_m\right)$ ззовні впливати не можна.

***Враховуючи** викладене, всі системи можна розділити на такі чотири категорії: що спостерігаються і керовані; що спостерігаються але некеровані; що не спостерігаються, але керовані; що не спостерігаються і некеровані.*

Поняття керованості і спостережності мають принципове значення при дослідженні систем будь-якої природи. Неврахування некерованості і неспостережності може привести до помилкових висновків.

Умови керованості і спостережності можна зв'язати з видом матриць, що описують систему. Для приклада розглянемо, при яких умовах може виникнути неспостережність чи некерованість у найпростішому випадку, коли матриця *A* діагональна, тобто $A = \operatorname{diag} a_{ij}$.

Нехай система має вигляд, показаний на рис. 6.3, де i - вектори розмірності 2, j - вектор розмірності 3.

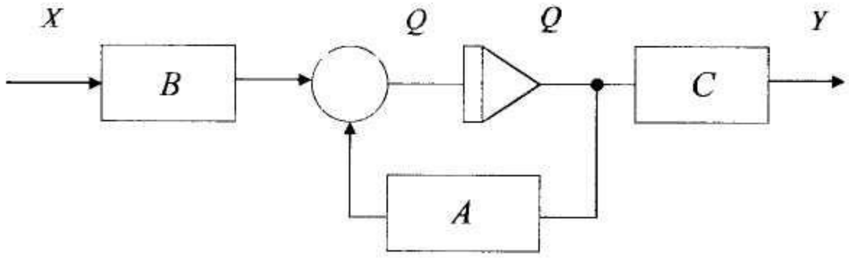

Рис. 6.3. **Схема системи**

Управління системи в матричному вигляді записується так:

$$\dot{Q} = A_{(2x2)}Q + B_{(2x3)}X;$$
$$Y = C_{(2x2)}Q + D_{(2x3)}X,$$

$$A = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix}; \quad B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix}; \quad C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}; \quad D = 0.$$

Якщо один з рядків у матриці ***B*** (наприклад, перший) складається цілком з нульових елементів, тоді відповідна координата (перша) буде некерованою, тому що жодна з трьох керуючих дій не чинить керуючого впливу на q_1.

Аналогічно, якщо один із двох стовпців матриці ***C*** складається з нульових елементів, то відповідна координата вектора стану не вчинить впливу ні на один із двох сигналів, що спостерігаються - y_1 і y_2. Її поводження ніяк не буде виявлятись зовні, координата неспостережна.

Таким чином, для системи найпростішого вигляду з діагональною матрицею ***A*** можна було б зв'язати умови керованості і спостережності з виглядом матриць ***B*** і ***C***: *керованість означає відсутність нульового рядка у B, спостережність - відсутність нульового стовпця у C.*

У загальному випадку матриця ***A*** недіагональна, а самі перемінні стану можуть впливати один на другий. Тому, навіть якщо немає безпосереднього впливу управління на дану координату стану ***q***, такий вплив може виникнути більш складним чином: управління ***X*** впливає на якусь іншу координату, а вже ця координата через матрицю ***A*** впливає на дану координату. У такому випадку роль матриці ***B*** відіграє добуток матриць ***AB***. Якщо й у цьому випадку впливу ***X*** на координату ***q***, немає, тоді може виявитись, що такий вплив здійснюється ще більш опосередкованим чином - через матрицю $A\ (AB) = A^2 B$ та ін.

Тоді умову повної керованості можна записати так: *система є цілком керованою, якщо ранг матриці* **[B, AB, A^2 B, A^{m-1}, B]** дорівнює ***m.***

Рангом матриці називають максимальний розмір її мінорів, відмінних від нуля. Мінор k-того порядку матриці розмірністю (m -1) виходить викреслюванням будь-яких (m - k) рядків і (l – k) стовпців матриці/

Аналогічна і умова спостережності системи/ *Система цілком спостережна, якщо ранг матриці* $\left[C^T, A^T, C^T, AT^2, C^T ... A^{T^{m-1}}, C^T\right]$ *дорівнює m (тут індекс T означає транспортування).*

6.3 Показники функціонування систем управління

Різноманітний характер процесу управління вимагає вибрати такий варіант, що забезпечує максимальну ефективність управління, досягнення мети найкращим способом. Це завдання стає розв'язаним лише тоді, коли існує кількісна характеристика, що дозволяє об'єктивно зіставити результати управління. Таку характеристику називають ***показником*** якості чи ефективності управління. Вибір показника якості управління диктується призначенням системи, метою й умовами її функціонування і розвитку. Звичайно він задається як функція чи функціонал вхідних (вихідних) перемінних, параметрів об'єктів управління, часу. Таку функцію часто називають ***цільовою,*** тому що вона дає кількісну міру мети управління.

При конструюванні об'єктів управління в техніці показники якості функціонування майбутніх систем можуть використовуватись з метою додання об'єкту управління визначених конструктивних характеристик. Це означає, що структуру об'єкта підбирають під деякий оптимальний режим функціонування системи в цілому. У цих випадках намагаються створити конструкцію, адекватну функції, що виконується Надалі конструкція об'єкта управління не змінюється.

Біологічні системи управління в процесі еволюції змінювались і за структурою і за функціями. У процесі еволюції виробилась єдність структури і функції всіх частин біологічних систем управління. Тому, якщо при розгляді показників функціонування біологічних систем можна використовувати результати, отримані в теорії автоматичного управління, тоді при оцінці якості структури можна спиратись лише на загальні положення, вироблені біологією.

Формальні показники якості - сукупність прийнятих чи постульованих кількісних характеристик, що дозволяють оцінити якість роботи системи.

Для оцінки якості найчастіше використовують інтегральний показник:

$$I = \int_0^T f(x)dt,$$

де $f(x)$ - функція перемінних, що характеризують стан системи.

З цього показника залежно від вигляду функції $f(x)$ можна отримати оцінки для різних окремих випадків. Якщо $f(x)=1$, тоді інтегральний показник оцінює час перехідного процесу:

$$I = \int_0^T dt$$

Якщо $f(x) = x(t)$ - погрішність перехідного процесу в системі управління чи $f(x) = |x(t)|$, тоді інтегральні показники оцінюють відхилення реального перехідного процесу від бажаного на всьому інтервалі часу управління:

$$I = \int_0^T x(t)dt,$$

$$I = \int_0^T |x(t)|dt.$$

Якщо $f(x) = x^2(t)$, тоді інтегральний показник оцінює якість перехідного процесу незалежно від знака:

$$I = \int_0^T x^2(t)dt.$$

Цей показник називають квадратичною погрішністю. При дії в системі управління випадкових збуджень розповсюдженим показником якості є середня квадратична погрішність, яка є характеристикою розсіювання можливих значень випадкової величини щодо їхнього середнього значення.

Поряд з цими оцінками при синтезі систем з випадковими впливами використовують питомий ризик, загальний ризик та ін.

Звичайно при побудові системи автоматичного управління ставиться задача мінімізації розглянутих інтегральних показників. Мінімум цих показників розглядають як ознаку **(критерій) оптимальності** системи управління. Як критерії оптимальності може використовуватись швидкодія системи, маса, витрати енергії, коефіцієнт корисної дії та ін.

Задача ***оптимального управління*** формулюється таким чином: задано об'єкт, координати якого описуються n-мірним вектором $x = \{x_1,\ldots,x_n\}$.

Координати об'єкта змінюються в часі за законом

$$\dot{x}_i = f_i(x,u), \qquad i = 1,\ldots,n,$$

де $f_i(x,u)$ - функція х і r-мірного вектора управління $u = \{u_1,...,u_r\}$.

Вектор **х** характеризує положення об'єкта у фазовому просторі і називається **вектором фазових координат.**

Необхідно вибрати управління, для якого значення функціонала мінімально

$$I = \int_{t_0}^{t_1} f_0(x(t), u(t))dt.$$

Управління і траєкторію, що відповідають рішенню цієї задачі, називають відповідно **оптимальним управлінням і оптимальною траєкторією.**

Не менш важливим показником функціонування системи управління є її ***надійність.*** *Під надійністю систем управління розуміють їхню здатність зберігати найбільш суттєві властивості на заданому рівні протягом визначеного проміжку часу і за певних умов функціонування.* Ступінь надійності системи визначають показниками, пов'язаними з явищем відмови. ***Відмова*** *- подія, що полягає в порушенні працездатності.* Відмови поділяють на поступові і раптові. Поступові відмови проявляються у вигляді поступового виходу параметрів системи за межі встановлених допусків, раптові - у вигляді різкої зміни параметрів, що визначають якість системи.

Показниками надійності є: ймовірність безвідмовної роботи, інтенсивність відмовлень, середній час безвідмовної роботи.

Необхідність зрівноважування розмаїтості системи, що управляє, й об'єкта управління оцінюється за допомогою показників ***адекватності*** за складністю і за рівнем відносної організації.

При встановленні адекватності за складністю між системою, що управляє, й об'єктом управління здійснюється мінімізація функціоналу:

$$I = \int_0^T f(H_m)dt,$$

де [0, T] - інтервал розвитку системи управління, H_m - складність.

Складність для об'єкта управління H_m^0 і системи управління H_m^s визначається так:

$$H_m^0 = \log n_0, \qquad H_m^s = \log n_s,$$

де n_0, n_s - число станів об’єкта і системи управління.

У цьому випадку

$$I = \int_0^T \left(\log n_0 - \log n_s\right)^2 dt.$$

Функціональне пристосування системи, що управляє, до об'єкта управління може характеризуватись аналогічним функціоналом, який беруть за рівнем відносної організації

$$\min I = \min \int_0^T \left[\frac{H^s}{H_m^s} - \frac{H^0}{H_m^0} \right] dt,$$

де

$$H^0 = \sum_{i=1}^{n_0} P_i^0 \log P_i^0,$$

$$H^s = \sum_{i=1}^{n_s} P_i^s \log P_i^s,$$

P_i^0, P_i^s - імовірності прийняття об’єктом і системою управління і-того стану.

Для біологічних систем, що еволюціонують, визначити складність і відносну організацію не завжди можливо. Тому, спостерігаючи тільки за системою управління, можна розглядати інтегральний показник, використовуючи принцип ***самоорганізації:***

$$I = \int_0^T \frac{\Delta H_k^Q}{\Delta H^Q} \varphi(Q) dQ.$$

де ΔH_k^Q - зміна невизначеності контрольних послідовностей; ΔH^Q - зміна невизначеності функціонування біологічної системи управління; $\varphi(Q)$- деяка функція послідовностей; Q - набір послідовностей, що навчають.

Конкретний вигляд підінтегрального вираження невідомий, і можна говорити, що для самоорганізації біологічної системи управління необхідно, щоб даний функціонал досягав свого максимального значення за мінімальний час і з мінімальною витратою речовини и енергії.

7 ПРОЕКТУВАННЯ І КОНСТРУЮВАННЯ СИСТЕМ

Проектуванням і проект.
Конструювання і конструкція.
Часткове і системне проектування.
Етапи системного проектування.
Область відносної цілісності.
Основні елементи проектування і конструювання як науки про технічну творчість
ТТС, МП і МК.
ТЗК, МДКР і СД та НК.
Ефект від нормалізації.
Методологія проектування.
Сутність і розходження методу і методики.
Евристичний метод проектування.
Алгоритмічний метод проектування.
Систематичний метод проектування.
Основні методологічні принципи проектування.
Принцип максимізації математичного очікування та субоптимізації.
Принцип явищ з малою ймовірністю та усунення слабких ланок.
Принцип максимізації довгострокової ефективності, централізації та здійсненності проекту.
Основні методичні принципи проектування.
Принцип автономності та найменшої взаємодії.
Принцип взаємо доповнення та забезпечення динамічної достатності.
Принцип ієрархії у задоволенні вимог, узгодження норм і цілей функціонування елементів і системи в цілому та динамічної рівноваги.
У чому полягає відмінність між методологічними і методичними принципами.
Принципи конструювання.
Принцип оптимального навантаження.
Принцип оптимального матеріалу.
Принцип оптимальної стабільності та оптимальних співвідношень взаємозалежних величин.
Соціальна потреблема.
Сукупність техніко-економічних дій, спрямованих на задоволення соціальних потреб при вирішенні завдань проектування.
Фактори, що є визначальними в моделях попиту і споживання.
Різниця між кінематичними і динамічними моделями попиту і споживання.
Функція корисності.
Кількісна характеристика корисності.
Визначення кривої байдужості.
Визначення нормативу і норми споживання.

7.1. Сутність проектування і конструювання

Проектування в широкому смислові - це вибір способу дій людини в процесі взаємодії із середовищем. У вузькому смислові проектування є процес створення проекту системи як логічної основи наступної діяльності людини. Об'єктом проектування є система.

Проект - модель системи, представлена у вигляді креслень, графіків, формул, пояснювальної записки та ін. Проект є результатом розумової діяльності проектувальника в сфері інформації. Система, створена на основі проекту, - результат діяльності людей у сфері матеріальних об'єктів, тобто у сфері маси й енергії.

Ядром процесу проектування є *конструювання.* Об'єктом конструювання є конструкція. *Конструкція* - комплекс структур і станів системи.

Якщо проектування, кінець кінцем, зводиться до обґрунтування необхідності створення системи і її розрахункових характеристик, тоді конструювання - процес підбору необхідних конструктивних характеристик, що визначають логічну основу конструкції. Так, наприклад, якщо на етапі проектування автомобільної дороги обґрунтовують необхідну міцність конструкції дорожнього одягу, тоді на етапі конструювання дорожнього одягу підбирають такі її конструктивні характеристики, які забезпечують її необхідну (задану) міцність. Конструкція, також як і проект, є моделлю системи і протиставляється конкретності.

Прийняття вирішень у проектуванні базується на критеріях соціальної адекватності. Конструювання виходить з техніко-економічних критеріїв.

Розрізняють часткове і системне проектування.

Часткове проектування - проектування частини цілого без урахування властивостей цілого. Часткове проектування найчастіше грунтується на суб'єктивно прийнятих критеріях.

Системне (інтегральне) проектування - проектування частини цілого з точки зору цілого (надсистеми).

Системне проектування включає таю етапи:

1) установлення цілого;
2) визначення зв'язків між частиною і цілим;
3) установлення вимог, що пред'являє ціле до його частин;
4) оптимізація частин цілого за системними вимогами.

Проблема частини і цілого має важливе значення. Поняття частина і ціле відносні. Для натураліста абсолютну цілісність має космос. Але із практичних позицій абсурдно для кожної технічної задачі визначати властивості космосу та їхні відносини до того, що повинно складати вирішення виниклої задачі. Проте суттєві цілісності можуть мати великі розміри.

Як критерій вичленовування частини з цілого можна використовувати специфічну особливість матеріальних комплексів - *здатність бути елементом.* Тут під елементом будемо розуміти надалі неподільну частину системи, що може брати участь у різних системах без зміни своїх властивостей. Здатність бути елементом реалізується у тим більшому ступені, чим менше зв'язків існує між елементом і складним комплексом.

Виділення частини з цілого може ґрунтуватись на розходженні *підсистем* з відповідним числом зовнішніх зв'язків і при бажаному числі внутрішніх зв'язків, тому що очевидно, чим менше число зовнішніх зв'язків (між підсистемами), тим у меншому ступені зміни в одній з підсистем впливають на змни в інших підсистемах.

Важливим моментом у проектуванні систем є визначення *області відносної цілісності.* Цю область часто називають областю значимої взаємодії (рис 7.1).

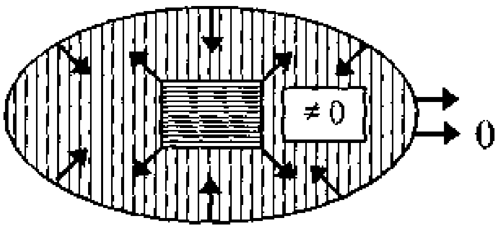

Рис 7 1 Модель задачі визначення відносної цілісності

На рис 7.1 еліпс утворює границю області відносної цілісності Прямокутник позначає об'єкт проектування. Стрілки характеризують взаємні впливи розглянутого об'єкта та інших систем.

Практично область відносної цілісності визначають на основі таких правил:

- границя області повинна охоплювати ті елементи, вплив яких на проектований об'єкт не дорівнює нулю;
- поза цією областю значення дії проектованого об'єкта має прагнути до нуля.

Основними елементами проектування і конструювання, як науки про технічну творчість, є:

- теорія технічних систем;
- методологія проектування;
- методологія конструювання;
- теорія запису конструкції;
- методологія дослідно-конструкторських робіт і системних досліджень;
- нормалізація конструкцій.

Теорія технічних систем і теорія конструювання сприяють раціональному розумінню суттєвих властивостей технічних засобів і тих виробів, що мають стати цими засобами.

Методологія проектування і методологія конструювання включають:

- описання творчої діяльності з виділенням процедур, з яких складаються проектування і конструювання;
- методи творчої діяльності чи їхньої основи.

Теорія запису конструкції - система знань про загальні закони і принципи запису конструктивного вигляду і комплексу розмірів конструкції. Те, що звичайно іменують *кресленням,* складає лише один із способів запису.

Конструктивний вигляд - це якісна властивість конструкції. Комплекс розмірів - спосіб визначення кількісних властивостей конструкції.

Методологія дослідно-конструкторських робіт у взаємозв'язку із системними дослідженнями дає представлення про способи побудови конструкцій і порядок їхнього застосування на кожному етапі конструювання.

Під *нормалізацією* в системному проектуванні розуміють обмеження різноманіття. Частковим випадком нормалізації є *стандартизація.* Нормалізація дає такий ефект:

- поліпшення комунікативності;
- збільшення можливості заміни деталей машин за рахунок їхньої стандартизації;
- зниження матеріальних і операційних витрат;
- спеціалізація виробництва;
- оптимальна концентрація виробництва.

Похідними від нормалізації є *уніфікація* і *типізація.*

Уніфікація заснована на обмеженні й упорядкуванні. Із сукупностей випадкових величин створюють обмежені й упорядковані набори (комплекси) величин меншої чисельності.

Типізація іде за уніфікацією і пов'язана з уведенням конструктивних властивостей в уніфікований комплекс величин, що стають при цьому конструктивними характеристиками. Такими властивостями можуть бути системні властивості, наприклад, передатні відносини.

7.2. Методологія проектування і конструювання

Методологія в широкому смислові - це навчання про методи і способи їхнього застосування при вирішенні різних задач. *Метод* - раціональна основа способу дії.

Для існування методу необхідні:

- правила (принципи) поведінки як описання способу дії;
- усвідомлення використання методу як основи дії;
- дисципліноване підпорядкування правилам поведінки;
- описання ситуації, в якій доцільне застосування даного методу.

Таким чином, метод представляє собою систему відтвореного способу дії при усвідомленому використанні правил поведінки для найбільш імовірного досягнення наміченої мети в даних обставинах. У практиці проектування і конструювання використовують такі методи: евристичний, алгоритмічний, систематичний.

Евристичний метод базується на використанні інтуїції проектувальника, загальних правил і рекомендацій. Даний метод характеризується тим, що побудову задовільного проекту розглядають не як вибір з безлічі проектів, а як вдалий вибір з альтернативних вирішень на кожному кроці проектування. У рамках евристичного методу не ставиться запитання про оптимальний проект, а лише про досить гарний.

Алгоритмічний метод базується на послідовностях указівок що стосуються процедур (операцій), які дозволяють вирішити проектне чи конструкторське завдання. Такі вказівки називають алгоритмами. Наявність твердих алгоритмів дозволяє широко використовувати обчислювальну техніку і засоби прикладної математики. Проблема оптимізації формулюється як проблема вирішення екстремальних завдань. Застосування алгоритмічного методу обмежено завданнями конструювання.

Систематичний метод аналогічний методу послідовного поліпшення вирішень чи послідовного скорочення нев'язань у математичному програмуванні. Але відсутність безлічі всіх можливих проектів для здійснення вибору не дозволяє отримати оптимальне вирішення. Тому перехід від одного вирішення до іншого, кращого, заміняють перетворенням вихідного проекту. Таким чином, у початковому пункті систематичного методу є не область можливих проектів, а один чи кілька початкових проектів.

Оптимальність у випадку застосування систематичного методу носить відносний характер. Оптимальним вважають всякий допустимий проект, що є найкращим з усіх проектів, які є до моменту завершення проектування. При цьому, коли ми говоримо «найкращий», ми маємо на увазі не відповідність проекту деякому оптимуму, а відсутність недоліків, виявлених в інших проектах (із збереженням їхніх переваг).

Коли початковий проект виявляється недопустимим, його перетворюють, що робить його допустимим. Для цього здійснюють аналіз проектувальних процедур. У процесі аналізу виявляють помилки, що послужили причиною недопустимості початкового проекту. Потім формулюють методичні принципи, що дозволять надалі уникнути помилок і одержати допустимий проект. Ці принципи вносять у бібліотеку принципів. Отримані принципи використовують для поліпшення початкового проекту. Усі названі операції повторюють, доки не буде отриманий допустимий проект.

На кожному кроці поліпшення проекту реалізують кілька напрямків поліпшення. Можливий рух за одним з них веде у тупик. Тоді виникає необхідність повернення до одного з проміжних варіантів, і тому необхідно зберігати всі проміжні варіанти. Інший спосіб - одночасний рух за всіма можливими напрямками. Рух за кожному з напрямків вимагає контролю здійсненості проекту в технологічному і фінансовому планах, а також можливостей і способів ліквідації створеної системи після закінчення терміну її служби.

У структурі правил проектування і конструювання можна виділити дві групи принципів: *методологічні* і *методичні.*

Методологічні принципи спрямовані на підвищення ефективності дій проектувальника при досягненні поставленої мети. Ці принципи часто називають *праксеологічними. Праксеологія* - наука про ефективність дій.

До методологічних принципів відносять: принцип максимізації математичного очікування; принцип субоптимізації; принцип явищ з малою ймовірністю; принцип усунення слабких ланок; принцип максимізації довгострокової ефективності; принцип централізації; принцип здійсненності проекту.

Принцип максимізації математичного очікування стверджує, що при створенні систем необхідно досліджувати відношення «вартість / ефективність» і в процесі дослідження прагнути максимізувати ефективність системи (чи її математичне очікування) при фіксованій вартості чи мінімізувати вартість при фіксованому значенні ефективності.

Принцип субоптимізації означає, що незалежна оптимізація кожного з елементів системи в загальному випадку не приводить до оптимізації системи в цілому. Так, наприклад, при проектуванні автомобільних доріг розширення проїзної частини дороги без відповідного забезпечення видимості і розширення мостів може не поліпшити умови руху, а навпаки погіршити їх і привести до зростання дорожньо-транспортних пригод. Тому система повинна оптимізуватись за системними критеріями. У результаті за рахунок компромісів усередині системи її елементи можуть виявитись не оптимальними, але система в цілому буде оптимальною.

Принцип явищ з малою ймовірністю говорить, що основне завдання системи не повинно переглядатись, а основні характеристики системи не повинні значно змінюватись, для того щоб система виявилась придатною також у ситуаціях, що мають малу ймовірність. Наприклад, проектування снігозахисту автомобільної дороги повинно вестись не на максимально можливий обсяг снігоприносу до дороги з повторністю 1 раз у 300-400 років, а на оптимальний обсяг, що забезпечує мінімум сумарних витрат на снігоочищення проїзної частини дороги, на посадку пришляхових снігозахисних лісонасаджень і догляд за ними.

Принцип усунення слабких ланок стверджує, що резерв підвищення ефективності системи слід шукати в її слабкій ланці. Так, якщо на автомобільній дорозі є міст з вузькою проїзною частиною, тоді підвищення пропускної здатності дороги в цілому виявляється можливим після розширення моста.

Принцип максимізації довгострокової ефективності означає, що оптимізація системи має вестись за критерієм ефективності, розрахованому за весь термін її служби, а не для окремих моментів часу.

Принцип централізації вимагає централізації керівництва і прийняття вирішень. Іншими словами, він вимагає централізації інформації, не вимагаючи при цьому централізації фізичних пристроїв.

Принцип здійсненності проекту. Технологічні і фінансові можливості суспільства мають забезпечувати реалізацію вимог до системи.

Методичні принципи - це правила, що відображують приватні закономірності проектування. Методичні принципи містять вказівки на деяку процедуру проектування, на безліч моментів процедури, що дозволяють одержувати якісний проект.

До основних методичних принципів відносять: принцип автономності; принцип найменшої взаємодії; принцип взаємодоповнення; принцип забезпечення динамічної достатності; принцип ієрархії у задоволенні вимог; принцип узгодження норм і цілей функціонування елементів і системи в цілому; принцип динамічної рівноваги.

Принцип автономності стверджує, що компоненти системи мають функціонувати автономно Залежність функціонування одних компонентів системи від інших має бути мінімізована Цей принцип широко використовують у практиці організації дорожнього руху Наприклад, для виключення впливу тихохідних автомобілів на швидкохідні виділяють смуги руху під вантажні і легкові автомобілі, будують транспортні розв'язки у різних рівнях та ін.

Принцип найменшої взаємодії означає, що взаємодія між системою і середовищем має бути мінімізована. Це дозволить при найменшій дії одержати найбільший ефект.

Принцип взаємодоповнення (синергетизму) встановлює, що компоненти системи повинні взаємно компенсувати недоліки кожного з них. Між компонентами системи мають формуватись синергетичні відносини, що передбачають взаємодопомогу в процесі вирішення поставлених завдань.

Принцип забезпечення динамічної достатності. Умови функціонування системи та її компонентів повинні забезпечувати підтримку керованих перемінних функціонування у допустимих межах. У противному випадку система утратить властивість пристосовності до умов функціонування.

Принцип ієрархи у задоволенні вимог. Для отримання якісних проектів необхідно послідовне задоволення вимог у порядку їхньої важливості (ієрархи) Відносна важливість вимог має встановлюватись на основі змістовного аналізу функціонування системи.

Принцип узгодження норм і цілей функціонування елементів і системи в цілому. Відповідно до цього принципу характер пропонованих середовищем чи закладених у середовище шлей повинен збігатись з індивідуальними нормами поводження системи. Цілі функціонування системи і середовища мають бути несуперечливими.

Принцип динамічної рівноваги - забезпечення рухливої рівноваги між системою і середовищем, а також між окремими компонентами системи. Для людино-машинних систем цей принцип трансформується у принцип забезпечення рухливої рівноваги між внутрішніми (психічні процеси, стани і властивості) і зовнішніми (знаряддя праці) засобами діяльності людини.

Найважливішими *принципами конструювання* є принцип оптимального навантаження, принцип оптимального матеріалу, принцип оптимальної стабільності, принцип оптимальних співвідношень взаємозалежних величин.

Принцип оптимального навантаження стверджує, що навантаження тим менше відрізняється від оптимального, чим більше конструкція системи відповідає обраним критеріям. Іншими словами, оптимізація навантаження веде до оптимізації конструкції.

Оптимізація навантаження можлива при:

1) рівномірному розподілі навантажень чи напруг,
2) збільшенні числа способів передачі навантажень,
3) зменшенні можливостей появи ударних навантажень,

4) зменшенні енергетичних утрат

Принцип оптимального матеріалу. Витрати матеріалу тим менше відрізняються від оптимальних, чим більше конструкція системи відповідає оптимальному комплексу критеріїв.

Оптимізація матеріалу здійснюється виходячи з часткових критеріїв, що характеризують:

- доступність (можливість одержання) матеріалу;
- виробничі обмеження і пріоритети;
- механічні властивості, що обумовлюють можливість сприйняття навантажень;
- можливість обробки матеріалу;
- особливості відкритих поверхонь (тертя, прилипання, зчеплення);
- чутливість до впливу зовнішніх факторів (корозійна стійкість, жароміцність, розчинність та ін.);
- фізичні параметри (питома вага, питома теплоємність, теплове розширення, тепло- і електропровідність та ін.);
- магнітні властивості;
- радіоактивні властивості;
- формостійкість.

Принцип оптимальної стабільності. Стабільність системи тим менше відрізняється від оптимальної, чим більше конструкція відповідає критеріям надійності.

Надійність системи виражає статистичні відносини між критичними навантаженнями для компонентів системи.

Принцип оптимальних співвідношень взаємозалежних величин. Відносини між конструктивними характеристиками, а також іншими характеристиками системи тим менше відрізняються від оптимальних, чим більше конструкція відповідає прийнятим критеріям. Серед взаємозалежних величин особливу увагу заслуговують:

- конструктивні геометричні характеристики;
- стереомеханічні властивості;
- динамічні характеристики;
- кінематичні властивості;
- пружні властивості;
- маса.

7 з Обґрунтування необхідності створення системи.

Основне завдання проектування зводиться до дослідження необхідності створення нової системи та її соціального значення. Проблема необхідності пов'язана з потребами і процесом їхнього задоволення.

Потреба в загальному значенні може розглядатись як відображення в суспільній свідомості зсувів життєво важливих параметрів стану галузі народного господарства щодо оптимального рівня. Існування потреби викликано невідповідністю у чому-небудь. Це може бути як відсутність, так і надлишок. Звичайну невідповідність розуміють як недостатність (відсутність) предмета чи дії. Але невідповідність виникає не тільки тоді, коли чого-небудь немає, але також і тоді, коли що-небудь є в надмірній кількості І в першому і в другому випадках з'являються напружені ситуації, що утруднюють життя.

Потреби можуть бути актуальними і потенційними. До актуальних відносяться ті, які починають задовольнятись з моменту їхнього виявлення. До потенційних потреб відносяться такі, які починають вимальовуватись у світлі перспектив розвитку.

На рис 7 2 представлена модель техніко-економічних дій у процесі задоволення потреб. Процес створення системи починається з прогнозування потреб. Термін «прогнозування потреб» пов'язаний із завданням розпізнавання і вибору і представляє собою дуже значиму соціальну процедуру. У результаті прогнозування формується повний перелік потреб, що містить як дійсно соціально визнані, так і суб'єктивні всілякі потреби. Потреби упорядковують відповідно до їхньої важливості для життя і розвитку суспільства.

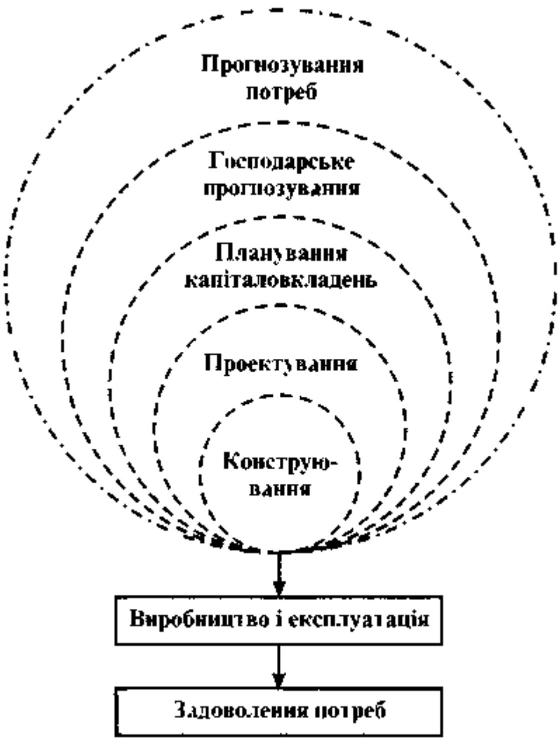

Рис 7 2 Модель області техніко-економічних дій у процесі задоволення потреб

Господарські прогнози охоплюють весь діапазон установлених потреб. На цій стадії за допомогою критеріїв відбирають потреби, що мають бути задоволені у першу чергу.

Планування капіталовкладень є практичним результатом господарського прогнозу. На даному етапі враховують конкретні обставини задоволення потреб.

Для кількісного і якісного аналізу потреб використовують статистичний підхід. Він складається у вивченні поведінки споживачів як статистичної сукупності. Поведінку споживачів оцінюють за допомогою моделей попиту і споживання.

У моделях попиту основними факторами є доходи споживачів ***Д*** и тарифи на послуги T_i. Статистичну модель попиту для фіксованого моменту часу представляють у вигляді двічі безперервно-диференційованої функції

$$C_i = \varphi_i(Д, T_i),$$

де C_i - попит на і-ту послугу.

При формуванні статистичних моделей попиту доход ***Д*** або тарифи на послуги T_i, вважають незмінними. Звичайно графік залежності попиту від тарифу будують як сімейство кривих для різних фіксованих значень доходу споживачів $Д_i$ (рис. 7.3).

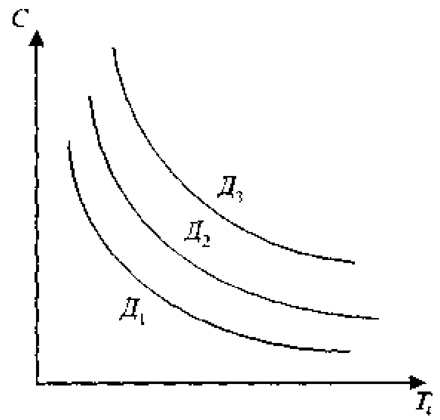

Рис. 7.3. Залежність попиту на транспортні послуги від тарифу і доходу підприємства

Кінематичну модель попиту будують у вигляді рівнянь множинної регресії за даними тимчасових рядів. Вона враховує зміну тарифу та інші тенденції попиту. Так, наприклад, одна з багатофакторних кінематичних моделей попиту має вигляд

$$C_i(t) = \alpha_{0i}[Д_i(t)]^{\alpha_1}[T_i(t)]^{\alpha_2} e^{l(t)},$$

де $Д_i(t)$ - середня витрата доходу на і-тий компонент попиту; $T_i(t)$ - тариф компонента попиту; l - показник тренда; α_1, α_2 - показники еластичності попиту по витраті доходу і тарифу відповідно.

Динамічна модель попиту враховує «інерцію» реакції споживачів на зміну факторів і її представляють у вигляді

$$C(\tau) = C_0(\tau - \theta),$$

де θ - лаг стану (θ = 1, 2, ...).

Моделі споживання звичайно представляють у вигляді двічі диференційованої функції

$$\Pi_{0i} = \varphi_i(Д, \psi),$$

де Π_{0i} - споживання і-тої послуги;

$$\psi = \frac{\Pi_{ід}}{\Pi_{ін}},$$

$\Pi_{ід}$ - норматив споживання і-тої послуги; $\Pi_{ід}$ - норма споживання і-тої послуги.

Як моделі споживання часто використовують степеневі функції вигляду

$$\Pi_{0i} = \alpha_i Д^{\beta_i} \quad \text{чи} \quad \Pi_{0i} = \alpha_i Д^{\beta_i} \psi^{z_i},$$

де α_i, β_i, z_i - коефіцієнти.

Для деяких видів послуг використовують багатофакторні моделі:

$$\Pi_{0i} = \alpha_i + \beta_i Д + z_i \psi.$$

Функції попиту і споживання є головним інструментом планування і прогнозування структури споживання. Але вони не дають повного представлення про поведінку споживача, для якого характерно комплексне відношення до всієї сукупності благ.

Побудова системи функцій попиту і споживання є першим кроком у вивченні поведінки споживача.

Другим кроком є визначення функції корисності споживання у вигляді

$$U_{no} = f(\Pi_{01}, \Pi_{02}, \ldots, \Pi_{0n}),$$

де U_{no} - функція корисності.

Для ув'язування першого і другого кроків приймають, що компоненти вектора попиту C_i цілком відображують і відповідні компоненти споживання Π_{0i}. Іншими словами, передбачають, що функції попиту і споживання ідентичні.

При побудові функції корисності припускають, що зростання інтенсивності споживання будь-якої послуги (при незмінній інтенсивності інших) підвищує рівень задоволення потреб, але її зростання безперервно сповільнюється в міру насичення даної потреби.

Через те, що різні потреби неоднакові за їх важливістю і ступенем задоволення, то при побудові функції корисності вважають, що самі потреби співвідносяться одна з одною і суспільне споживання в даних умовах як би порівнює інтенсивності споживання різних послуг між собою. Це формує поняття взаємозамінності послуг, що дозволяє зіставляти їх за суспільною корисністю. Таке зіставлення і приймають як початкову передумову в моделі, що описує комплексну поведінку споживачів. Воно визначає і спосіб оцінки показника корисності.

Для кількісної характеристики корисності часто використовують бальну шкалу, що формується на основі ранжирування послуг за корисністю споживачам. Як модель поведінки споживача приймають таку функцію корисності, що кожному сполученню інтенсивностей споживання послуг ставить у відповідність їхньої корисності в бальній шкалі.

Функцію корисності вибирають такою, щоб її перші і другі часткові похідні були безперервні. Часто при виборі з двох послуг у якості функції корисності використовують залежність вигляду

$$U_{по} = \frac{1}{2}\left(\alpha_{11}\Pi_{01}^{2} + 2\alpha_{12}\Pi_{01}\Pi_{02} + \alpha_{22}\Pi_{02}^{2}\right),$$

де $\alpha_{11}, \alpha_{12}, \alpha_{22}$ - коефіцієнти регресії.

Функцію корисності графічно представляють у вигляді деякої поверхні. Переріз цієї поверхні горизонтальною площиною на висоті $U_{под}$ відображається на площині $\Pi_{01}\Pi_{02}$ кривою $U_{по} = U_{под}\Pi_{02}^{2}$, яку називають ***кривою байдужості*** (рис. 7 4).

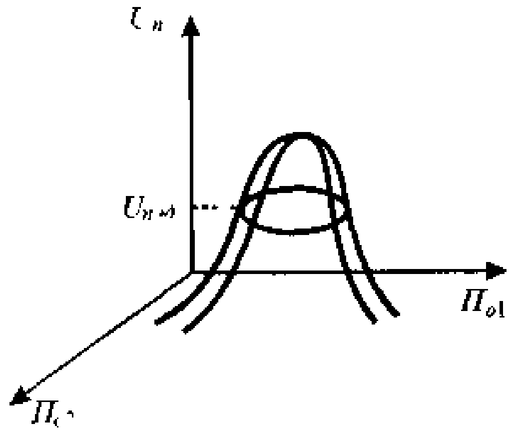

Рис 7 4 **Функція корисності**

Для визначення оптимального споживання Π_{0i} досить знайти значення $\tilde{\Pi}_{0i}$, що забезпечує максимум функції корисності $U_{по}$. Через те, що під оптимальним розуміють нормальне (норма) споживання, то, максимізуючи функцію корисності, знаходять норму споживання.

Для визначення нормативу споживання $П_{іД}$ знаходять $П_{0i}$, що максимізує функцію корисності при обмеженнях у вигляді балансового співвідношення

$$\sum_{i=1}^{m} T_i П_{0i} = Д.$$

Умова наявності екстремуму функції корисності при даному обмеженні представляється у вигляді

$$\frac{dU_{no}}{dП_{01}} = \alpha_{11}П_{01} + \alpha_{12}П_{02} = \lambda T_1,$$
$$\frac{dU_{no}}{dП_{02}} = \alpha_{12}П_{01} + \alpha_{22}П_{02} = \lambda T_2,$$

з балансовою умовою, що визначає розподіл доходу

$$T_1 П_{01} + T_2 П_{02} = Д,$$

де λ - невизначений коефіцієнт Лагранжа.

Вирішення отриманої системи рівнянь дає

$$П_{1Д} = \frac{Д}{J}\left(\alpha_{22}T_1 - \alpha_{12}T_2\right);$$
$$П_{2Д} = \frac{Д}{J}\left(\alpha_{11}T_2 - \alpha_{12}T_1\right);$$
$$\lambda = \frac{Д}{J}\left(\alpha_{11}\alpha_{22} - \alpha_{12}^2\right);$$
$$J = \alpha_{22}T_1^2 - 2\alpha_{12}T_1T_2 + \alpha_{11}T_2^2.$$

8 ЕКСПЛУАТАЦІЯ ТЕХНІЧНИХ СИСТЕМ

Експлуатація технічних систем.
Відмовлення технічних систем.
Принципові схеми виникнення відмовлень.
Експлуатаційні характеристики технічних систем.
Надійність систем.
Ресурс систем.
Експлуатаційна надійність системи.
Коефіцієнт готовності системи.
Симптом відмовлення системи.
Методи побудови програм відшукування відмовлень.
Стратегія.

8.1. Експлуатаційні характеристики систем

Експлуатація технічної системи — це процес її використання за призначенням і підтримки в технічно справному стані. Використання системи за призначенням включає організацію п функціонування. Підтримка системи в технічно справному стані включає технічне обслуговування (утримання), відновлення працездатності (ремонт), збереження, підготовку до роботи та ін. Забезпечення справного стану неможливе без чіткої організації, тобто без перспективного планування і управління.

Для того, щоб управляти процесом експлуатації, необхідно передбачати можливі стани системи **в** майбутньому. Передбачення забезпечується науковим прогнозуванням.

Прогнозування включає передбачення розрахункових характеристик (вихідних даних) і поточних станів системи.

Динаміка розрахункових характеристик залежить від структурних і функціональних змін системи, п якісної еволюції.

Поточний стан системи визначається сукупністю значень п технічних характеристик У загальному випадку можна вважати, що в процесі експлуатації технічні характеристики системи змінюються безперервно, тому і станів системи може бути нескінченна безліч. Але для організації експлуатацц важливо розрізняти стани, які відповідають крайнім чи допустимим (граничним) значенням технічних характеристик. Такі крайні чи допустимі значення технічних характеристик відповідають робочому стану, відмовленню, стану технічного обслуговування, збереження, відновлення та ш Тут під ***відмовленням*** розуміють подію, після якої система перестала виконувати (цілком чи частково) свої функції.

Розрізняють такі **принципові схеми виникнення відмовлень** миттєвих ушкоджень; змін, що накопичуються, релаксації, дій декількох незалежних причин

Схема миттєвих ушкоджень поєднує випадки, коли відмовлення системи чи її елемента викликане перевищенням навантаження на неї понад деякого допустимого значення. Такого роду «пікове» навантаження — явище випадкове, і відмовлення елемента настає незалежно від того, скільки часу до цього він знаходився в експлуатації, і яким був при цьому його стан.

Схема змін, що накопичуються, відноситься до випадків, коли відмовлення утворюються внаслідок поступового старіння чи зношення. Типовими прикладами таких відмовлень є випадки корозії залізобетонних опор, фундаментів, металевих конструкцій та ін.

Схема релаксації передбачає ті випадки, коли в результаті поступового накопичення ушкоджень відбувається стрибкоподібна зміна стану того чи іншого елемента або системи в цілому. Як приклад, можна навести відмовлення, що виникають під впливом циклічних навантажень. Тут поступове накопичення ушкоджень може бути лише непрямою причиною відмовлення. Накопичення ушкоджень приводить до зростання ймовірності відмовлення.

Схема дій декількох незалежних причин відповідає ситуації, коли одночасно діють кілька причин відмовлень.

Найважливішими експлуатаційними характеристиками системи є: надійність, працездатність, довговічність, збереженість, відновлюваність (ремонтованість), ремонтоздатність, термін служби, ресурс.

Під ***надійністю*** розуміють властивість системи виконувати задані функції, зберігаючи свої експлуатаційні показники в заданих межах протягом необхідного проміжку часу у визначених умовах експлуатації.

Працездатність — стан системи, при якому вона здатна виконувати задані функції з параметрами, встановленими вимогами технічної документації.

Властивість системи зберігати працездатність з необхідними перервами для технічного обслуговування і ремонтів до граничного стану, обговореного в технічній документації, називається ***довговічністю.***

Збереженістю називається властивість системи мати обумовлені експлуатаційні показники протягом і після термінів збереження і транспортування, встановленого в технічній документації.

Відновлюваністю (ремонтованістю) системи називається її здатність відновлювати працездатність у випадку виникнення відмовлення за допомогою **ремонту.**

Ремонтоздатність — властивість системи, шо полягає в пристосованості до попередження, виявлення й усунення відмовлень і несправностей у процесі технічного обслуговування і проведення ремонтів.

Ресурсом системи називається тривалість функціонування чи обсяг роботи, виконаний системою до граничного стану, обговореного в технічній документації.

Терміном служби називається календарна тривалість експлуатації системи до моменту виникнення граничного стану, обговореного в технічній документації.

Розходження між ресурсом і терміном служби обумовлене тим, що ресурс враховує фактичний наробіток системи, а термін служби — сумарну тривалість як роботи, так і простою з будь-яких причин.

Одним з основних завдань, розв'язуваних у процесі експлуатації технічних систем, є забезпечення їхньої надійної роботи. **В якості кількісних характеристик надійності використовують:** імовірність безвідмовної роботи, інтенсивність відмовлень, середній час безвідмовної роботи (наробіток на відмовлення), щільність розподілу часу безвідмовної роботи, характеристики і параметри потоку відмовлень та ін.

Інтенсивність відмовлень — це швидкість появи відмовлень неремонтованої системи за одиницю часу після даного моменту часу за умовою, що відмовлення до цього моменту не виникло. Інтенсивність відмовлень позначають латинською буквою λ.

Наробіток на відмовлення — середнє значення наробітку ремонтованої системи між відмовленнями.

Робота технічної системи в процесі експлуатації поділяється на три основних періоди: приробітки, нормальної експлуатації, зношення.

У період приробітки спостерігаються підвищені приробітні відмовлення, потім інтенсивність відмовлень суттєво знижується і тримається на якомусь мінімальному рівні, обумовленому ймовірністю раптових відмовлень. Мінімальний рівень інтенсивності відмовлень характерний для періоду нормальної експлуатації. Після закінчення цього періоду починає впливати фактор зношення, і з'являються так звані зносові відмовлення, інтенсивність яких з часом монотонно наростає.

Крива інтенсивності відмовлень у різні періоди експлуатації для технічної системи з великою кількістю однотипних елементів приведена на рис. 8.1.

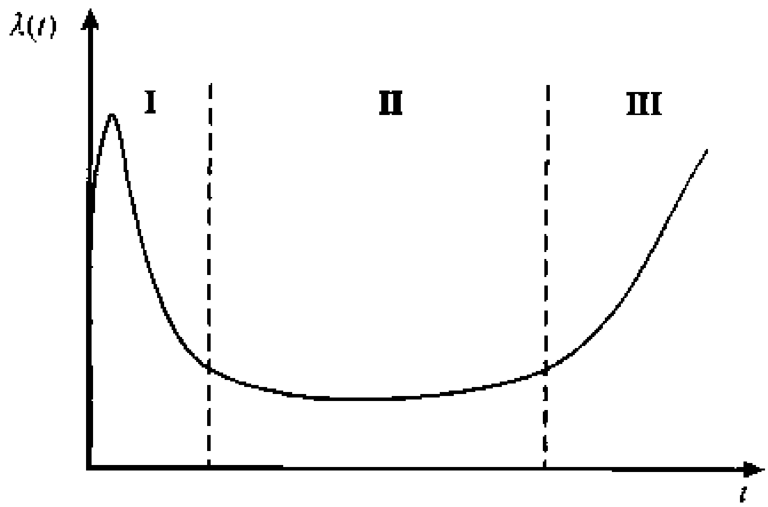

Рис. 8.1. Залежність інтенсивності відмовлень від часу:
I — період приробітки; II — період нормальної експлуатації;
III — період старіння.

У період приробітки на ділянці поступового зниження інтенсивності відмовлень для характеристики надійності може бути використаний розподіл Вейбула. При цьому розподілі щільність імовірності відмовлень $f(t)$ визначають за формулою

$$f(t)=\frac{a}{b}\left(\frac{t}{a}\right)^{b-1}\exp\left[-\left(\frac{t}{a}\right)^{b}\right].$$

де *a* і *b* - параметри розподілу, пов'язані між собою вираженням

$$\lambda=\frac{1}{a^{b}},$$

λ - інтенсивність відмовлень.

При b = 1 розподіл перетворюється в експонентний, а при b = 2 - у розподіл Релея. Таким чином, закон розподілу Вейбула є узагальненням експонентного і релеївського розподілів.

Для періоду прироботки системи у розподілі Вейбула параметр b < 1.

Імовірність безвідмовної роботи

$$p(t)=\exp\left[-\left(\frac{t}{a}\right)^{b}\right].$$

Інтенсивність відмовлень

$$\lambda(t)=\frac{b}{a}\left(\frac{t}{a}\right)^{b-1}.$$

У період нормальної експлуатації справедлива умова $\lambda(t)=\lambda=\text{const}$. Розподіл часу виникнення відмовлень характеризується експонентним законом, за яким щільність імовірності

$$f(t)=\lambda e^{-\lambda t}=\lambda\exp[-\lambda t].$$

У ряді випадків (при експонентному законі) замість Я користуються зворотною величиною *a* — середнім наробітком на відмовлення. Тоді формула для оцінки щільності ймовірності здобуває вигляд

$$f(t)=\frac{1}{a}e^{-\frac{t}{a}}=\frac{1}{a}\exp\left[-\frac{t}{a}\right].$$

Імовірність безвідмовної роботи в період нормальної експлуатації визначають за формулою

$$p(t) = e^{-\lambda t} = \exp[-\lambda t]$$

чи

$$p(t) = e^{-\frac{t}{a}} = \exp\left[-\frac{t}{a}\right].$$

У період старіння (зносових ушкоджень) системи розподіл часу виникнення відмовлень може бути приблизно охарактеризований нормальним законом. Але треба враховувати, що при нормальному розподілі випадкова величина може приймати і негативні значення, тоді як наробіток на відмовлення завжди величина позитивна. Тому можна використовувати або усічений нормальний закон, або просто нормальний закон за умовою, що коефіцієнт варіації менше 1/3, тобто $\frac{\sigma}{a} < \frac{1}{3}$, або $a = 3\sigma$, і тоді ймовірність отримання негативних значень настільки мала, що нею можна зневажити.

Використовуючи нормальний закон розподілу, знаходимо щільність імовірності

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(t-a)^2}{2\sigma^2}\right],$$

де σ - середнє квадратичне відхилення. Імовірність безвідмовної роботи

$$p(t) = \Phi_0\left(\frac{a-t}{\sigma}\right),$$

де Φ_0 - нормальна функція розподілу

$$\Phi_0(t) = \frac{1}{\sqrt{2\pi}} \int_{\infty}^{t} e^{-\frac{t^2}{2}} dt.$$

Звичайно ця функція табулюється.

Інтенсивність відмовлень

$$\lambda(t) = \frac{1}{\sigma} f_1\left(\frac{a-t}{\sigma}\right),$$

де f_1 - табульована функція,

$$f_1 = \frac{\varphi_0(t)}{\Phi_0(t)}.$$

Для нормованого і централізованого розподілу

$$\varphi_0(t)=\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}.$$

Розглянуті характеристики надійності справедливі для неремонтованих систем. На практиці після кожного відмовлення технічної системи, вона відновлюється або способом заміни несправного елемента, або способом його ремонту. Відновлення теоретично цілком повертає технічній системі ті властивості, які вона мала до відмовлення, так що її неможливо відрізнити від системи на початку експлуатації. Звичайно, при такому допущенні тривалість роботи системи з моменту її відновлення до чергового відмовлення не залежить від того, скільки разів вона відмовила у минулому. Послідовності подій при виникненні відмовлень у випадкові моменти часу $t_1, t_2, ..., t_n$ утворюють потік подій чи ***потік відмовлень.***

Цей потік можна розглядати як характеристику надійності відновлюваних технічних систем. Для кількісної оцінки надійності відновлюваних технічних систем приймають імовірність того, що випадкове число відмовлень $\nu(t)$ за розглянутий проміжок часу t менше заданого чи допустимого числа k. За показник надійності відновлюваної системи часто приймають імовірність появи за час t рівно k відмовлень. Таку ймовірність визначають за формулою

$$q(k, t) = q_k(t) - q_{k+1}(t),$$

де $q(k,t)$ - імовірність появи за час t рівно k відмовлень; $q_k(t)$ - імовірність того, що за час t виникне k або більше число відмовлень $\{\nu(t)\geq k\}$; $q_{k+1}(t)$ - імовірність того, що за час t виникне $\{\nu(t)\geq k+1\}$ відмовлень.

Більш універсальною характеристикою відновлюваних систем є середнє число відмовлень, що відбуваються до моменту t. Таке число називають ***функцією відновлення H(t):***

$$H(t)=\sum_{k=1}^{\infty} q_k(t).$$

У свою чергу для оцінки щільності відновлення h(t) використовують формулу

$$h(t)=\sum_{k=1}^{\infty} f_k(t),$$

де

$$f_k(t)=\frac{dq_k(t)}{dt}.$$

При взаємному накладанні *n* довільних потоків подай їхні функції - $H_i(t)$, а отже, і $h_i(t)$ складаються алгебраїчно

$$H_\Sigma(t) = \sum_{i=1}^{n} H_i(t);$$

$$h_\Sigma(t) = \sum_{i=1}^{n} h_i(t)$$

Для найпростішого потоку подій імовірність появи рівно k відмовлень на відрізку часу тривалістю t не залежить від положення цього відрізка на осі часу і виражається формулою Пуассона

$$q(k,t) = \frac{[H(t)]^k}{k!} e^{-\lambda t} = \frac{(\lambda t)^k}{k!} e^{-\lambda t},$$

де $H(t) = \lambda t$ і $h(t) = h = \lambda$.

Для підтримки надійності на належному рівні потрібно здійснення ряду заходів Мета цих заходів — попередити випадки появи відмовлень систем у процесі виконання ними робочих функцій До складу заходів включають перевірку через визначені інтервали часу стану системи, заміну деяких елементів, регулювання параметрів і усунення виявлених несправностей, тобто будь-яких ушкоджень чи відхилень від норм за допустимі межі Усі ці заходи поєднуються під загальною назвою *«обслуговування системи»*

Експлуатаційна надійність системи, що обслуговується, може бути визначена за формулою

$$P_s = P_k P_m P_n,$$

де P_k - імовірність безвідмовної роботи системи при раптових відмовленнях, які поки що не вдається прогнозувати, P_m - імовірність безвідмовної роботи системи при поступових змінах параметрів вихід яких за допустимі межі не завжди вдається попередити за допомогою існуючих методів обслуговування і прогнозування відмовлень; P_n - імовірність безвідмовної роботи системи, обумовлена наявністю елементів відмовлення яких є раптовими і можуть бути попереджені поліпшенням якості обслуговування.

Використання даної формули виходить із припущення незалежності відмовлень. Але, якщо відмовлення елементів системи викликані зміною умов експлуатації, тоді цю формулу застосовувати не можна.

Співмножник P_k може бути визначений за формулою

$$P_l = \int_{x}^{x} \int_{x}^{x} (q_{разн}) \int f(x_1, x_2, \ldots, x_q) \left[\prod_{j=1}^{k} p_j (x_1, x_2, \ldots, x_q) \right] dx_1 dx_2 \ldots dx_q,$$

де $f(x_1, x_2, \ldots, x_q)$ - багатомірна щільність імовірності появи величин навантажень $(x_1, x_2, \ldots, x_q)$; q - параметр навантажень; $p_j(x_1, x_2, \ldots, x_q)$ - імовірність того що раптове відмовлення елемента j не відбудеться при навколишніх умовах $(x_1, x_2, \ldots, x_q)$. Така імовірність залежить від конструктивних характеристик елемента.

Співмножник P_m знаходять за формулою

$$P_m = \int \int (m_{ра\,в}) \int_{y_m}^{y'} f(y_1, y_2, \ldots, y_m, t, \xi) dy_1 dy_2 \ldots dy_m,$$

де y_1, y'_1; y_2, y'_2; y_m, y'_m - відповідно нижні і верхні границі зміни кожного з *m* робочих параметрів системи; $f(y_1, y_2, \ldots, y_m; t; \xi)$ - щільність імовірності величин *m* робочих параметрів системи; ξ - зовнішні умови; $(y_1, y_2, \ldots, y_m)$- вихідні характеристики системи.

Співмножник P_n розраховують так:

$$P_n = \prod_{j=1}^{n} p_j,$$

де p_j - імовірність безвідмовної роботи j-того елемента, пов'язаного з якістю обслуговування.

Через те, що імовірність відмовлення j-того елемента $q_j = I - p_j$ і системи мають імовірність безвідмовної роботи практично прийнятну для їхньої експлуатації, тоді $q_j << 1$

$$P = \exp\left[-\sum_{j=1}^{n} q_j \right]$$

Вважаючи, що середнє значення імовірності відмовлення

$$\overline{q}_j = \frac{1}{n} \sum_{j=1}^{n} q_j,$$

одержимо

$$P_r = \exp\left[-n\overline{q}_j \right].$$

В міру удосконалення обслуговування, значення імовірностей безвідмовної роботи P_m і P_n у процесі експлуатації в принципі можуть бути доведені до одиниці. Таким чином, імовірність безвідмовної роботи системи в процесі експлуатації не може бути більше P_k.

Рівність $P_s = P_k$ може бути досягнута лише теоретично при ідеальній системі обслуговування.

При визначенні стратегії обслуговування мають бути добре відомі стратегії надійності елементів системи і вигляд розподілу часу їхньої безвідмовної роботи. Крім того, суттєвою є наявність резервування. Під *резервуванням* розуміють створення в конструкції системи тим чи іншим способом надмірності (резерву), що дозволяє з менш надійних елементів сформувати надійну систему Надмірність може бути структурною чи функціональною.

Структурне резервування передбачає створення надлишкових елементів у структурі системи. Для цього використовують паралельне з'єднання декількох однакових елементів у системі. Функціональне резервування допускає використання здатності елементів виконувати додаткові функції Наприклад, в автомобілі основне призначення зупиночного гальма — утримання автомобіля нерухомим. Але, крім того, він може використовуватись і при русі. При збільшенні періоду часу між оглядами системи зростає ймовірність виходу п елементів з ладу. Якщо стало помітним збільшення інтенсивності відмовлень, тоді час огляду чи заміни окремих елементів має бути в інтервалі

$$M - 3\sigma - M - 4\sigma,$$

де M — середня довговічність елементів, враховуючи зношення; σ — дисперсія часу між відмовленнями.

Періодичність і трудомісткість позапланового обслуговування є функцією величини, зворотній середньому наробітку на відмовлення системи a. Тут мають на увазі, що ймовірність безвідмовної роботи підкоряється експонентному закону, тобто розглядають нормальний період експлуатації.

Для якогось часу наробітку системи, рівного t, середнє число відмовлень складає $\frac{t}{a}$ і відповідно має місце деяке середнє число $\frac{t}{a}$ позапланових обслуговувань.

Для кожного виду відмовлень відомий час, необхідний для відновлення того чи іншого елемента системи T_1, T_2, T_3 та ін.

Загальне число позапланових обслуговувань протягом часу наробітку системи / в загальному випадку

$$\frac{t}{a} = \frac{t}{a_1} + \frac{t}{a_2} + \frac{t}{a_3} + \ldots + \frac{t}{a_n} = \sum_{i=1}^{n} \frac{t}{a_i} = \sum_{i=1}^{n} \lambda_i t;$$

а відповідний загальний середній час T_{n0} позапланового обслуговування для заданого наробітку розглянутої системи складе

$$T_n = \frac{t}{a_1}T_1 + \frac{t}{a_2}T_2 + \frac{t}{a_3}T_3 + \quad + \frac{t}{a_n}T_n = t\sum_{i=1}^{n}\frac{T_i}{a_i} = t\sum_{i=1}^{n}(\lambda_i T_i).$$

Розрахунок ведуть припускаючи, що трудомісткість позапланових обслуговувань визначається стосовно до експлуатаційного персоналу, який має достатній досвід. Слід врахувати, що загальний час позапланових обслуговувань містить у собі, крім самих ремонтних робіт, також виявлення ушкоджень і наступні перевірки системи.

Якщо позначити $T_{проф}$ - час, затрачуваний на профілактичне обслуговування системи, тоді повний час обслуговування $T_{обсл}$

$$T_{обсл} = T_{проф} + T_{п0}$$

З приведених формул ясно, що чим більше наробіток системи на відмовлення а, тобто вище її надійність, тим менше величина $T_{п0}$ і відповідно $T_{обсл}$.

Резерв системи може бути навантаженим чи не навантаженим. Якщо є два елементи системи, включені паралельно з однаковою інтенсивністю відмовлень λ, тоді середній наробіток на відмовлення такої системи складе - $\frac{3}{2\lambda}$, тобто в 1,5 рази вище, ніж при відсутності навантаженого резерву. Передбачають при цьому, що при відмовленні одного з двох елементів заміну його не виконують, тобто система працює доти, поки не відмовить і другий елемент. Якщо елемент, який відмовив, негайно заміняють, у той час як інший елемент знаходиться безперервно у роботі, тоді при цьому можуть бути створені системи із середнім наробітком на відмовлення, близького до нескінченності.

Дуже важливо досягти правильного співвідношення між часом обслуговування і наробітком системи. Для цього необхідно знати середній час обслуговування, що залежить від відмовлень системи під час робіт і від відмовлень елементів у резервованих системах. Воно може бути обчислене за надійностями елементів, що входять у систему, а середня тривалість простою — за планами обслуговування.

Середній час обслуговування на наробіток системи

$$\bar{T}_{обсл} = T_{обсл}\frac{a}{t} = \frac{aT_{обсл}}{8760 - T_{обсл}}, \text{ годин,}$$

де *a* — середній наробіток на відмовлення для заданої практики обслуговування

Відповідно може бути обчислений і *коефіцієнт готовності* системи

$$A = -\frac{a}{a + T_{обсл}}$$

що представляє собою відношення відпрацьованих системою годин наробітку й обслуговування

Можна дати й інше визначення коефіцієнта готовності системи, розглядаючи його як імовірність того, що система із середнім наробітком на відмовлення, рівного *a*, потребує час на обслуговування $\overline{T}_{обсл}$, для кожного наробітку *a,* буде готова до експлуатації в будь-який заданий момент часу в майбутньому.

Ефективність резервування (введення надмірності) характеризується відношенням

$$\rho = \frac{\mu}{S},$$

де μ - виграш надійності за ймовірністю появи відмовлення, *S* - збільшення вартості пристрою.

Виграш надійності за ймовірністю появи відмовлення

$$\mu = \frac{1 - P}{1 - P_{изб}},$$

де $P, P_{изб}$ - імовірності безпомилкової роботи ненадлишкового і надлишкового елемента чи системи.

Збільшення вартості елемента чи системи

$$S = \frac{C_{изб}}{C},$$

де $C, C_{изб}$ - вартості ненадлишкового і надлишкового елемента чи системи.

8. 2. Діагностика станів системи

Діагностичний процес - це процес обробки вихідної інформації для отримання висновку про стан досліджуваної системи.

Як вихідну інформацію при вирішенні діагностичного завдання використовують симптоми відмовлення и ознаки нормального функціонування системи

Симптомом відмовлення системи є інформація про відхилення від норм параметрів що характеризують п працездатність чи стан, а також про зміну цих відхилень у часі. Мова йде про ті параметри, що контролюються в процесі роботи системи, у процесі її технічного обслуговування та ін. Важливе значення мають зведення про те, як, скільки, в яких умовах працювала система і при яких умовах наступило відмовлення. Якщо попередніми спостереженнями можна встановити, що параметри, які характеризують працездатність системи, наближались до границі допуску поступово, тоді це свідчить про можливість появи відмовлення, викликаного зношенням і старінням. Це змушує звертати увагу в першу чергу на елементи, найбільш піддані таким змінам. Симптом, що враховує поступову зміну параметра за часом, називають ***інтегральним.***

Раптова поява відмовлення може свідчити про недотримання правил експлуатації системи чи про помилки, допущені при її виготовленні. Тоді при діагностиці стану звертають увагу насамперед на «слабкі» місця в конструкції і на можливість помилок, що допускає обслуговуючий персонал. Додатковий симптом, що враховує можливість відмовлення, називають ***диференціальним.***

Часто для визначення стану системи недостатньо інформації, яка міститься в симптомах відмовлення й ознаках нормального функціонування. Тоді для отримання додаткової інформації звертаються до різних дослідів. Програми дослідів можуть бути ***жорсткими*** **і** ***гнучкими***. При жорсткій програмі відшукування послідовність перевірок визначена заздалегідь і в процесі відшукування елемента, що відмовив, не змінюється. При гнучкій програмі характер (зміст) чергової перевірки встановлюється в ході діагностичного процесу, тобто вирішення про проведення наступного досліду приймається після аналізу результати? попереднього досліду.

Найважливішим компонентом діагностичного процесу є розпізнавання інформації, отриманої в результаті дослідів. Для досвідченого фахівця інформація більш помітна, ніж для некваліфікованого. Недосвідчений фахівець може взагалі невірно витлумачити результати дослідів і прийняти неправильне рішення. Тому велике значення мають методи навчання фахівців, їхні індивідуальні здібності, досвід експлуатації техніки та ін.

Діагностичні процеси характеризуються різною ефективністю. Оцінку ефективності здійснюють за такими критеріями: тривалістю діагностичного процесу; загальним числом дослідів (перевірок), необхідних для відшукування елемента, що відмовив; вартістю реалізації діагностичного процесу та ін. У ході розробки програм діагностичних процесіів виникає завдання оптимізації за одним чи декількома критеріями. Програма, що відповідає оптимальному значенню одного чи декількох критеріїв, називається ***оптимальною.***

Для деяких систем першорядне значення має підтримка їх у стані, готовому до негайного застосування за призначенням. Тоді основним критерієм оптимізації є тривалість діагностичного процесу. Процес відшукування несправного елемента оптимальний, якщо його тривалість мінімальна. Оптимальний діагностичний процес припускає негайне використання отриманої інформації, тобто проведення дослідів за гнучкою програмою.

Для одержання мінімальної тривалості діагностичного процесу необхідно на першому етапі прагнути до збільшення швидкості отримання інформації. Під швидкістю отримання інформації розуміють відношення

$$V_j = \frac{I_j}{t_j} \cdot$$

де V_j - швидкість отримання інформації j-тому етапі дослідів; I_j - кількість інформації, отриманої на j-тому етапі дослідів; t_j - тривалість j-того етапу дослідів.

Таким чином, на першому етапі дослідів домагаються виконання умови

$$V_1 \to \max$$

V_1 - швидкість отримання інформації на першому етапі дослідів.

Враховуючи результати першого етапу, визначають другий етап пошуку. З усіх можливих варіантів дослідів вибирають той, який знову забезпечує максимальне значення швидкості отримання інформації. Аналогічно вибирають варіанти дослідів на наступних етапах. У результаті оптимальний діагностичний процес характеризується таким рядом:

$$V_{1\max},\ V_{2\max},\ \ldots,\ V_{n\max},$$

де *n* — число етапів до виявлення елемента, що відмовив.

Отриманий ряд максимальних швидкостей одержання інформації визначає принцип діагностування за цими швидкостями.

Використання принципу максимальної швидкості отримання інформації для побудови оптимального діагностичного процесу дозволяє по-різному будувати програми відшукування відмовлень. **Для побудови програми можуть використовуватись:** метод поелементних перевірок, метод групових перевірок, метод логічного аналізу симптомів відмовлення.

Метод поелементних перевірок передбачає перевірку елементів по одному у визначеній заданій послідовності. Кожна перевірка має два результати: або елемент справний, або ні. Якщо елемент, який перевіряють, виявився справним, тоді приступають до перевірки наступного елемента і так до виявлення несправного. Передбачають наявність у системі одного відмовлення.

Послідовність елементів, які перевіряють, відповідає принципу максимальної швидкості отримання інформації і залежить від коефіцієнта відмовлення і-того елемента β_i і середнього часу перевірки кожного *елемента* $t_{сер1} = (i = 1,2,\ldots,N)$. *Під коефіцієнтом відмовлення розуміють відношення середнього числа відмовлень і-того елемента за час t до середнього числа відмовлень системи в цілому за цей же проміжок часу.*

Якщо $t_{сер1} = t_{сер2} = \ldots = t_{серN}$, тоді максимум інформації буде отриманий у випадку перевірки елемента, для якого β_i = 0,5. Перевірка на першому етапі будь-якого іншого елемента дала б менше інформації.

На другому етапі має бути перевірений елемент, у якого коефіцієнт відмовлення має найбільше значення в порівнянні з іншими, ще неперевіреними елементами, тому що в цьому випадку інформації буде отримано більше, ніж при будь-якому іншому варіанті перевірки.

На третьому етапі перевіряється елемент із найбільшим коефіцієнтом відмовлення серед неперевірених ще елементів та ін.

Таким чином, якщо середній час перевірки будь-якого елемента той самий $\left(t_{сері} = const\right)$, тоді відповідно з принципом максимальної швидкості одержання інформації елементи слід перевіряти у послідовності

$$\beta_1 \geq \beta_2 \geq \beta_3 \geq \ldots \geq \beta_j \geq \ldots \beta_N .$$

Зважаючи на те, що середній час перевірки кожного елемента $t_{сері}$ різний, оптимальна послідовність перевірки елементів відповідно з принципом максимальної швидкості одержання інформації визначається рядом:

$$\frac{\beta_1}{t_{сер1}} \geq \frac{\beta_2}{t_{сер2}} \geq \ldots \geq \frac{\beta_j}{t_{серj}} \geq \ldots \geq \frac{\beta_N}{t_{серN}} .$$

Якщо коефіцієнти відмовлень елементів рівні між собою $(\beta_1 = \beta_2 = \ldots = \beta_j \ldots \beta_N)$, а середній час $t_{сері}$ перевірки елементів різний, тоді для отримання оптимального діагностичного процесу елементи необхідно перевіряти в такій послідовності:

$$t_{сер1} \leq t_{сер2} \leq \ldots \leq t_{серj} \leq \ldots \leq t_{серN} .$$

Якщо $t_{сер1} = t_{сер2} = \ldots = t_{серj} = \ldots = t_{серN}$, $\beta_1 = \beta_2 = \ldots = \beta_j \ldots \beta_N$, тоді послідовність перевірок не має значення, тобто перевірку можна починати з будь-якого елемента.

Метод групових перевірок передбачає одночасну перевірку деякої групи елементів, у якій може знаходитись елемент, що відмовив. Якщо перевірка дає позитивний результат, тобто з'ясовується, що несправний елемент знаходиться в групі, яку перевіряють, тоді цю групу знову розбивають на дві підгрупи і відшукування несправності ведуть серед елементів цих підгруп. При негативному результаті перевірці піддають групу елементів, що залишилася (неперевірена). Такий процес розподілу продовжують до виявлення елемента, що відмовив. Тому часто цей метод називають *методом половинного розподілу чи методом середньої точки.*

У цьому методі оптимальну програму відшукування елемента, що відмовив, розробляють заздалегідь, тобто ще до настання відмовлення системи. Покладають, що система складається з послідовно з'єднаних елементів і в системі можлива наявність тільки одного відмовлення. Крім того, допускають, що на початку відмовлення будь-якого елемента з однаковою ймовірністю може бути причиною відмовлення системи і що середній час перевірки будь-якої групи елементів однаковий. Тоді принцип максимальної швидкості отримання інформації вироджується в принцип максимуму інформації на кожному етапі діагностичного процесу.

Для отримання максимальної кількості інформації на першому етапі діагностичного процесу систему поділяють на дві такі підсистеми, коефіцієнти відмовлення яких були б однакові Якщо коефіцієнти відмовлення всіх елементів рівні між собою $\left(\beta_i = \frac{1}{N}\right)$ необхідно розділити систему на рівні підсистеми, коефіцієнт відмовлення кожної з який дорівнює 0,5 (якщо число елементів у системі парне) Для першої підсистеми коефіцієнт відмовлення B_1 дорівнює

$$B_1 = \sum_{1}^{\frac{N}{2}} \beta_i = \frac{1}{N}\frac{N}{2} = 0{,}5$$

Отже, коефіцієнт відмовлення другої підсистеми також дорівнює 0,5.

Перевіривши сигнал на виході першої підсистеми, можна установити, чи містить ця підсистема елемент, що відмовив, чи ні. Якщо елемент, що відмовив, знаходиться у першій підсистемі, тоді друга підсистема справна і коефіцієнти відмовлення п елементів дорівнюють нулю. Подальшій перевірці мають бути піддаш елементи першої підсистеми.

Після першої перевірки коефіцієнти відмовлення першої підсистеми нормують Величини коефіцієнтів відмовлення елементів першої підсистеми визначають із співвідношення

$$B_1^1 = \sum_{i=1}^{\frac{N}{2}} \beta_{i1}^1 = 1{,}0$$

звідси

$$\beta_{i1}^1 = \frac{2}{N}$$

де β_{i1}^1 - коефіцієнт відмовлення і-того елемента першої підсистеми після першого доспілу.

На другому етапі перевірки системи першу підсистему поділяють на дві рівні підгрупи і виконують перевірку стану кожної з них.

Такий розподіл елементів на групи виконують доти, поки не буде встановлено елемент, що відмовив.

Якщо коефіцієнти відмовлення β_i неоднакові і середній час перевірки стану будь-якої групи елементів різний, тоді вибір оптимального варіанта розбивки системи на підсистеми зводиться до наступного.

Для проведення першого досліду знаходять суму коефіцієнтів відмовлення різних варіантів розбивки елементів системи на групи

$$B_1 = b_1,\ B_2 = b_1 + b_2, \quad ,\ B_k = b_1 + b_2 + \quad + b_k, \quad ,$$

$$B_{N-1} = b_1 + b_2 + \quad + b_{N-1}$$

Для кожного варіанта підраховують кількість інформації, що може бути отримана. Так, якщо система розбита на дві групи, що включають такі елементи:

- перша група - елемент №1;

- друга група - елементи з другого до N-го;

- тоді кількість інформації, що може бути отримана у першій груш, дорівнює

$$I_1 = -\left[B_1 \log B_1 + (1 - B_1) \log (1 - B_1)\right].$$

У другій групі

$$I_2 = -\left[B_2 \log B_2 + (1 - B_2) \log (1 - B_2)\right]$$

Аналогічно визначають кількість інформації, що може бути отримана при всіх інших можливих розбивках елементів

Для кожного варіанта розбивки на першому етапі визначають швидкість отримання інформації

$$V_1 = \frac{I_1}{t_{сер1}},\ V_2 = \frac{I_2}{t_{сер2}} \quad ;\ V_k = \frac{I_k}{t_{серk}},\ .\ .\ ,\ V_{N-1} = \frac{I_{N-1}}{t_{серN-1}},$$

де $t_{сер1}$ - середній час перевірки стану першого елемента; $t_{сер2}$ - середній час перевірки стану першого і другого елементів та ін.

Усі результати розрахунків зводять у матрицю оцінок (табл.. 8. 1).

Таблиця 8.1 Матриця оцінок

Сума коефіцієнтів відмовлення у групі	B_1	B_2		B_k		B_{N-1}
Кількість отриманої інформації	I_1	I_2		I_k		I_{N-1}
Швидкість отримання інформації	V_1	V_2		V_k		V_{N-1}

Аналізуючи матрицю оцінок, вибирають для проведення першого досліду (перевірки) такий варіант розбивки системи на групи, що забезпечує найбільшу швидкість отримання інформації.

При переході до другого етапу пошуку, групи елементів, що утворилися, розглядають як самостійні Коефіцієнти відмовлення в кожній групі нормують.

Так, якщо на першому етапі розбивка проводилась за k-*тим* елементом (елемент із номером k був віднесений до першої групи), тоді при переході до другого етапу, нормувальні коефіцієнти відмовлення визначають за формулами

$$\beta_{i1}^{1} = \frac{\beta_i}{\sum\limits_{i=1}^{k} \beta_i} \quad \text{(група 1);}$$

$$\beta_{i2}^{1} = \frac{\beta_i}{\sum\limits_{i=k+1}^{N} \beta_i} \quad \text{(група 2).}$$

Потім для кожної з груп складають можливі варіанти розбивки елементів на підгрупи й обчислюють B_i^1, I_i^1, V_i^1, де i - номер розбивки ($1 \le i \le k$ для групи 1 і $(k+1) \le i \le N$ - для групи 2), для того, щоб знайти максимум V_i^1 і номери елементів розбивки кожної з груп на другому етапі пошук.

Який із двох варіантів дослідів доведеться проводити на другому етапі пошуку, залежить від того, в якій із груп виявиться несправний елемент. Це з'ясується тільки після проведення досліду на першому етапі. Можливих перевірок третього етапу буде чотири, четвертого - вісім, п'ятого - шістнадцять та ін.

Обов'язковою умовою застосування методу групових перевірок є наявність функціональних зв'язків між елементами Тому при слабких функціональних зв'язках чи їхній відсутності перевагу слід віддавати методу поелементних перевірок.

Метод логічного аналізу симптомів відмовлення складається у швидкому й однозначному визначенні елемента системи, що відмовив, на підставі наявних чи отриманих у результаті проведення додаткових дослідів симптомів відмовлення.

Взаємозв'язок між елементами і симптомами представляють у вигляді матриці (табл. 8. 2.), що має *N* рядків і *n* стовпців. Тут *N* - число елементів системи, *n* - число симптомів. У клітинці на перетинанні рядків і стовпців ставиться одиниця, якщо спостерігається даний симптом, або нуль, якщо симптом не спостерігається.

Таблиця 8.2. **Матриця симптомів відмовлення елементів системи**

Симптоми / Елементи	C_1	C_2	C_3	C_4	Кодові числа	
$Э_1$	1	1	1	1	1234	134
$Э_2$	0	1	1	1	234	34
$Э_3$	0	0	1	0	3	3
$Э_4$	0	0	0	1	4	4

Для проведення наступного аналізу складають кодове число, властиве кожному елементові. Цифри кодового числа складають з номерів симптомів, які спостерігаються при відмовленні елемента. Кодове число записують у праву частину таблиці.

Якщо є симптоми, які входять у декілька кодових чисел, і після їхнього виключення як і раніше виконується умова розрізнення відмовлень, тоді такі симптоми можна виключити з розглядання.

Відшукування елемента, що відмовив, у системі при використанні даного методу здійснюють за кодовим числом.

Відшукування елемента, що відмовив, за симптомами можна здійснювати за заздалегідь складеними схемами Для складання таких схем використовують апарат математичної логіки.

Використання пошукових схем пояснимо на прикладі.

Нехай система складається з N елементів. Позначимо через $A_1; A_2;...; A_N$ елементи системи, а через $a_1; a_2;...; a_n$ - симптоми відмовлень цих елементів. Відмовлення кожного елемента, наприклад A_1, супроводжує деяка група симптомів. Зв'язок між симптомами і відмовленням елемента називають *диз'юнктивним* (логічне додавання) і позначають так:

$$A_1 = a_1 \vee a_2 \vee a_3 \vee ... \vee a_7,$$

де $\vee$ - знак логічного додавання (це зв'язок типу «або»).

Такий запис виражає необхідні і достатні умови відмовлення елемента A_1 і читається так: у системі відмовив елемент A_1, якщо має місце a_7, або симптом a_2, або a_3 та ін.

Якщо при відмовленні елемента A_1 одночасно спостерігаються симптоми $a_1; a_2; a_3;...; a_7$, тоді такий зв'язок називають *кон'юнктивним* і позначають

$$A_1 = a_1 \wedge a_2 \wedge a_3 \wedge ... \wedge a_7 = a_1 \cdot a_2 \cdot a_3 \cdot ... \cdot a_7,$$

де $\wedge$ (чи «.») - знак логічного множення (зв'язок типу «і»).

Таке вираження читається так: у системі відмовив елемент A_1, якщо одночасно спостерігаються симптоми $a_1; a_2; a_3$ та ін., тобто всі разом.

Поява будь-яких інших сполучень симптомів не свідчить про відмовлення елемента A_1.

Пошукові елементи дозволяють установлювати логічні зв'язки, на підставі яких будують гіпотези про можливе розташування елемента, що відмовив. Вірогідність гіпотези встановлюють за допомогою функціональних зв'язків.

При відсутності даних про надійність елементів системи, програму діагностичного процесу будують, припускаючи елементи рівнонадійними. У цьому випадку для встановлення елемента, що відмовив, використовують методи **раціональної технічної діагностики**: метод контролю схеми; метод заміни елементів на заздалегідь справні; метод використання таблиць несправностей.

Метод контролю схеми зводиться до перевірки правильності збірки деталей, монтажу схеми, справності запобіжних пристроїв, правильності робочого положення (перемикачів, тумблерів та ін.).

Метод заміни елементів на заздалегідь справні зводиться до заміни передбачуваного несправного елемента, блоку, вузла на заздалегідь справний. Після заміни систему перевіряють на функціонування. Якщо ознаки нормального функціонування системи цілком відновлюються, тоді це свідчить про те, що замінений елемент був несправний.

Метод використання таблиць несправностей зводиться до відшукування елементів, що відмовили, за допомогою спеціальних таблиць, які пов'язують причину з наслідком. Іншими словами, такі таблиці пов'язують симптоми відмовлень з елементами.

8.3 Оптимальне управління експлуатаційними процесами

Збільшення показників надійності і готовності систем здійснюється за рахунок своєчасного проведення регулювальних робіт, замін елементів і перевірок працездатності системи.

Для планування і управління експлуатаційними процесами застосовують методи динамічного і лінійного програмування, статистичних дослідів (метод Монте-Карло), сіткового планування та ш. Для формального описування моделей експлуатації використовують математичний апарат теорії керованих випадкових процесів, теорії відновлення, мінімаксні методи, правила припинення спостережень та ін.

У рамках теорії випадкових процесів для описування моделей профілактики широко використовують марківські і напівмарківські процеси.

Випадковий процес є марківським, якщо усі імовірнісні характеристики його в майбутньому залежать тільки від того, в якому стані цей процес знаходиться в даний момент часу, і не залежать від того, яким чином цей процес протікав у минулому. Такий процес називають випадковим без післядії.

Якщо випадковий марківський процес має не безперервний, а дискретний характер переходів з одного стану в інший, тоді він називається марківським ланцюгом, чи процесом з дискретним часом.

При описуванні марківських ланцюгів основними є поняття ***стану і переходу від одного стану в інший.*** Система знаходиться в деякому стані, якщо вона цілком описується значеннями перемінних, які задають цей стан. Система переходить з одного стану в інший, якщо перемінні, які її описують, змінюються від значень, що задають один стан, до значень, що визначають інший.

Можливі стани системи можуть бути зображені за допомогою графів станів. Імовірності переходу системи з одного стану в інший представляють звичайно у вигляді матриць перехідних імовірностей.

У загальному вигляді для марківського ланцюга зі станами a_1, a_2 і a_3

перехідні ймовірності P_i можуть бути записані в матричній формі:

$$P = \begin{Vmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{Vmatrix}$$

чи в загальному вигляді: $P = |P|$.

Нехай задані деякі числові значення ймовірностей, зазначені в нижченаведеній матриці, що має вигляд:

$$P = \begin{array}{c} \\ a_1 \\ a_2 \\ a_3 \end{array} \begin{array}{c} \begin{array}{ccc} a_1 & a_2 & a_3 \end{array} \\ \begin{Vmatrix} 0 & 1 & 0 \\ 0 & 1/2 & 1/2 \\ 1/3 & 0 & 2/3 \end{Vmatrix} \end{array}.$$

Суми ймовірностей у кожному рядку матриці обов'язково дорівнюють одиниці, тому що елементи кожного i-того рядка представляють імовірності всіх можливостей даного процесу, що знаходиться у стані a_i. Нулі у відповідних рядках чи стовпцях матриці вказують на можливість відповідних переходів. Якщо нулі стоять на головній діагоналі матриці, це означає неможливість переходів зі стану a_1 в a_1, чи з a_3 в a_3, чи в загальному випадку з a_i в a_i.

Інтервал часу між сусідніми переходами називають кроком.

Крім матриць імовірностей переходів *P,* марківський ланцюг має бути визначений ще і матрицею розподілу часу переходів $\tau(i, j)$ з одного стану в інший, тобто деякою матрицею

$$F = \| P\{\tau(i, j)\} \|,$$

в якій кожному ненульовому елементу матриці *P* відповідає свій розподіл.

Що стосується напівмарківських процесів, то, зберігаючи основну марківську властивість — не мати наслідку, вони мають більш загальний характер, ніж звичайні марківські процеси. Зокрема, розподіл часу $\tau(i, j)$ переходів з одного стану в інший може бути довільним.

Крім того, допускають випадки, коли $P_{ij} = 0$, тобто система чи її елемент може повертатись у той же стан, причому такий перехід триває протягом часу $\tau(i, j)$ з функцією розподілу часу $F_{ij}(t)$.

Напівмарківський процес може бути представлений і однією матрицею *Q* замість двох *P* і *F*

$$Q = \| Q_{ij}(t) \|,$$

де $Q_{ij}(t)$ - імовірність події, що при вихідному стані E_i, процес ξt перейде за один крок у стан E_j, причому час перебування ξt у стані E_1 не перевищить величини t. Імовірність події

$$Q_{ij}(t) = P_{ij} F_{ij}(t).$$

Іноді може мати місце і вироджений розподіл $\tau(i, j)$, при якому випадкова величина $\tau(i, j)$ з імовірністю одиниці дорівнює деякій константі *T*.

Напівмарківські процеси є дуже зручним апаратом для формалізованого описування моделей профілактики.

Прикладами напівмарківських процесів можуть бути:

1) зміна станів за схемою: робота — ремонт — робота;

2) зміна станів при профілактиці за віком: робота — ремонт при відсутності відмовлення (ремонт після відмовлення) — робота;

3) зміна станів при однорідному процесі з безперервним часом;

4) зміна станів за віком з миттєвим відновленням, при якому можлива реалізація процесу, наприклад, такого вигляду

$$a_1 \to a_1 \to a_2 \to a_1.$$

Для описування динаміки імовірнісного процесу переходів зі стану в стан використовують діаграми інтенсивностей переходів і диференційно-різницеві рівняння А.Н. Колмогорова.

Як приклад розглянемо граф станів, до якого можна звести перебування автомобілів у змінний і міжзмінний час (рис. 8 2).

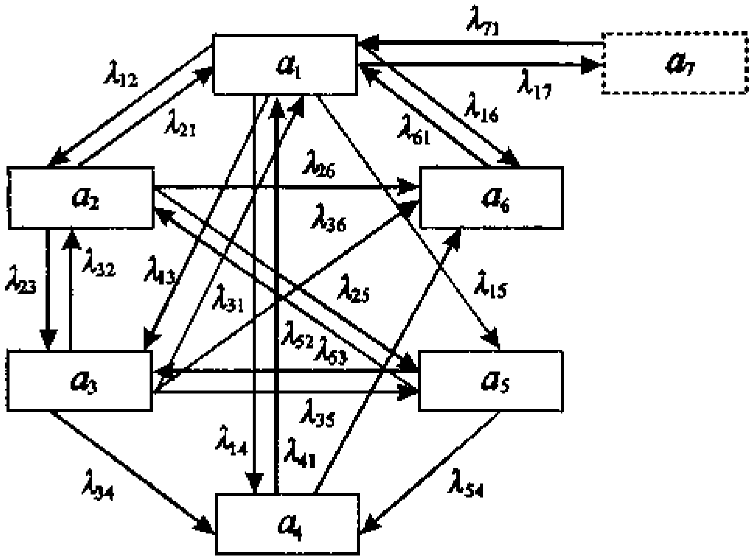

Рис 8 2 Узагальнений граф станів автомобіля в автотранспортному підприємстві.

Стани a_1 — робота на лінії, a_2 — ТО-2, a_3 — поточний ремонт, a_4 — капітальний ремонт. a_5 — простоювання з організаційно-технічних причин, a_6 — простоювання з організаційних причин, a_7 — ТО-1 у міжзмінний час

На рис. 8.2. $\lambda_{ij}(t)$ - інтенсивності переходів автомобіля з і-того стану в j-тий, де I, j = 1, ..., 7/

Якщо в заданий момент часу *t* автомобіль знаходиться в деякому стані *і*, тоді через проміжок *dt* він може виявитись в кожному із семи станів з імовірністю $\lambda_{ij}dt$, тобто залишитись в тому ж стані або перейти в кожний із шести.

Стан автомобіля в будь-який момент часу характеризується однією функцією — інтенсивністю потоку подій, яка може бути виражена будь-якою не негативною функцією часу Якщо автомобіль має сім станів, то вичерпною характеристикою його функціонування є квадратна матриця інтенсивності порядку 7х7. У цій матриці λ_{ij} = 0.

Матриця інтенсивностей дозволяє описати процес функціонування автомобіля диференціальними рівняннями А Н Колмогорова в тому загальному випадку, коли швидкість переходу з одного стану в інший велика Маємо систему рівнянь, в якій імовірностями, що відшукують, є імовірності перебування автомобіля в одному із семи станів $P_i(t)$:

$$\frac{dP_1(t)}{dt} = \lambda_{21}P_2(t) + \lambda_{31}P_3(t) + \lambda_{41}P_4(t) + \lambda_{61}P_6(t) + \lambda_{71}P_7(t) -$$

$$-(\lambda_{12} + \lambda_{13} + \lambda_{14} + \lambda_{15} + \lambda_{16} + \lambda_{17})P_1(t),$$

$$\frac{dP_2(t)}{dt} = \lambda_{12}P_1(t) + \lambda_{32}P_3(t) + \lambda_{52}P_5(t) - (\lambda_{21} + \lambda_{23} + \lambda_{25} + \lambda_{26})P_2(t);$$

$$\frac{dP_3(t)}{dt} = \lambda_{13}P_1(t) + \lambda_{23}P_2(t) + \lambda_{53}P_5(t) - (\lambda_{31} + \lambda_{32} + \lambda_{34} + \lambda_{35} + \lambda_{36})P_3(t);$$

$$\frac{dP_4(t)}{dt} = \lambda_{14}P_1(t) + \lambda_{34}P_3(t) - (\lambda_{41} + \lambda_{46})P_4(t),$$

$$\frac{dP_5(t)}{dt} = \lambda_{25}P_2(t) + \lambda_{35}P_3(t) - (\lambda_{52} + \lambda_{53} + \lambda_{54})P_5(t);$$

$$\frac{dP_6(t)}{dt} = \lambda_{16}P_1(t) + \lambda_{26}P_2(t) + \lambda_{36}P_3(t) + \lambda_{46}P_4(t) - \lambda_{61}P_6(t);$$

$$\frac{dP_7(t)}{dt} = \lambda_{17}P_1(t) + \lambda_{71}P_7(t).$$

Вирішуючи ці рівняння аналітично можна прогнозувати стани автомобілів в автотранспортному підприємстві за межами проміжку спостережень.

При експлуатації систем здійснюються заходи, що приводять до порушення природного ходу випадкового процесу. У результаті випадковий процес може перетворюватись в керований.

Розглянемо для простоти одномірний випадковий процес Θ_t, що приймає значення з кінцевої, рахункової чи незліченної множини *X*. Можливі випадки, коли параметр часу *t* є або безперервним $[t \in (0, \infty)]$, або дискретним $\left[t = t_i = i\Delta \left(i \in \overline{0, \infty} \right) \right]$. Задамо безліч $U = \{U_t\}$ прийнятих управлінь, чи вирішень. Покажемо зв'язок між процесом Θ_t і управліннями U_t для дискретного часу. Нехай при $t_0 = 0$ маємо значення процесу Θ_0 і приймаємо вирішення $U_{t_0}(\Theta_0) \in U$. При відомому управлінні $U_{t_0}(\Theta_0)$ і стані процесу Θ_0 процес Θ_t у момент $t = t_1 = \Delta$ переходить у стан Θ_1 з імовірністю

$$P\left\{\frac{\Theta_1}{\Theta_0}, U_{t_0}(\Theta_0)\right\},$$

якщо *X*— дискретна рахункова і кінцева множина.

Якщо *X* - незліченна множина, тоді перехід у стан Θ_1 характеризується щільністю

$$\rho\left(\frac{\Theta_1}{\Theta_0}, U_{t_0}(\Theta_0)\right).$$

У момент часу $t = t_1$ приймають вирішення

$$U_{t_1} = U_{t_1}(\Theta_0, \Theta_1),$$

і процес переходить при $t = t_1 = 2\Delta$ у стан Θ_2 відповідно з імовірностями

$$P\left\{\frac{\Theta_2}{\Theta_0}, \Theta_1, U_{t_1}(\Theta_0, \Theta_1)\right\} \text{ чи щільностями } \rho\left(\frac{\Theta_2}{\Theta_0}, \Theta_1, U_{t_1}(\Theta_0, \Theta_1)\right) \text{ та ін.}$$

Правило, відповідно з яким при ході процесу Θ_Z, що спостерігається, до моменту *t* з безлічі вирішень $U = \{U_t\}$ вибирається єдине вирішення U_t визначає стратегію δ.

Стратегією $\delta = \{U_t\}$ *називають набір управлінь* $\{U_t\}$, *прийнятих залежно від значень керованого процесу* Θ_Z *до моменту* $t(Z \le t)$.

У загальному випадку управління в момент t_i залежить від усіх попередніх значень процесу Θ_i

$$U_{t_i} = U_{t_i}(\Theta_0, \Theta_1, \ldots, \Theta_i).$$

Якщо обране управління в момент *t* залежить від часу *t* і значення процесу в цей момент $U_t = U_t(\Theta_t)$, тоді таке управління називають марковським, а стратегія, що складається з подібних управлінь - *марковською стратегією.*

Якщо, до того ж, обране управління безпосередньо не залежить від часу, тоді його називають *однорідним марковським управлінням.*

Стратегія δ, що складається з однорідних марковських управлінь, називається *однорідною марковською стратегією.*

Введемо далі критерії управління. Нехай $W(\Theta_{t_i}, U_{t_i})$ - утрати в момент t_i при перебуванні процесу в стані і при прийнятті в момент t_i управління U_{t_i}. Нехай для фіксованої стратегії $\delta = \{U_t\}$ і заданого початкового значення процесу Θ_0 середні утрати за час $T = N\Delta > 0$ такі:

$$V_T^{\delta}(\Theta_0) = M_{\Theta_0}^{\delta}\left\{\sum_{i=1}^{N} W\left(\Theta_{t_i}, U_{t_i}\right)\right\},$$

де N - число можливих станів системи.

Для оптимальної стратегії $(\tilde{\delta})$ маємо

$$V_T^{\delta}(\Theta_0) = \inf_{\delta \in U} V_T^{\delta}(\Theta_0),$$

де U — безліч усіх мислимих управлінь, за допомогою яких можна забезпечити $(\tilde{\delta})$.

Останніми можуть бути нерандомізовані чи рандомізовані однорідні марковські управління.

Якщо процес управління розглядають на досить великому інтервалі часу $(T \to \infty)$, тоді середні сумарні утрати можуть необмежено зростати:

$$\lim_{T \to \infty} V_T^{\delta}(\Theta_0) = \infty.$$

У цьому випадку середні утрати при нескінченно великому часі T представляють у вигляді

$$V_{\beta}^{\delta}(\Theta_0) = M_{\Theta_0}^{\delta}\left\{\sum_{i=0}^{N} e^{-i\beta\Delta} W\left(\Theta_{t_i}, U_{t_i}\right)\right\}, \; t_i = i\Delta,$$

де β - параметр $(0 \le \beta < \infty)$.

У цьому випадку для оптимальної стратегії $(\tilde{\delta})$, що мінімізує середні втрати при фіксованому β, маємо

$$V_{\beta}(\Theta_0) = V_{\beta}^{\tilde{\delta}}(\Theta_0) = \inf_{\delta} V_{\beta}^{\delta}(\Theta_0).$$

У деяких задачах виявляється більш зручним оптимізувати питомі витрати при тривалому управлінні (при $(T \to \infty)$):

$$\tilde{V}^{\delta} = \lim_{T \to \infty} \frac{1}{T} M^{\delta}_{\Theta_0}\left\{\sum_{i=0}^{T}\left(\Theta_i, U\right)\right\},\ t_i = i\Delta$$

Тоді для оптимальної стратегії $(\tilde{\delta})$ маємо

$$V(\Theta_0) = \tilde{V}^{\delta}(\Theta_0) = \inf \tilde{V}^{\delta}(\Theta_0),$$

Звичайно для обчислення оптимальних марковських однорідних стратегій застосовують метод динамічного і лінійного програмування. У випадку застосування методу динамічного програмування задача побудови оптимальної стратегії зводиться до вирішення рівняння Беллмана, до доказу одиничності отриманого вирішення і до побудови оптимальної стратегії за отриманими оптимальними утратами.

Рівняння Беллмана при фіксованому β представляється у вигляді

$$V(\Theta_0) = \inf_{u}\left[W(\Theta, U) + e^{-\beta\Delta} \sum V_{\beta}(\nu) P\left\{\frac{\nu}{\Theta}, U\right\}\right],$$

де $V_{\beta}(\Theta)$ - вечичина оптимальних утрат.

Усе вищевикладене дозволяє виробити оптимальну стратегію профілактичних ремонтів, що максимізує час безвідмовної роботи при найменших витратах