Міністерство освіти і науки України
Українська інженерно-педагогічна академія
Полтавський національний технічний університет
імені Юрія Кондратюка

Математичні методи та технічні засоби АСУ

Підручник для студентів вищих навчальних закладів

Під загальною редакцією
В. І. Барсова

Полтава – 2013

ББК 32.973.2
УДК 621.391
0-75

Затверджено
Міністерством освіти і науки України як підручник для студентів вищих навчальних закладів.
Лист № 1.4/18-Г-1339 від 11.06.2008 року.

Рецензенти:

В. М. Ілюшко, доктор техн. наук, професор, зав. кафедри виробництва радіоелектронних засобів літальних апаратів Національного аерокосмічного університету ім. М. Є. Жуковського ("ХАІ");

О. Д. Черенков, доктор техн. наук, професор, професор кафедри загальної електротехніки Харківського національного технічного університету сільського господарства ім. Петра Василенка.

0-75 В. І. Барсов, В. А. Краснобаєв, О. І. Тиртишніков, С. А. Орищенко, І.І. Слюсарь
ISBN 978-966-616-113-3

Математичні методи та технічні засоби АСУ / Під загальною редакцією В.І. Барсова: Підручник для студентів ВНЗ. – Харків, 2013. – 301 с.

Наведено загальні відомості про АСУ. Розглянуто основні положення теорії технічних систем, теорії масового обслуговування, теорії розкладів, теорії ухвалення рішень, теорії ігор, математичні методи розв’язання завдань оптимізації. Також розглянуто основні різновиди промислових АСУ, розкриті варіанти структурної організації АСУ, універсальні та спеціалізовані технічні засоби, питання надійності системи управління. Надано перелік контрольних питань.

Підручник призначений для студентів і аспірантів інженерних, інженерно - педагогічних спеціальностей вищих навчальних закладів України і інженерно-технічних працівників у галузі управління виробництвом, які займаються проектуванням сучасних АСУ широкого призначення.

© Барсов В. І., Краснобаєв В. А., Тиртишніков О. І., Орищенко С. А., Слюсарь І. І. 2013
ISBN 978-966-616-113-3

Зміст

Передмова

Безпосередньо завдання управління технологічними процесами (ТП) виникло одночасно із появою матеріального виробництва, тобто процесів цілеспрямованого перетворення матерії або енергії. У міру ускладнення виробничих процесів було потрібно більш розвинене і більш точне управління.

Тому стали застосовуватися різні контрольно - вимірювальні пристрої, які були встановлені безпосередньо на устаткуванні і працювали у прямому контакті із матеріальними потоками. Контрольно - вимірювальні пристрої давали можливість точніше і, головне, об'єктивно оцінювати роботу технологічного об'єкту і, отже, покращувати його використання.

Це дозволило організувати пости дистанційного контролю і управління, широко застосовувати автоматичні регулятори. У результаті були значного покращені умови роботи обслуговуючого персоналу: зменшилося фізичне навантаження, зручнішим стало робоче місце.

Із введенням уніфікованих сигналів, які можна було передавати на великі відстані, інформація про хід технологічного процесу, почала концентруватися на центральному пункті управління, де були встановлені відповідні прилади збору, відображення та обробки інформації. Метою процесу обробки інформації є ведення виробничого процесу у деякому оптимальному режимі, що дозволяє отримати максимально можливий позитивний ефект.

Важливо також відзначити, що автоматизовані системи управління промисловими установками часто належать до так званих великих систем. Такі системи характеризуються участю значного числа людей, різноманітних пристроїв і агрегатів, наявністю зв'язаних між собою достатньо складних підсистем, що володіють своїми особистісними цілями і критеріями, наявністю розвиненої ієрархії рівнів управління: агрегат – виробництво - підприємство.

Аналіз подібних промислових об'єктів і систем управління показує, що для них характерні наступна тенденція - інтенсивно зростає необхідна "потужність" систем контролю і управління. Це викликано підвищенням ступеня автоматизації виробництва, залучення все нових і нових агрегатів і ділянок у сферу дії централізованого управління. Оптимізація роботи окремого агрегату або окремої установки не гарантує максимального економічного ефекту для виробництва у цілому, оптимум для нього найчастіше досягається при деякому компромісі між особистісними критеріями оптимізації. У результаті цього зростає ступінь взаємозв'язків окремих агрегатів і ускладнюються алгоритми управління об'єктом у цілому; виникають завдання створення інтегрованих систем управління, що призводить до різкого ускладнення завдань управління.

У таких умовах і виникла необхідність автоматизації безпосередньо управління, тобто процесу прийняття рішень, яка вимагала залучення сучасних математичних методів і нових технічних засобів. У результаті з'явилися автоматизовані системи управління, тобто розвинені людино-машинні системи, що реалізовують автоматизований процес збору і переробки інформації необхідний для прийняття рішень по управлінню об'єктом (процесом, виробництвом) у цілому. При цьому роль людини у будь-якій АСУ вельми істотна оскільки ряд відповідальних завдань прийняття рішень через їх складність, багатогранність і невивченість не піддається формалізації. Їх виконання не може бути повністю автоматизоване і залишається за людиною. Очевидно, що застосування обчислювальної техніки може не тільки розвантажити людину від виконання рутинної нетворчої роботи, але і надати йому допомогу у виконанні творчих завдань (прийняття рішень з розподілу обмежених ресурсів, оптимізації технологічного процесу і т. п.).

На всіх етапах розвитку методів і засобів автоматизації технологічних процесів (від локальних пристроїв і систем автоматичного управління окремими агрегатами до сучасних ієрархічних систем із використанням комп'ютерно-інтегрованих технологій) найважливішими задачами, що підлягають вирішенню є:

- збільшення продуктивності технологічного устаткування;
- підвищення якості вироблюваної продукції;
- досягнення оптимальних режимів роботи технологічного устаткування;
- економія всіх видів ресурсів;
- зниження трудомісткості технологічних операцій;
- - створення умов праці, що зберігають фізичні і інтелектуальні сили людини при виконанні монотонних операцій;
- максимально можливе виключення умов, що шкідливо впливають на здоров'я людини.

Для ефективного вирішення указаних задач і використовуються автоматизовані системи управління технологічними процесами (АСУ ТП), у яких оптимальним чином поєднуються досягнення в області системного аналізу, сучасних математичних методів і новітніх досягнень в області інформаційних комп'ютерних технологій.

У зв'язку із означеним представляється необхідним створення підручника в якому були б викладені основи теорії побудови, математичні методи оптимізації та технічні засоби сучасних АСУ ТП.

Підручник містить 9 розділів.

У 1-му розділі наводяться загальні відомості про АСУ, розглянуті основні поняття і визначення автоматизованого управління, види управління, типи об'єктів управління, основні операції управління технологічним процесом, класифікація АСУ. Також розкриті методи автоматизованого управління.

У 2-му розділі розглянуті основні поняття і визначення теорії технічних систем, методи опису структур технічних систем, наведено класифікацію технічних систем, розглянуто прості та складні, статичні та динамічні технічні системи, методи композиції та управління технічними системами. Також розкриті основні властивості технічних систем.

У 3-му розділі сформульована загальна постановка задачі оптимізації та розглянуто методи вирішення задач оптимізації АСУ.

У 4-му розділі розглянуті загальні поняття, характеристики та визначення теорії масового обслуговування. Дано визначення найпростішому потоку подій та розглянуто його властивості. Розглянуто та дано визначення апарату марковських випадкових процесів. Розглянуто час обслуговування замовлень у системі. Досліджено властивості систем масового обслуговування із відмовами, чергою та змішаного типу.

У 5-му розділі розглянуто основні поняття і визначення теорії ухвалення рішень, загальну схему процесу ухвалення рішень. Надана класифікація задач ухвалення рішень. Також розкриті переваги особи, що ухвалює рішення.

У 6-му розділі розглянуто основні поняття і визначення теорії ігор, рішення гри у змішаних стратегіях, основна теорема теорії матричних ігор. Також розкрито принципи максиміна й мінімакса, формальний опис рішення гри, розглянута гра як модель конфліктної ситуації. Надано приклади застосування теорії ігор.

У 7-му розділі розглянуто основні поняття і визначення, загальну постановку задачі оптимізації із застосуванням теорії розкладів, комбінаторні задачі на складання розкладу, задачу комівояжера та приклади задач, що зводяться до задачі комівояжера. Також розглянуто методи рішення задачі оптимізації.

У 8-му розділі розглянуто основні різновиди промислових АСУ, склад і функції АСУ ТП, основні операції управління технологічним процесом. Також розкриті варіанти структурної організації та методи автоматизованого управління.

У 9-му розділі розглянуто універсальні та спеціалізовані засоби обробки, зберігання, візуалізації інформації, технічні засоби регулювання, що використовуються в АСУ ТП. Також розглянуто питання надійності системи управління.

Ще у підручнику наведено перелік контрольних запитань. Приведено докладне рішення типових задач, що повинно істотно полегшити виконання студентами контрольних завдань та розуміння теоретичних питань дисципліни у цілому.

При самостійній роботі над підручником бажано скласти короткий конспект проробленого матеріалу, після вивчення кожної теми необхідно дати відповіді на контрольні запитання.

Розділ1

Загальні положення

Розглянуті основні поняття та визначення автоматизованого управління, види управління, типи об'єктів управління, основні операції управління технологічним процесом, класифікація АСУ. Також розкриті методи автоматизованого управління. Наведено контрольні питання.

1.1. Основні поняття та визначення

Розглянемо основні поняття та визначення автоматизованого управління**.**

Автоматизована система управління (АСУ) - це людино - машинна система, що забезпечує автоматизований збір і обробку інформації, необхідної для оптимізації управління у різних сферах людської діяльності.

Автоматизація **-** комплекс заходів щодо створення, впровадження і застосування інформаційних, технічних, програмних і інших засобів для часткової або повної заміни інтелектуальних зусиль людини у різних областях її діяльності.

Автоматизація виробництва - це процес у розвиток машинного виробництва, при якому функції управління і контролю, що раніше виконувалися людиною, передаються приладам і автоматичним пристроям. При цьому мета автоматизації виробництва полягає у підвищенні ефективності праці, поліпшенні якості продукції, що випускається, у створенні умов для оптимального використання всіх ресурсів виробництва.

Система - це сукупність цілеспрямовано взаємодіючих елементів. Безліч елементів, що складають систему, завжди можна розбити за деякими ознаками на підмножини, тим самим, виділяючи із системи її складові частини ***– підсистеми*****.**

Об'єкт, що управляє, призначений для вироблення та здійснення управляючих дій – тобто дій системи управління на процес, який управляється, що змінює його хід у бажаному напрямі.

Відповідно, ту фізичну систему, в якій ми управляємо процесами, та стан якої контролюємо, за значеннями контрольованих величин, називатимемо ***об'єктом управління***.

Об'єктами управління можуть бути як окремі об'єкти, виділені за певними ознаками (наприклад, конструктивними, функціональними), так і сукупність об'єктів (комплекси, системи). Залежно від властивостей або призначення об'єктів управління можуть бути виділені технічні, технологічні, економічні, організаційні, соціальні та інші об'єкти управління, комплекси та системи.

Основні типи промислових об'єктів управління можна ***класифікувати*** наступним чином (рис. 1.1):

1) ***За характером протікання технологічних процесів*** об'єкти управління поділяються на:

- циклічні;
- неперервно-циклічні;
- неперервні.

2) ***За характером сталого значення вихідної величини*** об'єкту при дії на його вхід ступінчастого сигналу виділяють об'єкти:

- із самовирівнюванням;
- без самовирівнювання.

3) ***За кількістю вхідних і вихідних величин і їх взаємозв'язку*** об'єкти поділяється на:

- одновимірні (один вхід і один вихід):
- багатовимірні, які можуть бути:

а) багатозв'язковими – коли спостерігається взаємний вплив каналів регулювання один на одного;

б) незв'язковими – взаємозв'язок між каналами малий.

4) ***За виглядом статичних характеристик*** об'єкту управління встановлюють зв'язок між сталими значеннями входу і виходу об'єкту та за виглядом статичних характеристик діляться на:

- лінійні:
- нелінійні: статична характеристика може бути гладкою, яка лінеаризується у околі заданої точки, або носити істотно нелінійний характер, за наявності в об'єкті декількох нелінійностей, тоді його сумарна нелінійна характеристика визначається графічним методом.

5) ***За ознакою мобільності*** об'єкти можуть бути поділені таким чином:

- стаціонарні;
- пересувні;
- нестаціонарні.

Рис. 1.1. Основні типи об'єктів управління

6) ***Залежно від інтенсивності випадкових обурень***, що діють на об'єкт, вони поділяються на:

- стохастичні;
- детерміновані;

7) ***За фізичною природою*** об'єкти можна класифікувати як:

- технічні пристрої;
- біологічні системи;
- виробничі підприємства;
- соціальні системи;
- економічні системи.

Управління **-** вид діяльності, що є сукупністю дій, направлених на досягнення певної мети, тобто бажаного протікання процесу, що управляється.

У подальшому будемо вживати терміни керування і управління як синоніми, але при цьому термін керування будимо вживати в сенсі стосовно організаційної діяльності, яка спрямована на досягнення визначених цілей, а термін управління у сенсі стосовно управляючого впливу на технічні засоби, тобто деякої фізичної величини, зміна якої впливає на характер процесів в об'єкті управління.

Процес оптимізації припускає вибір такого варіанту управління, при якому досягається мінімальне або максимальне значення деякого критерію, що характеризує якість управління.

Види управління:

- ***ручне*** - оператор вручну змінює положення робочих органів технологічного агрегату;
- ***механізоване*** - зміна положення робочих органів з використанням механізмів, що полегшують працю оператора;
- ***автоматичне*** - управління технологічним процесом - автоматичне, оператор виконує функції контролю протікання ТП, запуску і зупинки технологічного устаткування;
- ***автоматизоване*** - управління технологічним процесом здійснюється в основному автоматично, але функція прийняття рішень виконується за участю оператора.

Система управління - система (технічна, біологічна, соціальна), що впливає на процес, що управляється так, щоб його протікання змінювалося в бажаному напрямі.

Дія, що управляє - дія системи управління на процес, що управляється, яка змінює його хід у бажаному напрямі.

Обурююча дія (обурення) - дія на процес, що управляється, яка змінює його хід у небажаному напрямі.

Регулювання - вид управління, метою якого є підтримка на необхідному значенні (значеннях) однієї або декількох величин, що характеризують стан об'єкту управління.

Зауваження. Більшість систем регулювання відноситься до класу систем автоматичної стабілізації режиму роботи об'єкту відносно його робочої точки (щодо номінального режиму роботи). В цьому випадку у процесі роботи відхилення змінних щодо робочої точки будуть малими, що дозволяє використовувати лінійні моделі об'єкту управління. Реальні об'єкти займають у просторі деякий об'єм, тому регульована величина залежить не тільки від часу, але і від поточних координат точки вимірювання. Тому повний опис об'єкту управління складатиметься із системи диференціальних рівнянь з особистісними похідними.

Критерій управління - співвідношення, що характеризує якість роботи технологічного об'єкту управління в цілому і що приймає числові значення залежно від використовуваних дій, що управляють.

Ефективність АСУ - здатність системи забезпечувати досягнення корисних технічних, технологічних, економічних, соціальних та інших результатів шляхом поліпшення якості управління об'єктом.

Комплекс технічних засобів АСУ - достатня для виконання всіх функцій АСУ сукупність: обчислювальних пристроїв, засобів перетворення, відображення і реєстрації сигналів, пристроїв передачі і обробки сигналів (даних), а також виконавчих пристроїв, що управляють.

Технологічний процес (ТП) представляє собою сукупність технологічних операцій, що виконуються у певній послідовності за допомогою технологічного устаткування, що перетворює вхідний матеріал (вхідний продукт) у готовий виріб (готовий продукт), який відповідає заданим вимогам (стандартам або іншим технічним документами).

Усі технологічні процеси умовно поділяються на три основні класи:

- ***неперервні***, в яких і вхідні і вихідні параметри являють собою функцію часу;
- ***дискретні*** – процеси, параметри та стани яких вимірюються у дискретні моменти часу;
- ***неперервно-дискретні*** – процеси, що являють собою комбінацію безперервних та дискретних.

Управляємий технологічний процес - це процес, для якого встановлені (визначені) вхідні ($\boldsymbol{x_i}$) і вихідні змінні ($\boldsymbol{y_j}$), а також залежності між ними, визначені методи вимірювання вхідних і вихідних змінних і методи управління процесом.

За характеристиками вхідних і вихідних змінних розрізняють наступні технологічні процеси:

- ***детерміновані*** – процеси, вихідні параметри яких однозначно визначаються множиною (комбінацією) вхідних змінних і не містять неконтрольованих збурень;
- ***стохастичні*** – процеси, які містять неконтрольоване джерело випадкових збурень, і тому їх вихідні параметри неоднозначно визначаються вхідними змінними.

Параметр ТП – величина, яка кількісно характеризує якусь небудь властивість або ознаку цього процесу.

Параметри ТП бувають ***якісні*** та ***кількісн*і**.

По відношенню до ТП параметри підрозділяються на ***вхідні, проміжні*** та ***вихідні.***

У залежності від мети здобуття інформації параметри ТП можна класифікувати за 5 категоріями: ***контроль, облік, вимірювання*, *ідентифікація*** і ***управління.***

За характером дискретизації параметри підрозділяються на ***неперервні*** (аналогові) і ***дискретизовані*** (дискретні).

Основні операції управління технологічним процесом.

Підготовчі операції:

- контроль готовності устаткування;
- контроль готовності операторів – диспетчерів;
- контроль наявності сировини та матеріалів;
- контроль наявності тари для готової продукції;
- контроль готовності інших елементів ТП.

Пускові операції:

- пуск (включення);
- виконання визначеної послідовності пускових операцій.

Збір даних:

- вимірювання (визначення кількісних значень величин);
- визначення положення, стану;
- опитування джерел інформації;
- передача даних.

Накопичення та огляд даних:

- фіксація, запам'ятовування даних;
- реєстрація;
- сигналізація;
- індикація;
- визначення відхилень від норми.

Аналіз ситуації:

- відокремлення головних на даний момент фактів та подій;
- виявлення напряму наступних дій у залежності від ситуації.

Підготовка та прийняття рішень:

- попередній вибір можливих рішень;
- прийняття остаточного рішення.

Реалізація рішень (регулювання та управління):

- регулювання позиційне (включення-виключення виконавчих механізмів, відкриття-закриття регулюючих органів);
- регулювання складне (рух виконавчих механізмів за заданими законами регулювання).

Операції з зупинки технологічного процесу:

- зупинка (виключення) устаткування;
- виконання визначеної послідовності операцій з зупинки устаткування.

У загальному випадку до АСУ ТП слід відносити усі системи, в яких функції управління технологічними процесами автоматизовані. Більш точним вважається таке визначення ***АСУ ТП*** - складна автоматизована ієрархічна система, що призначена для формування та реалізації управляючих впливів на технологічний процес згідно з прийнятими критеріями та обмеженнями.

АСУ ТП призначена для управління технологічним процесом, що реалізується на технологічному обладнанні за заданою технологією.

Сукупність обладнання та технологічного процесу являє собою ***технологічний об'єкт управління*** (ТОУ).

Сукупність ТОУ і взаємодіючої з ним АСУ ТП являє собою ***автоматизований технологічний комплекс*** (АТК). Складові АТК показані на рис. 1.2.

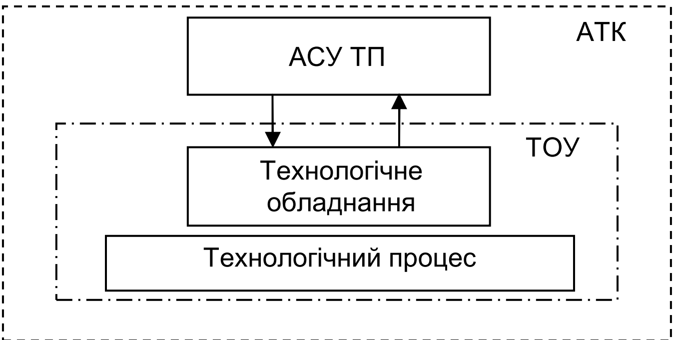


Рис. 1.2. Складові автоматизованого технологічного комплексу

1.2. Класифікація АСУ

Велика розмаїтість існуючих АСУ, розходження їхнього функціонального призначення й розв'язуваних завдань, характеристик і розміщення елементів приводять до необхідності класифікації АСУ. Класифікація дозволяє методологічно, більш послідовно підійти до вибору основних технічних засобів, в залежності від цільового призначення АСУ. Класифікація АСУ може бути досить різноманітною, будуватися за всілякими ознаками і тому є досить умовною.

Дуже важливою й досить аргументованою є класифікація АСУ за наступними ознаками:

- за реактивністю АСУ;
- за територіальною ознакою;
- за характером процесу, що управляється;
- за ступеню мобільності засобів автоматизації;
- за числом рівнів (ієрархічній структурі) управління;
- за ступеню автоматизації;
- за пропускною здатності;
- за характером протікання технологічного процесу;
- за серійністю.

Розглянемо класифікацію АСУ за перерахованими вище ознаками.

Класифікація АСУ за реактивністю.

Найбільш важливою характеристикою, що впливає на вибір технічних рішень при побудові АСУ, є ***реактивність системи***, або час, затрачений системою на доведення інформації про стан керованого процесу до пунктів управління, на обробку вхідної інформації, прийняття рішення на основі цієї інформації й доведення прийнятого рішення до виконавчих органів. Всі існуючі у цей час АСУ ТП за розглянутою ознакою можна розбити на три класи.

1) ***АСУ жорстко регламентованого режиму*** реального часу або АСУ із високою реактивністю на вхідний потік інформації. У таких АСУ реальний час, вимірюється з високою точністю і є найважливішим параметром, від якого залежать цінність вхідної інформації та її вплив на рішення, що приймається. Прикладом цього класу АСУ є автоматизовані системи управління повітряним рухом (АСУПР).

2) ***АСУ не жорстко регламентованого режиму*** реального часу або АСУ із середньою реактивністю на вхідний потік

інформації. Для цього класу АСУ час доведення інформації про стан процесу, що управляється становить десятки секунд, час на обробку вхідної інформації - одиниці секунд, час на прийняття рішень - десятки секунд (у окремих випадках хвилини), час доведення рішень до виконавців становить десятки секунд.

Як приклади АСУ другого класу можна привести АСУ енергосистемами, електромережами, теплоелектроцентралями, тепловими мережами, АСУ роботою нафтопроводів та ін.

3) ***АСУ* зі слабкою реактивністю** або із відсутністю реактивності на вхідний потік інформації. Для таких АСУ час на обробку вхідної інформації становить хвилини або години, а сам процес збору інформації досить тривалий і може досягати декількох діб.

У подібного роду АСУ найчастіше відсутній замкнутий автоматизований контур управління, АСУ обробляє й аналізує вхідну інформацію, видає відповідні довідкові дані, а рішення, як правило, приймає людина. Доведення рішення до виконавців також може досягати декількох діб. Такі системи без замкнутого автоматизованого контуру управління не є АСУ в загальноприйнятому розумінні й перетворюються в інформаційно-довідкові або інформаційно-розрахункові системи. До систем розглянутого класу належать АСУ підприємств, галузі й підгалузі промисловості, різні інформаційно-довідкові системи, АСУ банківськими операціями, АСУ із продажу квитків на різні види транспорту й ін.

Класифікація АСУ за територіальною ознакою.

1. ***Локальні АСУ***. До цього класу відносяться АСУ, елементи яких зосереджені на обмеженій території: у межах одного будинку, інженерно - будівельного спорудження, території підприємства або установи й т.п. Як правило, комплекси засобів автоматизації й інших технічних засобів цих АСУ зв'язані між собою кабельними з'єднаннями за різного роду інтерфейсам, або через канали зв'язку. Прикладом локальних АСУ є аеродромна АСУПР, елементи якої - командо-диспетчерський пункт, радіолокаційний комплекс, посадковий радіолокатор, радіопеленгатор - розташовані на території аеродрому, що обслуговується, тобто в обмеженій зоні.

2. ***Територіальні АСУ***. Це АСУ, технічні засоби яких розосереджені на великій території. Елементи територіальних АСУ можуть бути віднесені один від одного на значні відстані (до 200 - 400 км і більше) і з’єднані між собою каналами передачі даних. До територіальних АСУ належать, наприклад, районні енергетичні

управління, до яких входять АСУ електромережами, АСУ тепловими мережами, АСУ ТЕЦ, АСУ електростанціями розосереджені на значній території й з'єднані із комплексом засобів автоматизації (КЗА) районного енергетичного управління.

Класифікація АСУ за характером процесу, що управляється.

За цією ознакою доцільно виділити три великих класи АСУ, хоча за необхідністю можна визначити істотно більше класів.

1. ***АСУ технологічними процесами*** (АСУ ТП). Тут у поняття ***технологічного процесу*** включається досить широке коло процесів функціонування у різних областях господарства країни, різноманітних технічних засобів і людей, що виконують певну роботу із досягнення поставленої мети, виробництва тієї або іншої продукції, перевезенню вантажів і пасажирів, забезпеченню енергією, організації зв'язку й т.п. Таким чином, до технологічних процесів відносяться, зокрема, повітряний рух, рухи залізничного, водного й автомобільного транспорту, різні перевезення, електро, тепло, водопостачання і т.п. Загальною закономірністю для цього класу АСУ є те, що технологічний процес у більшості випадків має динамічний і порівняно швидко мінливий характер і найчастіше піддається впливу різних випадкових факторів. Управління технологічними процесами пов'язане з оперативним аналізом інформації про керований процес і оперативне прийняття рішень, що коректують окремі компоненти або і хід самого процесу. АСУ ТП завжди функціонують у режимі реального часу.

2. ***Організаційні АСУ***, тобто всі інші види АСУ, такі як АСУ підприємствами, галузями, АСУ банківськими й фінансовими операціями й т.п.

3. ***АСУ спеціальними технологічними процесами***. Автоматизовані системи спеціального призначення служать для обробки і аналізу інформації, що одержується від відповідних джерел, як правило, автоматично чи напівавтоматично. До їхнього числа входять системи попередження, збору й обробки даних, гідрометеорологічні системи й ін.

Класифікація АСУ за ступенем мобільності засобів автоматизації.

За цією ознакою розрізняють чотири класи АСУ.

I. ***Стаціонарні АСУ***. У цьому випадку елементи АСУ розміщуються в стаціонарних спорудженнях. Це найпоширеніший варіант виконання АСУ. При такому варіанті КЗА розміщуються в спеціальних приміщеннях або інженерно-будівельних комплексах, що забезпечують виконання вимог із монтажу обладнання, а також мають необхідну вентиляцію й кондиціонування, енерго, тепло і водопостачання й ін. Перевага цього варіанта АСУ в тому, що КЗА виготовляють із найменш жорсткими вимогами зі стійкості до механічних і кліматичних впливів, тобто на більше дешевих елементах і конструктивних базах, що в цілому приводить до їх меншої вартості. Однак істотним недоліком АСУ із КЗА, розташовуваними у стаціонарних спорудах, є необхідність більших обсягів капітального будівництва, що є серйозною перешкодою, при впровадженні цих АСУ. Крім того, такі системи практично на можуть бути перебазовані.

2. ***Мобільні АСУ.*** До цього класу належать такі АСУ, елементи яких розміщуються на рухомих об'єктах. Подібні АСУ при бажанні можуть бути перебазовані на інші об'єкти аналогічного функціонального призначення. При такому варіанті виконання АСУ апаратура може розміщатися в кузовах автомашин, спеціальних причепах, напівпричепах, контейнерах, фургонах, а також на водному, залізничному й повітряному транспорті. До апаратури таких АСУ пред'являються більш жорсткі вимоги у відношенні механічних і кліматичних впливів, а також до займаного обсягу. Умови роботи оператора в рухомих об'єктах незмірно більше складні, ніж у стаціонарних приміщеннях. Тому розміщення засобів АСУ на рухомих об'єктах розглядається як виняткове.

3. ***Комбіновані (змішані) АСУ***. До них належать АСУ, одна частина апаратури яких розміщується у стаціонарних спорудженнях, а інша - на рухомих об'єктах. Комбіновані АСУ дозволяють у ряді випадків мати переваги й позбутися від недоліків стаціонарних і мобільних АСУ.

4. ***Рухомі АСУ***. Це АСУ, які розміщуються на об'єктах, що рухаються, таких, як літаки, вертольоти, кораблі, пароплави, автомобілі й ін. Основна вимога до подібних АСУ - забезпечення можливості роботи на ходу. Технічні засоби рухомих АСУ, у порівнянні з іншими АСУ, мають найбільш жорсткі конструктивні вимоги й, отже, за інших рівних умов мають .

Класифікація АСУ за кількістю рівнів.

За ознакою ієрархії управління доцільно розглядати два класи АСУ.

I. ***Однорівневі (централізовані) АСУ***. У таких АСУ є КЗА центра управління, що за інформацією від відповідних джерел здійснює безпосереднє управління об'єктами, тобто управління об'єктами здійснюється прямо без проміжних інстанцій.

2, ***Багаторівневі (ієрархічні) АСУ***. У структуру ієрархічних АСУ входять КЗА для пунктів управління різних рівнів. Число рівнів управління може досягати 4-5, а в деяких випадках й більше. На КЗА центрів управління вищих рівнів у більшості випадків за отриманою узагальненою інформацією про керований процес виробляються укрупнені рішення з управління КЗА, що розташовані нижче які, в свою чергу, управляють підлеглими їм КЗА. Остання інстанція - КЗА найнижчого рівня - здійснює безпосереднє управління об'єктами. Іноді КЗА високих або середніх рівнів безпосередньо управляють найбільш важливими об'єктами. Однак у цих випадках багаторівнева структура АСУ не порушується.

Класифікація, АСУ за ступенем автоматизації.

За ступені автоматизації завдань прийнято розглядати два класи АСУ.

І. ***Автоматичні системи***. В автоматичних системах всі процеси збору, обробки, аналізу, зберігання й відображення інформації, виробітку або скасування рішень з управління, доведення прийнятих рішень до виконавців, контроль виконання рішень і інші операції здійснюються автоматично без участі людини. Роль людини в таких системах зводиться до спостереження за ходом процесу автоматичного управління з використанням засобів відображення, документування й контролю технічного стану апаратури, виявленню несправностей і прийняття, за необхідністю, заходів з відновлення працездатності АСУ. Однак і ці операції можуть виконуватися автоматично без участі оператора.

2. ***Автоматизовані системи.*** У системах цього класу частина завдань (як правило, найбільш трудомістких) вирішується автоматично за допомогою ЕОМ, а частина - за участю людини. У автоматизованих системах на людину покладають функції й завдання, які більш ефективно вирішуються не чисто формальними методами, а евристичними, шляхом інтелектуального аналізу із оцінкою сукупності всіх, нерідко неформалізованих факторів і обставин, з використанням досвіду й інтуїції оператора - учасника й головної особи, процесу управління. При цьому людині надається можливість приймати рішення самій або коректувати рішення, які вироблені засобами автоматики. Автоматизовані системи, у свою чергу, можна поділити на АСУ з високим, середнім і малим ступенем автоматизації завдань.

Класифікація АСУ за пропускною здатністю.

Під ***пропускною здатністю*** АСУ прийнято розуміти узагальнений обсяг обчислювальних робіт у одиницю часу, спрямованих на досягнення кінцевої мети функціонування АСУ. Звичайно стосовно до конкретних АСУ або сукупності АСУ однакового функціонального призначення пропускну здатність намагаються вимірювати у категоріях, що відображають реальний фізичний зміст функціонування і призначення АСУ. Однак характеристика (одиниця виміру) пропускної здатності АСУ повинна бути не залежною від призначення й складу конкретної АСУ. Тому, такою характеристикою доцільно обирати сумарну пропускну здатність обчислювальних засобів АСУ, необхідну для реалізації покладених на систему завдань.

Як додаткову характеристику пропускної здатності АСУ можна прийняти сумарну швидкість передачі даних між елементами АСУ, необхідну для забезпечення циркуляції в АСУ необхідного потоку інформації у одиницю часу.

За ***характером протікання технологічного процесу*** *розрізняють:*

- АСУ дискретним ТП;
- АСУ неперервно-дискретним ТП;
- АСУ неперервним ТП.

Класифікація АСУ за серійністю.

При класифікації АСУ за ознакою серійності необхідно враховувати наступні обставини.

Наявність об'єктів, що оснащуються АСУ одного й того же самого функціонального призначення. Об’єкти, що мають однакове функціональне призначення, можуть відрізнятися за складом, числом й характеристиками джерел інформації, за складом, характеристикам і взаємному розташуванню об'єктів управління.

За ознакою серійності можна виділити три класи АСУ.

1. ***Унікальні АСУ***. Це такі АСУ, які у силу тих або інших причин виробляються тільки в єдиному екземплярі (космічні станції).

2. ***Малосерійні АСУ***, До них відносяться такі АСУ, виробництво яких із тих або інших причин реалізується в обмеженій кількості (АЕС).

Якщо вважати, що період існування таких АСУ, до їхнього морального старіння, дорівнює 20 рокам, то максимальний обсяг серії не перевищує 5-7 екземплярів АСУ.

3. ***Багатосерійні АСУ***. Це АСУ, виробництво яких за час їхнього існування носить досить масовий характер, тобто реалізується у десятках, сотнях екземплярів.

Зауваження. Приналежність АСУ до того або іншого класу значною мірою визначає вибір і застосування обчислювальних *засобів, комплексів технічних засобів зв'язку й передачі даних, відображення й документування інформації, особливості побудови математичного забезпечення й технології його розробки, зміст основних технічних рішень, специфіку організації АСУ й забезпечення її працездатності із заданими характеристиками.*

1.3. Основні методи автоматизованого управління

До основних методів автоматизованого управління, що застосовуються при створенні та експлуатації АСУ, відносяться:

- програмно-логічне управління;
- оптимальне управління;
- адаптивне управління;
- екстремальне управління;
- ситуаційне управління.

1.3.1. Програмно-логічне управління

У загальному випадку ***логічним управлінням*** називається управління, що здійснюється у строгій відповідності із правилами формальної логіки. Для опису процесу логічного управління об'єктами промислового призначення застосовується апарат так званих логічних функцій, що, як і їхні аргументи, можуть приймати тільки два значення "*0*" і "*1*". Для реалізації логічного управління використовують логічні управляючі автомати (ЛУА), що можуть працювати за принципом "жорсткої" або програмованої логіки.

Незалежно від елементної бази „***жорсткою***” називають логіку, що передбачає наявність фіксованого електричного монтажу і необхідність виконання перемонтажу при змінах у управляючих алгоритмах, що значно обмежує здатність адаптації таких ЛУА до можливих змін алгоритмів управління у зв'язку, наприклад, із модернізацією чи вдосконалюванням технологічного устаткування.

Програмованою називають логіку, що реалізується програмним шляхом, а ЛУА, що її реалізують, називають програмованими ЛУА. У програмованих ЛУА алгоритм управління представлений у вигляді програми, що зберігається в оперативній чи постійній пам'яті. Можливість оперативної заміни і програмне коректування алгоритмів управління без зміни електричного монтажу сприяла появі терміна "гнучка" (програмована) логіка.

Логічне управління, що здійснюється за принципом програмованої логіки, називається ***програмно-логічним управлінням***.

У даний час абсолютна більшість пристроїв управління промисловим устаткуванням, що застосовуються у промисловості і створюється є пристроями програмно-логічного управління.

Під ***алгоритмом управління*** промисловим устаткуванням будемо розуміти набір команд і контрольних операцій, які виконуються у визначеній послідовності, що забезпечують функціонування устаткування відповідно до заданого циклу і режиму його роботи.

Під ***алгоритмом програмного управління*** слід розуміти алгоритм управління, що реалізується програмним шляхом.

Первинним (вихідним) описом алгоритму управління промисловим устаткуванням є неформальний (вербальний) змістовний опис задачі управління, що на практиці, як правило, фіксується у виді розділів технічного завдання на розробку управляючого пристрою.

Такий опис обов'язково містить перелік заданих режимів роботи, перелік виконавчих механізмів, схему їхньої взаємодії та послідовність вмикання-вимикання, перелік органів управління, перелік контрольних датчиків, перелік необхідних інформаційних і діагностичних сигналів про роботу устаткування, необхідні вказівки у частині забезпечення електричних блокувань при спрацюванні окремих механізмів і вузлів, а також перелік параметрів технологічного процесу, що реалізується на даному устаткуванні, які підлягають контролю і регулюванню.

Неформальний опис алгоритмів управління є основою усіх формалізованих способів їх завдання для реалізації як за допомогою пристроїв із "жорсткою" логікою, так і програмним шляхом.

До основних способів формалізації алгоритмів, що застосовуються в інженерній практиці, відносяться:

- часові діаграми;
- технологічні циклограми; технологічні тактограми;
- мова таблиць рішень;
- блок-схеми (операторні схеми) алгоритмів;
- автоматні способи опису алгоритмів.

Часові діаграми, технологічні такто і циклограми відносяться до слабо формалізованих способів представлення алгоритмів. Часові діаграми і циклограми називають також "технологічними" мовами опису, тому що вони дозволяють одержати найбільш наочне уявлення про технологічний процес, що управляється, про послідовність операцій, що виконуються а також про особливості технологічного циклу або циклу роботи устаткування, як системи циклічно повторюваного набору операцій і рухів робочих органів, що виконуються.

Часова діаграма і тактограма являють собою графіки вмикання-вимикання виконавчих механізмів, де на одну з координат наносяться конкретні часові інтервали (для часової діаграми) або такти (для тактограми), причому тривалість такту умовно показується однаковою.

Як приклад розглянемо тактограму фрагмента циклу роботи промислового робота (ПР), що здійснює перевантаження готового виробу з технологічного агрегату на інше устаткування. Розглянемо тільки дії ПР по опусканню пристрою схвату, захопленню виробу і підйому пристрою схвату у верхнє положення.

При наявності готового виробу на позиції розвантаження технологічного агрегату Неформальний опис обраного фрагменту виглядає так.

1. Включення електроприводу механізму розтиснення пристрою схвату.
2. Перевірка того, що розтиснення виконано.
3. Включення електроприводу механізму опускання пристрою схвату.
4. Виключення електроприводу механізму опускання пристрою схвату, при перебуванні схвату у нижньому положенні.
5. Включення електроприводу механізму стиснення пристрою схвату (забезпечення захоплення виробу).
6. Контроль того, що схват стиснуто.
7. Включення електроприводу механізму підйому пристрою схвату.
8. Відключення електроприводу механізму підйому пристрою схвату при досягненні схватом верхнього положення.

Відповідно до неформального опису задачі управління побудована тактограма рухів ПР і відповідних контрольних дій (рис. 1.3).

Якщо на осі часу такти замінити конкретними часовими інтервалами, що відповідають тривалості спрацьовування виконавчих механізмів, то тактограма стає часовою діаграмою.

Для зображення ***циклограми*** користуються двома основними способами. При використанні першого способу циклограма являє собою граф-схему циклу, на якій за допомогою стрілок (або інших графічних символів) показується послідовність виконуваних технологічних операцій спрацьовування виконавчих механізмів. Так, для розглянутого вище приклада циклограма може мати вид, представлений на рис. 1.4. Слід зазначити, що у загальному випадку циклограма повинна мати вид замкнутої схеми, що відповідає повному циклу роботи устаткування (повернення механізмів наприкінці циклу у вихідний стан).

Рис. 1.3. Тактограма рухів ПР

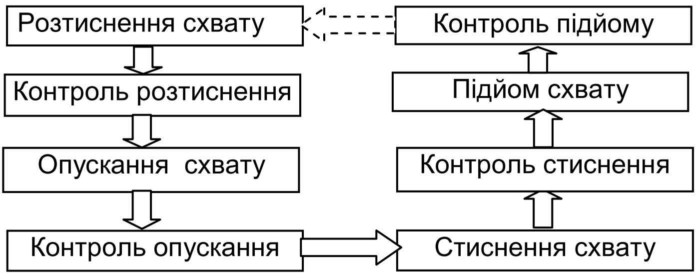


Рис. 1.4. Циклограма рухів ПР

Іншим способом представлення циклограм є таблиця, в якій елементи часової діаграми (тактограми) зображуються у виді чарунок таблиці, заповнюваних умовними знаками типу "+", "-", або "*1*","*0*", чи "*x*"," " – порожня чарунка і інш.

У такій таблиці число рядків відповідає числу тактів (кроків) циклу, а число стовпців – сумарній кількості керованих механізмів, контрольованих датчиків і органів управління. У табл. 1.1 приведена циклограма для розглянутого фрагмента циклу роботи ПР.

Таблиця 1.1

Циклограма рухів ПР

Такти циклу	Дії, контроль операції	Стани								
		датчиків					механізмів			
		Наявність виробу	Схват у горі	Схват унизу	Схват стиснутий	Схват розтиснутий	Підйом схвату	Опускання схвату	Стиснення схвату	Розтиснення схвату
1	Розтиснення схвата	X								X
2	Контроль розтиснення	X				X				X
3	Опускання схвату	X				X		X		X
4	Контроль опускання	X		X		X				X
5	Стиснення схвату	X		X		X			X	
6	Контроль стиснення	X		X	X				X	
7	Підйом схвату	X		X	X		X		X	
8	Контроль підйому		X		X				X	

Циклограми досить широко застосовуються в інженерній практиці тому, що вони є універсальною мовою спілкування між фахівцями різного профілю: технологами, механіками, електриками, спеціалістами з автоматизації технологічних процесів, особливо на етапі розробки технічних завдань на системи автоматизації. Крім того, як показує досвід, циклограми виявляються дуже зручним доповненням до складного програмного забезпечення управляючих систем при освоєнні їх обслуговуючим персоналом систем автоматизації.

Мова таблиць рішень за формою запису алгоритму дуже близька до другого варіанта представлення циклограм. Головна відмінність тут полягає у тому, що при використанні мови таблиць рішень не є обов'язковим строго послідовне виконання операцій. У загальному випадку таблиця рішень - це таблиця, що містить число рядків, яке повинно бути принаймні не менше числа прийнятих рішень чи їхніх комбінацій, число стовпців у лівій частині відповідає кількості умов, а у правій – кількості компонентів вектора рішень (вектора управлінь).

Як приклад розглянемо, як за допомогою мови таблиць рішень може бути описаний запуск в автоматичний режим роботи технологічної установки (табл. 1.2).

Таблиця. 1.2

Запуск у автоматичний режим роботи технологічної установки

Кнопка пуск	Реж. автоматика	Механізм 1 у реж. автоматика	Механізм 1 у реж. налагодження	Механізм 2 у реж. автоматика	Механізм 2 у реж. налагодження	Запуск установки в автоматичному режимі	Дозвіл роботи у налагоджувальних режимах
1	1	1		1		1	
1	0	1		1			1
0	1	1		1			1
1	1		1				1
1	1				1		1

З представленої таблиці випливає, що при комбінації станів "натиснута кнопка ПУСК", "обраний режим АВТОМАТИКА" і "механізми *1* і 2 знаходяться в режимі АВТОМАТИКА" - дозволяється запуск установки в автоматичний режим роботи. У випадку, якщо не обраний режим АВТОМАТИКА чи не натиснута кнопка ПУСК, чи хоча б один із механізмів *1* чи *2* переведений у режим НАЛАГОДЖЕННЯ, видається дозвіл на роботу установки тільки в налагоджувальному режимі.

Слід відмітити, що елементи мови таблиць рішень у явному або неявному вигляді практично завжди мають місце у всіх алгоритмах програмного управління, в яких існує задача автоматичного вибору одного з декількох можливих режимів експлуатації устаткування.

До числа широко застосовуваних на практиці способів формалізації алгоритмів і особливо алгоритмів програмного управління відносяться блок-схеми (операторні схеми) алгоритмів.

Блок-схема алгоритму являє собою набір геометричних фігур – операторів (блоків), з'єднаних певним чином. Кожен оператор відповідає кроку алгоритму, послідовність кроків визначається зв'язками між операторами. Державними стандартами визначені позначення різних операторів, серед яких можна виділити 4 основні типи:

- початковий оператор;
- кінцевий оператор;
- оператори, що призначені для опису тих чи інших дій та операцій (введення, підготовки, обробки, виведення даних і ін.);
- умовні (логічні) оператори, що визначають напрямок розгалуження алгоритму.

Початковий оператор не має входу і має єдиний вихід, кінцевий оператор не має виходу і може мати безліч входів, оператори, що описують дії, можуть мати кілька входів і один вихід, а умовні оператори можуть мати кілька входів, але завжди мають *2* виходи, що позначаються "*1*" (відповідає виконанню умови) і "*0*" (відповідає невиконанню умови). Умовні оператори зображуються у виді ромба.

На рис. 1.5 наведена блок-схема алгоритму управління для розглянутого в даному розділі приклада.

Цілий ряд способів опису алгоритмів називають "***автоматними мовами***" з тієї причини, що ці мови використовуються для формального опису функціонування цифрових автоматів. До автоматних мов відносяться таблиці станів, таблиці переходів, мова логічних схем алгоритмів (ЛСА), мова матричних схем алгоритмів (МСА), мова регулярних виражень, мова секвенцій і ін. Зазначені мови дуже ефективно використовуються не тільки для опису процесу функціонування цифрових автоматів, але і для рішення багатьох задач, зв'язаних з аналізом функціонування автоматів і, насамперед, із рішенням задач мінімізації схемної реалізації автоматів.

Однак, ефективність застосування автоматних способів опису алгоритмів управління обмежується об'єктами з малим числом можливих внутрішніх станів, тому при розробці алгоритмів програмного управління промисловим устаткуванням автоматні мови не знаходять широкого застосування.

У загальному випадку реалізація алгоритмів програмно-логічного управління може бути здійснена за допомогою будь якого стандартного пристрою програмного управління (програмованого логічного контролера, промислового комп'ютера), для чого за розробленим алгоритмом програмно-логічного управління складається програма, що управляє на відповідній мові програмування.

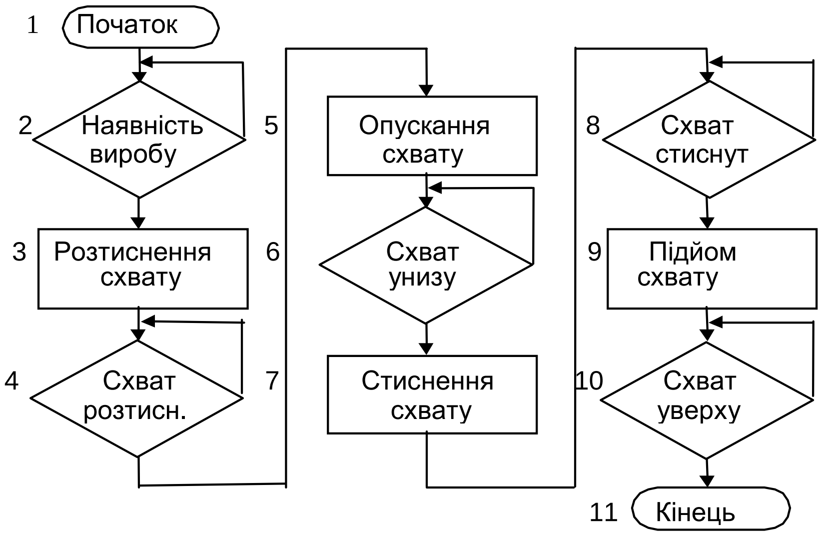


Рис. 1.5. Блок-схема алгоритму управління

Проміжним етапом між розробкою алгоритму програмно-логічного управління і створенням тексту програми, що управляє є складання так званих логічних рівнянь, які мають дуже наочну форму реалізації алгоритмів програмно-логічного управління в термінах мови програмування.

Логічне рівняння складається з двох частин: логічного виразу (ліва частина) та виконавчої (правої) частини. Логічний вираз записується за правилами булевої алгебри, являє собою множину логічних умов реалізації функції (аргументів), що пов'язані між собою знаками логічних функцій І, АБО, НІ, а права частина рівняння представляє собою результат рішення логічного рівняння – функцію (наприклад, включення – відключення якогось механізму, пристрою, приладу).

Логічні рівняння є основою для складання тексту робочої програми. Перетворення тексту робочої програми у машинний код конкретного управляючого пристрою (програмованого контролера, промислового комп'ютера) виконується за допомогою спеціального транслятора.

Програмно-логічне управління в основному орієнтовано на рішення задач управління детермінованими технологічними процесами та устаткуванням циклічної дії. При цьому використовують алгоритми програмно-логічного управління:

- строго послідовної дії (коли послідовність технологічних операцій, що виконуються є незмінною);
- алгоритмів управління з розгалуженнями (за певними умовами);
- алгоритмів управління з вкладеними циклами.

У детермінованих технологічних процесах при реалізації алгоритмів управління із розгалуженнями початок і саме виконання тієї чи іншої операції може носити випадковий характер, але ця випадковість у даному разі знаходиться у межах заздалегідь відомої детермінованої множини варіантів або ситуацій.

Із усього різноманіття розглянутих способів формалізованого представлення алгоритмів програмного управління в інженерній практиці перевага надається блок-схемам алгоритмів, причому навіть тоді, коли цей спосіб не є оптимальним. Відбувається це тому, що блок-схеми виявляються найбільш універсальною формою представлення алгоритму, зрозумілої сучасним фахівцям різного профілю. Слід зазначити, однак, що досить часто при розробці алгоритмів управління складними об'єктами користаються двоступінчастою схемою, а саме: на 1-м етапі по неформальному опису задачі управління складають часову діаграму або циклограму, чи таблицю рішень, а на 2-м етапі по розробленій циклограмі будують блок-схему алгоритму програмного управління.

1.3.2. Оптимальне управління

Для досягнення своїх цілей ми здійснюємо найрізноманітніші операції. Проте, у повсякденному житті ми рідко замислюємося над тим, що створюється для проведення операції й наскільки ефективно її здійснюють. Інша справа, коли однотипні операції здійснюються на регулярній основі у вигляді технологічного процесу, і від ефективності таких операцій залежать темпи розвитку підприємства й конкурентоздатність продукції. У цьому випадку ми прагнемо до того, щоб і операції, які здійснюються були максимально ефективними, найкращими або, що теж, оптимальними.

У загальному випадку оптимальне управління дозволяє при заданих умовах (часто суперечливих) досягти поставленої мети найкращим образом, наприклад за мінімальний час, із найбільшим економічним ефектом, з максимальною точністю. Якщо говорити коротко, оптимальне управління це технологічний процес, що складається з безлічі операцій із такими параметрами, які до певного моменту часу забезпечать одержання максимального за величиною цільового продукту.

Термін "оптимальне управління" історично пов'язаний із розвитком і вдосконаленням теорії автоматичного управління. Області її застосування надзвичайно розширилися від простих систем стабілізації окремих параметрів машин і механізмів до систем управління великими підприємствами. З'явилися нові системи: оптимальні, екстремальні, адаптивні і т.д. Нові теоретичні і практичні задачі, що виникли при цьому, зажадали для свого вирішення залучення іншого математичного апарату і, зокрема, варіаційного числення. Проте виявилося, що традиційні класичні методи варіаційного числення не завжди можуть бути застосовні до специфічних задач теорії управління. Тому спеціально для теорії управління були розроблені принцип максимуму і метод динамічного програмування. Принцип максимуму був запропонований радянським ученим Л.С. Понтрягіним і його школою в особі В.Г. Болтянського, Р.В. Гамкрелідзе, Е.Д. Міщенко. Метод динамічного програмування розвивався американською школою на чолі з Р. Беллманом. Нові ідеї, що містяться у цих методах, дали могутній поштовх до розвитку математичної теорії оптимального управління.

Під ***оптимальною системою управління*** розуміється система, що забезпечує найкраще (оптимальне) з деякого погляду функціонування об'єкту, що управляється.

Для вирішення задачі про оптимальне управління необхідно мати повну апріорну інформацію про об'єкт управління (його математичний опис), а також достатньо точне математичне формулювання мети управління (досягнення екстремуму деякої величини Q - критерію оптимальності).

Наприклад, для досягнення максимальної точності функціонування системи критерієм оптимальності може служити мінімум середньої квадратичної помилки регулювання, вираженої у вигляді інтеграла

$$Q = \int_0^T x^2(t)dt,$$

де ***x(t)*** - відхилення регульованої величини від необхідного значення.

Величина ***Q*** є функціоналом, залежним від виду функції ***x(t)***. Критерієм оптимальної точності може бути також мінімум статичної помилки при максимальній зовнішній дії. У інших випадках критерієм може бути мінімум витрати енергії на виконання процесу, що управляється максимум надійності, к.к.д., продуктивність і т.д.

Найчастіше критерій оптимальності задається у вигляді інтегрального квадратичного функціонала від декількох функцій $\varphi_i(t)$

$$Q = \int_0^T \sum_{i=1}^{n} r_i \varphi_i^2(t)\, dt \cdot$$

У формулі замість ***t***, взагалі кажучи, може бути будь-який інший фізичний параметр (наприклад, якщо ставиться задача отримання мінімуму нерівномірності розподілу температур в об'ємі ***V***, то замість ***t*** буде ***V***) або навіть умовна комбінована незалежна змінна.

У загальному випадку критерій ефективності процесу, що підлягає оптимізації залежить від стану системи, що характеризується векторами вихідних, $\overline{x} = (x_1, x_2 ... x_n)$, задаючих $\overline{x^*}$, збурюючих $\overline{z}$ і управляючих $\overline{u}$ дій, а також від часу ***t***. Вектори $\overline{x}, \overline{x^*}, \overline{z}$ і $\overline{u}$ також є функціями часу. Таким чином,

$$Q = \int_0^T F\left(\overline{x}, \overline{x^*}, \overline{z}, \overline{u}, t\right) dt \cdot$$

У задачах про оптимальне управління вимагається знайти таке допустиме управління $\overline{u}(t)$, що задовольняє умові $\overline{u} \in \Omega\left(\overline{u}\right)$ ($\overline{u}$ належить області допустимих рішень $\Omega\left(\overline{u}\right)$) і відповідну змінну стану системи $\overline{x}(t)$, щоб під час переходу системи зі стану $\overline{x}_0$ у стан $\overline{x}_T$ мінімізувався функціонал

$$Q = \int_0^T F\left(\overline{x}(t), \overline{u}(t), t\right) dt \cdot$$

Тут вектори $\overline{z}$ і $\overline{x^*}$ вважаються незмінними і тому не враховуються. В окремому випадку, коли ***F = 1***, одержимо ***Q = T***. Тоді, якщо вимагається мати ***Q = min***, одержуємо задачу про максимальну швидкодію, в якій вимагається знайти такий закон управління $\overline{u}(t)$, при якому система переводиться із стану $\overline{x}_0$ в $\overline{x}_T$ за мінімальний час ***T***.

Сформульована задача належить до класу варіаційних задач. Існують різні методи рішення таких задач. Найширше поширення набули:

- класичні варіаційні методи;
- метод динамічного програмування;
- принцип максимуму.

Класичні варіаційні методи доцільно застосувати до задач, у яких області зміни параметрів $\overline{x}(t)$ і $\overline{u}\ (t)$, не містять обмежень. Це буває у випадках, коли розглядаються малі відхилення $\overline{x}(t)$ і $\overline{u}\ (t)$, від сталих станів системи.

Метод динамічного програмування і принцип максимуму застосовуються, коли області $\overline{x}(t)$ і $\overline{u}\ (t)$, замкнуті, а координати векторів $\overline{x}(t)$ і $\overline{u}$ можуть знаходитися на межах цих областей.

1.3.3. Адаптивне управління

Проектуючи систему автоматизованого управління, конструктор завжди прагне зробити її можливо простіше. При цьому розрахунок параметрів системи управління ведеться в припущенні, що статичні і динамічні характеристики об'єкту управління і всіх елементів відомі і не змінюються із зміною зовнішніх умов і у часі. Проте у дійсності характеристики об'єкту і деяких елементів системи бувають відомі лише приблизно, вони змінюються у результаті фізичного старіння і, крім того, залежать від зовнішніх умов. Завдяки запасам стійкості система управління, що спроектована для певних розрахункових характеристик, задовільно управлятиме об'єктом і у тому випадку, коли його дійсні характеристики дещо відрізняються від розрахункових. Проте, у реальних випадках діапазон зміни статичних і динамічних характеристик настільки великий, що управління об'єктом за допомогою простої системи з постійними параметрами виявляється або незадовільним, або взагалі неможливим (втрата стійкості управління). У таких випадках необхідно застосувати систему управління з властивостями, що змінюються.

Процес зміни властивостей системи, що дозволяє їй досягти найкращого або, принаймні, задовільного функціонування в умовах, що змінюються, називається ***адаптацією***. Системи, що здійснюють процес адаптації, називаються ***адаптивними*** (що пристосовуються) ***системами*****.**

Необхідно відзначити, що адаптивні системи давно існують у природі. Властивість адаптації виразно виявляється у механізмі **гомеостазиса**, який полягає у тому, що живі організми володіють здатністю утримувати свої життєво важливі характеристики (наприклад, температуру) у допустимих фізіологічних межах, при значних змінах умов, у яких існує організм.

Характерною ознакою адаптивних систем є відсутність повної апріорної інформації про об'єкт управління, зовнішні збурення і граничні умови, тобто адаптивній системі властивий індетермінізм (невизначеність). Функціонування системи направлено на зниження рівня невизначеності, тобто на знаходження такого стану, при якому задовольняється певний критерій.

Усунення невизначеності в адаптивних системах забезпечується завдяки:

- надмірності системи, яка виявляється в багатоступінчатості, багатоконтурності і т. д.;
- підвищенню рівня логічного "мислення" системи;
- аналізу накопичуваної інформації з метою самонавчання і прогнозуванню можливого стану системи.

Оптимальне функціонування адаптивної системи, у відомому значенні, може розраховуватися на підставі аналізу інформації про стан системи. Такі адаптивні системи називаються ***аналітичними***.

Якщо ж оптимальний режим роботи визначається у результаті пошуку умов екстремуму критерію ефективності, то такі адаптивні системи називають ***пошуковими***. У цьому випадку система як би ставить серії експериментів і отримує з них дані, необхідні для поліпшення своєї поведінки.

Зміну стану системи можна проводити за рахунок зміни управляючих дій, параметрів настройки і структури системи. Ці зміни називаються ***контрольованими змінами*** стану системи. Залежно від об'єму цих змін адаптивні системи підрозділяються на:

- екстремальні, у яких можна здійснювати зміни тільки управляючих дій;
- такі, що самоналагоджуються, у яких, крім того, змінюються параметри системи;
- такі, що самоорганізовуються, у яких, окрім управляючих дій і параметрів, змінюється ще і структура системи;
- такі, що навчаються, у яких на додаток до всього може змінюватися алгоритм дії, а у разі самонавчання і критерій ефективності.

За способом здійснення контрольованих змін адаптивні системи підрозділяються на:

- пасивні, у яких ці зміни здійснюються за наперед розробленою програмою, наприклад, система управління автопілотом;
- активні, уяких контрольовані зміни наперед не визначені, а визначаються ситуацією, що склалася.

Як і будь-які інші автоматичні системи, адаптивні системи можуть працювати за замкнутим і розімкненим циклах. У першому випадку проводиться аналіз контрольованих змін, у другому - не проводиться. Самими довершеними є аналітичні активні замкнуті системи, що навчаються.

1.3.4. Екстремальне управління

У цілісному вигляді теорія екстремальних систем була розроблена радянськими вченими Козакевічем В. В. і Хлібцевічем Ю. С. у 1948 р. Задачею екстремальної системи управління є забезпечення якнайкращого, у відомому значенні, статичного режиму роботи об'-єкту. Природно, що при цьому передбачається обов'язкова наявність екстремуму критерію ефективності Q (функції якості) контрольованих управляючих діях $\overline{u}$. У гіршому випадку знання лише того, що такий екстремум існує, може служити достатньою апріорною інформацією для побудови системи екстремального управління (СЕУ).

Таким чином, якщо статична характеристика $Q = f\left(\overline{u}\right)$ має екстремум, то СЕУ повинна вивести і утримати робочу точку в глобальному екстремумі.

Нехай об'єкт має статичну характеристику $Q = f\left(\overline{u}\right)$ з явним екстремумом, координати якого змінюються (дрейфують) у часі. У випадку якщо $u = u^*$, при якому досягається екстремум, фіксовано (має місце тільки вертикальний дрейф статичної характеристики (рис. 1.6)), або змінюється за наперед заданим законом, то достатньо застосувати або систему стабілізації, або систему програмного управління.

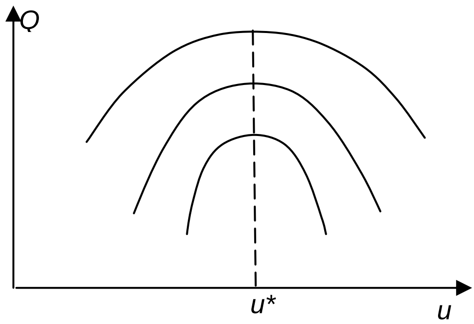


Рис. 1.6. Вертикальний дрейф статичної характеристики

У тому випадку, якщо крім вертикального спостерігається і горизонтальний дрейф статичної характеристики (рис. 1.7), викликаний зміною зовнішніх і внутрішніх збурень за наперед невідомим законом, то названі системи вже не зможуть забезпечити підтримку екстремального значення ***Q***. Саме в цьому випадку і потрібне вживання СЕУ, яка організує такі зміни управляючих дій, при яких забезпечується рух системи до екстремуму і утримання її в точці екстремуму.

$$Q_{\Sigma} = c_1 G - c_2 |K|^n$$

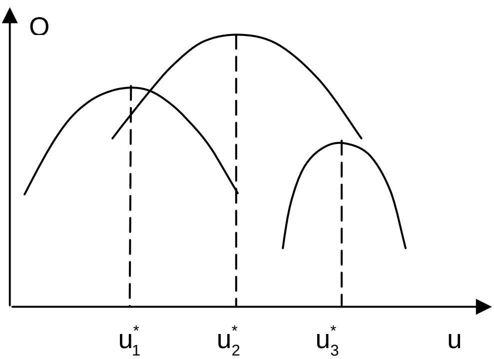


Рис. 1.7. Горизонтальний дрейф статичної характеристики

Оскільки положення точок екстремуму у просторі управлінь $\bar{u}$ залежить від збурюючих дій, то необхідно підібрати таке значення управлінь, при яких величина Q_Σ максимізується. Цю задачу і вирішує система екстремального управління.

Прикладом об'єкту з екстремальною характеристикою може служити промислова пічна установка. Статична характеристика каналу витрат повітря на горіння u - температура топочних газів Q має екстремум при заданій витраті палива певної якості, який досягається при певному значенні $u = u^*$. При $u < u^*$ маємо неповне згорання палива, при $u > u^*$ надлишок повітря знижує температуру топочних газів. При зміні витрат або якості палива, що подається, значення u, при якому $Q = max$, змінюється. Задача СЕУ в даному випадку зводиться до зміни подачі повітря до значень, що забезпечують якнайкращі умови згорання палива.

Системи екстремального управління можна класифікувати за різними ознаками, зокрема:

- за способом формування пошукового сигналу: з безпосереднім вимірюванням похідної; із запам'ятовуванням екстремуму; крокового типу; із зовнішнім пошуковим сигналом і інш.;
- за виглядом сигналів: безперервні і дискретні, детерміновані і стохастичні;
- за числом управляючих дій: одно, двох і багатоканальні.

1.3.5. Ситуаційне управління

Метод ситуаційного управління використовується для рішення задач управління слабо структурованими процесами та об'єктами, для яких складно або практично неможливо розробити формальні моделі.

У загальному випадку під ***ситуаційним*** розуміють управління, при якому використовують спеціальні, так звані семіотичні (знакові) моделі за допомогою яких встановлюються зв'язки між множиною ситуацій, яка має місце на об'єкті управління в ***i***-й момент часу і множиною необхідних відповідних управляючих дій.

Розглянемо основні поняття, що застосовуються при використання методів ситуаційного управління.

Поточна ситуація - сукупність всіх відомостей про структуру об'єкта управління і його функціонування в даний момент часу.

Повна ситуація - сукупність, що складається з поточної ситуації, знань про стан системи управління в даний момент і знань про технологію управління.

Будемо позначати повні ситуації через S_n, а поточні ситуації - через Q_i.

Допустимо, що у розпорядженні системи управління є m різних варіантів (однокрокових рішень). Кожне таке рішення будемо позначати як U_k (k - номер варіанта). Елементарний акт управління можна представити в наступному вигляді:

$$\begin{vmatrix} S_n; Q_i \Rightarrow Q_l \\ U_k \end{vmatrix}$$

Значення цього співвідношення полягає в наступному. Якщо на об'єкті управління склалася ситуація Q_i і стан системи управління і технологічна схема управління, що визначаються S_n, припускають використання дії U_k, то вона застосовується і поточна ситуація Q_i перетворюється на нову ситуацію Q_l. Подібні правила перетворення надалі називаються ***логіко-трансформаційними правилами*** (ЛТП) або ***кореляційними правилами***. Повний список ЛТП задає можливості системи управління впливати на процеси, що протікають на об'єкті.

Очевидно, що вся множина можливих повних ситуацій може бути поділена на n класів, кожному із яких буде відповідати одна з можливих дій на об'єкт управління. Тобто, повинні існувати такі процедури класифікації, які дозволили би класифікувати можливі повні ситуації таким чином, щоб із них можна було утворити стільки класів, скільки різних однокрокових рішень є у розпорядженні системи управління. Якщо для якихось повних ситуацій неможливо вказати єдине однокрокове рішення, то можливо включення цих ситуацій у декілька класів.

Для здійснення вибору того або іншого рішення з числа можливих, для даної повної ситуації потрібні спеціальні процедури екстраполяції наслідків прийняття того або іншого рішення. З їхньою допомогою можна, на підставі знань про об'єкт управління і його функціонування, наперед оцінити результати застосування вибраної дії і порівняти отримані прогнози для всіх можливих дій у даній повній ситуації. Тоді виявляється можливим побудувати загальну схема рішення задачі управління (рис. 1.8), де БВВ – блок випадкового вибору.

Опис поточної ситуації, що склалася на об'єкті управління, подається на вхід аналізатора. Його задача полягає в оцінці повідомлення і визначенні необхідності втручання системи управління в процес, що протікає в об'єкті управління. Якщо поточна ситуація не вимагає такого втручання, то аналізатор не передає її на подальшу обробку. Інакше опис поточної ситуації поступає до класифікатора.

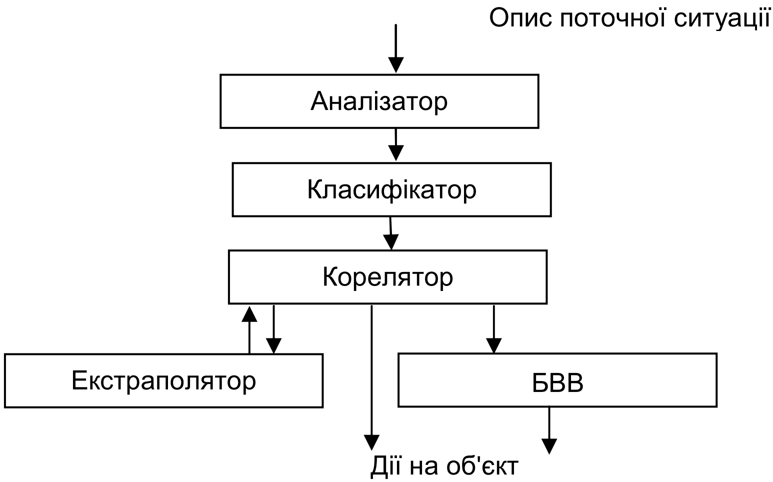


Рис. 1.8. Загальна схема рішення задачі управління

Класифікатор, використовуючи інформацію, що зберігається в ньому, відносить поточну ситуацію до одного або декількох класів, яким відповідають однокрокові рішення. Ця інформація передається до корелятора, в якому зберігаються всі ЛТП. Корелятор визначає те ЛТП, яке повинне бути використане. Якщо таке правило єдине, то воно видається для виконання. Якщо ж таких правил декілька, то вибір кращого з них проводиться після обробки попередніх рішень у екстраполяторі, після чого корелятор видає рішення про дію на об'єкт. Якщо корелятор або класифікатор не можуть прийняти рішення з опису поточної ситуації, що надійшов, то спрацьовує блок випадкового вибору і вибирається одна з дій, що забезпечує не дуже великий вплив на об'єкт, або ж система зовсім відмовляється від якої-небудь дії на об'єкт. Це говорить про те, що система управління не має в своєму розпорядженні необхідної інформації про свою поведінку у даній ситуації.

Фактично через складність об'єктів управління початкові знання про них і способи управління ними не завжди бувають достатньо повні. Тому система управління подібного типу принципово повинна бути "відкритою" системою, яка повинна мати можливість коректувати свої знання про об'єкт і методи управління їм. У роботі такої системи управління можна виділити два етапи:

- етап навчання і налагодження;

- етап функціонування.

У початковий період створення системи управління збираються відомості про об'єкт управління за допомогою яких формуються класи ситуацій і ЛТП. Також визначається доцільність використання тих або інших впливів на об'єкт управління в тій або іншій ситуації. При цьому думки експертів можуть не співпадати, що і приводить до попадання однієї і тієї ж ситуації в різні класи. За допомогою експертів формуються і процедури екстраполяції, способи оцінки ситуацій, що витікають з бажаного функціонування об'єкта управління.

Після етапу накопичення знань і формування процедур система може починати працювати. Але, у початковий період експлуатації через неповноту інформації і неточності процедур, система може приймати невірні рішення. Тому необхідний додатковий етап – етап "донавчання" системи управління.

В АСУ ТП методи ситуаційного управління застосовують при вирішенні питань управління складними технологічними процесами та роботизованими технологічними комплексами, для яких на даний час поки ще не існує (або їх у принципі розробити неможливо) достатньо точних формальних моделей їх функціонування.

1.4. Контрольні питання

1. Дайте визначення поняттю АСУ.
2. Дайте визначення поняттю технологічний процес, що управляється.
3. Дайте визначення поняттю автоматизація виробництва.
4. Дайте визначення поняттю ефективність АСУ.
5. Дайте визначення поняттю критерій управління.
6. Дайте визначення поняттю регулювання.
7. Перерахуйте основні операції управління ТП.
8. Назвіть основні види управління ТП.
9. Наведіть класифікацію типів об'єктів управління.
10. Дайте визначення поняттю блок-схема алгоритму.
11. Дайте визначення поняттю технологічна циклограма.
12. Дайте визначення поняттю часова діаграма.
13. У чому полягає сутність програмно-логічного управління?
14. У чому полягає сутність оптимального управління?
15. У чому полягає сутність адаптивного управління?
16. У чому полягає сутність екстремального управління?
17. У чому полягає суть ситуаційного управління?

Розділ 2

Застосування елементів теорії технічних систем для опису структури АСУ

Розглянуті основні поняття і визначення теорії технічних систем, методи опису структур технічних систем, наведено класифікацію технічних систем, методи композиції та управління технічними системами. Також розкриті основні властивості технічних систем. Наведено контрольні питання.

2.1. Основні поняття та визначення теорії технічних систем

У основі теорії систем лежить сформульована біологом Л. фон Берталанфі теза, згідно з якою для системи, що складається із елементів, можуть бути вказані певні загальні принципи незалежно від того, яка фізична сутність цих елементів. Теорія систем стала потужно розвиватись у зв'язку із появою кібернетики, де сама система, що управляється, стала об'єктом дослідження. У центрі уваги теорії систем лежить дослідження причинно-наслідкових взаємозв'язків.

Дамо визначення поняттю ***теорія систем*** - це науковий напрямок, що вивчає загальні властивості систем, способи їхньої організації і процеси, що в них протікають. Об'єктом вивчення теорії систем є системи різної фізичної природи і різного призначення. Так, для теорії технічних систем об'єкт дослідження – це механізми й агрегати, силові установки, автоматизовані й автоматичні системи управління та ін. У цьому випадку ***теорія технічних систем*** – це розвинута форма організації наукового знання: систематизовані й узагальнені знання про закономірності й особливості розвитку та удосконалювання технічних і інформаційно-управляючих систем. Метою створення теорії технічних систем (системотехніки) можна вважати оптимізацію системотехнічних комплексів при заданих обмеженнях на умови їхнього функціонування.

Складність об'єктів, які досліджуються у теорії систем обумовило використання як формальних, так і неформальних (евристичних) методів. Тому, теорія систем широко використовує як математичні методи, так і суб'єктивні думки кваліфікованих фахівців.

Основним методом вивчення технічних систем є системний підхід, який припускає вивчення об'єкта - систем із обліком, як внутрішніх процесів, так і взаємодії досліджуваного об'єкта з іншими. Дослідження, що базуються на системному підході, є системними дослідженнями.

Система - одне з розповсюджених, основних понять, що широко використовується в даний час при вивченні різних об'єктів і явищ навколишнього світу. До систем можна віднести: виробниче підприємство і живий організм, електронну обчислювальну машину і навчальний заклад, сукупність диференціальних рівнянь і підсилювач електричних сигналів, транспортний цех і цукровий завод. Системами є транспортна мережа міста і послідовність деяких дій, наприклад, по контролю якості продукції, що випускається, таблиця множення, сукупність проблем, що підлягають рішенню та ін.

Навіть простого перерахування цих об'єктів і явищ досить, щоб помітити, наскільки вони відрізняються один від одного. Серед них є об'єкти реальні й абстрактні, природні і штучні, різної фізичної природи (біологічні, технічні, соціальні) і різного цільового призначення.

Очевидно, найбільш загальним є визначення системи яке дано професором Месаровичем: системою називається відображення на непорожніх (абстрактних) множинах

$$S \subset x\{V_i : i \in I\}, \tag{2.1}$$

де x - символ прямого (декартового) добутку;

V_i - елемент системи з індексом i;

I - множини індексів.

Для кінцевої множини елементів відображення (2.1) можна переписати у вигляді

$$S \subset V_1 \times V_2 \times . \;\; \times V_n . \tag{2.2}$$

Виходячи з (2.2) стає очевидним визначення системи як множини елементів V_i, що знаходяться у взаємодії один з одним.

Для відкритих систем, що мають входи $X = x\{V_i : i \in I\}$, через які можуть надходити впливи, і виходи $y = x\{V_i : i \in I_y\}$ на яких можуть спостерігатися реакції на ці впливи, можна дати більш конкретне визначення:

$$S \subset X \times y\,. \qquad (2.3)$$

Систему (2.3) називають системою "вхід-вихід" або "чорною шухлядою". Останнє пов'язане з тим, що у визначенні (2.3) може не конкретизуватися внутрішня структура системи, а цікавить лише реакція на виході такої "чорної шухляди" на вплив, що надходить на його вхід. Така абстракція доречна у загальній теорії систем і при моделюванні цих систем, оскільки тут предметом вивчення є властивості, що спостерігаються, і їх взаємозв'язки, а не те, що вони (ці властивості та явища) насправді являють собою.

Близькими до поняття система є таке поняття, як: ***елемент системи*** - це найпростіша неподільна частина системи. Неподільність в даному випадку умовна, оскільки при бажанні це ділення можна необмежено здійснювати аж до мікросвіту. Однак неподільність тут означає, що подальше ділення не доцільне, бо руйнує властивості елемента або не дає додаткової інформації при вивченні конкретних властивостей і структури системи, що розглядається.

Елементи системи можуть знаходитися в різній взаємозалежності. Якщо вони незалежні один від одного, їх спільна (взаємна) невизначеність, що характеризується ентропією, є сумою окремих невизначеностей:

$$H_{V_1V_2} = H_{V_1} + H_{V_2}\,, \qquad (2.4)$$

де H_{V_1} - невизначеність одного елемента системи.

Загальна невизначеність системи з незалежних елементів

$$H_{\Sigma} = \sum_{i=1}^{n} H_{V_i}\,, \qquad (2.5)$$

Якщо ж елементи V_i залежні, то

$$H_{V_1V_2} = H_{V_1} + H_{V_2/V_1}\,, \qquad (2.6)$$

де H_{V_2/V_1} — умовна ентропія.

Очевидно, що $H_{V_2} \geq H_{V_2/V_1}$, тобто взаємна невизначеність залежних елементів є меншою, ніж незалежних. Отже, і невизначеність системи із залежними елементами менша, ніж з незалежними. За наявності взаємозв'язку між елементами система стає більш організованою, із більш упорядкованими відношеннями.

Технічна система (ТС) – створена людиною чи автоматом, реально існуючий пристрій, призначений для задоволення визначених життєвих потреб. До ТС можна віднести окремі машини й агрегати, апарати, прилади, ручні знаряддя праці, будинки, спорудження та ін. Тобто пристрої, що виконують визначену функцію (операцію) по перетворенню речовини, енергії й інформації.

Як видно з визначення, ТС може являти собою дуже широке поняття. Так, наприклад, до ТС можна віднести трактор і літак, борошномельний завод і лопату, ЕОМ і тракторний завод та ін. Незважаючи на істотні розходження, усім їм властиві деякі загальні ознаки, що дозволяють розглядати перераховані об'єкти як системи.

Слід зазначити, що система виступає як цілісний об'єкт, що представляє сукупність окремих частин. Електронна обчислювальна машина, наприклад, складається з окремих вузлів і блоків, складовою частиною живого організму є клітина, а система рівнянь містить у собі ті чи інші змінні.

З іншого боку, система - це об'єкт комплексний, окремі частини якого зв'язані і тісно взаємодіють між собою. Очевидно, при відсутності обміну інформацією між оперативним запам'ятовуючим пристроєм, пристроєм управління й інших блоків, електронну обчислювальну машину важко було б розглядати як систему для вирішення науково-технічних і задач управління. Функціонування таких систем спрямовано на досягнення визначених цілей. Але цілі формулюються і потім досягаються під впливом якої-небудь особи, чи групи осіб, що прагнуть задовольнити свої потреби. Тому, говорячи про цілеспрямовані системи, необхідно ввести у розгляд фігуру суб'єкту, розуміючи під суб'єктом яку-небудь особу, або групу осіб (рис. 2.1).

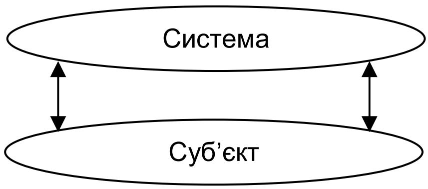


Рис. 2.1. Схема взаємодії система-суб'єкт

Будь-яка технічна система, як правило, має декілька критеріїв свого розвитку. Виходячи з того, що сукупність критеріїв розвитку для різних класів (видів) ТС у більшому ступені збігаються, можна виділити чотири основні групи критеріїв (рис. 2.2):

- функціональні критерії, що характеризують найважливіші показники ТС;
- технологічні критерії, що зв'язані тільки з можливістю та про-

стотою проектування і виготовленням ТС;

- економічні критерії, що визначають тільки економічну доцільність проектування й експлуатації ТС;
- антропологічні критерії, зв'язані з питаннями людського фактора.

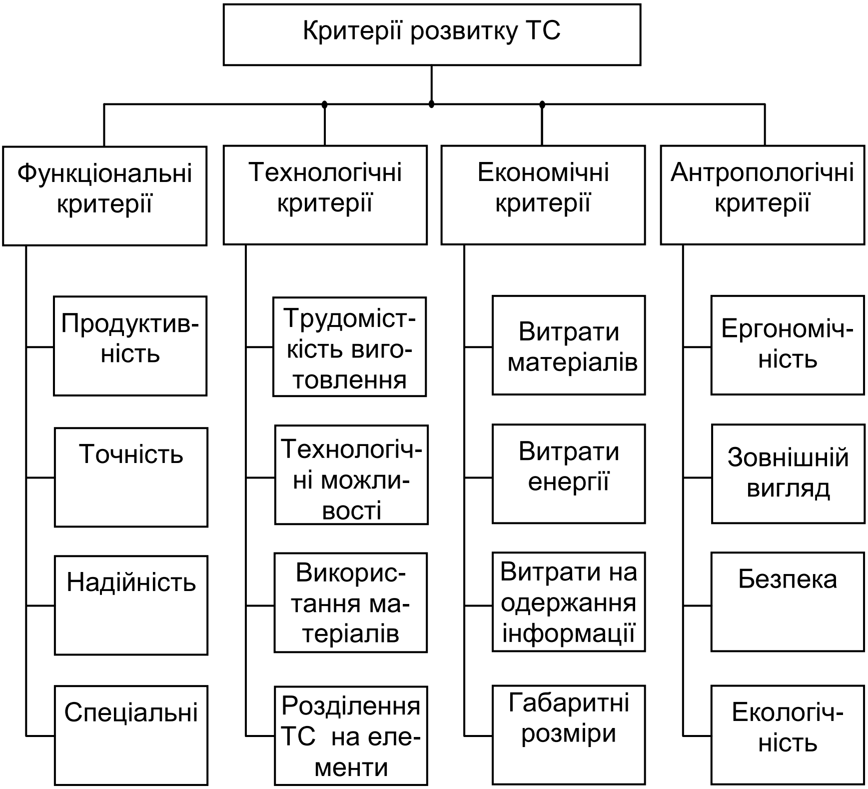


Рис. 2.2. Система критеріїв розвитку технічних систем

Сучасні складні технічні системи характеризуються великою кількістю показників для оцінки ефективності їхнього функціонування. Так, наприклад, сучасне підприємство характеризується великою різноманітністю продукції, що випускається, досить великим набором засобів виробництва, розгалуженими зв'язками, як між внутрішніми підрозділами (ділянка, поточна лінія виробництва, цех, лабораторія, відділи управління та ін.) виробництва, так і зовнішніми підприємствами (постачальники сировини, замовники готової продукції та ін.).

Виникає питання, як організувати діяльність цих об'єктів, щоб забезпечити якісне виконання виробничих, соціальних і інших завдань? З огляду на складний і відповідальний характер задач покладатися лише на досвід, інтуїцію, передбачення керівника досить ризиковано. Необхідно залучати наукові методи, які дозволяють уникнути, чи послабити наслідки цих суб'єктивних помилок. Такі наукові методи дає теорія систем.

Вивчення будь-якої системи починається з її опису. Опис системи - це модель, що відображає групу властивостей (чи навіть одну властивість) системи. В даний час при вивченні систем широке поширення одержали наступні види опису (рис. 2.3):

- морфологічний;
- функціональний;
- інформаційний;
- процесуальний.

Морфологічний опис системи використовується при дослідженні структури ТС. При такому описі важливо знайти всі зв'язки між елементами системи і визначити структурно-топологічні характеристики, що, у свою чергу, є основою формалізованого опису структури ТС. У цьому випадку система – це сукупність елементів і зв'язків між ними.

Функціональний опис виходить з того, що будь-яка система виконує деякі функції. Функціонування цілеспрямованої системи приводить до досягнення цілей. Основними поняттями функціонального опису є: стан, вплив, поводження (рух), рівновага, стійкість.

Стан - це сукупність існуючих властивостей (характеристик як кількісних, так і якісних) системи у розглянутий момент часу. Наприклад, стан людини можна охарактеризувати температурою тіла, тиском, частотою пульсу (кількісні характеристики), стомленістю, настроєм (якісні характеристики). Стан якої-небудь харчової машини можна охарактеризувати продуктивністю, числом обертів елементів, тиском (кількісні характеристики), смаком, запахом, коляром продукції та ін. (якісні характеристики).

Під ***впливом*** розуміється процес взаємодії між елементами системи, або між системою і навколишніми об'єктами. У першому випадку впливи є внутрішніми, у другому - зовнішніми.

Поводженням систем називається процес переходу системи з одного стану в інший.

Так, наприклад, процес зміни настрою людини, логіка його вчинків дозволяє говорити про його поведінку. Поводження ЕОМ характеризується послідовністю змін режимів роботи. У загалі, термін поводження використовується тоді, коли невідомі закономірності переходів. Якщо ж ці закономірності відомі, то говорять про рух системи.

Рис.2.3. Види опису систем

Рівновага - це здатність системи у відсутності зовнішніх впливів зберігати свій стан як бажано довго.

Стійкість - це здатність системи повертатися у стан рівноваги після того, як вона була виведена з цього стану зовнішніми впливами. Прикладом стійкої системи є звичайний маятник.

Функціональний опис дозволяє визначити ступінь досягнення мети функціонування. Як правило, цей ступінь намагаються оцінити кількісно. Функціонал, що описує процес функціонування (дії) системи по досягненню поставлених цілей, називається ***функціоналом ефективності***. Значення цього функціоналу наприкінці інтервалу функціонування системи оцінюється показником ефективності і його називають просто ефективністю системи.

Інформаційний опис визначає залежність властивостей системи від якості і кількості інформації. Відповідно до цих основних понять, при інформаційному описі системи, є поняття **–** ентропії, організації, кількості інформації, цінності інформації й інш.

Ентропія - це ступінь невпорядкованості (невизначеності) стану системи.

Організація (організованість, упорядкованість, негентропія) - це ступінь упорядкованості системи. Організованість системи досягається управлінням, що визначається також і інформаційними процесами.

Процесуальний опис системи характеризує "життєвий цикл" системи, тобто її розвиток (чи вгасання) у часі. Найважливішим поняттям при цьому описі є поняття етапу існування системи. Для будь-якої цілеспрямованої системи можна виділити наступні етапи:

- етап створення;
- етап використання;
- етап ліквідації.

У свою чергу етапи поділяються на фази та ін. Процесуальний опис дуже важливий на початкових етапах існування системи, коли необхідно передбачити (спрогнозувати) майбутнє системи, аж до ліквідації й утилізації (наприклад утилізація відходів АЕС).

Елементи системи можуть бути однотипні і різнотипні, більш того, можуть мати і різну фізичну природу (технічні засоби, люди). Якщо обчислювальний центр (ОЦ) укомплектований тільки однаковими ЕОМ, то елементи системи ОЦ можна вважати однотипними,

За допомогою ***зв'язку*** здійснюється взаємодія елементів системи між собою. Зв'язки відіграють дуже важливу роль, тому що забезпечують функціонування системи як єдиного цілого. Зв'язки характеризуються спрямованістю, призначенням, силою (вагою) і місцем прикладення.

Зв'язки призначені для передачі від елемента до елемента відповідної речовини, енергії або інформації. У залежності від призначення, зв'язки підрозділяються на:

- речовинні;
- енергетичні;
- інформаційні.

Прикладами речовинних і енергетичних зв'язків є нафто і газопроводи, лінії електропередач паливно-енергетична система країни.

За ознакою сили розрізняють сильні і слабкі зв'язки. За спрямованістю виділяють прямі і зворотні зв'язки. Нарешті, за місцем прикладення розрізняють внутрішні і зовнішні зв'язки.

Навколишнє середовище - це сукупність об'єктів, що не входять у систему, але, впливають на неї, або находяться під впливом з боку системи. Навколишнє середовище і система завжди знаходяться в єдності і взаємодії. Система, що не взаємодіє із навколишнім середовищем, називається ***закритою***. Відзначимо, що серед цілеспрямованих систем закриті не зустрічаються. Але у ряді випадків слабко взаємодіючі із середовищем системи можна віднести до закритих. На противагу закритим ***відкриті системи*** завжди взаємодіють із середовищем.

Розглянемо також таке важливе поняття теорії технічних систем, як машина. ***Машина*** - технічний пристрій, що виконує перетворення енергії, матеріалів і інформації з метою полегшення фізичної й розумової праці людини, підвищення її якості й продуктивності.

Існують наступні види машин.

1. ***Енергетичні машини*** - перетворюють енергію одного виду в енергію іншого виду. Ці машини бувають двох різновидів:

- двигуни, які перетворюють будь-який вид енергії в механічну (наприклад, електродвигуни перетворять електричну енергію, двигуни внутрішнього згоряння перетворюють енергію розширення газів при згорянні в циліндрі);
- генератори, які перетворюють механічну енергію в енергію іншого виду (наприклад, електрогенератор перетворить механічну енергію парової або гідравлічної турбіни в електричну).

2. ***Робочі машини*** – машини, які використовують механічну енергію для здійснення роботи із переміщення й перетворення матеріалів. Ці машини мають два різновиди:

- транспортні машини, які використають механічну енергію для зміни положення об'єкта (його координат);
- технологічні машини, що використають механічну енергію для перетворення форми, властивостей, розмірів і стану об'єкта.

3. ***Інформаційні машини*** - машини, що призначені для обробки й перетворення інформації. Вони підрозділяються на:

- математичні машини, що перетворять вхідну інформацію у математичну модель об'єкту, що досліджується;
- контрольно-управляючі машини, що перетворять вхідну інформацію (програму) у сигнали управління робочою або енергетичною машиною.

4. ***Кібернетичні машини*** - управляють робочими або енергетичними машинами, здатні змінювати програму своїх дій залежно від стану навколишнього середовища (тобто машини, що мають елементами штучного інтелекту).

Машинним агрегатом називається технічна система, що складається з однієї або декількох з'єднаних послідовно або паралельно машин і яка призначена для виконання яких-небудь заданих функцій. Звичайно до складу машинного агрегату входять: двигун, передавальний механізм і робоча або енергетична машина. Також до складу машинного агрегату включається управляюча або кібернетична машина. Передавальний механізм у машинному агрегаті необхідний для узгодження механічних характеристик двигуна із механічними характеристиками робочої або енергетичної машини.

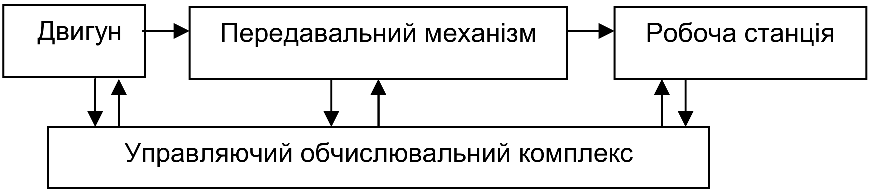


Рис. 2.4. Схема машинного агрегату

2.2. Методи опису структур технічних систем

Структура системи віддзеркалює ті властивості системи, що залишаються незмінними протягом усього проміжку часу функціонування системи чи досить тривалої його частини. Під структурою розуміється організація системи з ***N*** окремих елементів з їхніми взаємозв'язками ***A*** відповідно до функцій, що виконуються системою. Структура є поняттям морфологічного опису системи і містить у собі загальносистемні властивості (наявність елементів, існування зв'язків між ними і т.д.).

Структура технічної системи може бути описана різними методами.

Словесний метод опису структури технічної системи (на вербальному рівні) передбачає, що елементи системи та їх зв'язки описують словесно без елементів формалізації. Даний метод опису не дозволяє проводити математичні дослідження складних ТС.

Якщо зобразити систему як сукупність блоків, що здійснюють деякі функціональні перетворення, то отримаємо ***структурну схему***, що в узагальненому вигляді описує структуру системи. Крім функціональних блоків, у структурну схему можуть включатися логічні блоки, що дозволяють змінювати характер функціонування ТС у залежності від того, виконуються або ні деякі, заздалегідь задані умови. Під ***блоком*** розуміють функціонально-завершений і оформлений об'єкт у вигляді окремого цілісного пристрою.

Приклад спрощеної структурної схеми пристрою системи автоматичного вимикання механізму надано на рис. 2.5. Датчик часу здійснює вимір поточного часу. При перевищенні деякого, заздалегідь установленого, терміну спрацьовує реле, що відключає від мережі живлення механізм і датчик часу. Недолік даного опису - не можливо проводити глибокі математичні дослідження складних ТС.

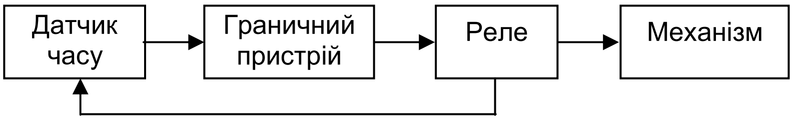


Рис. 2.5. Структурна схема пристрою

Структура ТС може бути задана у вигляді ***алгоритму роботи системи*** (логічної схеми алгоритму функціонування системи) Такі структурні схеми наочні і вміщають у себе інформацію про велике число структурних властивостей системи (рис. 2.6).

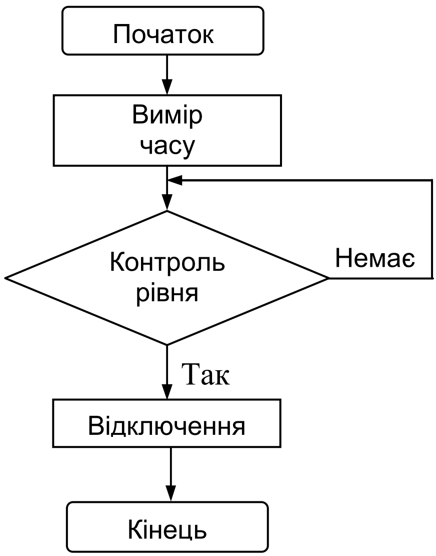


Рис. 2.6. Схема алгоритму роботи пристрою

Однак вони не дозволяють скористатися математичним апаратом при дослідженні структури систем. Тому існує ще один опис структури – метод представлення структури ТС у вигляді ***графа*** (аналітичний опис), тобто у формалізованому вигляді.

Принцип представлення системи у вигляді графа досить простий. Для цього елементи ***N*** системи нумеруються, зображуються вершинами і розташовуються у довільному порядку. Якщо між двома елементами існує односпрямований зв'язок, то відповідні вершини (вузли) з'єднуються орієнтованим ребром зі стрілкою. Якщо зв'язок двоспрямований, елементи з'єднуються неорієнтованим ребром (рис. 2.7).

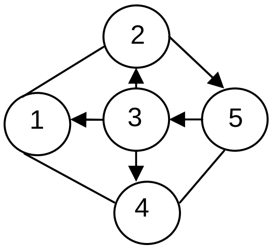


Рис. 2.7. Схема графа

Опис системи у вигляді графа дає уявлення про її структуру (взаємне розташування елементів, спрямованість зв'язків та ін.).

У той же час, граф не дає можливості вказати природу елементів (машини, механізми, люди) і характер зв'язків (речовинні, енергетичні, інформаційні). Графічне представлення системи є зручним із погляду наочності, однак, таке представлення не єдине.

Аналітично граф може бути представлений матрицею суміжності (зв'язності) це матриця ***A*** розміром $N \times N$, де ***N*** - число вершин графа, елементи якої a_{ij} приймають значення *1*, якщо з вершини ***i*** можна безпосередньо перейти у вершину ***j,*** і приймають значення *0* у протилежному випадку.

Розглянутий вище опис структури системи у вигляді графа дозволяє перейти безпосередньо до аналізу структурних властивостей. Цей аналіз, що називається структурно-топологічним, передбачає аналіз елементів і зв'язків, що утворять структуру, а також і визначення структурних характеристик системи.

Структурна надмірність є однією з найважливіших структурних характеристик і визначає перевищення загального числа зв'язків над мінімально необхідним.

***Ранг елементу*.** Дана характеристика дозволяє розподілити елементи у порядку їхньої значущості. Очевидно, чим більш значущій елемент, тим більше зв'язків він має з іншими елементами. У загальному випадку для визначення рангу елемента необхідно використовувати матрицю, що характеризує всі зв'язки між елементами.

Усе різноманіття структур складних технічних систем може бути розбите на два великих класи:

- клас часових структур;
- клас просторових структур.

Часові структури характеризують порядок функціонування системи у часі. Найбільше поширення на практиці знайшли так звані мережеві моделі. Так, наприклад, на етапі проектування система може бути представлена мережевим планом проектування, при експлуатації - мережевим графіком експлуатації та ін.

Просторові структури відповідають розташуванню елементів і зв'язків у просторі. Основними видами просторових структур є:

- лінійна структура, яку можна представити, наприклад, багатокаскадним підсилювачем електричних сигналів;
- кільцева структура, за допомогою якої можуть бути описані структури систем електрозв'язку, структури виробництва й ін.;
- радіальна структура за допомогою якої можуть бути описані деякі системи зв'язку і управління (наприклад, локальні автоматичні системи регулювання, що виконують функції підтримки технологічних параметрів у необхідних межах, які забезпечують, при заданих рівнях навантаження технологічного об'єкту, економічне ведення технологічного процесу і безпечну роботу основного і допоміжного устаткування АСУ ТП);
- ієрархічна структура - є найбільш широко розповсюдженою для ТС, а також для складних систем управління.

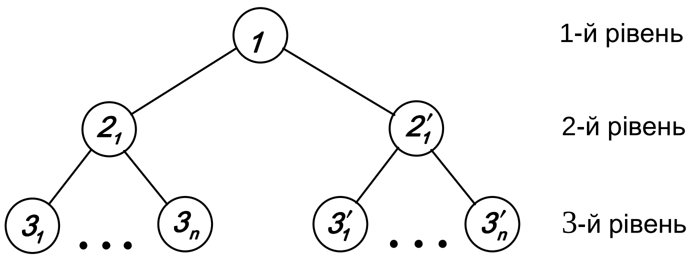


Рис. 2.8. Ієрархічна схема ТС

З використанням перерахованих типових структур можуть бути створені більш складні структури. Так, наприклад, об'єднання радіальної і кільцевої структур приведе до радіально-кільцевої структури та ін.

2.3. Класифікація технічних систем

Приведемо класифікацію технічних систем за тими ознаками, що особливо характерні для систем автоматизації виробничих процесів. Метою такої класифікації є виявлення основних ознак складних технічних систем. Відзначимо, що будь-яка класифікація за ознаками дуже умовна.

За кількістю елементів системи.

1. Малі - мають невелику кількість елементів у системі.

2. Великі - відрізняються наявністю великої кількості елементів, підсистем, компонентів, мають великий рівень значимих зв'язків між елементами і підсистемами.

Однак відзначимо, що в деяких літературних джерелах велика система (ВС) визначається як система, для актуалізації моделі якої з метою управління бракує матеріальних ресурсів (машинного часу, ємності пам'яті, інших матеріальних засобів моделювання). Це економічні, організаційно-управлінські, нейрофізіологічні, біологічні й т.п. системи. Елементами ВС є складні системи.

Для великих систем характерні специфічні властивості:

- ➢ різноманітність можливих виконуваних функцій, великі розміри системи за числом частин і елементів, а також входів і виходів, висока вартість;
- ➢ складна ієрархічна структура організації ВС, наявність функціональних підсистем;
- ➢ складність поводження ВС;
- ➢ наявність речовинних, енергетичних і інформаційних зв'язків (можливо, усіх трьох одночасно) між підсистемами ВС, а також зв'язків із зовнішнім середовищем і людиною, що визначають якість функціонування;
- ➢ цілеспрямованість і управляємість ВС, високий рівень автоматизації й ін.

За складністю зв'язків у структурі ТС.

1. Прості - мала кількість і неістотність зв'язків.

2. Складні **-** багаторівневість побудови системи.

За походженням.

1. Природні **-** виникли у природі без участі людини.

2. Штучні - виникли в результаті людської діяльності.

За способом управління.

1. Ручне - за участю людини.
2. Автоматизоване - за участю людини і засобів автоматизації.
3. Автоматичне - без участі людини.

За характером зв'язків із зовнішнім середовищем.

1. Відкриті - які взаємодіють із зовнішнім середовищем. Як правило всі штучні ТС є відкритими.
2. Закриті - які не взаємодіють із зовнішнім середовищем.

За видом структур ТС.

1. Лінійні.
2. Кільцеві.
3. Радіальні.
4. Ієрархічні.
5. Комбіновані.

За поводженням у часі.

1. Статичні.
2. Стаціонарні - реакція ТС на вхідний вплив не залежить від того, у якому проміжку часу здійснений цей вплив.
3. Динамічні - характеризуються у першу чергу наступним:

- система функціонує у реальному часі;
- із часом система під впливом зовнішніх або внутрішніх впливів змінює свій стан.

За ступенем організації ТС.

1. Добре організовані.
2. Погано організовані.
3. Що самоорганізуються (розвиваються у часі).

2.4. Статичні та динамічні технічні системи

Система ***S*** називається ***статичною*** тоді і тільки тоді, коли значення її вихідної величини ***y(t)*** у будь-який момент часу залежать виключно від поточного значення вхідного впливу ***x(t)*** і стану $z_0(t)$, із якого почалася еволюція системи. З використанням логічних операцій статична система визначається наступним виразом

$$(x,y)\in S \Leftrightarrow \exists Z_0, \forall t \;\; y(t)=K_t(z_0, x(t)),$$

що інтерпретується наступним чином: система, у якої визначені значення входів та виходів (***x*** і ***y***) буде тоді й тільки тоді статичною системою ***S***, коли існує початковий стан z_0, який належить до безлічі можливих початкових станів Z_0 і для всіх моментів часу ***t*** вихідна реакція $y(t)$ визначається початковим станом z_0 і вхідними впливами $x(t)$, котрі забезпечують відображення K, в цю вихідну реакцію.

Іншими словами, статична система S є безінерційною. В ній відсутні перехідні режими при дії впливів, що збурюють цю систему на вході. Не слід плутати зі статичним рівноважним станом інерційної або динамічної системи, яка знаходиться у стані спокою після перехідних процесів. Статичні системи є певною абстракцією реальних систем, яким притаманні інерційність та динамічні перехідні режими. Часто статичні системи є водночас і системами без пам'яті, тобто системами, у яких початковий стан z_0 однаковий і відповідно

$$K_t\,(z_0, x\,(t)) = K_t\,(x\,(t)).$$

Динамічною називається інерційна система, у якої визначені функції переходу станів $f\,(t)$ і вихідної реакції $g(t)$.

Стаціонарними динамічними називають клас динамічних систем, стан та структура яких не залежить від того, у який момент часу буде розглядатися вплив. Про них говорять, що ці системи інваріанті щодо часового зсуву:

$$\forall t' F^t\,(x\,(t,t')) = F^t\,(x\,(t'-t)),$$

тобто для кожного моменту часу t можна визначити оператор зсуву часу F^t такий, що реакція системи на вхідний вплив у момент часу t' залежить лише від різниці між часом його початку і поточним часом, а не від поточного часу, при цьому $t \le t \Leftrightarrow t - t \ge 0$.

Для стаціонарної системи $S \subset X \otimes Y$, де $x \in X, y \in Y$ впливи і реакції є стаціонарними, якщо

$$\begin{Bmatrix} (\forall t)\quad (X_t = F^t\,(X)), \\ (\forall t)\quad (Y_t = F^t\,(Y)). \end{Bmatrix}.$$

Важливою властивістю стаціонарних (інваріантних у часі) систем є те, що функцію переходу стану для будь-якого моменту часу можна одержати як результат застосування оператора зсуву до початкової реакції системи.

Адекватним описом математичної моделі динамічної системи є диференційне рівняння

$$\frac{dx\,(t)}{dt} = F\,(x\,(t), t, u\,(t)),\; x\,(0) = x_0,$$

де $x(t) \in X$ — множина станів системи;

$u\,(t) \in U$ — множина збурюючих впливів.

Перша похідна $dx(t)/dt$ яка є не чим іншим, як швидкістю зміни станів системи, може дорівнювати нулю, що відповідає стану спокою системи, вона може дорівнювати від'ємній або позитивній величині, бо це рівняння може мати змінну праву частину:

$$\frac{dx(t)}{dt} = Fx(t), x(0) = x_0. \tag{2.7}$$

При $F > 0$ система поводить себе збуджено, не стало, при $F < 0$ вона повертається у стан спокою і її поведінка стала:

$$x(t) = x_0 e^{-Ft}.$$

Очевидно, при $F \equiv \infty$ динамічна система стає статичною, при відході від ∞ до 0 інерційність зростає, при $F = 0$ система стає нерухомою, у стані спокою.

Очевидно, вся теорія диференційних рівнянь може бути прикладена до теорії динамічних систем і цей математичний апарат є досить продуктивним, добре опрацьованим. Так диференційне рівняння першого порядку типу (2.7) з вільним членом у вигляді збурюючого зваженого білого гаусового шуму $\xi(t)$ носить назву ***стохастичного*** і рівняння стану і записується так:

$$\xi(t)\frac{dx(t)}{dt} = F(x(t),t) + G(x(t),t)\,\xi(t), x(0). \tag{2.8)]}$$

На рис. 2.9 наведено структурну схему формуючого фільтра, процес на виході якого відповідає стану динамічної системи, аналогічної (2.8), де ***G(t)*** - збурюючий коефіцієнт, який впливає на величину дисперсії стану ***x(t)*** даної динамічної системи.

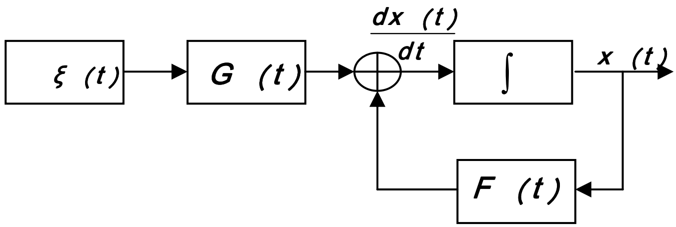


Рис. 2.9. Структурна схема формуючого фільтра

2.5. Прості та складні технічні системи

Під ***складною системою*** розуміється система, що складається із великого числа взаємозалежних і взаємодіючих між собою елементів різної природи, що має розгалужену структуру, яка функціонує в умовах невизначеності, у моделі якої недостатньо інформації для ефективного керування цією системою і яка здійснює цілеспрямований вибір свого поводження. З цього визначення випливають наступні основні ознаки складної системи:

- багатомірність системи, що обумовлена великим числом елементів;
- розгалуженість структури і велике число взаємозв'язків між елементами;
- різноманіття природи елементів системи (елементами однієї системи можуть бути люди, автомати, машини, агрегати й ін.);
- відносна автономність функціонування елементів системи, тобто здатність елементу складної системи автономно виконувати своє коло задач, незалежно від того, функціонують інші елементи цієї системи чи ні;
- слабка передбачуваність поводження, що виражається у тому, що як завгодно точне знання поводження системи у поточний момент часу не дозволяє точно передбачити її поводження у наступні моменти чи інтервали часу;
- цілеспрямованість, тобто здатність вибирати поводження, яке приводить до досягнення поставленої мети.

На відміну від складної системи ***проста система*** складається з невеликого числа елементів, має невелике число внутрішніх зв'язків і характеризується передбаченим поводженням.

Слід зазначити, що чіткої границі між простими і складними системами провести не можна. Та сама система в залежності від повноти опису, цільового призначення і задач дослідження може бути віднесена як до простих, так і до складних систем.

Необхідно відзначити особливості складних систем, які умовно можна розбити на чотири групи (рис. 2.10):

- особливості складу, структури і властивостей складних систем;
- особливості взаємодії складних систем одна з одною і з зовнішнім середовищем;
- особливості функціонування складних систем;
- особливості управління складними системами.

Особливості складу, структури і властивостей складних систем:

- складні системи складаються із великої кількості елементів;
- елементи складних систем відрізняються різноманіттям природи, різнорідністю виконуваних ними функцій;

- складні системи відрізняються розгалуженою, найчастіше ієрархічною структурою;

елементи, що входять у систему, відрізняються високим ступенем зв'язності і взаємодії один з одним;

1.1. Велика кількість елементів.
1.2. Різноманіття природи елементів, різнорідність виконуваних ними функцій.
1.3. Структурованість системи.
1.4. Високий ступінь зв'язаності елементів.
1.5. Емерджентність.

2.1. Різноманіття форм і видів зв'язку.
2.2. Велика кількість каналів взаємодії.
2.3. Випадковий характер взаємодій.

1. Особливості складу, структури і властивостей складних систем

2. Особливості взаємодії складних систем між собою і з зовнішнім середовищем

Особливості складних систем

3. Особливості функціонування складних систем

4. Особливості управління складними системами

3.1. Унікальність.
3.2. Слабка передбачуваність.
3.3. Негентропійність чи цілеспрямованість.
3.4. Різноманіття цілей.
3.5. Змінюваність у часі.
3.6. Різнорідність циркулюючої інформації.

4.1. Ієрархічність управління.
4.2. Самоорганізація.
4.3. Різноманіття і велике число управляючих впливів.

Рис. 2.10. Класифікація особливостей складних систем

- вивчення об'єкта у цілому, як системи, буває дуже складним, тому об'єкт розбивають (зі збереженням зв'язків) на кінцеве число більш малих підсистем, таких, більш простих і зручних для безпосереднього вивчення;
- складна система відрізняється наявністю ***емерджентних властивостей,*** тобто таких властивостей системи, що не характерні для окремих її складових елементів, а породжуються системою у цілому за рахунок наявності та взаємодії цих елементів між собою.

Особливості взаємодії складних систем одна з одною і з зовнішнім середовищем:

- складні системи відрізняються різноманіттям можливих форм і видів зв'язків;
- взаємодія між частинами системи і середовищем може відбувається по великому числу каналів;
- взаємодія середовища і системи носить випадковий характер.

Особливості функціонування складних систем:

- унікальність - складна система не має повних аналогів поводження, ця властивість є зовнішньою стосовно системи;
- слабко передбачуваність - ніяке, як завгодно докладне знання морфології і функцій елементів (підсистем) не дозволяє однозначно визначити функції об'єкта;
- негентропійність (цілеспрямованість): система спроможна (у визначених межах) управляти своєю ентропією (зменшувати її, зберігати, гальмувати збільшення) при випадковому і несприятливому впливі середовища або здатна здійснювати поводження, що переслідує досягнення визначеної мети;
- для складних систем характерне різноманіття цілей і складність функцій, що виконуються системою і спрямованих на досягнення заданої мети функціонування;
- реальні складні системи функціонують в умовах великої кількості випадкових факторів, джерелами яких є вплив зовнішнього середовища, помилки, шуми і відхилення різних величин, що виникають всередині системи, тобто стан складної системи неможливо докладно і точно описати;
- сучасний рівень знань недостатній для проникнення у суть зв'язків складної системи (мозок людини);
- функціонування ТС характеризується розгалуженістю інформаційної мережі, різноманітністю і високою інтенсивністю потоків інформації в системі.

Особливості управління складними системами:

- у складних системах завжди має місце управління, яке найчастіше має ієрархічну структуру з більшим чи меншим ступенем централізації;
- багатьом складним системам властиві риси самоорганізації, тобто ТС здатна на підставі оцінки впливів зовнішнього середовища, шляхом послідовної зміни своїх властивостей, прийти до деякого стійкого стану, коли вплив зовнішнього середовища проявляється у припустимих межах;
- різноманіття і велике число управляючих впливів.

У процесі вивчення і дослідження складних технічних систем були виявлені закономірності, що дозволяють глибше зрозуміти побудову і функціонування системи. Розглянемо основні, найбільш фундаментальні з них.

Закономірність цілісності (емерджентності) - відбиває діалектику взаємодії частини і цілого в системі і формулюється в такий спосіб: система як ціле має особливі, системні властивості, яких немає в елементах, її складових. Суть цієї закономірності полягає у тому, що:

- властивості системи, як цілого, не є простою арифметичною сумою властивостей її елементів;
- властивості системи залежать від властивостей елементів;
- системні властивості об'єкта формуються при взаємодії елементів шляхом накопичення, посилення і прояви одних властивостей елементів із одночасним ослабленням інших.

Нехай система S складається із сукупності елементів, e_i тобто $S = \{e_1, e_2, \ldots, e_i, \ldots, e_n\}$, $e_i \in E$. Введемо функцію $F(e_i) = Q_i$, що характеризує властивості одного елементу e_i. Очевидно, що $F(S) = Q_S$ характеризує властивості системи. Закономірність цілісності затверджує, що:

$$Q_S \neq \sum_{e_i \in E}^{n} Q_i ,$$

або

$$F(S) \neq \sum_{e_i \in E}^{n} F(e_i) .$$

Таким чином, закономірність цілісності визначає, що при об'єднанні елементів у систему відбувається якісний стрибок - перехід кількості в якість. Наприклад, окремо взята жива клітина має властивості обміну, розподілу та ін. У сукупності ж клітина утворить живий організм властивості якого якісно відрізняються від властивостей окремої клітини.

Закономірність ієрархічності характеризує побудову світу за принципом супідлеглості.

Розглядаючи який-небудь об'єкт як систему, можна виділити у ньому підсистеми чи системи більш низького порядку. У той же самий час розглянута система є підсистемою іншої більш складної системи, тобто системи більш високого порядку (надсистеми). Прояву закономірності ієрархічній упорядкованості можна спостерігати як у цілеспрямованих, так і у нецілеспрямованих системах. Цілеспрямована система, наприклад на підприємстві, цех містить підсистеми - верстати та одночасно є підсистемою заводу. Нецілеспрямована система, наприклад, Сонячна, як об'єкт вивчення, містить підсистеми - планети, але у той же час вона є підсистемою більш високого порядку - Галактики.

Відзначимо деякі важливі особливості ієрархічної упорядкованості:

- кожен рівень ієрархії має складні взаємини між вищестоящим і нижче лежачим рівнями, стосовно нижче лежачих рівнів вищестоящий рівень ієрархії виступає як система, що функціонує автономно;
- об'єднання компонентів на кожнім рівні ієрархії приводить до появи нових властивостей не тільки даного рівня ієрархії, але й у підлеглих рівнів;
- закономірність ієрархічної упорядкованості при дослідженні систем дозволяє розбити «велику» проблему на більш «дрібні», що легше піддаються вивченню і розв’язанню.

Закономірність додатковості говорить: складна система у взаємодії із середовищем може виявляти різні властивості у різних ситуаціях (несумісні ні з однією із них).

Класичним прикладом, що пояснює закономірність додатковості, є характер поводження електрона. При одних видах взаємодії електрон поводиться як частинка, при інших як хвиля. Важко знайти такий вид взаємодії, при якому корпускулярні і хвильові властивості електрона виявлялися б одночасно. Фізичний принцип додатковості був сформульований Н. Бором.

Закономірність додатковості дає концептуальну основу при дослідженні систем. Вона змушує шукати у різних ситуаціях прояв відповідних властивостей системи, виявляючи нові, раніше не відомі. Власне кажучи так і відбувається при створенні нових складних технологічних комплексів і технічних об'єктів. Так відбувається і при пошуку нових форм господарювання, управління і інш.

Закон необхідної різноманітності також є одним з фундаментальних законів теорії систем. Будь-яка цілеспрямована складна система служить для досягнення визначених цілей. Досягнення цілей здійснюється вирішенням цілої низки задач, сукупність яких має визначену різноманітність. Тобто, для досягнення цілей, система повинна мати різноманітність способів рішення задач.

Закон необхідної різноманітності стверджує, що різноманітність λ_c можливостей системи, призначеної для рішення визначеного кола задач, повинна бути більше (чи, принаймні, дорівнювати) різноманітності $\lambda_з$ необхідної для рішення задач систем. У протилежному випадку система не зможе досягти поставлених цілей, тобто повинна виконуватися умова: $F(S) = Q_S$, $\lambda_c \geq \lambda_з$.

Використання цього закону допомагає побачити причини недоліків системи і намітити шляхи їхнього усунення.

Закономірність вибору поводження. Ця закономірність полягає у наступному: складні системи мають здатність вибору поводження, тобто визначеної реакції на впливи з боку середовища у залежності від внутрішніх критеріїв цілеспрямованості. Таким чином, система має у наявності кілька альтернативних рішень. При цьому ніякі як завгодно глибокі попередні знання про побудову системи і законів її функціонування не дозволяють однозначно визначати цей вибір.

Закономірність вибору поводження особливо яскраво виявляється у людино-машинних (ергатичних) системах, до яких відносяться, наприклад, АСУ різного призначення.

Закономірність історичності систем. Будь-яка система не залишається незмінною у часі. Вона розвивається, змінюється, старіє. У цьому і полягає сутність закономірності історичності. Використання цієї закономірності призвело до необхідності процесуального опису системи, де, як вже згадувалося вище, вводяться поняття «життєвий цикл», «етапи розвитку» і т.д. Урахування закономірності історичності дуже важливе при створенні нових технічних систем, системотехнічних комплексів. Вже на стадії проектування повинні враховуватися не тільки питання створення і розвитку системи, але і питання того, коли і як її ліквідувати, передбачення механізмів ліквідації, утилізації залишків системи і т.д.

2.6. Властивості технічних систем

Розглянемо загальні властивості, які притаманні усім технічним системам, це:

- цілісність;
- причинність;
- управляємість;
- спостережуваність;
- здатність систем до ідентифікації;
- адаптованість;
- стійкість;
- зв'язність;
- складність.

Цілісність - є однією із найхарактерніших властивостей усіх технічних систем, яка проявляється у виникненні нових інтегральних якостей, не властивих для утворюючих систему компонентів. Властивості цілісності проявляються у системі з двох основних сторін:

- властивості системи як цілого не зводяться до суми властивостей її елементів або частин;
- властивості системи як цілого залежать від властивостей її елементів та частин, модифікація однієї частини викликає модифікації в усіх інших частинах і у всій ТС у цілому.

Суттєвим проявом властивостей цілісності є нові взаємовідносини системи як цілого з довкіллям, які відмінні від взаємодії окремих елементів ТС з цим довкіллям. Властивість цілісності пов'язана з метою, для виконання якої створена система.

Системи можуть мати властивість альтернативну цілісності - фізичну адитивність (незалежність елементів). Очевидно, що у випадку, коли всі елементи стають незалежними, то говорити про систему не має сенсу. Таким чином, будь-яка система знаходиться між двома крайніми станами: абсолютною цілісністю, що досягається при максимальному зв'язку між елементами, і абсолютною адитивністю, коли ці зв'язки відсутні.

Причинність (каузальність) - властивість систем, що визначає залежність вихідної реакції у будь-який момент часу виключно від вхідних впливів. Тобто властивість системи реагувати лише на існуючі, а не на майбутні впливи.

Управляємість - визначається як умова можливості переведення системи з одного стану в інший за заданий час або при виконанні заданих обмежень (за один крок, з мінімальною витратою енергії та ін.). У якості впливу, що управляє ***u(t)*** *за*звичай виступає сигнал, що впливає на регулятор, чим власне і досягається бажаний кінцевий стан системи. Разом із тим, впливи, що управляють можуть бути сформовані й особою, що приймає рішення (ОПР). У цьому випадку говорять про ергатичне чи ситуаційне управління. Якщо вплив, що управляє формується у відповідному приладі, то говорять про автоматичне управління. При цьому багато залежить від того, з якого початкового стану $x(0)$ і в який кінцевий стан $x(t) = x_0 + \Delta x(t)$ переводиться система, і яким чином завдяки оптимальному управлінню $\Delta u(t)$ досягається кінцева мета.

Спостережуваність - характеризує пряму і непряму можливість вимірів параметрів, що беруть участь у формуванні управляючих впливів. Іншими словами, спостережуваність досягається, коли до управляючого органу надходить необхідна інформація про стан технічної системи.

Для спостерігача, який вимірює відповідні сигнали чи стан системи, інформація, що його цікавить, може надходити у вигляді неперервної $y(t)$ або дискретної $y(к)$ ($к = 1, 2, 3,..., n$) функцій. Зокрема ця інформація може мати разовий вимір y_0. З величиною, що вимірюється, або функцією $x(t)$ інформація, яка спостерігається може бути зв'язана лінійно

$$y(t) = H(t)\,x(t) + \xi(t),$$

або нелінійно

$$y(t) = h(x(t), t) + \xi(t),$$

для дискретного лінійного випадку

$$y(k) = H(k)\,x(k) + \xi(k),$$

для дискретного нелінійного випадку

$$y(k) = h(x(k), k) + \xi(k)$$

де $H(t)$, $h(\bullet)$ - функції спостереження, які враховують особливості засобу вимірювання;

$\xi(t)$, $\xi(к)$ - похибки спостереження.

Функції $y(t)$ або $y(к)$ можуть набувати як скалярного так і векторного вигляду, при цьому умови спостережуваності, як і умови управляємості пов'язані із відповідністю розміру координат системи і спостерігача.

Результати спостереження (вимірювання) використовуються в АСУ, де ОПР на основі цих даних приймає рішення про стан системи. У цьому разі результати вимірювання слід оцінити і дати їм статистичну оцінку.

Здатністю ***ідентифікуватись*** називається властивість системи, що характеризує можливість визначення параметрів системи за результатами спостережень. Для задач ідентифікації використовують різноманітні статистичні методи: точкові або інтервальні методи оцінки параметрів, стохастичної апроксимації, регресивні, прогнозу та ін.

Успішність ідентифікації багато у чому залежить від того, яким чином змінюються параметри $\{a\}$, що ідентифікуються у системі $S(x)$. Для системи, яка описується рівнянням стану

$$\frac{dx}{dt} = f(x,u,a,t),$$

І рівнянням спостереження

$$y = h(x,u,a,t),$$

найбільш придатними необхідними умовами для ідентифікації є

$$\frac{da(t)}{dt} = 0,$$

що відповідають випадку системи із постійними параметрами, які належить ідентифікувати.

Самим простим і поширеним випадком є такий, коли задача ідентифікації вирішується разом із оцінкою стану системи $\hat{x}(t)$. Разом з оцінкою $\hat{x}(t)$ за допомогою аналогічних процедур можна оцінювати і $\hat{a}(t)$. Об'єднавши обидві оцінки $\hat{x}(t)$ і $\hat{a}(t)$, одержимо вектор

$$\hat{\hat{x}}(t) = \left\{ \frac{\hat{x}_i(t)}{\hat{a}_i(t)} \right\}, i = \overline{1,n},$$

який і належить оцінювати. Для одержання такої оцінки треба, щоб відповідно до вектора $\hat{\hat{x}}$ виконувались умови спостереження.

У задачах адаптації часто використовують рівняння Уідроу – Хоффа:

$$\hat{w}(k) = \hat{w}(k-1) + z(k)\,(y_{оп}(k) - H(k)\,\hat{w}(k-1))\,H(k), \quad (2.9)$$

де $y_{оп}(к)$ - опорний, еталонний, бажаний сигнал;

$\hat{w}(k)$ - ваговий коефіцієнт.

Здатність системи ***адаптуватись*** - властивість, що визначає спроможність забезпечувати необхідний режим функціонування в умовах невизначеності по відношенню до зовнішніх впливів. Здатність адаптуватись часто інтерпретується як властивість самоорганізації системи. Адаптивна система повинна бути управляємою, спостережуваною і здатною ідентифікуватись по відношенню до самої себе, а також спостережуваною і здатною ідентифікуватись по відношенню до зовнішніх впливів.

Існує два види адаптивних систем. Адаптивність систем першого виду досягається за рахунок такого вибору внутрішніх станів та режимів окремих елементів, що при різноманітних непередбачених, небажаних впливах або випадкових, невизначених модифікаціях якихось характеристик забезпечується необхідне цілеспрямоване функціонування даної системи. При цьому структурні модифікації не припускаються. Така система носить назву ***гомеостатичної адаптивної системи***. Якщо ж по відношенню до різних впливів чи модифікацій у системі припустима наявність її структурних модифікацій, то така система називається ***морфогенетичною адаптивною системою***.

Можливі адаптивні режими у системі передбачаються на етапі її створення, тобто система повинна створюватись як адаптивна. У даному випадку можуть бути використані не тільки одноконтурні засоби адаптації, але й багатоконтурні, коли адаптивною виконується сама підсистема адаптації.

На рис. 2.11 показано адаптивну процедуру оцінки вагових коефіцієнтів $\hat{w}\ (k-1)\ , i=1,2,3,\ldots,n,$ число яких може бути різним у залежності від розмірності сигналів на вході $H(к)$. Але вихідний сигнал $y = H\ (k)\ \hat{w}\ (k-1)$- скалярний. У цій процедурі є два контури управління. Контур I - управління "вперед" (за принципом Понселє), контур II - управління за відхиленням від опорного сигналу $y_{оп}$ по нев'язці $v(k)$ (управління за принципом Уайта). Процедура виконується стало при $z(k)$, що відповідає вказаним вище обмеженням і набуває відповідно до критерію МСКВ (мінімального середньоквадратичного відхилення) значень, що задаються еталонним сигналом $y_{оп}(к)$, МСКВ $(y - y_{оп}\ (k)\ )^2 \to 0$. Такі алгоритми використовуються в адаптивних антенах, наприклад в системах GSM-900, антени цих систем носять назву «інтелектуальних».

Здатність адаптуватися в систему закладають у залежності від очікуваних впливів або модифікацій якихось характеристик. Якщо ці впливи або модифікації виходять за рамки застосованих програм, то така система не може вважатися адаптивною. Тобто, коли говорять про здатність адаптуватися, то необхідно уточняти, до якого класу впливів або модифікацій ця здатність належить. Так, система (2.9), відображена на рис. 2.11, адаптивна щодо сигналів (завад), які не збігаються з еталонним сигналом $y_{оп}(к)$, а також щодо деяких відхилень у поширенні радіохвиль. Але вона не може вважатись адаптивною відносно швидкості передачі інформаційних сигналів чи перепадів у енергоспоживанні.

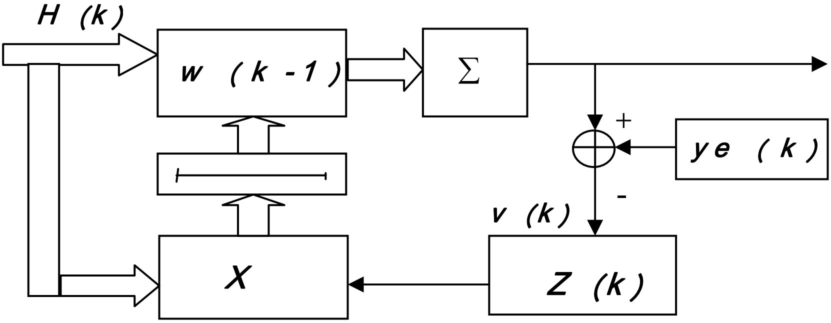


Рис. 2.11. Адаптивна процедура Уідроу-Хоффа

Стійкість є однією із найважливіших властивостей системи. Для простих систем характерні пасивні форми стійкості, вони пов'язані із такими властивостями, як міцність, збалансованість, гомеостатизм (повернення в рівноважний стан при виході з нього). Для складних організаційно-технічних систем характерні активні форми стійкості: надійність, живучість, завадо захищеність, які слід розглядати з точки зору вразливості системи під дією зовнішніх впливів. Таким впливам можуть піддаватися як окремий елемент (група елементів), так і відповідні зв'язки між елементами системи.

Виходячи з причинно-наслідкового аспекту зв'язків, стійкість трактують як властивість малих модифікацій причин викликати відповідно малі модифікації наслідків. Або система вважається стійкою, якщо малі вхідні впливи викликають малі вихідні її реакції. З даного визначення витікає, що стійкість характерна для динамічних систем.

Для стійких систем характерна наявність рівноважних станів, до яких система переходить внаслідок дії зовнішніх впливів. Ці стани характеризують можливості складних систем розв'язувати задачі, що стоять перед ними після надходження зовнішніх впливів. До таких рівноважних станів відносяться:

- ентропійний, до нього система переходить внаслідок витрат наявного запасу ресурсів чи руйнування своєї структури;
- гомеостатичний, припускає збереження структури системи за будь-яких зовнішніх впливів;

- морфогенетичний, пов'язаний із перебудовою структури системи під час будь-яких впливів для збереження необхідних властивостей.

Інтерпретація поняття стійкості залежить від класу системи. Так, для лінійних систем, що описуються звичайними диференційними рівняннями, частіше всього використовують поняття стійкості і асимптотичної стійкості за Ляпуновим. Наприклад, для системи, що описується диференційними рівняннями

$$x_i\,(t) = f_i\,(t, x_1, \ldots, x_n\,)\ ,\ \ i = \overline{1,n},\ \ x_i\,(t_0\,) = x_{i0}\,,\ \ \frac{dx}{dt} = x\ \ (t)\ ,$$

незбурений рух називається стійким за Ляпуновим по відношенню до змінних y_i що визначають відхилення (варіації) величини x_i, якщо при будь-якому довільно заданому позитивному числі ϵ, яким би малим воно не було, можна вибрати інше таке позитивне число δ (ϵ), що при всяких збурюваннях y_{i0}, які задовольняють умові

$$\sum_{i=1}^{n} y^2{}_{i0} \le \delta\,,$$

і при будь-якому $t > t_0$ буде виконуватися нерівність

$$\sum_{i=1}^{n} x^2{}_i \le \epsilon\,.$$

Якщо ж незбурений рух стійкий за Ляпуновим і при цьому будь-який збурений рух при достатньо малих початкових збурюваннях наближається до незбуреного руху, тобто

$$\lim_{t \to \infty} x_i\,(t) = 0\,,$$

то збурений рух називається ***асимптотично стійким*** за Ляпуновим.

Стійкість за Ляпуновим є, взагалі кажучи, локальною стійкістю (стійкістю у вузькому розумінні). У літературі часто використовуються поняття стійкості у широкому розумінні. Так, стан рівноваги системи називають ***стійким у широкому розумінні*** тоді, коли безліч можливих у даній системі відхилень цілком належить області притягання, тобто множині початкових станів (відхилень), при яких забезпечується стійкість за Ляпуновим.

У свою чергу, якщо область притягання охоплює увесь простір станів, то відповідну стійкість називають ***стійкістю у цілому***.

Стосовно до нелінійних систем використовують поняття абсолютної стійкості. Так, нелінійна система виду

$$\left\{\begin{array}{l} x = Ax + b\xi \\ \xi = j\ (\delta) \\ \sigma = C^T x \end{array}\right\},$$

де x - n-мірний вектор стану;

A - постійна ($n \times n$) матриця;

C - постійний n-мірний вектор;

$\varphi(\sigma)$- деяка нелінійна функція,

називається абсолютно стійкою на заданому класі функцій $\varphi(\bullet)$ у тому випадку, коли її положення рівноваги стійке в цілому при будь-якій нелінійній функції $\varphi(\bullet)$ заданого класу.

У реальних умовах функціонування багатьох систем здійснюється при дії великого числа різноманітних чинників, які є непередбаченими, випадковими, слабко контрольованими, взаємодія із якими має складний характер. Наприклад, стосовно системи управління та зв'язку до числа таких чинників відносять радіоелектронний вплив інших систем, електромагнітну сумісність, несправності та відмови радіоелектронних засобів, відхилення їх конструктивних параметрів, зміни стану зовнішнього середовища та ін. Однією із важливих причин неповноти інформації є запізнення, викликане обмеженістю часу, необхідного для проведення спостережень і обробки їх результатів.

Для визначення стійкості такого виду систем використовується імовірнісний підхід. У багатьох випадках, наприклад для систем, що задаються стохастичними диференційними рівняннями, питання про стійкість деякого руху, відповідно до теорії стійкості, може бути зведене до дослідження стійкості тривіального розв'язку рівняння (2.10).

$$dx(t) = f(t, x(t)), dt + b(t, x(t)), d\xi, t \geq 0, \qquad (2.10)$$

$$x(0) = x_0,\ x \in R_n,\ \xi \in R_m,$$

де x_o — детермінований або випадковий вектор початкових умов;

$f(t, x) \in R_n$ - вектор-функція, що задана;

b(t, x(t)) - матриця вимірності ($n \times m$), що задана;

ξ **-** стандартний векторний вінерівський процес.

Під ***стійкістю тривіального рішення*** рівняння (2.10) розуміється його властивість мало змінюватися при малій варіації початкових

кових умов. У залежності від конкретного розуміння виразу "мала варіація рішення" можливі різноманітні визначення стійкості.

Зв'язність (або структурна зв'язність) системи є однією з найбільш істотних її якісних характеристик. Дійсно, із зникненням структурної зв'язності зникає і сама система, вона розпадається на взаємонезалежні елементи. Математичний опис зв'язності може бути здійснений із застосуванням різноманітних підходів, причому найбільш вдалі з них побудовані на основі теорії графів та алгебричної топології.

Суть дослідження зв'язності полягає у тому, щоб усвідомити і виявити ті математичні конструкції, які описують характер зв'язку між окремими елементами системи. Розглянемо деяку систему ***S****,* у якій можна виділити ***n*** різних елементів. Тоді її структуру можна уявити у вигляді деякого графа, у якому ***n*** вершин зображають ***n*** елементів системи ***S****,* а дуга, що з'єднує елементи ***і*** та ***j,*** показує, що ці два елементи знаходяться у деякому співвідношенні між собою.

Однак, застосування теоретико - графових засобів, особливо для візуального аналізу, обмежено труднощами теоретичного і аналітичного характеру, особливо якщо враховується структура самих компонентів зв'язку. Практика вирішення інженерних завдань показала, що більш продуктивним для аналізу зв'язності є підхід, що базується на топологічних ідеях.

***Складність*,** як властивість системи – має два тлумачення. Перше, чисто суб'єктивне, характеризує відношення спостерігача до об'єкту, що спостерігається, коли один і той же об'єкт може різними спостерігачами сприйматися як простий чи достатньо складний. Друге тлумачення - об'єктивна характеристика, не зв'язана зі спостерігачем. Відносно самих систем об'єктивна складність виявляється по-різному. Виділяють такі сторони складності:

- структурну;
- динамічну;
- обчислювальну.

Очевидно, будь-яка система може бути складною, із однієї сторони, та простою, з іншої, або ж ця складність може виявлятися в декількох відношеннях.

Структурна складність визначається властивостями зв'язків між елементами системи. Вона може характеризувати ієрархічну структуру, схему зв'язності, різноманітність елементів, рівень або силу взаємодії між елементами системи.

Структурна складність припускає наявність ієрархії елементів системи. Така ієрархія по відношенню до систем управління встановлюється не тільки відповідною підпорядкованістю, але й тим, що існують первинні і вторинні мережі, різні рівні перетворення сигналів.

Структурна складність, віднесена до схеми зв'язності, визначається способом, яким елементи системи об'єднуються в єдине ціле. Ця складність може бути визначена чисто геометрично, через вимірність зв'язків. Разом із тим, вона може бути визначена також і з алгебричних позицій. Так, задання системи лінійним диференційним рівнянням

$$\vec{x} = (x_1,...,x_N),\ i = \overline{1,N},\ \frac{d\vec{x}}{dt} = F(\vec{x},\vec{x}_0) = c,$$

припускає наявність у матриці F різноманітних недіагональних елементів f_{ij}, що визначають зв'язки між окремими компонентами вектора $\vec{x} = (x_1,...,x_N)$, $i = \overline{1,N}$. Таким чином, складність у даному випадку визначається не тільки вимірністю N, але ще у більшій мірі наявністю міжкомпонентних зв'язків, що визначені за рахунок f_{ij}. При цьому може виявитися, що навіть система великої вимірності насправді виявляється достатньо простою, наприклад у випадку, якщо матриця F діагональна. Система при цьому подається набором взаємонезв'язаних елементів, і, по суті, не може бути системою.

Структурна складність є однією із властивостей, за допомогою яких система набуває здатності перетворювати інформацію, що до неї надходить. Ця складність пов'язана з ***принципом необхідної різноманітності Ешбі***: різноманітність може бути усунена або зруйнована тільки різноманітністю. Наприклад, різноманітності впливів на систему управління різних завад можна запобігти лише відповідною різноманітністю управління її параметрами чи іншими різноманітними заходами, у результаті яких перейдемо до ентропійного, гомеостатичного чи морфологічного рівноважних станів.

Динамічна складність або складність поведінки системи визначається тим, наскільки складна реакція системи на простий вплив. Динамічна складність також пов'язана із характеристиками стійкості, коли малі вхідні впливи призводять до значних вихідних реакцій.

Структурна складність системи, очевидно, впливає на динамічну складність, але зворотна залежність не виконується. Так, система може бути структурно простою, хоча її поведінка виявиться достатньо складною.

Обчислювальна складність може характеризуватися кількістю кроків, необхідних для обчислення відображення системи. Якщо система відображена яким-небудь алгоритмом, то обчислювальна складність може характеризуватися за допомогою машини Тюрінга. Обчислювальна складність системи у вигляді управляючого автомата може визначатися кількістю операторів, що використовуються.

Аксіоми системної складності є математичною основою, за допомогою якої складність може бути уявлена й, відповідно, порівняна. Як міру складності використовують реальну величину $\Theta(S)$. Аксіомами складності є:

- аксіома ієрархії;
- аксіома паралельного сполучення підсистем;
- аксіома послідовного сполучення;
- аксіома сполучення із зворотним зв'язком;
- аксіома нормалізації.

Аксіома ієрархії: якщо S_o підсистема S, то

$$\Theta(\Sigma_0) \le \Theta(\sigma),$$

тобто підсистема не може бути складнішою, ніж система у цілому.

Аксіома паралельного сполучення підсистем: якщо $S = S_1 \oplus S_2 \oplus ... \oplus S_N$, то

$$\Theta(S) = \max_{1 \le i \le N} \Theta(S_i),$$

тобто складність системи визначається тією паралельною підсистемою, що має максимальну складність.

Аксіома послідовного сполучення: якщо $S = S_1 \oplus S_2 \oplus ... \oplus S_n$, то $\Theta(S) \le \Theta(S_1) + \Theta(S_2) + ... + \Theta(S_n)$, тобто складність системи визначається сумою складностей усіх послідовних підсистем.

Аксіома сполучення із зворотним зв'язком: якщо має місце зворотний зв'язок Θ із системи S_2 у систему S_1, то $\Theta(S_1 \oplus S_2) \le \Theta(S_1) + \Theta(S_2) + \Theta(S_2 \oplus S_1)$

Очевидно, попередня аксіома є окремим випадком даної, якщо зворотний зв'язок відсутній.

Аксіома нормалізації: у класі систем, що задовольняють цим аксіомам, може бути виділена підмножина систем Σ, для яких $\Theta(S).= 0$ для всіх $S \in \Sigma$.

На основі даних аксіом можна побудувати відповідну алгебру, що дасть змогу об'єктивно аналізувати та синтезувати системи різної складності.

2.7. Управління технічними системами

Процес функціонування системи міститься у послідовній зміні її станів під дією вхідних впливів. При цьому одні вхідні впливи цілеспрямовано змінюють стан системи, інші - перешкоджають цьому. У загальному випадку забезпечення виконання задач, поставлених перед системою, здійснюється за допомогою управління (управляючих впливів).

При розгляданні, управління технічними системами необхідно мати на увазі наступні обставини.

1. Управління припускає наявність цілі. Ціль виступає як найважливіша характеристика управління. Безцільні перетворення не є управлінням, вони позбавлені будь-якого сенсу. Безглуздо говорити про процеси управління у системі, утвореної сукупністю молекул газу в замкнутому об'ємі, хоча вона постійно змінює свої стани. Неможливо також управляти, наприклад, системою, що утворена сукупністю зерна на продовольчому складі, у якому немає устаткування для контролю його параметрів і засобів для зміни стану зерна (сушіння, знезаражування і т.п.), хоча зерно при збереженні також постійно змінює свій стан.

2. Управляти можна системою, що має більш ніж один стан. Власне кажучи, управління - це і є цілеспрямована зміна стану системи, що і приводить до досягнення поставленої мети. Іншими словами, управління можливе тільки у динамічних системах.

3. Управління полягає у виборі і реалізації таких управляючих впливів, що змінюють стан системи у бажаному напрямку. Але хоча екологічна система нашої планети є динамічною системою, людина ще не навчилася управляти такими явищами природи як землетрус, виверження вулканів та ін.

Раніше було відзначено, що при управлінні будь-якою системою обов'язково фігурує суб'єкт, у інтересах, якого здійснюється управління. Саме суб'єкт, у якості якого може виступати одна людина чи група осіб:

- формулює цілі управління, що повинні бути досягнуті у результаті функціонування системи;
- визначає систему, яку необхідно застосувати для досягнення цих цілей;
- формулює і реалізує управляючі впливи, що змінюють стан системи в напрямку досягнення мети.

Управління можливе при наявності суб'єкту, органу, що управляє (УО) і системи, якою управляють об'єкта управління (ОУ). При цьому, для формулювання і коректування цілей, формування і реалізації управляючих впливів суб'єкт може залучати різні групи людей і різні технічні засоби. Наприклад, директор із заводоуправлінням, є УО, а об'єктом управління - відповідно цех (ділянка), підприємство, технологічна лінія й ін.

Визначимо деякі характерні риси, що властиві процесам управління на виробництві.

1. Процес управління є циклічним.

2. Формування управляючих впливів засновано на порівнянні інформації про бажаний стану ОУ з інформацію про його поточний стан. Тобто процес управління виступає як інформаційний процес, зв'язаний з перетворенням і передачею інформації від одного об'єкта до іншого. При цьому, інформацію про стан ОУ прийнято називати ***інформацією стану*,** а відомості про те, що слід зробити ОУ - ***командною інформацією***.

3. У кожнім циклі управління можна виділити наступні елементи процесу управління:

- одержання УО інформації про цілі і задачі управління;
- збір УО інформації про стан ОУ (інформації стану);
- аналіз отриманої інформації й прийняття рішення;
- формування управляючих впливів і передача їх на об'єкт управління (командна інформація);
- виконання ОУ дій, відповідно до отриманої командної інформації.

Процес управління може включати не всі перераховані елементи, тому усі системи управління можна розділити на два великих класи.

Для першого класу систем характерно те, що інформація про поточний стан ОУ не використовується при управлінні. У таких системах здійснюється тільки лише передача командної інформації від УО до ОУ. Структурна схема такої системи приведена на рис. 2.12. Зв'язок між УО та ОУ називається прямим зв'язком, а сама система - ***розімкнутою***. Типовими представниками розімкнутих систем є системи програмного управління верстатами, технологічними лініями, ділянками і т.д.

Другий клас систем характеризується тим, що при управлінні використовується інформація про стан ОУ. Структурна схема такої системи наведена на рис. 2.13. Крім прямого зв'язку між УО й ОУ має місце і зворотний зв'язок для передачі інформації стану. Така система одержала назву ***замкнутої***. У замкнутій системі реалізуються всі елементи процесу управління. Відзначимо, що практично всі системи управління складними технічними системами є замкнутими системами.

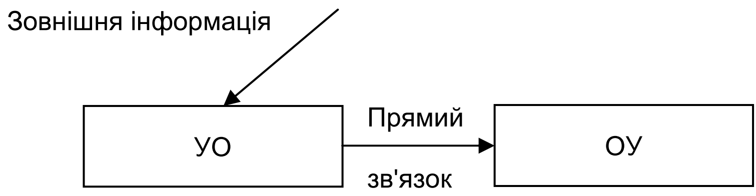


Рис. 2.12. Розімкнута система управління

4. У процесі управління можуть застосовуватися, а можуть і не застосовуватися технічні засоби. Наприклад, при управлінні діями навчальної групи староста групи не застосовує технічних засобів. Така система управління називається ***неавтоматизованою*.** Якщо в системі процеси передачі і перетворення інформації (обробка, аналіз, представлення) здійснюються за допомогою технічних засобів (ЕОМ, апаратура зв'язку, пристрої документування), а рішення приймає людина - такі системи називаються ***автоматизованими***. А коли всі процеси, аж до прийняття рішення, здійснюються тільки технічними засобами (без втручання людини-оператора), то системи відносяться до класу ***автоматичних*** систем. Нас надалі будуть в основному цікавити автоматичні й автоматизовані системи управління.

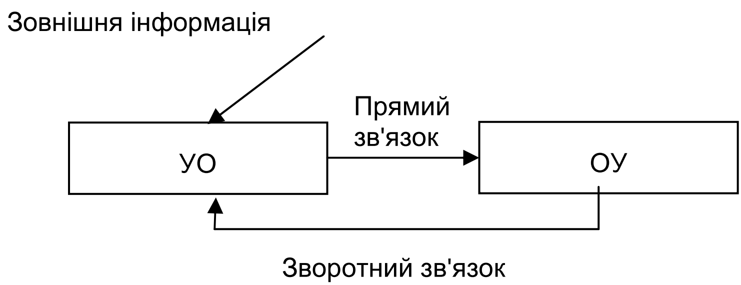


Рис. 2.13. Замкнута система управління

2.8. Методи композиції технічної системи

Під ***композицією*** будемо розуміти об'єднання складових частин об'єкта в єдине ціле (у систему) з урахуванням їх властивостей.

Очевидно, процедура композиції є, як би зворотна процедурі декомпозиції.

Об'єднання складових частин (елементів) у одне ціле (систему) здійснюється за допомогою системоутворюючих факторів, які виявляються у вигляді зв'язків між елементами. Тому при композиції системи важливо виділити ці фактори і співвіднести їх зі зв'язками між елементами (речовинними, енергетичними, інформаційними), а також їх напрямком і силою.

Як правило, процедура композиції припускає:

- точне формулювання поставленої задачі (створення нової системи, удосконалення існуючої);
- визначення сукупності елементів, що поєднуються у систему, у ряді випадків ця сукупність може бути визначена декомпозицією;
- вибір методу об'єднання, що представляє систему правил використання системоутворюючих факторів;
- оцінку варіантів побудови системи, якщо процедура композиції приводить до декількох альтернативних варіантів;
- вибір оптимального варіанта композиції.

Розглянемо основні методи. ***Метод вирішальних матриць***. Для конкретності будемо вважати, що дерево цілей розглядається, як ієрархічна структура ТС. Нехай, у результаті декомпозиції цілей, отримано ієрархічний граф-дерево (рис. 2.14), ребра якого характеризують взаємовідносини між глобальними і локальними цілями різних рівнів.

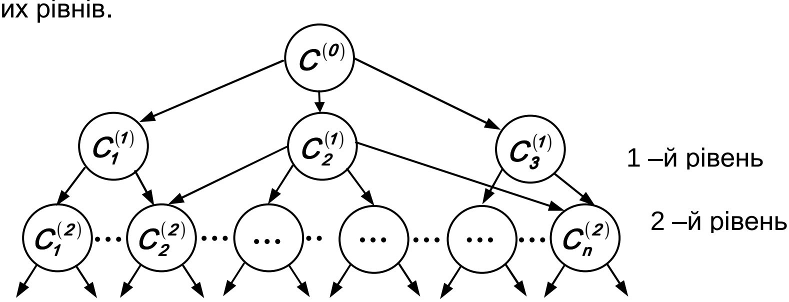


Рис. 2.14. Ієрархічна структура дерева цілей

Де $C^{(0)}$ - глобальна мета (нульовий рівень ієрархії);

$C_1^{(1)}$, $C_2^{(1)}$,.... - підсистеми 1 рівня;

$C_1^{(2)}$, $C_2^{(2)}$, ... - підсистеми 2 рівня і т.д.

Позначимо кожне ребро через $r_{jн}^{(i)}$, де верхній індекс означає приналежність елемента (вершини), з якого виходить ребро до i - го рівня, перший нижній індекс - це номер вершини, з якої виходить ребро, а другий нижній індекс - номер вершини нижчележачого рівня, у яку входить ребро. Тоді дерево цілей можна представити як

$$C_g = (C\ R)\ , \tag{2.11}$$

де: C - безліч вершин дерева;

R - безліч ребер (відносин).

Композиція системи йде у зворотному напрямку від нижчележачих рівнів до вищестоящих. Тоді "зворотне дерево" має вигляд (рис. 2.15):

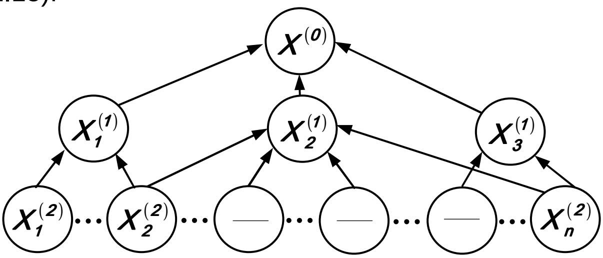


Рис. 2.15. Ієрархічна структура “зворотного дерева ”цілей

В залежності від того, якими засобами досягаються цілі буде визначатися фізичний зміст вершин графа 2.15.

Якщо такими засобами є елементи системи, то граф 2.15 буде відображати її структуру, якщо засобами виступають операції (наприклад, технологічні операції виготовлення якого-небудь продукту) - то граф 2.15 буде відображати план, зокрема мережевий, випуску продукції і т.д. За аналогією з (2.11) граф зображений на рис. 2.15 можна представити як

$$G_k = \left(X, R^{-1}\right)\ , \tag{2.12}$$

де: X - безліч вершин графа;

R^{-1} - інверсія відношення R $\left(R \times R^{-1} = 1\right)$.

Отже, композиція системи за методом вирішальних матриць починається з нижнього рівня. Нехай нижній рівень містить n цілей. Для досягнення цілей існує m засобів, причому, та сама мета може бути досягнута декількома засобами, а один і той же засіб може служити для досягнення декількох цілей. Тоді можна скласти матрицю $m \times n$ виду (2.13)

$$Q = \begin{vmatrix} q_{11} & q_{12} & \dots & q_{1n} \\ q_{21} & q_{22} & \dots & q_{2n} \\ \times & \times & \times & \times \\ q_{m1} & q_{m2} & \dots & q_{mn} \end{vmatrix}, \tag{2.13}$$

де елемент q_{ij} визначає значимість чи доцільність i - го засобу для досягнення j - ї мети. Така матриця одержала назву ***вирішальної матриці***.

Елементи матриці Q, як правило, визначаються експертними оцінками. Послідовно застосовуючи матриці Q,, на кожному рівні "зворотного дерева", реалізують композицію системи. Завдання дослідника полягає у оцінці кожного з варіантів і вибору найкращого.

Крім розглянутого вище методу вирішальних матриць широке застосування одержали ***морфологічні методи***. Як відомо, терміном "морфологія" визначається науковий напрямок, що вивчає внутрішню структуру систем, що досліджуються. При морфологічному підході до процедури композиції система утворюється шляхом об'єднання елементів за допомогою відповідних зв'язків за визначеними правилами.

Розглянемо один з морфологічних методів композиції - ***метод морфологічної шухляди***. Припустимо, необхідно створити систему, що складається з n елементів. Кожний i - й елемент, де $i = \overline{1\ n}$, може бути реалізований l_i варіантами. Передбачається, що той самий елемент (навіть у різних варіантах) не може двічі використовуватися у системі.

Позначимо i -й елемент у j - му варіанті виконання через $x_j^{(i)}$, де $j = \overline{1, l_n}$ і назвемо його ***морфологічною змінною***. Оскільки той самий елемент не може двічі використовуватися у системі, то змінні $x_j^{(i)}$ і $x_s^{(q)}$ (де $i \neq q$) є незалежними. Тоді можемо побудувати матрицю

$$X = \begin{vmatrix} x_1^{(1)} & x_2^{(1)} & \ldots & x_{l_1}^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \ldots & x_{l_2}^{(2)} \\ \ldots & \ldots & \ldots & \ldots \\ x_1^{(n)} & x_2^{(n)} & \ldots & x_{l_n}^{(n)} \end{vmatrix},$$

яка називається ***морфологічною шухлядою***. Правило композиції за методом морфологічної шухляди полягає у формуванні набору значень змінних $x_j^{(i)}$ по одному з кожного рядка, що і дає один з варіантів побудови системи, тобто наприклад:

$$S_k = x_1^{(1)} \cup x_1^{(2)} \cup \ldots \cup x_j^{(i)} \cup \ldots \cup x_{l_n}^{(n)},$$

де S_k - k -й варіант системи.

Таким чином, ідея морфологічної шухляди базується на переборі незалежних змінних і одержанні розміщень з повтореннями з l_i по n. Очевидно, загальна кількість таких розміщень (загальна кількість одержуваних варіантів системи) дорівнює

$$N = l_1 \times l_2 \times \ldots \times l_n.$$

Якщо число значень кожної незалежної змінної (число елементів рядків матриці шухляди) однаково, тобто $l_i = l_q = l$, то тоді

$$N = l^n.$$

При великих значеннях l і n число розміщень виходить настільки великим, що перебір стає важким навіть з використанням ЕОМ. З метою скорочення числа варіантів вводять різні обмеження. Наприклад, за критерієм сумісності елементів у системі. Це дозволяє звузити кількість варіантів перебору і скоротити час одержання варіанту системи.

Метод морфологічної шухляди можна проілюструвати на наступних прикладах. Нехай необхідно створити систему виміру часу (годинник). Як можливі елементи системи виступають: пружина, електричний двигун, маятник, термопара, зубчаста передача, генератор електричних коливань, гідравлічний двигун, стрілочний індикатор, акумуляторна батарея, електрична передача, ланцюгова передача, цифровий індикатор.

У першу чергу виділимо незалежні змінні $x^{(i)}$ (проведемо класифікацію елементів). Тоді можна виділити: $x^{(1)}$ - джерело енергії; $x^{(2)}$ - привід; $x^{(3)}$ - датчик часу; $x^{(4)}$ - передача; $x^{(5)}$ - індикатор.

Знайдемо значення незалежних змінних: $x_1^{(1)}$ - пружина; $x_2^{(1)}$ - акумуляторна батарея; $x_3^{(1)}$ - термопара; $x_1^{(2)}$ - електричний двигун; $x_2^{(2)}$ - гідравлічний двигун; $x_1^{(3)}$ - маятник; $x_2^{(3)}$ - генератор електричних коливань; $x_1^{(4)}$ - зубчаста передача; $x_2^{(4)}$ - ланцюгова передача; $x_3^{(4)}$ - магнітна передача; $x_1^{(5)}$ - стрілочний індикатор; $x_2^{(5)}$ - цифровий індикатор. Тоді морфологічна шухляда прийме вигляд таблиці 2.1.

Таблиця 2.1

Зміст морфологічної шухляди

$x^{(i)}$ \ $x_j^{(i)}$	$x_1^{(i)}$	$x_2^{(i)}$	$x_3^{(i)}$
$x^{(1)}$ - джерело енергії	$x_1^{(1)}$ - пружина	$x_2^{(1)}$ - акумуляторна батарея	$x_3^{(1)}$ - термопара
$x^{(2)}$ - привід	$x_1^{(2)}$ - електричний двигун	$x_2^{(2)}$ - гідравлічний двигун	-
$x^{(3)}$ - датчик часу	$x_1^{(3)}$ - маятник	$x_2^{(3)}$ - генератор електричних коливань	-
$x^{(4)}$ - передача	$x_1^{(4)}$ - зубчаста передача	$x_2^{(4)}$ - ланцюгова передача	$x_3^{(4)}$ - магнітна передача
$x^{(5)}$ - індикатор	$x_1^{(5)}$ - стрілочний індикатор	$x_2^{(5)}$ - цифровий індикатор	-

Тут поряд зі значеннями незалежних змінних вказано і їхній зміст.

Якщо композицію елементів провести у вигляді

$$S_1 = x_1^{(1)} \cup x_1^{(3)} \cup x_1^{(4)} \cup x_1^{(5)} ,$$

то система S_1 - являє собою годинник зі звичайним механічним заводом і стрілочним циферблатом.

Якщо

$$S_2 = x_2^{(1)} \cup x_1^{(2)} \cup x_2^{(3)} \cup x_1^{(4)} \cup x_1^{(5)} ,$$

то система S_2 являє собою електричний годинник.

А композиція

$$S_3 = x_3^{(1)} \cup x_1^{(2)} \cup x_2^{(3)} \cup x_1^{(4)} \cup x_1^{(5)} ,$$

може являти собою нову систему (новий тип годинника) і т.д.

Необхідно відзначити, що об'єднання елементів у систему здійснювалося за критерієм сумісності. Дійсно, варіант композиції

$$S_k = x_2^{(1)} \cup x_2^{(2)} \cup x_1^{(3)} \cup x_3^{(4)} \cup x_2^{(5)} ,$$

очевидно, неприйнятний, тобто сполучити акумуляторну батарею з гідравлічним двигуном навряд чи є найкращім рішенням.

Таким чином, композиція за методом морфологічної шухляди дає безліч варіантів побудови системи, з яких можна вибрати найкращий, що відповідає поставленій цілі.

У самому загальному вигляді морфологічна шухляда може бути представлена і у вигляді табл. 2.2.

Де $a_{jq}^{(ik)}$ - вказують на наявність чи відсутність зв'язку між елементами $x_j^{(i)}$ і $x_q^{(k)}$ (наприклад, за принципом сумісності). Якщо зв'язок є, то у відповідній клітинці записується *1*, якщо немає - записується *0*. Можна вказати і значимість зв'язку числами *1,2, ...* . Така таблиця одержала назву «Матриці взаємодій».

Таблиця 2.2

Загальний вигляд морфологічної шухляди

		$x^{(1)}$			…	$x^{(k)}$			…
		$x_1^{(1)}$	…	$x_{l1}^{(1)}$		$x_1^{(k)}$	…	$x_q^{(k)}$	
$x^{(1)}$	$x_1^{(1)}$								
	…								
⋮	$x_l^{(1)}$					$a_{l1}^{(1k)}$			
$x^{(i)}$	$x_1^{(i)}$								
	…								
	$x_j^{(i)}$							$a_{jq}^{(ik)}$	

Сформулюємо етапи проведення процедури композиції ТС за методом морфологічної шухляди.

1. Точне формулювання задачі композиції.

2. Визначення сукупності можливих елементів, з яких створюється система.

3. Класифікація елементів (визначення незалежних змінних).

4. Проведення операції співвідношення елементів до відповідного класифікатору (визначення незалежних перемінних).

5. Формування матриці - морфологічної шухляди.

6. Формування набору значень змінних по одному з кожного рядка.

7. Оцінка варіантів і вибір оптимального з них за яким-небудь критерієм (вартість, економічність і т.д.).

Розглянемо ***просторову композицію*** елементів ТС. До таких задач відносяться, наприклад, задачі розміщення елементів (мікросхем) радіоелектронної апаратури на печатній платі, розміщення блоків у стійках АСУ, раціонального розміщення ЕОМ і допоміжного устаткування на ОЦ, раціонального розміщення технологічного устаткування на корисних площах, наприклад, підприємств і т.д. Для рішення таких задач можна скористатися так називаною мережею взаємодії.

Ідея використання мережі взаємодії полягає в тому, що відповідно до матриці взаємодії будується граф - ***мережа взаємодії***, що представляє собою просторову композицію елементів. Проводячи операції з мережею взаємодії, одержують оптимальне (чи, у всякому разі, раціональне) просторове розміщення елементів.

Використання мереж взаємодії передбачає побудову, у першу чергу, матриці взаємодії яка є матрицю розміром $n \times n$, де рядки і стовпці відповідають елементам системи, а елементи матриці - зв'язкам між ними.

Нехай необхідно вирішити задачу розміщення n елементів $x_1, x_2, x_3, \ldots, x_n$, тобто провести просторову композицію елементів. Складемо матрицю взаємодії (табл. 2.3). Якщо елементи a_{ij} матриці взаємодії приймають значення *0* чи *1*, то ця матриця являє собою звичайну матрицю суміжності. Однак на практиці у матриці взаємодії вказується не тільки наявність чи відсутність зв'язку, але і значимість такого зв'язку, Значимість зв'язку може виражатися, наприклад, числами *0* - небажаний зв'язок, *1* - бажаний зв'язок, *2* - істотно необхідний зв'язок. Як правило, значимість зв'язку визначається експертами. У цьому складається принципова відмінність матриці взаємодії від матриці суміжності чи матриці повних зв'язків.

Таблиця 2.3

Матриця взаємодії

	x_1	x_2	...	x_n
x_1	a_{11}	a_{12}	...	a_{1n}
x_2	a_{21}	a_{22}	...	a_{2n}
...				
x_n	a_{n1}	a_{n2}	...	a_{nn}

Відповідно до матриці взаємодій будується граф-мережа взаємодій, на якій суцільними лініями показуються істотно необхідні зв'язки, а пунктирними - бажані зв'язки. Змінюючи розташування вершин графа, домагаються максимального виключення перетинань ребер. Отриманий граф дає раціональну структуру просторового розміщення елементів.

Наприклад, у цеху підприємства необхідно розмістити 6 верстатів. Матриця з'єднань (взаємодії) представлена таблицею 2.4.

Таблиця 2.4

Матриця з'єднань

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	*0*	*2*	*2*	*0*	*2*	*0*
x_2	*2*	*0*	*2*	*2*	*0*	*2*
x_3	*2*	*2*	*0*	*1*	*1*	*0*
x_4	*0*	*2*	*1*	*0*	*0*	*1*
x_5	*2*	*0*	*1*	*0*	*0*	*0*
x_6	*0*	*2*	*0*	*1*	*0*	*0*

Елементи матриці приймають значення:

0 - з'єднання відсутні;

1 - з'єднання є, але вимоги до довжини з'єднань не жорсткі;

2 - з'єднання є, вимоги до довжини жорсткі (з'єднання повинне бути найкоротшим).

Побудуємо граф-мережу взаємодії (рис. 2.16), де суцільні лінії відповідають *2*, пунктири – *1*.

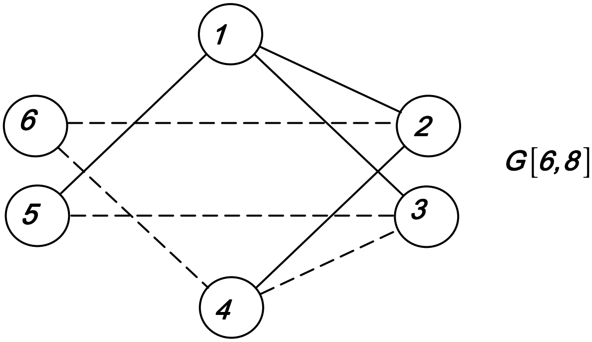


Рис. 2.16. Граф-мережа взаємодії

Видозмінимо граф (рис. 2.16). У центрі розташуємо вершини, що мають найбільше число зв'язків, а навколо - інші, але так, щоб ребра не перетиналися. Тоді граф (рис. 2.17) дає раціональне розміщення, наприклад, верстатів.

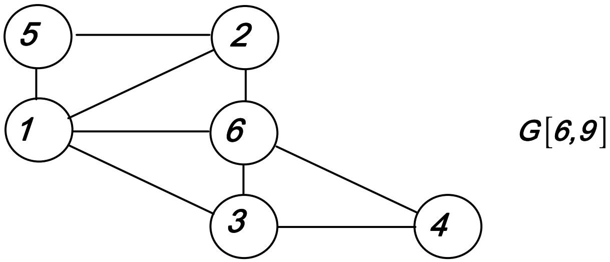


Рис. 2.17. Варіант раціонального розміщення верстатів

2.9. Контрольні питання

1. Дайте визначення поняттю елемент системи.
2. Дайте визначення поняттю ТС.
3. Дайте визначення поняттю відкрита ТС.
4. Дайте визначення поняттю закрита ТС.
5. Дайте визначення поняттю стан ТС.
6. Дайте визначення поняттю функціональний опис ТС.
7. Дайте визначення поняттю інформаційний опис ТС.
8. Дайте визначення поняттю процесуальний опис ТС.
9. Дайте визначення поняттю рівновага ТС.

10. Дайте визначення поняттю стійкість ТС.
11. Дайте визначення поняттю морфологічний опис ТС.
12. Дайте визначення поняттю машинний агрегат.
13. Охарактеризуйте словесний метод опису структур ТС.
14. Дайте визначення поняттю структурна надмірність ТС.
15. Дайте визначення поняттю часової структури складних ТС**.**
16. Дайте визначення поняттю просторової структури складних ТС**.**
17. Наведіть класифікацію ТС за кількістю елементів системи.
18. Наведіть класифікацію критеріїв розвитку ТС.
19. Наведіть класифікацію основних видів машин.
20. Наведіть загальну класифікацію ТС.
21. Наведіть класифікацію особливостей складних ТС.
22. Перерахуйте загальні властивості ТС.
23. Назвіть основні методи композиції ТС.
24. Дайте визначення поняттю статична ТС.
25. Дайте визначення поняттю динамічна ТС.
26. Дайте визначення поняттю складна ТС.
27. Дайте визначення поняттю емерджентних властивостей ТС.
28. У чому полягає закономірність цілісності складних ТС?
29. У чому полягає закономірність ієрархічності складних ТС ?
30. У чому полягає закон необхідної різноманітності ТС?
31. Дайте визначення поняттю цілісність ТС.
32. Дайте визначення поняттю причинність ТС.
33. Дайте визначення поняттю управляємість ТС.
34. Дайте визначення поняттю адаптованість ТС.
35. Дайте визначення поняттю стійкість ТС.
36. Дайте визначення поняттю зв'язність ТС.
37. Дайте визначення поняттю складність ТС.
38. Дайте визначення поняттю розімкнута система управління.
39. Дайте визначення поняттю замкнута система управління.
40. У чому полягає сутність управління ТС?
41. У чому полягає композиція ТС за методом вирішальних матриць?
42. У чому полягає композиція ТС за методом морфологічної шухляди?

Розділ 3

Математичні методи розв'язання задач оптимізації

Сформульована загальна постановка задачі оптимізації та розглянуто методи вирішення задач оптимізації АСУ. Наведено контрольні питання.

3.1. Загальна постановка задачі оптимізації

У загальному випадку процес оптимізації припускає вибір такого варіанта управління, при якому досягається мінімальне або максимальне значення деякого критерію, що характеризує якість управління. Вибір належного критерію є основною проблемою правильної постановки задачі оптимізації. Обов'язково повинна бути сформульована мета. Якщо є різні шляхи досягнення цієї мети, то необхідно знайти найкращий з них.

Постановка задачі оптимізації містить у собі безліч припустимих рішень $X = \{x_1, x_2, ..., x_n\}$ та визначену на цій безлічі числову функцію $f(x)$, що називається цільовою функцією.

Цільова функція – це аналітична залежність між критерієм (критеріями) оптимальності та параметрами, що підлягають оптимізації, із вказівкою напрямку екстремуму.

Не можна ототожнювати критерій (критерії) оптимальності й цільову функцію.

Відмінність понять «критерій» і «цільова функція» полягає у такому.

1. Цільова функція може містити в собі більше одного критерію.
2. Для цільової функції завжди обов'язково вказується вид екстремуму:

$$f(x) \rightarrow max \quad (f(x) \rightarrow min).$$

У загальній задачі оптимізації необхідно знайти вектор $x=(x^{(1)},..., x^{(n)})$ з допустимої області X, який обертає в мінімум цільову функцію $f(x)$, тобто такий вектор $x^* \in X$, для якого виконується умова

$$f(x^*) \leq f(x) \quad \text{для усіх} \quad x \in X.$$

Якщо такий вектор ***x****** існує, то він визначає ***слабкий глобальний (абсолютний) мінімум f*(x)*** у допустимій області ***X***. Цей мінімум називається слабким, оскільки задовольняє слабкій нерівності. Він називається глобальним або абсолютним тому, що нерівність справедлива для будь-якого $x \in X$. Мінімум при $x = x^*$ називається ***сильним***, коли має місце $f(x^*) \ll f(x)$ для $x \neq x^*$. Якщо поміняти знаки нерівностей, що розглядалися, одержимо слабкий і сильний максимум.

Існування слабкого глобального мінімуму допускає наявність декількох оптимальних точок, оскільки будь-який ***x***, що задовольняє умові $f(x) = f(x^*)$, також є оптимальною точкою. Сильний глобальний мінімум завжди єдиний.

Хоча метою задачі оптимізації є отримання глобального мінімуму цільової функції, проте при її розв'язанні важливе значення має поняття локального або відносного мінімуму.

Мінімум в точці $x = x^*$ називається ***локальним*** (відносним), якщо знайдеться така окраїна точки **x***, що для всіх **x,** які належать цій окраїнні має місце $f(x^*) \leq f(x)$.

Наприклад нехай $a \leq x \leq b$ і $f(x) \to min$, а цільова функція має вигляд, як показано на рис. 3.1.

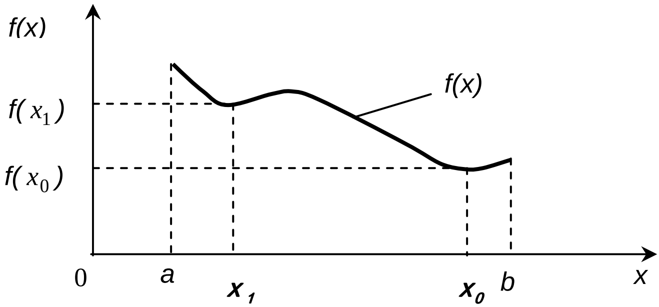


Рис. 3.1. Ілюстрація понять «локальний мінімум» і «глобальний мінімум» цільової функції

Тут точка x_0- глобальний мінімум, а точка x_1 - локальний мінімум цільової функції.

Якщо функція ***f(x)*** диференційована, то задача відшукання локальних мінімумів зводиться до знаходження стаціонарних точок, у яких звертаються в нуль особистісні похідні функції ***f(x)***:

$$\frac{\partial f(x)}{\partial x^{(i)}} = 0,\ i = \overline{1,n}.$$

Слід зазначити, що ця умова є лише необхідною, але недостатньою умовою існування відносного мінімуму. Стаціонарна точка може бути і точкою відносного максимуму, і сідловою точкою, у якій не досягається ні відносний мінімум, ні відносний максимум. Достатньою умовою існування у стаціонарній точці відносного мінімуму є позитивна визначеність квадратичної форми.

Відзначимо також, що хоча абсолютний мінімум є в той же час і відносним, але відносний мінімум не обов'язково буде абсолютним. Більш того, на допустимій множині ***X*** функція ***f(x)*** може мати не один, а декілька відносних мінімумів. Більшість же відомих методів оптимізації дозволяє знаходити лише точки відносного мінімуму. При цьому виникає дуже складна задача знаходження того з відносних мінімумів, який є абсолютним.

Приведене вище завдання оптимізації має рішення не при будь-яких цільових функціях і допустимих множинах. Існують завдання, у яких неможливо знайти оптимальне рішення і екстремум цільової функції. Наприклад, не існує точок мінімуму функції однієї змінної ***f*** на множині ***X*** у випадках, наведених нижче на рис. 3.2, 3.3 і 3.4.

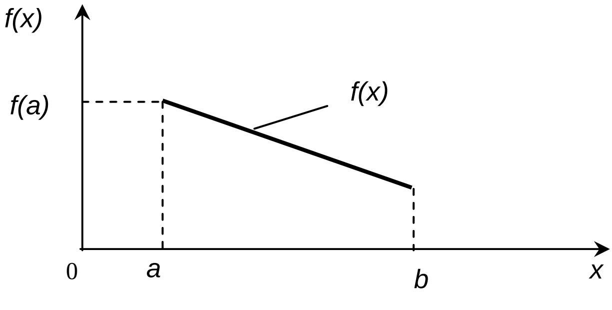


Рис. 3.2. Ілюстрація випадку, коли безліч допустимих рішень не замкнута

Тут межа ***a*** безлічі допустимих рішень в інтервал входить, а межа ***b*** відсутня. Тобто множина ***X*** ***не замкнута***, отже, $f(b)$ – не існує.

У випадку, представленому на рис. 3.3, визначена лише одна ліва межа безлічі допустимих рішень. $X = [a,\infty)$, тобто безліч допустимих рішень ***необмежена***.

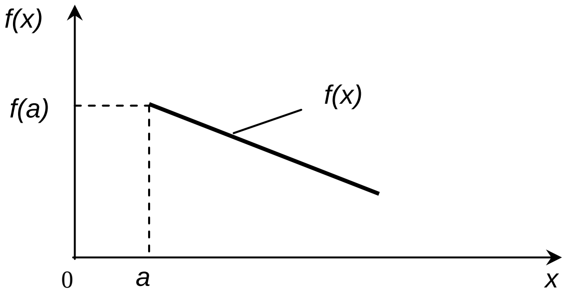


Рис. 3.3. Ілюстрація необмеженості безлічі допустимих рішень.

На рис. 3.4. показаний ще один випадок, коли завдання оптимізації не має однозначного рішення.

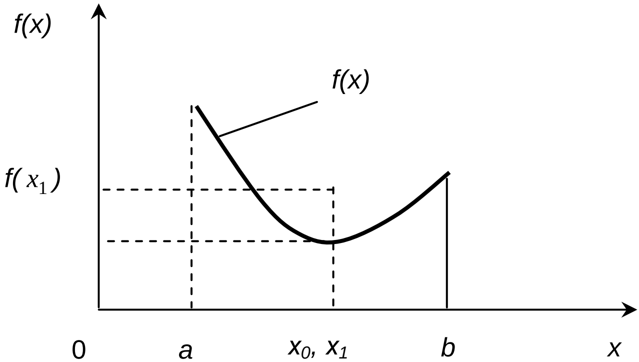


Рис. 3.4. Ілюстрація випадку, коли функція ***f(x)*** не є безперервною.

Тут функція ***f(x)*** не є неперервною, оскільки в т. $\boldsymbol{x_0}$ ***(***$\boldsymbol{x_1}$***)*** існують два значення функції – $\boldsymbol{f(x_0)}$ і $\boldsymbol{f(x_1)}$.

Отже, завдання оптимізації має рішення, якщо виконуються наступні три умови.

1. Безліч допустимих рішень ***X*** замкнута, тобто якщо граничні точки належать цій множині.
2. Множина ***X*** обмежена.
3. Цільова функція ***f(x)*** неперервна.

Це нестроге формулювання ***теореми Вейерштрасса.***

3.2. Методи вирішення задач оптимізації

3.2.1. Метод множників Лагранжа.

Раніше було визначено, що задача оптимізації полягає у знаходженні екстремуму цільової функції ***f(x),*** де $x=(x^{(1)},\ldots, x^{(n)})$ - точка у просторі раціональних чисел за наявності обмежень типу рівності

$$q_i(x) = 0,\ i = \overline{1,m},\ m < n. \tag{3.1}$$

Зауваження. Наявність обмежень на параметри або їх взаємозв'язки у технологічному процесі пов'язана із реалізацією конкретних механічних, хімічних, теплових і інших процесів, які не можуть мати параметри довільної величини.

Якщо обмеження (3.1) мають місце, то мінімум функції ***f(x)*** називають ***умовним мінімумом***. Якщо обмеження (3.1) відсутні, то говорять про ***безумовний мінімум***, знаходження якого зводиться до визначення і дослідження стаціонарних точок функції ***f(x)***.

Класичний спосіб рішення даної задачі полягає у тому, що рівняння (3.1) використовуються для виключення з розгляду ***m*** змінних. При цьому цільова функція приводиться до вигляду

$$f(x^{(1)},\ldots,x^{(n)}) = f_1(y^{(1)},\ldots,y^{(n-m)}),$$

де через $y^{(1)},\ldots, y^{(n-m)}$ позначені змінні, що невиключені.

Задача полягає тепер у знаходженні значень $y^{(1)},\ldots, y^{(n-m)}$, які обертають у мінімум (максимум) функцію ***f(x)*** і на які не накладено ніяких обмежень, тобто до задачі на безумовний екстремум.

Відмітимо, що за відсутності обмеження $q(x) = 0$ мінімум цільової функції ***f(x)*** досягається у точці $x = (1,0)$ і рівний нулю. Таким чином, обмеження приводить до збільшення значення цільової функції, тобто до погіршення якості оптимуму.

Якщо рівняння (3.1) мають складний вигляд, то виключення з їх допомогою ***m*** змінних з функції ***f(x)*** представляє значні труднощі. У зв'язку з цим велике практичне значення набув метод зведення задачі на умовний екстремум до задачі на безумовний екстремум, оснований на використанні функції Лагранжа.

Введемо у розгляд вектор $\lambda = (\lambda_1, \ldots, \lambda_m)$ і дослідимо властивості функції

$$L(x,\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i q_i(x), \tag{3.2}$$

де λ_i - множниками Лагранжа.

Функція $L(x, \lambda)$ називається функцією Лагранжа $n + m$ змінних $x^{(1)},..., x^{(n)}$, $\lambda = \lambda_1, ..., \lambda_m$.

Розглянемо стаціонарні точки функції $L(x, \lambda)$, які одержимо, прирівнявши нулю особистісні похідні по i по λ_j $j = \overline{1,m}$ и по $x^{(i)}, i = \overline{1,n}:$

$$\frac{dL(x,\lambda)}{d\lambda_j} = q_j(x) = 0,\ j = \overline{1,m}; \tag{3.3}$$

$$\frac{dL(x,\pi)}{dx^{(i)}} = 0,\ i = \overline{1,n}.$$

Відзначимо, що рівняння (3.3) співпадають із обмеженнями (3.1) і, як випливає з (3.2), при їх виконанні має місце $L(x, \lambda) = f(x)$. Тому, якщо у стаціонарній точці (x^*, λ^*) функція $L(x, \lambda)$ досягає мінімуму, то x^* забезпечує і мінімум функції $f(x)$ при виконанні обмежень (3.1), тобто дає розв'язання задачі.

Таким чином, задача на умовний екстремум цільової функції $f(x)$ за наявності обмежень типу рівності зводиться до задачі на визначення стаціонарних точок функції Лагранжа $L(x, \lambda)$.

У тих випадках, коли отримується тільки одне рішення, потрібна додаткова перевірка (з використанням других похідних): чи є це рішення точкою локального екстремуму або точкою перегину. У зв'язку з тим, що перебір безлічі рішень системи рівнянь операція, що практично не реалізовується, метод множників Лагранжа не набув широкого поширення при рішенні складних практичних завдань. Подібні складні екстремальні завдання вирішуються зазвичай іншими методами, у основі яких лежить ідея поступового наближення до екстремуму.

3.2.2. Ітеративні методи пошуку оптимуму

Обмежені можливості вживання класичних методів у задачах зі складними видами обмежень і довільним видом цільової функції привело до широкого використання ітеративних методів пошуку оптимального рішення, в основі яких лежать поняття градієнта цільової функції $f(x)$.

Градієнтом функції $f(x)$, що позначається **grad** $f(x)$ або $\nabla f(x)$, називається вектор, величина якого визначає швидкість зміни функції $f(x)$, а напрям співпадає із напрямом найбільшого зростання цієї функції. Вектор $\nabla f(x)$, який вказує напрям найбільшого зменшення функції $f(x)$, називається ***антиградієнтом*** функції $f(x)$.

Припустимо, що $x = (x^{(1)}, \ldots, x^{(n)})$. Тоді градієнт функції $f(x)$ буде вектором-стовпцем вигляду

$$\nabla f(x) = grad\ f(x) = \left(\frac{\partial f(x)}{\partial x^{(1)}}, \ldots, \frac{\partial f(x)}{\partial x^{(n)}} \right)'.$$

Умова того, що точка x є стаціонарною точкою функції $f(x)$, за допомогою градієнта запишеться у вигляді

$$\nabla f(x) = 0.$$

Для обґрунтування ітераційних алгоритмів оптимізації розкладемо функцію $f(x)$ у ряд Тейлора в околиці точки оптимуму x^*, яку вважаємо стаціонарною точкою.

$$f(x) = f(x^*) + \frac{1}{2}\sum_k \sum_j a_{kj} \Delta x^{(k)} \Delta x^{(j)}, \tag{3.4}$$

де $\Delta x = x^* = x$ - відхилення від точки оптимуму;

$$a_{kj} = a_{jk} = \left. \frac{\partial^2 f(x)}{\partial x^{(k)} \partial x^{(j)}} \right|_{x = x^*}. \tag{3.5}$$

Беручи особистісні похідні від $f(x)$ по $x^{(i)}$, $i=1, \ldots, n,$ знаходимо складові градієнта функції $f(x)$ в даній точці x:

$$\frac{\partial f(x)}{\partial x^{(i)}} = -\sum_k a_{ki} \Delta x^{(k)}, \; i = \overline{1,n}.$$

Для градієнта функції $f(x)$ при цьому отримуємо вираз

$$\nabla f(x) = -A \Delta x = -A(x^* - x), \tag{3.6}$$

де A - квадратна симетрична матриця розміром $n \times n$ з елементами a_{ki}, визначуваними співвідношенням (3.5). Розв'язуючи систему рівнянь (3.6) щодо x^*, отримуємо:

$$x^* = x - A^{-1} \nabla f(x). \tag{3.7}$$

Якби розкладання (3.4) було точним поданням $f(x)$ і існували б прості методи визначення матриці A, то значення x^* можна б було знайти безпосередньо за формулою (3.7). У більшості практичних задач ці умови звичайно не виконуються.

Проте, якщо замінити невідому матрицю A^{-1} на матрицю Γ з елементами γ_{kj}, то можна сподіватися, що величина $x \nabla f(x)$ при відповідному виборі коефіцієнтів дає значення, ближче до оптимального, ніж x. При цьому відкривається можливість багатокрокової процедури пошуку рішення.

Позначимо через x_n значення x на n - му кроці і вважатимемо, що елементи матриці Γ можна вибирати на різних кроках по-різному.

Тоді відповідно до (3.7) багатокрокова процедура пошуку оптимуму функції ***f(x)*** запишеться у вигляді

$$x_{n+1} = x_n - \Gamma_n \nabla f\ (x_n\ )\ .$$

Для того, щоб послідовність значень x_n приводила до оптимального рішення, коефіцієнти матриці Γ_n повинні задовольняти певним умовам, які мають назву умовами збіжності. При цьому різні способи вибору Γ_n призводять до різних пошукових алгоритмів оптимізації.

Іноді алгоритм оптимізації записують у вигляді

$$x_{n+1} = x_n + \Delta x_n, \tag{3.8}$$

де

$$\Delta x_n = -\Gamma_n \nabla f\ (x_n\ )\ , \tag{3.9}$$

є вектором поправок.

На рис. 3.5 представлена структурна схема алгоритму пошуку оптимуму з використанням виразів (3.8) і (3.9).

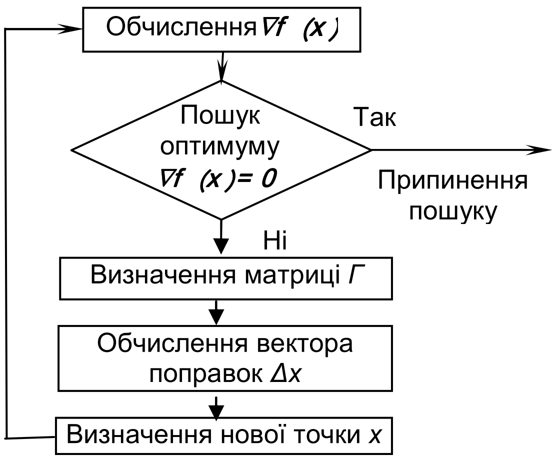


Рис. 3.5. Структурна схема алгоритму пошуку оптимуму

Якщо стоїть задача мінімізації, то цільова функція ***f(x),*** де $x = (x^{(1)}, \ldots, x^{(n)})$ може бути зображена у вигляді деякої поверхні, що має вид котловини, зображеної на рис. 3.6 а, і має назву ***поверхня відгуку***.

На рис. 3.6 б профіль цієї котловини зображений на площині $(x^{(1)}, x^{(2)})$ у вигляді ліній рівних значень ***f***, що є перетинами котловини горизонтальними площинами. Біля кожної такої лінії ставиться відповідне їй значення ***f***.

Наочну інтерпретацію задачі пошуку оптимуму можна одержати, якщо уявити собі людину, що стоїть на схилі котловини, зарослої лісом. Людина повинна опуститися на дно котловини, але не має нагоди оглянути котловину цілком, а може вести огляд лише вельми обмеженої ділянки поблизу місця свого розташування.

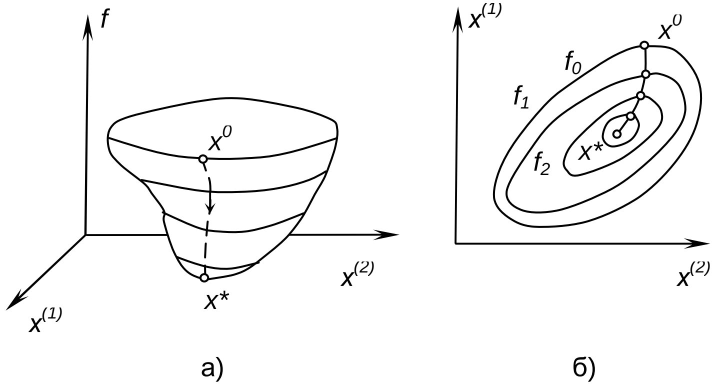


Рис. 3.6. Поверхня відгуку

У цих умовах найприроднішим напрямом руху є напрям у бік найбільшої крутизни спуску, тобто у напрямі, протилежному напряму градієнта функції ***f(x).*** Одержувана при цьому стратегія, має назву ***градієнтного методу*****,** буде послідовністю кроків, кожний з яких містить дві операції.

1) Визначення напряму найбільшої крутизни спуску, тобто напряму антиградієнта функції ***f(x)***.

2) Переміщення у вибраному напрямі на задану відстань.

Характер траєкторії при градієнтному методі приведений на рис. 3.6, б. Її особливість полягає в тому, що кожний крок відбувається у напрямі, перпендикулярному лінії постійного рівня.

Математично стратегію градієнтного методу можна реалізувати, якщо переміщення $\Delta x^{(i)}$ на кожному кроці уздовж кожної з осей буде пропорційне складовій градієнта $\dfrac{\partial f}{\partial x^{(i)}}$ у напрямі цієї осі:

$$\Delta x^{(i)} = -\gamma \frac{\partial f\ (x)}{\partial x^{(i)}},\ i = \overline{1,n}.$$

Це означає, що у виразі (3.9) матриця Γ_n буде діагональною з елементами γ на головній діагоналі

$$\Gamma_n = \gamma I.$$

При цьому поправка на ***n*** - м кроці може бути представлена у вигляді

$$\Delta x_n = -\gamma I \nabla q \ (x_n) = -\gamma \nabla f \ (x_n). \qquad (3.10)$$

Стратегія, у відповідності до співвідношення (3.10), визначає рух із змінним кроком, оскільки значення кроку визначається значенням градієнта $\nabla f(x)$. Оскільки значення градієнта зменшується на пологих схилах і поблизу оптимуму, то на деяких ділянках кроки будуть дрібними, що подовжує час пошуку.

Цього недоліку можна позбавитися, якщо використовувати градієнтну стратегію із постійним кроком, значення якого дорівнює γ. У цьому випадку поправка на кожному кроці визначається за формулою

$$\Delta x_n = -\gamma D(x_n),$$

що витікає з (3.10) заміною вектора градієнта $\nabla f(x)$ на вектор напряму градієнта

$$D(x) = \frac{\nabla f(x)}{|\nabla f(x)|},$$

де $|\nabla f(x)| = \left[\left(\frac{\partial f}{\partial x^{(1)}}\right)^2 + \ldots + \left(\frac{\partial f}{\partial x^{(n)}}\right)^2\right]^{\frac{1}{2}}$ є модулем градієнта ***f(x)***.

Для того, щоб рухатися у напрямі градієнта, кроки повинні бути не дуже великими. Тому їх буде багато. Але оскільки на кожному кроці необхідно знову визначати напрям градієнта процес наближення до оптимуму буде повільним.

Додаткові труднощі виникають, якщо точка оптимуму лежить у ямі або на довгому вузькому гребені, як показано на рис. 3.7. У цьому випадку градієнтний метод може викликати стрибки через яму, так що траєкторія хоча і досягне точки оптимуму, але ціною багатьох неефективних кроків. Від цих недоліків вільні деякі інші методи.

На відміну від градієнтного методу, у якому градієнт визначається на кожному кроці, у методі ***найшвидшого спуску*** градієнт визначається тільки у початковій точці і рух у знайденому напрямі продовжується однаковими кроками доти, поки зменшуються значення функції ***f(x)***. Якщо на якомусь кроці ***f(x)*** зросло, то рух у даному напрямі припиняється, останній крок знімається повністю або наполовину і обчислюється новий градієнт функції ***f(x)***, а відповідно, і новий напрям руху. На рис 3.8 приведена структурна схема алгоритму методу найшвидшого спуску.

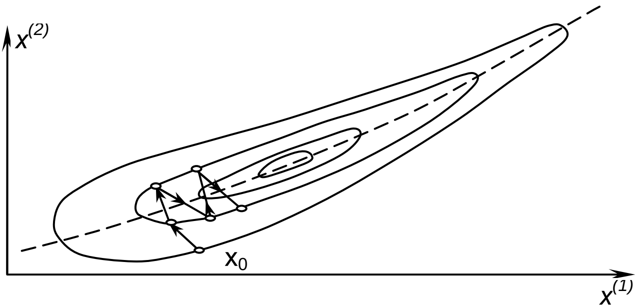


Рис. 3.7. Градієнтний пошук

У тих випадках, коли поверхня відгуку достатньо добре описується рівнянням другого порядку, різке зменшення числа кроків можна одержати, якщо скористатися ***алгоритмом Ньютона.***

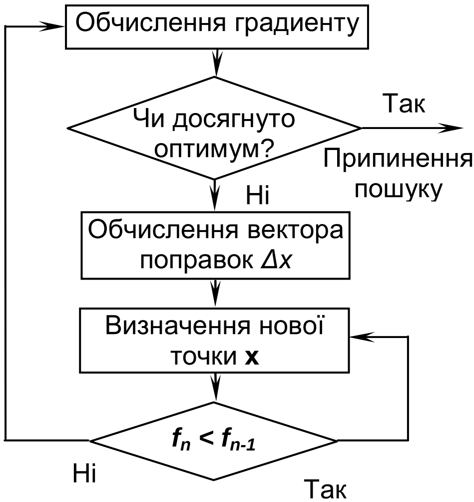


Рис. 3.8. Структурна схема алгоритму методу найшвидшого спуску

При цьому представлення $f(x)$ у вигляді (3.4) буде достатньо точним при значному видаленні від точки оптимуму і як матрицю Γ_n можна взяти безпосередньо матрицю A. Але елементи a_{kj} матриці

A, що розраховані у точці оптимуму, наперед невідомі. Проте при достатньо хорошій поверхні відгуку другі похідні функції ***f(x)***

$$\frac{\partial^2 f(x)}{\partial x^{(k)} \partial x^{(j)}}, \; k, j = \overline{1,n},$$

що розраховані у довільній точці $x = x_n$, будуть близькі до елементів a_{kj} матриці **A**. Позначуючи через $\left[\nabla^2 f\, x\, (x_n)\right]$ матрицю других похідних ***f(x)*** в точці x_n і використовуючи її як матрицю Γ_n, одержуємо вектор поправок для алгоритму Ньютона:

$$\Delta x_n = -\left[\nabla^2 f(x_n)\right]^{-1} \nabla f(x_n).$$

Якщо розкладання (3.4) є точним представленням функції ***f(x)***, то точка оптимуму досягається за один крок.

3.2.3. Графоаналітичний метод рішення задач оптимізації.

Цим методом вирішуються прості завдання оптимізації. Математичні моделі у цих завданнях не повинні бути складними, оскільки інакше потрібно багато часу для їх вирішення. Спершу розглянемо ***однопараметричне однокритеріальне завдання оптимізації***.

Постановка завдання: Даний один критерій ***y***. Об'єкт (процес) описаний рівнянням (рівняннями), що включають один параметр $y = f(x)$, який потрібно знайти. Є система обмежень типу:

- $x \geq a_1$;
- $a_2 \leq x \leq b_1$ і т.п.

Необхідно знайти оптимальне значення параметра $x - x_{опт}$, що обертає цільову функцію $f(x)$ в максимум або мінімум.

Завдання вирішується у два етапи.

1. Побудова області припустимих рішень (ОПР).
2. Знаходження в межах ОПР оптимального рішення.

При побудові ОПР ***на першому етапі*** розглядається система обмежень. Всі обмеження повинні бути виконані. Виконання першого обмеження означає, що значення параметра ***x***, яке потрібно знайти повинно знаходитися правіше a_1, причому, a_1 у дозволений інтервал входить (рис.3.9). Виконання другого обмеження означає, що значення параметра ***x***, яке потрібно знайти повинно знаходитися в інтервалі (на відрізку) $[a_2, b_1]$, слід мати на увазі, що межі інтервалу в інтервал входять.

На другому етапі застосовують метод перебору. Сутність його полягає у наступному. У межах ОПР через певний інтервал ***h*** вибирається ряд значень параметра x_i. У випадку, що розглядається нами, ОПР розбита на чотири відрізки, і вибрано п'ять значень параметра x. Для цих значень параметра x розраховуються відповідні значення цільової функції. Серед них знаходять мінімальне (максимальне) значення. Значення параметра x_i, що обертає цільову функцію в мінімум (максимум), є оптимальним. Якщо у випадку, що розглядається нами, $f(x)$ прагне до мінімуму, то $x_{опт} = x_3$, якщо до максимуму, то $x_{опт} = x_5$.

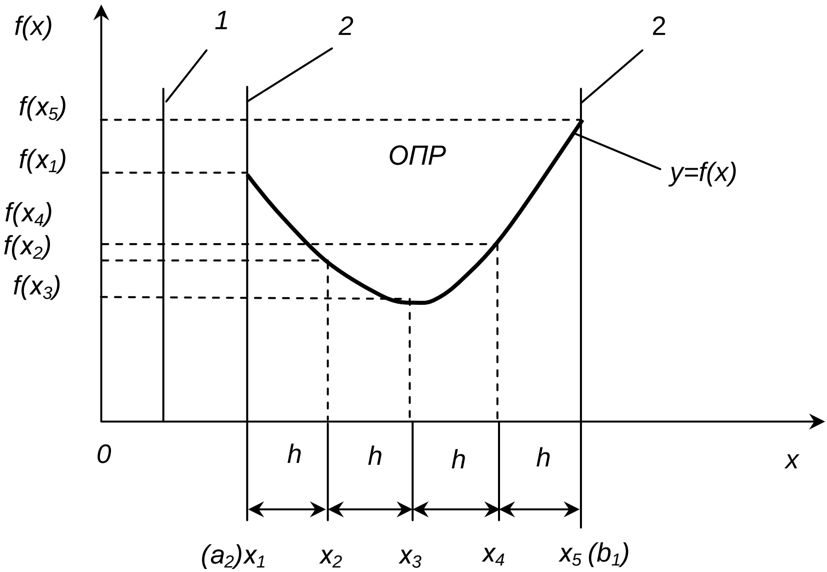


Рис. 3.9. Графічна ілюстрація рішення однопараметричної однокритеріальної задачі оптимізації

Розглянемо окремий випадок, коли цільова функція лінійна (рис. 3.10).

В даному випадку на другому етапі обчислюють значення цільової функції тільки на межах ОПР. Ці значення порівнюють і вибирають найменше або найбільше. Для прикладу, приведеного на рис. 3.10, якщо $f(x) \to min$, то $x_{опт} = b_1$, якщо $f(x) \to max$, то $x_{опт} = a_2$.

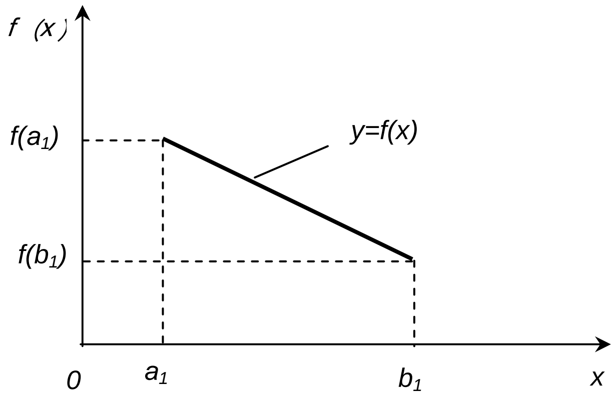


Рис. 3.10. Ілюстрація рішення однопараметричної однокритеріальної задачі оптимізації для випадку лінійної цільової функції

Розглянемо тепер графоаналітичний ***метод рішення багатопараметричної однокритеріальної задачі оптимізації.***

Постановка завдання. Задано один критерій y. Об'єкт (процес) описаний рівнянням (рівняннями), що включають ряд параметрів $y = f(x_1, x_2, ..., x_n)$. Є система обмежень:

$$a_1 \le x_1 \le b_1;$$
$$a_2 \le x_2 \le b_2;$$
$$....$$
$$a_n \le x_n \le b_n;$$
$$....$$
$$x_1 \times x_2 \le c_1;$$
$$x_2 \times x_3 \le c_2;$$
$$....$$

Потрібно визначити оптимальне значення ряду параметрів $x_{1опт}, x_{2опт}, ..., x_{mопт}$ $(m \le n)$, що обертають цільову функцію $f(x_1, x_2, ..., x_n)$ у максимум або мінімум.

Приклад 3.1. Розглянемо рішення багатопараметричної однокритеріальної задачі оптимізації на наступному прикладі. Заданий критерій $y = x_2/x_1$. Потрібно знайти $x_{1опт}$ і $x_{2опт}$, що обертають у максимум цільову функцію $y = x_2/x_1 \rightarrow max$. Обмеження: $1 \le x_1 \le 8, \quad 2 \le x_2 \le 12, \quad x_1 \times x_2 \ge 10$.

Завдання вирішується також у два етапи.

1. Побудова ОПР.
2. Знаходження в межах ОПР оптимального рішення.

Побудова ОПР у даному завданні на відміну від завдання однопараметричного полягає у тому, що працювати потрібно у двох напрямах. У результаті в площині x_1 x_2. ОПР буде багатогранником (рис. 3.11).

Для побудови нелінійного обмеження $x_1 \times x_2 \geq 10$ спочатку необхідно прирівняти ліву і праву частини нерівності і побудувати відповідну криву $x_1 \times x_2 = 10$, $x_2 = 10/x_1$ (табл. 3.1)

Таблиця 3.1

x_1	1	2	5
x_2	10	5	2

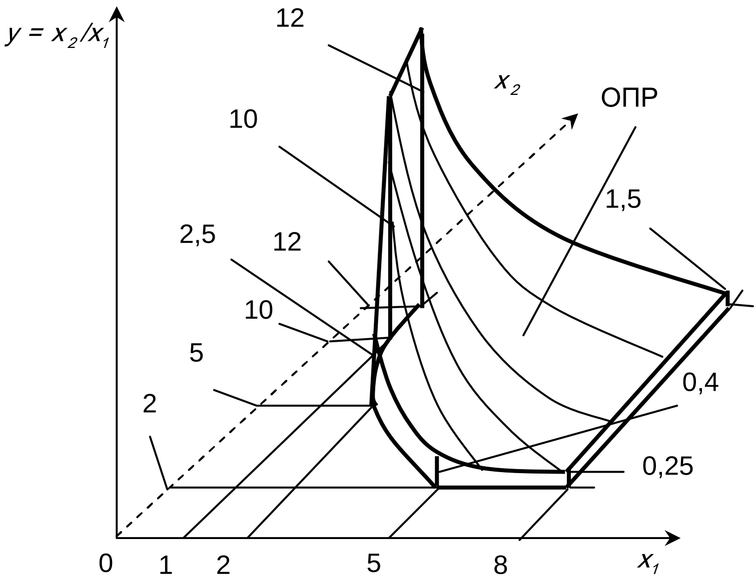


Рис. 3.11. Графічна ілюстрація рішення двопараметричної однокритеріальної задачі оптимізації

Після цього потрібно визначити напрям допустимості параметрів x_1 і x_2. Щоб не помилитися, для цього можна застосувати наступний прийом. Виберемо довільну точку на площині $x_1 x_2$ з будь-якого боку кривої. Наприклад, виберемо точку з координатами $x_1 = 5,\ x_2 = 5$, тобто «справа - вгорі» від кривої. Обчислимо значення лівої частини нерівності: $5 \times 5 = 25,$ а $25 > 10$, отже, нерівність виконується. Це означає, що вибрана точка знаходиться у допустимій області параметрів, які потрібно знайти. Тобто допустима область параметрів знаходиться «справа - вгорі» від кривої.

На другому етапі необхідно обчислити значення цільової функції у межах ОПР. У даному прикладі точка, що визначає оптимальні значення параметрів функції

$$y = x_2/x_1 \to max\ ,$$

які потрібно знайти знаходитися на межі ОПР:

$$x_{1опт} = 1,\ x_{2опт} = 12\ .$$

Якщо необхідно знайти екстремум

$$y = x_2/x_1 \to min\ ,$$

то

$$x_{1\,опт} = 8,\ x_{2опт} = 2\ .$$

3.2.4. Методи рішення багатокритеріальних задач оптимізації

На практиці багатокритеріальні завдання оптимізації виникають, коли об'єкт, що проектується не може бути описаний однокритеріальною залежністю, або об'єднати окремі критерії в єдиний критерій неможливо. Таке об'єднання критеріїв у єдиний критерій буде розглянуто нижче. Але це об'єднання, як правило, буває формальним, штучним. З математичної точки зору не існує ідеального способу чи методу вирішення таких завдань. Кожний з них має переваги й недоліки. Розглянемо деякі методи рішення багатокритеріальних завдань оптимізації.

Першим розглянемо метод пошуку ***Парето - ефективних рішень***. Розглянемо суть даного методу на прикладі використання двох критеріїв. Критерії при використанні даного методу є рівнозначними.

Нехай є безліч варіантів рішення. По кожному з варіантів визначені значення всіх критеріїв. Представимо множину оцінок варіантів рішення у просторі критеріїв (рис.3.12). де прийняті наступні позначення:

K_1 та K_2 – критерії оцінки варіантів рішення;

Y = {y₁, y₂, …, yₘ} - множина оцінок альтернативних варіантів рішень;

$K_{11}, K_{12}, \dots , K_{1m}$ - значення першого критерію для ***1, 2, … , m*** - го варіанту рішення;

$K_{21}, K_{22}, \dots , K_{2m}$ – значення другого критерію для ***1, 2, … , m*** - го варіанту рішення;

P(Y) – множина Парето – ефективних оцінок рішень.

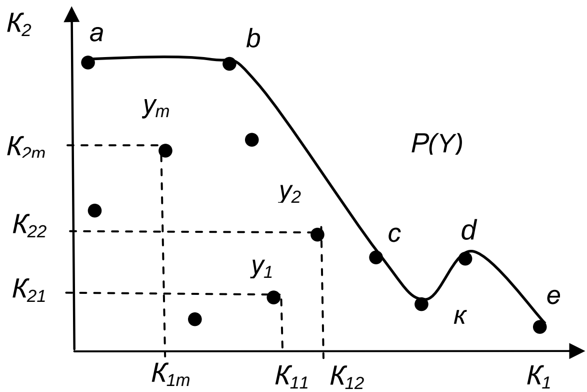


Рис. 3.12. Ілюстрація пошуку Парето - ефективних рішень

Правило. Безліч Парето - ефективних оцінок ***P(Y')*** являє собою "північно - східну" границю безлічі ***Y*** без тих її частин, які паралельні однієї із координатних осей або лежать у "глибоких" провалах.

Для випадку, що зображено на рис. 3.12, Парето - ефективні оцінки складаються із точок кривої (***bc***), крім точки (***c***), і лінії (***de***).

Переваги методу: критерії рівнозначні, метод є математично об'єктивним.

Недолік методу: одне остаточне рішення отримується тільки у окремому випадку, тобто кількість Парето - ефективних рішень, як правило, більше одного.

Приклад 3.2. Є 10 варіантів металорізальних верстатів, серед яких, для ділянки цеху, що проектується необхідно вибрати найкращий. Верстати оцінені експертами по двох показниках (критеріях): продуктивності й надійності. Оцінювання робилося по ***11*** - бальній шкалі від ***0*** до ***10***. Результати оцінки верстатів наведені в таблиці 3.2.

Представимо безліч оцінок варіантів металорізальних верстатів у просторі критеріїв (рис. 3.13). Парето - ефективними рішеннями тут є варіанти верстатів C_5, C_7 і C_9.

Таблиця .3.2

Експертні оцінки верстатів

Критерії	Оцінки експертів									
	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}
продуктивність	6	4	10	3	10	0	2	4	6	7
надійність	6	2	1	7	4	4	10	4	8	2

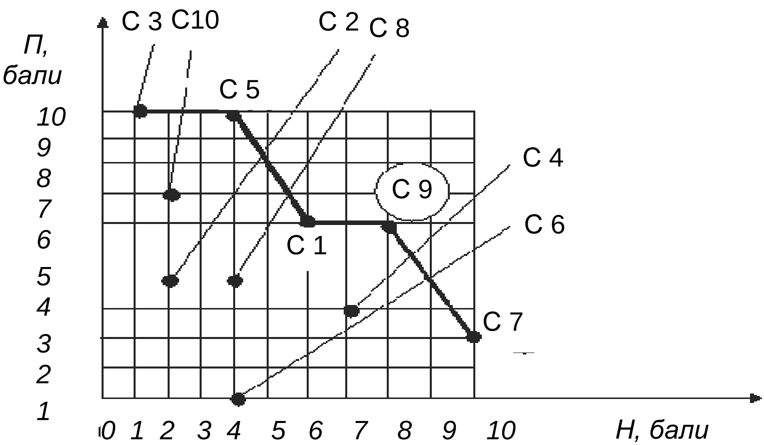


Рис. 3.13. Приклад пошуку Парето – ефективних рішень

Досить часто, при рішенні багатокритеріальних завдань оптимізації використається ***метод узагальненого (інтегрального) критерію.***

Сутність даного методу полягає у тому, що особистісні критерії $F_i(X)$, $i=\overline{1,n}$ якимось чином поєднуються у один інтегральний критерій

$$F(X)=\Phi(F_1(X),F_2(X),...,F_n(X)),$$

а потім знаходиться максимум або мінімум даного критерію.

Якщо об'єднання особистісних критеріїв здійснюється, виходячи з об'єктного взаємозв'язку особистісних критеріїв і узагальненого критерію, то тоді оптимальне рішення буде коректним. Але таке об'єднання здійснити вкрай складно або неможливо, тому, як правило, узагальнений критерій є результат чисто формального об'єднання особистісних критеріїв.

Залежно від того, яким чином особистісні критерії об'єднуються в узагальнений критерій розрізняють наступні види узагальнених критеріїв:

1. адитивний критерій;
2. мультиплікативний критерій;
3. максимінний (мінімаксний) критерій.

При застосуванні ***адитивного критерію*** цільова функція отримується шляхом додавання нормованих значень особистісних критеріїв. У загальному вигляді цільова функція має наступний вигляд:

$$F(X) = \sum_{i=1}^{n} C_i \frac{F_i(X)}{F_i^o(X)} = \sum_{i=1}^{n} C_i f_i(X) \to max\ (min), \quad (3.11)$$

де n – кількість особистісних критеріїв, що поєднуються;
C_i – ваговий коефіцієнт i – го особистісного критерія;
$F_i(X)$ - числове значення i - го особистісного критерія;
$F^o_i(X)$ – i – й нормуючий дільник;
$f_i(X)$ – нормоване значення i – го особистісного критерія.

Особистісні критерії мають різну фізичну природу й тому різну розмірність. Тобто просте їх підсумування не є коректним. У зв'язку із цим у формулі (3.11) числові значення особистісних критеріїв діляться на деякі нормуючи дільники, які призначаються у такий спосіб.

У якості нормуючих дільників приймаються директивні значення параметрів або критеріїв, що задані замовником. Вважається, що значення параметрів, які закладені в технічному завданні, є оптимальними або найкращими.

У якості нормуючих дільників приймаються максимальні (мінімальні) значення критеріїв, що можна досягти в області припустимих рішень.

Розмірність самих особистісних критеріїв і відповідних нормуючих дільників однакові, тому в підсумку узагальнений адитивний критерій виходить безрозмірною величиною.

Приклад 3.3. Визначити оптимальний варіант машини з використанням узагальненого (інтегрального) адитивного критерію. Особистісними критеріями, за допомогою яких оцінені варіанти машини, є її продуктивність і надійність (наробіток на відмову). Обидва критерії "працюють" на максимум, тобто найкращими варіантами машини є ті з них, які забезпечують найбільшу її продуктивність і надійність. Вихідні дані для рішення завдання наведені в таблиці 3.3.

Цільова функція на основі адитивного критерію запишеться в такий спосіб:

$$F(X) = C_1 \frac{F_1(X)}{F_1^o(X)} + C_2 \frac{F_2(X)}{F_2^o(X)} \to max.$$

У якості нормуючих дільників у даному завданні приймемо найкращі (максимальні) значення особистісних критеріїв:

$$F_1^o(X) = 4000 \text{ шт/год.}, F_2^o(X) = 1500 \text{ шт/год}.$$

Значення узагальненого адитивного критерію розраховуються для кожного варіанта машини.

Варіант 1. ***F(X) = 0,6(1000/4000) + 0,4(1500/1500) = 0,55.***

Варіант 2. ***F(X) = 0,6(2000/4000) + 0,4(1000/1500) = 0,558.***

Варіант 3. ***F(X) = 0,6(4000/4000) + 0,4(500/1500) = 0,732.***

Оптимальним є 3 варіант машини, тому що йому відповідає максимальне значення узагальненого адитивного критерію.

Таблиця 3.3

Вихідні дані для визначення оптимального варіанту машини

Критерій F_i	Ваговий коефіцієнт	Значення критеріїв варіантів машин		
		Варіант 1	Варіант 2	Варіант 3
продуктивність F_1 *(шт./год.)*	0,6	1000	2000	4000
надійність F_2 *(год.)*	0,4	1500	1000	500

Один з недоліків цього методу полягає у тому, що вагові коефіцієнти призначає проектувальник. Різні проектувальники можуть призначати різні вагові коефіцієнти. Нехай $C_1 = 0,4$; $C_2 = 0,6$. Визначимо значення адитивних критеріїв для варіантів машини:

Варіант 1. ***F(X) = 0,4·0,25+0,6·1=0,7.***

Варіант 2. ***F(X) = 0,4·0,5+0,6·0,67=0,602.***

Варіант 3. ***F(X) = 0,4·1+0,6·0,33=0,598.***

Тобто при такій зміні значень вагових коефіцієнтів оптимальним вже буде 1 варіант машини.

Перевага даного методу: як правило, завжди вдається визначити єдиний оптимальний варіант рішення.

Недоліки методу:

- труднощі (суб'єктивізм) у визначенні вагових коефіцієнтів;
- адитивний критерій не випливає з об'єктної ролі особистісних критеріїв і тому виступає як формальний математичний прийом;
- у адитивному критерії відбувається взаємна компенсація особистісних критеріїв, тобто зменшення одного з них може бути компенсовано збільшенням іншого критерію.

У випадку застосування ***мультиплікативного критерія*** цільова функція записується у такий спосіб:

$$F(X) = \prod_{i=1}^{n} C_i F_i(X) \to max \ (min),$$

де **П** – знак добутку;

C_i – ваговий коефіцієнт ***i*** – го особистісного критерія;

$F_i(X)$ - числове значення *i*-го особистісного критерія.

Переваги застосування мультиплікативного критерію:

- не потрібне нормування особистісних критеріїв;
- практично завжди визначається одне оптимальне рішення.

Недоліки:

- труднощі (суб'єктивізм) у визначенні вагових коефіцієнтів;
- необхідність перемножування різних розмірностей;
- взаємна компенсація значень особистісних критеріїв.

Максимінний (мінімаксний) критерій працює за принципом компромісу, що ґрунтується на ідеї рівномірності. Сутність принципу максиміна полягає у наступному. При проектуванні складних систем має місце наявність великої кількості особистісних критеріїв встановити між якими аналітичний взаємозв'язок дуже складно. Тому намагаються знайти такі значення змінних (параметрів) $X = (x_1, x_2, \ldots, x_n)$, при яких нормовані значення всіх особистісних критеріїв стають рівними між собою:

$$C_i f_i(X) = K,$$

де C_i – ваговий коефіцієнт ***i*** – го особистісного критерія;

$f_i(X)$ – нормоване значення ***i*** – го особистісного критерія;

К – константа.

При великій кількості особистісних критеріїв через складні взаємозв'язки досягти виконання зазначеного вище співвідношення дуже складно. Тому на практиці так підбирають значення змінних проектування $x_1, x_2, \ldots, x_m$, щоб послідовно "підтягувалися" ті нормовані критерії, чисельні значення яких у вихідному рішенні виявилися найменшими. Оскільки ця операція є компромісною, то "підтягування" "відстаючого" критерію неминуче приводить до зниження значень частини інших критеріїв. Але при проведенні ряду перетворень можна отримати певне зрівноважування суперечливих особистісних критеріїв, що і є метою принципу максиміна.

Формально принцип максиміна формулюється наступним чином: вибрати такий набір змінних $X^{(o)} \in X$, при якому реалізується максимум з мінімальних нормованих значень особистісних критеріїв, тобто

$$F\ (X^{(o)}) = \max \min f_i\ (X).$$

Такий принцип вибору $X^{(o)}$ має назву - гарантованого результату. Цей принцип запозичено з теорії ігор.

Якщо особистісні критерії необхідно мінімізувати, то самим "відстаючим" критерієм є той.жл, що приймає максимальне значення. У цьому випадку застосовують принцип мінімакса:

$$F\ (X^{(o)}) = \min \max\ f_i\ (X).$$

Вибір критеріїв оптимальності дуже складне завдання, тому що цілі, що ставляться при створенні АСУ, як правило, суперечливі (забезпечення мінімальної вартості й максимальної надійності, максимальної продуктивності й мінімальної енергоємності й т.д.).

Якщо потрібно оптимізувати один з показників якості АСУ, при дотриманні обмежувальних вимог на інші показники, то потрібно сформулювати один особистісний критерій. Завдання оптимізації при цьому зводиться до завдання максимізації (мінімізації) даного критерію з урахуванням обмежень, що задаються.

При можливості застосування декількох критеріїв вибирають:

а) адитивний критерій, якщо істотне значення мають абсолютні значення критеріїв при обраному векторі параметрів X;

б) мультиплікативний критерій, якщо істотну роль грає зміна абсолютних значень особистісних критеріїв при варіації вектора X;

в) максимінний (мінімаксний) критерій, якщо поставлене завдання досягнення рівності нормованих значень суперечливих (конфліктних) особистісних критеріїв.

3.3. Лінійне програмування

3.3.1. Загальна постановка задачі

Лінійне програмування (ЛП) виникло у зв'язку з розглядом питань про знаходження найвигідніших варіантів при рішенні різних планово-виробничих задач. У цих задачах необхідно знайти такі значення параметрів, які з деякої точки зору були б якнайкращими, при умові, що зміна різних параметрів можлива у великих межах і має місце низка обмежуючих умов. Завдання знаходження екстремуму у цьому випадку носить спеціальну назву завдання лінійного програмування. До таких задач відносяться задачі знаходження най раціональнішого способу використовування сировини і матеріалів, визначення найвигідніших виробничих режимів, підвищення ефективності роботи транспорту т.п.

Термін «лінійне програмування» зв'язується з дослідженням і розв'язанням наступної задачі. Знайти такі значення $\dot{x}_1, \dot{x}_2, \dot{x}_3, \dots, \dot{x}_n$ змінних $x_1, x_2, x_3, \dots, x_n$, які задовольняють наступній системі співвідношень

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \;\; R_1 \;\; a_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \;\; R_2 \;\; a_2 \\ \quad \dots\dots\dots\dots \\ a_{s1}x_1 + a_{s2}x_2 + \dots + a_{sn}x_n \;\; R_s \;\; a_s \end{cases}, \tag{3.12}$$

і при цьому визначають найбільше (найменше) значення функції

$$z\,(x_1, x_2, \dots, x_n) = c_1x_1 + c_2x_2 + \dots + c_nx_n\,, \qquad (3{,}13)$$

у порівнянні з її значеннями при всіх інших значеннях змінних $x_1, x_2, x_3, \dots, x_n$ які задовольняють системі (3.12),

де R_i $(i = 1,2,\dots s)$ – один зі знаків $=$, $\geq$, $\leq$;

a_i, c_i, a_{ij} $(i = 1, 2, \dots, s;\; j = 1, 2, \dots, n)$ – задані дійсні числа.

Істотно, що ліві частини всіх співвідношень (3.12) і функція (3.13) лінійні щодо перемінних $x_1, x_2, x_3, \dots, x_n$. Саме тому задача називається задачею лінійного програмування (ЗЛП).

Співвідношення (3.12) називають ***обмеженнями ЗЛП***, а функцію (3.13) – ***цільовою функцією*** цієї задачі.

Всякий упорядкований набір $\bar{x}_1, \bar{x}_2, \bar{x}_3, \dots, \bar{x}_n$ значень змінних, які задовольняють усім обмеженням (3.12) (тобто всяке рішення системи лінійних рівнянь і (або) нерівностей (3.12)), називається ***припустимим рішенням***.

Припустиме рішення $\dot{x}_1, \dot{x}_2, \dot{x}_3, \dots, \dot{x}_n$, при якому цільова функція приймає максимальне (мінімальне) значення, називається ***оптимальним рішенням*** або рішенням ЗЛП.

Таким чином, припустиме рішення $\dot{x}_1, \dot{x}_2, \dot{x}_3, \dots, \dot{x}_n$ задачі ЛП, є її оптимальним рішенням, якщо для будь-якого іншого припустимого рішення $\bar{x}_1, \bar{x}_2, \bar{x}_3, \dots, \bar{x}_n$ цієї задачі маємо

$$c_1\bar{x}_1 + c_2\bar{x}_2 + c_3\bar{x}_3 + \dots + c_n\bar{x}_n \leq c_1\dot{x}_1 + c_2\dot{x}_2 + c_3\dot{x}_3 + \dots + c_n\dot{x}_n\,,$$

якщо задача на знаходження максимуму, і

$$c_1\bar{x}_1 + c_2\bar{x}_2 + c_3\bar{x}_3 + \dots + c_n\bar{x}_n \geq c_1\dot{x}_1 + c_2\dot{x}_2 + c_3\dot{x}_3 + \dots + c_n\dot{x}_n\,,$$

якщо задача на знаходження мінімуму.

Значення цільової функції при кожному з оптимальних рішень задачі називається ***оптимумом*** цієї задачі.

Задачу ЛП, яка має припустимі рішення (тобто система обмежень якої сумісна) будемо називати ***припустимою***, у протилежному випадку ЗЛП будемо називати – ***неприпустимою***.

У більшості задач ЛП серед обмежень (3.12) маються обмеження виду:

$$x_j \geq 0, \quad x_j \leq 0, \qquad \text{або} \qquad b_j \leq x_j \leq d_j, \ \text{або} \ \ b_j \geq x_j \geq d_j,$$

для всіх чи деяких j ($j = 1, 2, \dots, n$). Такі обмеження будемо називати ***прямими обмеженнями на змінні***. Змінні, на які не накладені обмеження типу $x_j \geq 0$, $x_j \leq 0$ називаються ***вільними***, а змінні на які накладені такі обмеження відповідно ***невільними.***

3.3.2. Задача лінійного програмування у канонічній формі

Розглянемо задачу лінійного програмування:

$$c_1x_1 + c_2x_2 + \dots + c_nx_n \to max \ \ (min)\ ;$$

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = a_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = a_2, \\ \qquad\qquad \dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = a_m, \\ x_j \geq 0 \quad (j = 1, 2, \dots, n). \end{cases} \tag{3.14}$$

Задачу ЛП, надана у такому вигляді будемо називати ***канонічною задачею ЛП.***

Якщо задача ЛП задана не в канонічній формі, то вона може бути зведена шляхом нескладних математичних перетворень до еквівалентної їй канонічної задачі.

Наприклад, якщо помножити на -1 обидві частини (ліву і праву) обмежень, можна поміняти зміст нерівностей на протилежний. Якщо у задачі є прямі обмеження виду $x_j \geq 0$, то можна перенумерувати змінні так, щоб ці обмеження відносилися до перших n_1 змінних ($n_1 \leq n$) і тоді будь-яку задачу ЛП можна записати у вигляді

$$c_1x_1 + c_2x_2 + \ldots + c_nx_n \to max \quad (min)\ ;$$

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = a_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m_11}x_1 + a_{m_12}x_2 + \ldots + a_{m_1n}x_n = a_{m_1}, \\ a_{m_1+1,1}x_1 + a_{m_1+1,2}x_2 + \ldots + a_{m_1+1,n}x_n \leq a_{m_1+1}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \leq a_m, \\ x_j \geq 0 \quad \left(j = 1,\ 2, \ldots, n_1\right); \quad n_1 \leq n, \end{cases} \tag{3.15}$$

вважаючи особистісними випадки, коли серед обмежень немає рівнянь, чи навпаки, нерівностей.

Нехай нам потрібно перетворити дану задачу, що не є канонічною, до канонічного виду. Якщо серед обмежень є нерівності, то при підстановці будь-якого припустимого рішення $\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots, \bar{x}_n$ кожне з таких обмежень перетвориться або у числову рівність

$$a_{i1}\bar{x}_1 + a_{i2}\bar{x}_2 + a_{i3}\ \bar{x}_3 + \ldots + a_{in}\bar{x}_n = a_i\ ,$$

або у числову нерівність

$$a_{i1}\bar{x}_1 + a_{i2}\bar{x}_2 + a_{i3}\ \bar{x}_3 + \ldots + a_{in}\bar{x}_n < a_i\ .$$

Отже, для кожного припустимого рішення і кожного обмеження – нерівності мається своє не негативне число, додаток якого до лівої частини нерівності перетворює його у тотожність. Для тих обмежень – нерівностей, що при підстановці даного припустимого рішення перетворюються у рівність, ці числа дорівнюють нулю, а для інших вони позитивні.

Якщо у задачі (3.14) не на всі змінні накладено умову не негативності, то (3.15) не буде задачею канонічного виду. Є кілька способів переходу від задачі виду (3.15) до канонічної задачі.

Розглянемо один з них. При цьому для простоти приймемо, що (3.14) відразу має вигляд (3.15), тобто в (3.14) $m_1 = m$.

Замінимо кожну змінну, від якої не потрібно виконання умови не негативності, різницею двох інших змінних, на які накладемо умову не негативності:

$$x_{n1+1} = x_{n1+1}' - x_{n1+1}''; \qquad x_{n1+1}' \geq 0; \qquad x_{n1+1}'' \geq 0;$$
$$x_{n1+2} = x_{n1+2}' - x_{n1+2}''; \qquad x_{n1+2}' \geq 0; \qquad x_{n1+2}'' \geq 0;$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$x_n = x_n' - x_n''; \qquad x_n' \geq 0; \qquad x_n'' \geq 0.$$

Підставимо ці різниці замість відповідних змінних у задачу (3.14)

$$c_1x_1 + \ldots + c_{n1}x_{n1} + c_{n1+1}(x_{n1+1}{}' - x_{n1+1}{}'') + \ldots + c_n(x_n{}' - x_n{}'') \to \max,$$

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n1}x_{n1} + a_{1n1}(x_{n1}{}' - x_{n1}{}'') + a_{m1+1}(x_{n1+1}{}' - x_{n1+1}{}'') + $$
$$+ a_{1n}(x_n{}' - x_n{}'') = a_1;$$

……………………………………………………………..

$$a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn1}x_{n1} + a_{mn1+1}(x_{n1+1}{}' - x_{n1+1}{}'') +$$
$$+ a_{2n}(x_n{}' - x_n{}'') = a_m.$$

$$x_j \geq 0, \quad j = 1, 2, \ldots, n_1; \quad x_{n+i}{}' \geq 0; \quad x_{n+i}{}'' \geq 0, \quad i = 1, 2, \ldots, n - n_1.$$

У такому вигляді ЗЛП має канонічний вид.

3.3.3. Задача лінійного програмування у матричній і векторній формах

Введемо в ЗЛП (3.14) позначення

$$\vec{c} = (c_1, c_2, \ldots, c_n);$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}; \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}; \quad A_0 = \begin{pmatrix} a_1 \\ a_2 \\ \ldots \\ a_3 \end{pmatrix}.$$

Тоді канонічну ЗЛП можна записати у ***матричній формі***:

$$Cx \to \min(\max);$$
$$Ax = A_0; \quad x \geq 0,$$

або в більш компактній формі

$$\max_x \left(\min_x\right)\left(cx\,\Big|_{Ax = A_0;\, x \geq 0}\right).$$

Введемо позначення векторів A_i для стовпців матриці A

$$A_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \ldots \\ a_{m1} \end{pmatrix}; \quad A_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \ldots \\ a_{m2} \end{pmatrix}; \quad \ldots; \quad A_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \ldots \\ a_{mn} \end{pmatrix}.$$

Тоді ЗЛП (3.14) у ***векторній формі*** запису приймає вид

$$c_1x_1 + c_2x_2 + \ldots + c_nx_n \to \max(\min),$$
$$A_1x_1 + A_2x_2 + \ldots + A_nx_n = A_0, \quad x_j \geq 0, \quad j = 1, 2, \ldots, n.$$

Усі перераховані й інші види ЗЛП варто розглядати лише як різні форми запису загальної задачі лінійного програмування.

3.3.4. Графічний метод вирішення задач лінійного програмування

Спочатку розглянемо найпростіший випадок ЗЛП (для $n = 2$):

$$c_1x_1 + c_2x_2 \to max,$$
$$a_{11}x_1 + a_{12}x_2 \leq a_1,$$
$$\dots\dots\dots\dots\dots\dots$$
$$a_{m1}x_1 + a_{m2}x_2 \leq a_m.$$

Якщо ввести на площині декартову прямокутну систему координат то можна зіставити кожній парі чисел (x_1, x_2) точку площини з координатами x_1 і x_2 . З'ясуємо, що буде являти собою безліч точок, які відповідають припустимим рішенням задачі.

Кожне з обмежень визначає на площині одну з напівплощин на який пряма

$$a_{i1}x_1 + a_{i2}x_2 = a_i,$$

розбиває площину. При цьому відповідна напівплощина містить у собі граничну пряму (так звана “замкнута напівплощина“). Отже, кожне обмеження задачі ЛП задає на площині деяку напівплощину.

Щоб визначити припустиму множину рішень ЗЛП необхідно визначити всі точки площини, координати яких задовольняють усім обмеженням - нерівностям, тобто точки які належать усім m напівплощинам, які обумовлені окремими обмеженнями.

Отже геометричною інтерпретацією припустимої множини ЗЛП (для $n = 2$) є зображення перетинання напівплощин, які обумовлені окремими обмеженнями.

Будемо називати це перетинання ***припустимою областю ЗЛП***.

Припустима область і припустима множина ЗЛП може бути:

- порожньою;
- непорожньою і обмеженою;
- непорожньою і необмеженою.

При $n = 2$ ясно, що припустима область не може складатися з кінцевого числа точок (крім випадку, якщо це не одна точка, що теж можливо).

Якщо припустима область непорожня, то вона являє собою деякий багатокутник (може бути і необмежений). Будемо відносити до них і припустимі області, які виродилися у смугу, відрізок, точку.

Розглянемо геометричну інтерпретацію цільової функції

$$c_1x_1 + c_2x_2 = z\,(x_1, x_2).$$

Функція приймає одне саме значення a у всіх точках прямої

$$c_1 x_1 + c_2 x_2 = a,$$

де a – деяке дійсне число.

Якщо вважати a параметром, одержимо сімейство рівнобіжних прямих, які називають ***лініями рівня функції*** $z\,(x_1, x_2)$.

Значення a і значення цільової функції зростають необмежено, якщо переміщати пряму $c_1 x_1 + c_2 x_2 = a$ у напрямку її нормалі – вектора $\vec{c} = (c_1, c_2)$ і убувають, якщо переміщати в зворотному напрямку – напрямку вектора $\overrightarrow{-c} = (c_1, c_2)$.

Таким чином, щоб знайти оптимальне рішення задачі на максимум необхідно переміщати лінію рівня цільової функції $z\,(x_1, x_2) = c_1x_1 + c_2x_2$ у напрямку вектора $c = (c_1, c_2)$, починаючи з якого-небудь фіксованого положення, при якому вона перетинається із припустимою областю і доти поки вона не перестане перетинатися з нею. Перетинання припустимої області з лінією рівня у тім її положенні, коли подальше переміщення дає порожню множину і буде безліччю оптимальних рішень. У випадку задачі мінімізації лінію рівня необхідно переміщати у зворотному напрямку.

Отже, для ЗЛП можливі наступні випадки:

- вона має єдине рішення;
- вона не має рішень;
- вона має незліченну безліч рішень.

При $n = 2$ очевидно, що задача ЛП неодмінно має оптимальні рішення, якщо її припустима область непорожня й обмежена (тобто цілком міститься в колі з центром на початку координат досить великого радіуса). Однак можливий варіант, коли задача ЛП має оптимальні рішення тоді, коли припустима область необмежена.

Для $n = 2$ геометрично очевидно наступна необхідна і достатня умова існування оптимального рішення.

Задача ЛП на максимум (мінімум) тоді і тільки тоді має оптимальне рішення, коли її цільова функція обмежена зверху (знизу) у припустимій області.

Для $n = 2$ точки підозрілі на оптимум лежать на границі багатокутника (або вершина, або грань). Для випадку єдиного рішення дві нерівності – обмеження перетворюються у точну рівність.

Непорожня припустима область може мати вершини припустимої множини або не мати їх. Але якщо вершини є, то і серед оптимальних точок є хоча б одна вершина (якщо оптимальні точки взагалі є).

Отже, аналіз ЗЛП із двома змінними дозволяє сформувати наступні геометрично очевидні твердження:

- припустима область і множина оптимальних точок – об'ємні множини (якщо вони не порожні);
- обмеженість цільової функції у припустимій області (відповідно зверху або знизу) є необхідною і достатньою умовою можливості розв'язання задачі;
- оптимум задачі досягається у вершині припустимої області (якщо припустима область має вершини і задача може бути розв'язана).

Приклад 3.4. Визначити графічно максимум цільової функції

$$z = x_1 + 3x_2 + x_3 - 2, \quad (3.16)$$

при умовах

$$\begin{aligned} &-x_1 - 2x_2 + 2x_3 \geq -8, \\ &x_1 - 2x_2 - 4 \leq 0, \\ &x_1 - x_2 - x_3 + 2 = 0, \\ &x_j \geq 0 \ (j = 1, 2, 3). \end{aligned}$$

Розглянемо третє обмеження $x_1 - x_2 - x_3 + 2 = 0$.

Визначимо з нього x_1:

$$x_1 = x_2 + x_3 - 2. \quad (3.17)$$

Підставимо значення x_1 з (3.17) у цільову функцію (3.16):

$$z = 4x_2 + 2x_3 - 4. \quad (3.18)$$

З огляду на, що $x_1 \geq 0$, з (3.17) справедливо буде записати

$$x_2 + x_3 - 2 \geq 0.$$

Якщо підставити значення $x_1 = x_2 + x_3 – 2$ у друге і третє обмеження, то система обмежень приймає вид

$$\begin{aligned} &-3x_2 + x_3 + 10 \geq 0, \\ &-x_2 + x_3 - 6 \leq 0, \\ &x_2 + x_3 - 2 \geq 0, \\ &x_j \geq 0 \ (j = 2, 3). \end{aligned}$$

Введемо на площині прямокутну декартову систему координат і зіставимо кожної парі чисел (x_2, x_3) задачі точку площини з координатами x_2 і x_3.

Кожне з обмежень визначає на площині одну з напівплощин на які прямі

$$\begin{aligned} &-3x_2 + x_3 + 10 = 0, \\ &-x_2 + x_3 - 6 = 0, \\ &x_2 + x_3 - 2 = 0, \\ &x_2 = 0, \\ &x_3 = 0, \end{aligned}$$

розбивають площину. При цьому відповідна напівплощина містить у собі граничну пряму.

Побудуємо на площині ці прямі і визначимо припустиму область розглянутої задачі лінійного програмування, як перетинання напівплощин (рис.3.14). Припустима область задачі лінійного програмування в даному випадку не порожня й опукла.

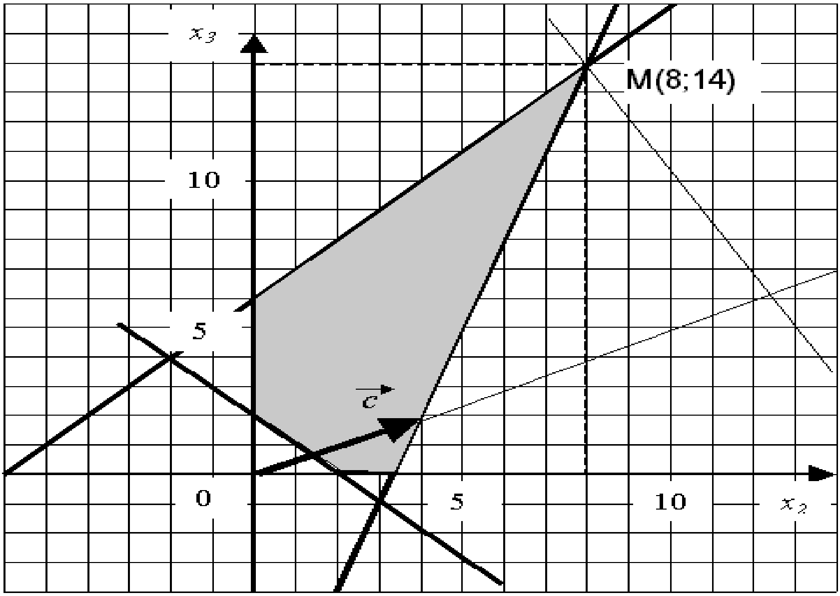


Рис. 3.14. Рішення ЗЛП графічним методом

Геометричною інтерпретацією цільової функції (3.18)

$$z = 4x_2 + 2x_3 - 4,$$

буде сімейство рівнобіжних прямих, так звані лінії рівня функції.

Щоб знайти максимальне значення цільової функції необхідно побудувати вектор $\vec{c} = (c_1, c_2)$ і переміщати нормаль до нього, починаючи з якого-небудь фіксованого положення, при якому вона перетинається із припустимою областю доти поки вона не перестане перетинатися з нею. Перетинання припустимої області із лінією рівня у тім її положенні, у якому подальше переміщення дає порожню множину і буде оптимальним рішенням.

На рис. 3.14 показано вектор $\vec{c} = (4;\ 2)$ і оптимальне рішення задачі точка ***M***. Для того, щоб визначити координати точки ***M*** (оптимальні значення x_2 і x_3) необхідно вирішити систему з двох рівнянь

$$-3x_2 + x_3 + 10 = 0,$$
$$x_2 + x_3 - 2 = 0.$$

До системи входять рівняння прямих, на перетинанні яких знаходиться точка ***M***. Рішенням системи є

$$x_2 = 8, \qquad x_3 = 14.$$

Підставимо ці значення змінних у (*3.17*) і одержимо значення $x_1 = 20$.

Далі необхідно підставити оптимальні значення $x_1 = 20$, $x_2 = 8$ і $x_3 = 14$ у вираження для цільової функції (3.16) або (3.18) і одержати оптимум $z = 56$. Дана задача лінійного програмування має одне рішення:

$$z = 56 \text{ при } x_1 = 20,\ x_2 = 8 \text{ і } x_3 = 14.$$

3.3.5. Рішення задачі лінійного програмування симплекс методом

Область припустимих рішень (ОПР) у двопараметрічній ЗЛП являє собою плоский багатогранник, а у загальному вигляді це опуклий багатогранник.

Теорема*:* екстремум цільової функції у ЗЛП, якщо він існує, завжди є абсолютним і досягається хоча б у одній крайній точці багатогранника, що визначає ОПР.

Примітка. Крайні точки - це точки перетинання границь ОПР.

Ідея симплекс методу полягає у спрямованому переборі крайніх точок ОПР. Цей метод є класичним у лінійному програмуванні.

Розглянемо алгоритм рішення задач ЛП симплекс-методом.
По перше необхідно записати задачу ЛП у вигляді:

$$z = c_1x_1 + c_2x_2 + \ldots + c_nx_n \to max,$$
$$0 = -a_{i1}x_1 - a_{i2}x_2 - \ldots - a_{in}x_n + a_i, \quad (i = \overline{1, n}),$$
$$y_i = -a_{i1}x_1 - a_{i2}x_2 - \ldots - a_{in}x_n + a_i \geq 0, \qquad (i = \overline{n+1, m}).$$

По друге - скласти та записати симплекс таблицю, у яку не входять обмеження типу $x_j \geq 0$.

Необхідно звернути увагу, що *0* - строки рівняння записуються у такому виді, щоб вільний член a_i був не негативний. Якщо у вихідній задачі $a_i < 0$, то рекомендується помножити на (*-1*) ліву та праву частини *0* – строки.

По третє - вилучити вільні змінні.

Для цього необхідно виконувати модифіковані жорданові виключення (МЖВ), вибираючи у якості дозволяючого будь-який елемент зі стовпця під вільною змінною, яку вилучають. Після виконання кроку МЖВ необхідно із отриманої нової симплекс-таблиці окремо виписати вирази для вільної змінної, яку вилучають, а рядок, що відповідає цій змінній із таблиці викреслити.

Четверте - виключити *0* – рядки.

Сипмлекс-таблиця має наступний вигляд (табл. 3.4.):

Таблиця 3.4.

		$-x_1$	…	$-x_s$	$-y_1$	…	$-y_{n-s}$	1
y_{n-s+1}	=	$b_{n-s+1,1}$	…	$b_{n-s+1,s}$	$b_{n-s+1,s+1}$	…	$b_{n-s+1,n}$	b_{n-s+1}
...		… … … … … … … … … … … …						
y_r	=	b_{r1}	…	b_{rs}	$b_{r,s+1}$	…	b_{rn}	b_r
0	=	$b_{r+1,1}$	…	$b_{r+1,s}$	$b_{r+1,s+1}$	…	$b_{r+1,n}$	b_{rn}
...		… … … … … … … … … … … …						
0	=	b_{m1}	…	b_{ms}	$b_{m,s+1}$	…	b_{mn}	b_m
Z	=	q_1	…	q_s	q_{s+1}	…	q_n	Q

Від ***k****-го 0* - рядка позбавляються в такий спосіб: береться стовпець, що містить позитивний елемент, якщо $b_{kl} > 0$, тоді дозволяючим вибирається ***l***-й стовпець. Для вибору дозволяючого рядка, обчислюють усі ненегативні співвідношення $\frac{b_i}{b_{il}} \geq 0$ вільного члена до коефіцієнтів дозволяючого стовпця (***l*** - го), знаходять серед них найменше, що досягається нехай при $i = i_0$. Тоді i_0 - й рядок беруть у якості дозволяючого, так що дозволяючим елементом буде $b_{i_0 l}$.

У випадку виродження, коли

$$\min_i \left\{ \frac{b_i}{b_{il}} \geq 0 \right\} = \frac{b_{i_0}}{b_{i_0 l}} = 0,$$

у якості дозволяючого елементу беруть $b_{i_0 l}$ лише якщо $b_{i_0 l} > 0$.

Якщо дозволяючим елементом виявився b_{kl}, тобто $i_0 = k$, то позбавляються і від ***k*** - го *0* - рядка, тобто викреслюють стовпець, що знаходиться під перекинутим наверх нулем і розмірність таблиці зменшується на одиницю.

Якщо ж $i_0 \neq k$, то продовжують виконувати кроки МЖВ, працюючи увесь час з ***k*** - м *0* - рядком (тобто вибираючи дозволяючий елемент зі стовпця, що містить позитивний коефіцієнт ***k***-го *0* - рядка) доти поки не позбудуться від ***k*** - го *0* - рядка, або у *0*-рядках не залишиться жодного позитивного елемента, при позитивному вільному члені, що означає несумісність системи.

На п'ятому кроці знаходиться опорне рішення. Для цього необхідно оцінити знаки елементів стовпця вільних членів. При цьому можливі два випадки.

Перший - всі елементи стовпця вільних членів ненегативні, необхідно дорівняти нулю всі незалежні змінні, а залежні відповідним елементам стовпця вільних членів. Підстановка отриманих значень у вираз для виключених вільних змінних дасть опорне рішення, що є вершиною багатогранника припустимих рішень.

Другий випадок – хоч би один елемент стовпця вільних членів негативний. У цьому випадку необхідно виконувати кроки МЖВ доти, поки усі вільні члени не стануть ненегативними. Для визначення дозволяючого елемента вибирають рядок з негативним елементом стовпця вільних членів (якщо їх декілька, то беруть найбільший по модулю) і оцінюють знаки елементів цього рядка. Якщо серед них немає негативного, то система несумісна. Якщо серед елементів рядка є негативні, то стовпець, у якому знаходиться цей елемент, вибирають дозволяючим. Наприклад, якщо в r-му рядку елемент $b_{rs} < 0$, тоді стовпець s вибирають дозволяючим. Обчислюють усі ненегативні відносини $\frac{b_i}{b_{is}} \geq 0$ вільних членів до відповідних, відмінним від нуля, коефіцієнтів дозволяючого стовпця. Знаходять серед них найменшу, яка досягається нехай при $i = i_0$. Тоді дозволяючим елементом буде $b_{i_0 s}$. У випадку виродження $\frac{b_{i_0}}{b_{i_0 s}} = 0$ у якості дозволяючого беруть тільки $b_{i_0 s} > 0$.

На шостому кроці визначається оптимальне рішення. Для цього необхідно оцінити знаки коефіцієнтів z – рядку. Можливі два варіанти. Перший - усі коефіцієнти z – рядку ненегативні $q_i \geq 0 \quad (i = \overline{1, n})$. У цьому випадку необхідно дорівняти нулю всі незалежні змінні, а залежні відповідним елементам стовпця вільних членів. Підстановка отриманих значень у вираз для виключених вільних змінних дасть оптимальне рішення, причому $\max z = Q$.

Другий випадок - деякі коефіцієнти z – рядку негативні, наприклад $q_s < 0$. Для переходу до наступної вершини багатогранника припустимих рішень необхідно виконувати кроки МЖВ і так доти, поки всі коефіцієнти z – рядку не будуть ненегативні.

Правило вибору дозволяючого елемента таке. Вибирають той стовпець, що містить негативний елемент z – рядка (якщо їх декілька, то береться більший по модулю). Складають співвідношення усіх вільних членів до позитивних елементів цього стовпця і вибирають з них найменше. Нехай воно досягається при $i = i_0$. Тоді i_0 – й рядок беруть дозволяючим, так що елемент $b_{i_0 s}$ - дозволяючий (якщо таких більш одиниці, те беруть будь-який з них). Виконують крок МЖВ. Після цього кроку або всі коефіцієнти z - рядку нової симплекс-таблиці будуть ненегативні й оптимальне рішення буде знайдено. Якщо оптимальне рішення не буде знайдено, то процедура повторюється, поки не прийдуть до випадку коли усі коефіцієнти z - рядку нової симплекс-таблиці будуть ненегативні, чи до випадку відсутності позитивних коефіцієнтів z – рядку, що означає необмеженість зверху припустимої множини рішень.

Приклад 3.5. Розглянемо приклад рішення симплекс-методом задачі лінійного програмування. Нехай необхідно визначити

$$z = x_1 + 3x_2 + x_3 \to max,$$

при умовах

$$-x_1 - 2x_2 + 2x_3 \geq -8,$$

$$x_1 - 2x_2 - 4 \leq 0,$$

$$x_1 - x_2 - x_3 + 2 = 0,$$

$$x_j \geq 0 \ (j = 1, 2, 3).$$

Запишемо вихідну задачу у виді:

$$z = x_1 + 3x_2 + x_3 - 2 \to max,$$

$$-x_1 - 2x_2 + 2x_3 + 8 \geq 0,$$

$$-x_1 + 2x_2 + 4 \geq 0,$$

$$x_1 - x_2 - x_3 + 2 = 0,$$

$$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0.$$

Складається симплекс-таблиця, до якої не входять обмеження типу $x_j \geq 0$.

	$-x_1$	$-x_2$	$-x_3$	1
$y_1 =$	1	2	-2	8
$y_2 =$	1	-2	0	4
$0 =$	-1	1	$\boxed{1}$	2
$z =$	-1	-3	-1	-2

Виконується виключення вільних змінних. У розглянутій задачі вільних змінних немає. На всі змінні накладено умову не негативності: $x_j \geq 0$ ($j = 1, 2, 3$).

Наступний етап рішення ЗЛП - виключення *0* – рядків. У вихідній симплекс-таблиці один *0*–рядок – третій. Для того, щоб позбутися від нього виберемо дозволяючим стовпцем третій - у ньому знаходиться позитивний елемент третього *0*-рядка ($b_{33} = 1 > 0$). Обчислимо всі ненегативні співвідношення елементів стовпця вільних членів до елементів цього третього стовпця. Всі інші елементи цього стовпця (крім $b_{33}= 1$) дорівнюють нулю. Тому дозволяючим елементом вибираємо $b_{33} = 1$. Дозволяючим елементом виявився елемент із *0*-рядка, тому позбавляються від третього *0*-рядка, тобто викреслюють стовпець, що знаходиться під перекинутим наверх нулем і розмірність таблиці зменшується на одиницю.

Симплекс-таблиця після виконання цього кроку МЖВ буде мати такий вигляд:

	$-x_1$	$-x_2$	*0*	1
$y_1 =$	1	4	2	12
$y_2 =$	1	-2	0	4
$x_3 =$	-1	1	1	2
$z =$	-2	-2	1	0

і після того, як викреслять третій стовпець отримуємо:

	$-x_1$	$-x_2$	1
$y_1 =$	1	4	12
$y_2 =$	$\boxed{1}$	-2	4
$x_3 =$	-1	1	2
$z =$	-2	-2	0

Далі визначається опорне рішення задачі лінійного програмування. Виконується оцінка знаків елементів стовпця вільних членів. Всі елементи стовпця вільних членів ненегативні. Для того щоб одержати опорне рішення необхідно дорівняти нулю всі незалежні змінні, а залежні відповідним елементам стовпця вільних членів. Опорне рішення задачі лінійного програмування має вигляд:

$$x_1 = -2,\ x_2 = -2 \text{ і } x_3 = 2.$$

Це рішення є вершиною багатогранника припустимих рішень. Цільова функція у цій вершині має значення $z = 0$.

Визначимо оптимальне рішення ЗЛП. Оцінимо знаки елементів z – рядка. Обидва елементи z–рядка симплекс-таблиці негативні. Для переходу до наступної вершини багатогранника припустимих рішень необхідно виконувати кроки МЖВ, і так доти, поки всі елементи z–рядка не стануть ненегативними.

Обидва стовпці z-рядка містять негативні елементи, які дорівнюють (**–2**), тому вибираємо один з них, наприклад перший. Складемо співвідношення усіх вільних членів до позитивних елементів цього стовпця і виберемо з них найменше. Не негативне співвідношення одне: **4/1**. Вибираємо у якості дозволяючого другий рядок і перший стовпець. Виконуємо крок МЖВ, після якого симплекс-таблиця придбає наступний вигляд:

	$-y_2$	$-x_2$	1
$y_1 =$	1	$\boxed{2}$	16
$x_1 =$	1	-2	4
$x_3 =$	1	-1	6
$z =$	2	-6	8

Після виконання цього кроку МЖВ один елемент ***z***–рядку усе ще негативний (елемент другого стовпця дорівнює (***–6***). Оберемо в якості дозволяючого другий стовпець і складемо співвідношення усіх вільних членів до позитивних елементів цього стовпця. Таке співвідношення в другому стовпці є одне (у першому рядку). Виконуємо крок МЖВ з першим дозволяючим рядком і другим дозволяючим стовпцем. Після виконання кроку МЖВ симплекс-таблиця буде мати наступний вид:

	$-y_2$	$-x_2$	1
$y_1 =$	0,5	0,5	8
$x_1 =$	2	1	20
$x_3 =$	1,5	0,5	14
$z =$	5	3	56

Після цього кроку всі елементи ***z***-рядка нової симплекс-таблиці будуть ненегативними й отже оптимальне рішення знайдене. Для цього необхідно дорівняти нулю всі незалежні змінні, а залежні - відповідним елементам стовпця вільних членів. Оптимальним рішенням буде $z = 56$ при $x_1 = 20$, $x_2 = 8$ і $x_3 = 14$.

Приклад 3.6. **З**адача про використовування ресурсів. Для здійснення різних технологічних процесів $T_1, ..., T_n$ заводу необхідно m видів ресурсів $S_1, ..., S_m$ (сировина, паливо, матеріали, і т. п.). Запаси ресурсів кожного виду обмежені і рівні $b_1, ..., b_m$. Відома витрата ресурсів на одиницю продукції по кожному технологічному процесу. Треба визначити, у якій кількості випускати продукцію кожного вигляду, щоб прибуток від реалізації цієї продукції був максимальним.

Позначимо через a_{ij} витрату ресурсів виду S_i на одиницю продукції вигляду T_j, а через c_j - дохід від реалізації одиниці продукції виду T_j. Всі наявні дані представимо у табл. 3.5, поклавши для конкретності $n = 3, m = 4$.

Позначимо через x_j кількість одиниць продукції виду T_j , що випускається. Обмеженнями у цій задачі є вимога, щоб витрата ресурсів виду S_i на випуск всіх видів продукції не перевищувала наявних запасів

$$\sum_j a_{ij} x_j \le b_i,\ i = \overline{1,m}. \qquad (3.19)$$

Ці обмеження легко перетворити на рівняння, ввівши змінні $x_{n+i} \ge 0$, що означають невикористані ресурси вигляду S_i. При цьому замість (3.19) одержимо:

$$\sum_j a_{ij} x_j + x_{n+i} = b_i,\ i = \overline{1,m}. \qquad (3.20)$$

Таблиця 3.5.

Запис початкових даних у задачі про використовування ресурсів

Види ресур-сів	Витрата ресурсів на одиницю продукції			Запаси ресурсів
	T_1	T_2	T_3	
S_1 S_2 S_3 S_4	a_{11} a_{21} a_{31} a_{41}	a_{12} a_{22} a_{32} a_{42}	a_{13} a_{23} a_{33} a_{43}	b_1 b_2 b_3 b_4
Прибутки від реалізації	c_1	c_2	c_3	-

Прибуток від реалізації випущеної продукції

$$q' = \sum_j c_j x_j. \qquad (3.21)$$

Оптимальним планом випуску продукції буде таке невід'ємне рішення системи рівнянь (3.20), при якому цільова функція (3.21) буде максимальна.

Приклад 3.7. Задача про розподіл випуску продукції по підприємствах**.** План галузі передбачає за час T випуск наступних видів продукції:

A_1 в кількості N_1 штук;
A_2 в кількості N_2 штук;
……………………………
A_n у кількості N_n штук.

Ці види продукції можуть випускатися на r однорідних підприємствах $П_1$, …, $П_r$. Припускаємо, що жодне підприємство не може одночасно випускати декілька видів продукції.

Крім того, задано:

a_{ij} - кількість продукції A_i, що випускається на підприємстві $П_j$ у одиницю часу;

b_{ij} - вартість одиниці продукції виду A_i випущеної на підприємстві $П_j$;

x_{ij} - час роботи підприємства $П_j$ по випуску продукції A_i.

Треба знайти такі значення x_{ij}, при яких вартість продукції, що випускається, буде мінімальною.

Обмеження:

1) час роботи кожного підприємства не повинен перевищувати T

$$\sum_i x_{ij} \le T,\ j = \overline{1,r}; \quad (3.22)$$

2) кількість продукції, що випускається, повинна відповідати номенклатурі

$$\sum_j a_{ij} x_{ij} = N_i,\quad i = \overline{1,n}. \quad (3.23)$$

Цільова функція буде загальною вартістю випущеної продукції. Якщо взяти до уваги, що величина $a_{ij}\ b_{ij}\ x_{ij}$ є вартістю частини продукції A_i, що випускається підприємством $П_j$, то загальна вартість продукції, що випускається

$$q = \sum_j \sum_i a_{ij} b_{ij} x_{ij}.$$

Згідно умовам задачі ця величина повинна бути мінімізована при виконанні обмежень (3.22) і (3.23).

3.4. Чисельні методи рішення задач нелінійного програмування

У задачі нелінійного програмування (НЛП) треба знайти значення багатовимірної змінної $x = (x^{(1)}, \ldots, x^{(n)})$, що мінімізує цільову функцію $f(x)$ за умов, коли на змінну x накладені обмеження типу нерівностей

$$f_i(x) \ge 0 \quad i = \overline{1,m}, \quad (3.24)$$

а змінні $x^{(i)}$, тобто компоненти вектора x, невід'ємні:

$$x^{(i)} \ge 0.$$

На відміну від лінійного програмування цільова функція $f(x)$ і обмеження $f_i(x)$ - нелінійні функції відносно x.

Такий підхід дозволяє істотно розширити клас задач, що вирішуються. Якщо в прикладі 1 покласти, що величина доходу c_i від продажу одиниці продукції виду T_i залежить від кількості продукції x_i, що продається, це часто зустрічається при оптовій торгівлі, то цільова функція $f(x)$ стане нелінійною функцією. Природно, що у цьому випадку для оптимізації цільової функції ***f(x)*** необхідно застосовувати методи нелінійного програмування, проте одержані результати можуть виявитися кращими у порівнянні із лінійним програмуванням.

Іноді у формулюванні задачі обмеження (3.24) мають протилежні знаки нерівностей, але завжди можна звести задачу НЛП до нерівностей одного знаку. Якщо деякі обмеження входять у задачу із знаком рівності, наприклад, $\varphi(x) = 0$, то їх можна представити у вигляді пари нерівностей $\varphi(x) \le 0$ та - $\varphi(x) \le 0$, зберігши тим самим типове формулювання задачі.

При рішенні задач оптимізації дуже важливо мати необхідні умови, яким повинна задовольняти точка, що оптимізує цільову функцію і задовольняє всім обмеженням. Хоча такі умови звичайно не є достатніми, але знання їх істотно полегшує пошук оптимального рішення.

Розроблено безліч чисельних методів рішення задачі нелінійного програмування, які можна класифікувати у наступним чином.

1. Чисельні методи пошуку екстремуму функції однієї змінної:

- класичний метод;
- метод рівномірного перебору;
- метод золотого перетину;
- метод Фібоначчі й т.д.

2. Чисельні методи пошуку екстремуму функції ***n*** – змінних.

2.1. Чисельні методи в завданнях без обмежень:

- метод покоординатного спуску;
- метод Хука – Дживса;
- градієнтний метод;
- метод Ньютона;
- метод сполучених напрямків і т.д.

2.2. Чисельні методи в завданнях з обмеженнями:

- метод покоординатного спуску;
- метод умовного градієнта;
- метод бар'єрних функцій;
- метод штрафних функцій;
- метод лінеаризації й т.д.

Універсального методу, за допомогою якого можна було б вирішити будь-яке завдання оптимізації, не існує. Тому для рішення конкретного завдання застосовують один або декілька чисельних методів.

3.4.1. Класичний метод мінімізації (максимізації) функції однієї змінної

Нехай $a \le x \le b$, функція $f(x)$ безперервна на цьому відрізку й має на ньому безперервну похідну. Необхідно обчислити значення похідної $f'(x)$ і визначають критичні точки, тобто такі внутрішні точки відрізку $[a, b]$, у яких похідна звертається у нуль або не існує. У околиці кожної такої критичної точки досліджують знак похідної й відбирають ті з них, при переході через які похідна міняє знак з мінуса на плюс (це точки локального мінімуму) або із плюса на мінус (це точки локального максимуму). Потім обчислюється значення цільової функції у цих точках і на границях відрізку $[a, b]$. Ці значення порівнюють між собою й визначають точку, у якій досягається мінімум (максимум) цільової функції. Ця точка є точкою глобального мінімуму (максимуму) функції $f(x)$ на відрізку $[a, b]$.

При вирішенні реальних завдань оптимізації даний метод застосовується рідко, тому що найчастіше похідну цільової функції визначити складно або неможливо.

3.4.2. Метод рівномірного перебору

Нехай дана функція $y = f(x) \rightarrow min$ (рис. 3.15). Відповідно до даного методу алгоритм пошуку оптимального значення багатовимірної змінної $x_{опт}$ полягає у наступному. Фіксується деяка величину кроку $h > 0$ після чого обчислюється значення цільової функції у точках $x_1 = a$ та $x_2 = x_1 + h$. Отримані значення порівнюють. Запам'ятовують менше із цих двох значень. Далі вибирається точка $x_3 = x_2 + h$ й у ній обчислюється значення цільової функції $f_3(x)$. Порівнюється значення, що залишилося на попередньому кроці, і значення $f_3(x)$. Найменше з них знову запам'ятовують. Так продовжують доти поки чергове значення x не буде більше за b. Останнє значення, що залишилося, є наближеним значенням глобального мінімуму.

Однак якщо при використанні даного методу цільова функція має вузьку западину, то її можна пропустити, і замість точки глобального мінімуму визначити точку локального мінімуму. Тобто. замість x' можна знайти x''. Ця проблема частково знімається, якщо вибрати дуже маленький крок, але це зажадає для рішення задачі більших витрат часу (у тому числі й машинного).

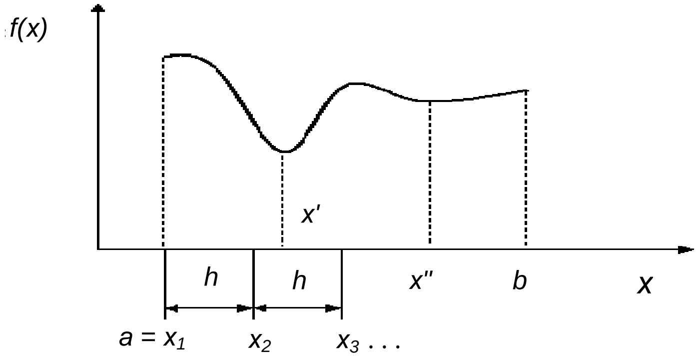


Рис. 3.15. Графічна ілюстрація методу рівномірного перебору

3.4.3. Метод лінеаризації (приведення задачі нелінійного програмування до задачі лінійного програмування)

Даний метод строго не відноситься до чисельних методів рішення задач оптимізації. Але він ефективний і часто використається для рішення практичних задач. Розглянемо суть даного методу на прикладі.

Приклад 3.8. Знайти $x_{1опт}$ й $x_{2опт}$, якщо цільова функція $y = x_1 / x_2, \rightarrow max$, при наступних обмеженнях:

$$1 \le x_1 \le 8;$$
$$2 \le x_2 \le 12;$$
$$x_1 \cdot x_2 \ge 10.$$

Приводимо дане завдання до задачі лінійного програмування. Для цього проводимо логарифмування обмежень і цільової функції:

$$lg\ 1 \le lg\ x_1 \le lg\ 8;$$
$$lg\ 2 \le lg\ x_2 \le lg\ 12;$$
$$(lg\ x_1 + lg\ x_2) \ge lg\ 10.$$

Після обчислень одержимо:

$$0 \le lg\ x_1 \le 0{,}903; \qquad (3.25)$$
$$0{,}301 \le lg\ x_2 \le 1,\ 079; \qquad (3.26)$$
$$lg\ x_1 + lg\ x_2 \ge 1. \qquad (3.27)$$

Після логарифмування цільової функції отримуємо:

$$lg\ y = lg\ x_2\ - lg\ x_1 \rightarrow max.$$

Далі задача вирішується із застосуванням симплекс - алгоритму або графо - аналітично (рис.3.16). Для побудови області допустимих рішень у логарифмічних координатах використовуємо обмеження (3.25) - (3.27). Обмеження (3.25) і (3.26) - це обмеження, що графічно представляють собою прямі лінії, паралельні відповідним осям $\lg x_2$ і $\lg x_1$. Причому, ліва обмежувальна лінія в обмеженні (3.25) збігається з віссю $\lg x_2$.

Обмеження (3.27) являє собою пряму лінію, що нахилена під кутом 45 градусів до осей, що має координати перетину осей "***0-1***". Для знаходження точки торкання лінії, що відповідає цільової функції, спочатку будується "довільна" лінія для цільової функції, її вираз дорівнюється до де якого довільного числа в даному масштабі. Дорівняємо вираз для цільової функції до числа ***1,2***:

$$\lg x_2 - \lg x_1 = 1{,}2 \quad \text{звідки} \quad \lg x_2 = 1{,}2 + \lg x_1.$$

Тоді задаючи довільні значення $\lg x_1$ отримуємо значення $\lg x_2$ (табл. 3.6).

Таблиця 3.6

$\lg x_1$	0	0,3
$\lg x_2$	1,2	1,5

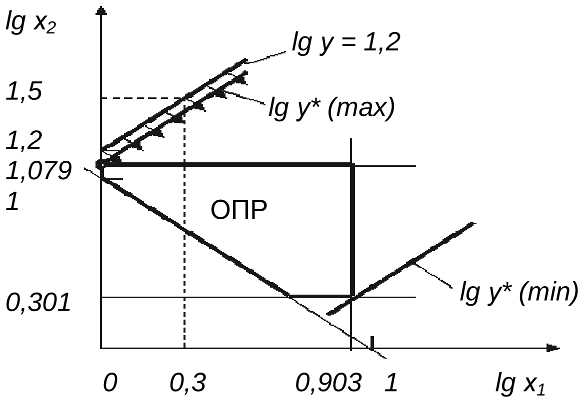


Рис.3.16. Графо - аналітичне рішення завдання оптимізації методом лінеаризації

Далі будується лінія, паралельна до даної лінії, що торкається границі ОПР. Знаходимо координати крапки торкання:

$$\lg x_{1опт} = 0 \text{ звідки } x_{1опт} = 1; \quad \lg x_{2опт} = 1{,}079 \text{ звідки } x_{2опт} = 12.$$

Якщо цільова функція

$$\lg y = \lg x_2 - \lg x_1 \rightarrow max,$$

то пряма лінія, що відповідає їй, торкнеться границі ОПР у точці з координатами:

$$\lg x_{1опт} = 0{,}903 \text{ звідки } x_{1опт} = 8 \text{ та } \lg x_{2опт} = 0{,}301 \text{ звідки } x_{2опт} = 2.$$

3.4.4. Метод покоординатного спуску у задачах без обмежень

Це задачі безумовної мінімізації, тобто мінімізації цільової функції $y = f(x_1, x_2, \ldots, x_n)$ на всьому просторі змінних (на всьому евклідовому просторі). Якщо потрібно вирішити задачу максимізації, то вираз цільової функції множать на (**-1**) і знову вирішується задача мінімізації.

Суть даного методу полягає у побудові послідовності точок $x^{(0)}$, $x^{(1)}, \ldots, x^{(n)}$, що монотонно зменшують значення цільової функції $f(x^{(0)}) \geq f(x^{(1)}) \geq \ldots \geq, f(x^{(n)})$.

Відповідно до цього методу напрямок спуску вибирається паралельно координатним осям, тобто спочатку спуск здійснюється уздовж першої вісі OX_1 потім уздовж другої осі OX_2 і т.д. до останньої осі OX_n.

Нехай $x^{(0)}$ - початкова точка (рис. 3.17), a - деяке позитивне число. Обчислюється значення цільової функції в цій точці - $f(x^{(0)})$. Далі обчислюється значення цільової функції при $x = x^{(0)} + a$ й перевіряється виконання нерівності

$$f(x^{(0)} + a) < f(x^{(0)}). \qquad (3.28)$$

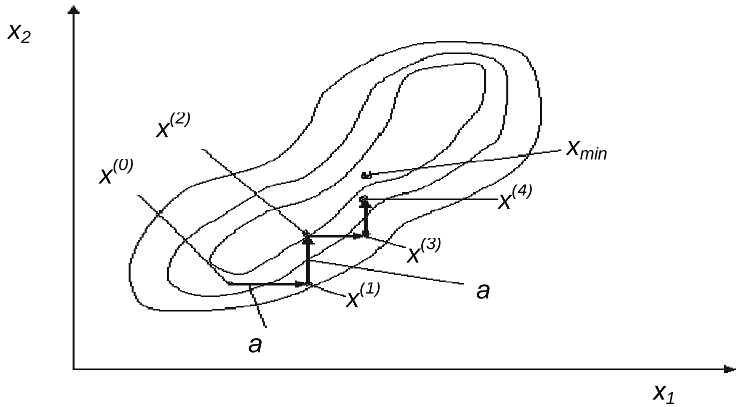


Рис.3.17. Графічна ілюстрація пошуку точки мінімуму методом покоординатного спуску

Якщо ця нерівність справедлива, то уздовж напрямку осі OX_1 значення функції f зменшилося, і тому $x^{(1)} = x^{(0)} + a$. Якщо нерівність (3.28) не виконується, то виконується крок у протилежному напрямку й перевіряють виконання нерівності

$$f(x^{(0)} - a) < f(x^{(0)}). \quad (3.29)$$

У випадку виконання цієї нерівності $x^{(1)} = x^{(0)} - a$. Якщо нерівності й (3.28), і (3.29) не виконуються, то $x^{(1)} = x^{(0)}$.

Далі уздовж координатної осі OX_2 обчислюється значення функції у точці ($x^{(1)} + a$), яке порівнюється із попереднім значенням, тобто перевіряється виконання нерівності

$$f(x^{(1)} + a) < f(x^{(1)}). \quad (3.30)$$

Якщо ця нерівність виконується, то вважають, що $x^{(2)} = x^{(1)} + a$. Якщо вона не виконується, то роблять крок у протилежному напрямку й перевіряють виконання нерівності

$$f(x^{(1)} - a) < f(x^{(1)}). \quad (3.31)$$

У випадку виконання нерівності (3.31) вважають, що $x^{(2)} = x^{(1)} - a$. Якщо нерівності (3.30) і (3.31) не виконуються, то вважають, що $x^{(2)} = x^{(1)}$.

Так перебирають всі n - напрямків координатних осей. На цьому перша ітерація закінчена. На n - м кроці буде отримана деяка точка $x^{(n)}$. Якщо $x^{(n)} \neq x^{(0)}$, то аналогічно, починаючи з $x^{(n)}$ здійснюють другу ітерацію. Якщо ж $x^{(n)} = x^{(0)}$ (це має місце, якщо на кожному кроці жодне з пари нерівностей не виявиться виконаним), то величину кроку потрібно зменшити, взявши, наприклад, $a_{n+1} = a_n/2$, і у наступній ітерації використати нове значення величини кроку.

Наступні ітерації виконуються аналогічно. На практиці обчислення припиняють при виконанні якої - не будь умови закінчення розрахунку, наприклад

$$\left| f(x)_{(k+1)} - f(x)_{(k)} \right| < \delta,$$

де $f(x)_{(k+1)}$ – значення цільової функції на ($к+1$) ітерації;

$f(x)_{(k)}$ – значення цільової функції на $к$ -й ітерації;

δ- деяке позитивне число, що характеризує точність рішення вихідного завдання мінімізації цільової функції.

3.4.5. Метод покоординатного спуску у задачах з обмеженнями

Даний метод поширюється на задачі, із простими обмеженнями типу:

$$a_1 \le x_1 \le b_1; \tag{3.32}$$
$$a_2 \le x_2 \le b_2; \tag{3.33}$$
$$\dots$$
$$a_2 \le x_2 \le b_2. \tag{3.34}$$

Основні процедури даного методу аналогічні попередньому методу. Розходження полягає в тому, що поряд з перевіркою виконання нерівностей

$$f(x^{(0)} + a) < f(x^{(0)})\ (f(x^{(0)} - a) < f(x^{(0)})),$$
$$f(x^{(1)} + a) < f(x^{(1)})\ (f(x^{(1)} - a) < f(x^{(1)}))\ \text{і т.д.}$$

здійснюють перевірку виконання нерівностей (3.32) - (3.34). Виконання або невиконання цих нерівностей приводить до тих же наслідкам, що й виконання або невиконання нерівностей, що наведені вище.

3.4.6. Квадратичне програмування

Задачею квадратичного програмування називається задача НЛП, у якій мінімізується сума лінійної і квадратичної форм при обмеженнях виду лінійних нерівностей і позитивності змінних. У матричній формі ця задача має вигляд:

$$\begin{cases} q(x) = cx + x'Dx = min \\ Ax \le b,\ x \ge 0. \end{cases}.$$

Припускається, що задача має m обмежень і n змінних, так що матриця A має розміри $m \times n$; матриця D - розміри $n \times n$; b є m - мірний вектор, x і c є n - мірними векторами. У розгорненій формі ці співвідношення мають вигляд:

$$q(x) = \sum_i c_j x^{(i)} + \sum_i \sum_k d_{jk} x^{(i)} x^{(k)} = min;$$

$$\sum_k a_{jk} x^{(k)} - b_j \le 0,\ j = \overline{1,m};\ x^{(i)} \ge 0,\ i = \overline{1,n}.$$

Для даної задачі, як і взагалі для задачі НЛП, ефективний обчислювальний метод можна знайти тільки в тому випадку, якщо цільова функція має єдиний оптимум, який і є глобальним. Звичайно обмежуються розглядом випадку, коли квадратична форма є позитивно визначеною, а значить, опуклою. Лінійна форма - також опукла функція. Тому цільова функція буде опуклою.

3.5. Контрольні питання

1. Наведіть класичну математичну постановку завдання оптимізації.
2. Дайте визначення поняттю цільова функція.
3. Дайте визначення поняттю глобальний мінімум (максимум) цільової функції.
4. Дайте визначення поняттю локальний мінімум (максимум) цільової функції.
5. Сформулюйте теорему Вейерштраса.
6. Наведіть алгоритм вирішення однопараметричного однокритеріального завдання оптимізації графоаналітичним методом.
7. Наведіть алгоритм вирішення багатопараметричного однокритеріального завдання оптимізації графоаналітичним методом.
8. Поясніть вирішення завдання оптимізації АСУ методом множників Лагранжа.
9. Поясніть вирішення завдання оптимізації АСУ методом найскорішого (градієнтного) спуску.
10. Що називається оптимальним рішенням задачі лінійного програмування?
11. Як перейти від канонічної форми ЗЛП до матричної?
12. Запишіть ЗЛП у векторно-матричній формі.
13. Як перейти до канонічної форми запису, якщо в вихідній задачі є відсутніми обмеження-рівняння, а є тільки обмеження-нерівності.
14. Як перевести до симетричної форми задачу лінійного програмування на максимум, яку записано у загальному виді?
15. Що є геометричною інтерпретацією обмежень двомірної задачі лінійного програмування?
16. Як визначається ОПР двомірної задачі лінійного програмування?
17. Що є геометричною інтерпретацією цільової функції двомірної задачі лінійного програмування?
18. Що є геометричною інтерпретацією цільової функції тримірної задачі лінійного програмування?
19. Перелічите послідовність дій при рішенні багатомірної задачі лінійного програмування ($n > 3$, $m - n = 2$) графічним методом.
20. Назвіть послідовність дій при виконанні кроку модифікованого жорданового виключення.
21. Назвіть послідовність дій при відшуканні опірного рішення системи лінійних нерівностей.

22. Як здійснити перехід від системи нерівностей до відповідній їй таблиці при використанні симплекс-методу?
23. Як одержати опорне рішення, якщо усі вільні члени таблиці – позитивні?
24. У чому полягає сутність класичного методу мінімізації (максимізації) функції однієї змінної.
25. Назвіть послідовність дій при відшуканні оптимального рішення системи лінійних нерівностей.
26. Надайте класифікацію чисельних методів рішення задачі нелінійного програмування.
27. Поясніть вирішення завдання оптимізації АСУ методом лінеаризації.
28. Поясніть вирішення завдання оптимізації АСУ методом рівномірного перебору.
29. Поясніть вирішення завдання оптимізації АСУ методом покоординатного спуску у задачах без обмежень .
30. Поясніть вирішення завдання оптимізації АСУ методом покоординатного спуску у задачах з обмеженнями.
31. Поясніть вирішення завдання оптимізації АСУ методом квадратичного програмування.

Розділ 4

Математичні методи аналізу імовірнісно – часових показників ефективності АСУ. Елементи теорії масового обслуговування

Розглянуті загальні поняття, характеристики та визначення теорії масового обслуговування, властивості найпростішого потоку подій. Розглянуто та дано визначення апарату марковських випадкових процесів. Досліджено властивості систем масового обслуговування з відмовами, чергою та змішаного типу. Наведено методи рішення задачі розрахунку СМО та контрольні питання.

4.1. Предмет теорії масового обслуговування

За останні десятиліття в самих різних областях практики виникла необхідність у рішенні своєрідних імовірнісних задач, що пов'язані із роботою так званих систем масового обслуговування (СМО). Більшість систем, з якими людина має справу, є стохастичними. Спроба їхнього математичного опису за допомогою детермінованих моделей приводить до перекручення дійсного стану речей. При рішенні завдань аналізу й проектування таких систем доводиться зважати на таке положення речей, коли випадковість є визначальною для процесів, що протікають у системах. При цьому зневага випадковістю, спроба "втиснути" рішення перерахованих завдань у детерміністичні рамки приводить до перекручування та до помилок у висновках і практичних рекомендаціях стосовно створення різноманітних систем управління різновидом яких є СМО.

Перші задачі теорії систем масового обслуговування (ТСМО) були розглянуті співробітником Копенгагенської телефонної компанії, датським вченим А.К. Ерлангом у період між 1908 і 1922 рр. Рішення цих задач припускало розробку методів, що дозволяють підвищити якість обслуговування споживачів телефонної мережі у залежності від числа пристроїв, що використовуються. Виявилося, що ситуації, які виникають на телефонних станціях, є типовими не тільки для телефонного зв'язку. Робота аеродромів, морських і річкових портів, магазинів, електронних обчислювальних комплексів, та інш. може бути описана в рамках ТСМО.

Розглянемо характерні приклади задач ТСМО.

Приклад 4.1. Система швидкої допомоги якогось міського району являє собою пункт (який приймає замовлення на виконання), який має деяку кількість автомашин швидкої допомоги й кілька лікарських бригад. Необхідно визначити кількість лікарів, допоміжного персоналу, автомашин, для того щоб час очікування виклику був для хворих оптимальним, за умови мінімізації витрат на експлуатацію системи й максимізації якості обслуговування.

Приклад 4.2. При організації морських перевезень вантажів необхідно забезпечити оптимальне використання судів і портових споруджень, певний обсяг перевезень при мінімальних витратах та скоротити простої судів при вантажно-розвантажувальних роботах.

Характерним для задач ТСМО, що вирішуються є:

- випадковий момент часу надходження замовлення на обслуговування (на телефонну станцію, на пункт швидкої допомоги);
- випадкова тривалість часу обслуговування;
- наявність черг: судів перед шлюзами, машин перед прилавками, завдань на вході процесорів обчислювального комплексу й інш.

Математично інформаційна структура системи масового обслуговування, до яких відноситься також і АСУ, задається її топологічною структурою:

- числом каналів зв'язку;
- пропускною здатністю каналів зв'язку;
- продуктивністю вузлів обробки інформації.

При цьому математичні методи, що використовуються для аналізу імовірнісно – часових показників ефективності різних систем масового обслуговування були запозичені із теорії телетрафику (теле – далеко, трафік – пересилати).

Теорія телетрафіку вивчає і розробляє методи аналізу і синтезу систем розподілу інформації й оперує не безпосередньо з такими системами, а із їх математичними моделями (рис. 4.1.), що включають у себе:

- вхідний потік замовлень на обслуговування;

- вихідні потоки замовлень;
- дисципліну обслуговування потоків замовлень.

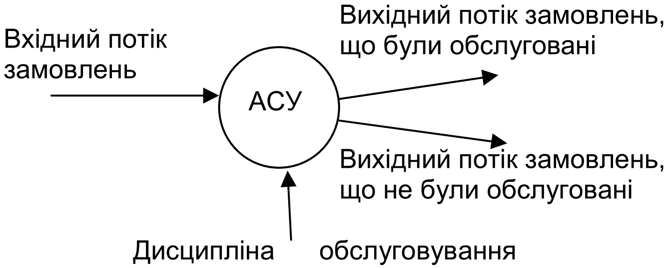


Рис.4.1. Модель вузла обробки інформації

Через незалежність функціонування джерел і споживачів інформації мережі АСУ (СМО) по її елементах передаються випадкові потоки інформації. Тому елементи мережі АСУ стосовно безлічі інформаційних повідомлень, що обробляються і передаються через них можуть розглядатися у якості ***однофазних СМО***. Сукупність елементів АСУ, які послідовно вступають у роботу при реалізації процесу управління можна розглядати, як ***багатофазну СМО*****.**

Задачі масового обслуговування виникають у тих випадках, коли вхідні потоки інформації (замовлення на виконання робіт) надходять на обробку у випадкові миті часу, а виконання цих робіт (***обслуговування***), здійснюється одним або декількома пристроями обробки інформації (***каналами обслуговування***). У якості каналів можуть бути: системи зв'язку; ЕОМ і т.п. Системи масового обслуговування можуть бути як одноканальними так і багатоканальними. На рис. 4.2. приведена загальна структура трьохканальної, однофазної системи масового обслуговування, а на рис. 4.3. багатоканальної, багатофазної СМО.

Таким чином, у теорії масового обслуговування розглядаються задачі планування процесу управління роботами по обробці потоків замовлень на обслуговування (повідомлень), що виникають випадково і вимагають різного, заздалегідь точно не передбачуваного часу для їхнього виконання. Типовими виробничими задачами такого роду є роботи з ремонту та налагодження устаткування.

Обробка повідомлення, що надійшло, продовжується якийсь час, після чого канал звільняється і знову готов для прийому наступного повідомлення. Кожна система масового обслуговування, у залежності від числа каналів і їхньої продуктивності, володіє якоюсь пропускною здатністю, що дозволяє їй більш-менш успішно справлятися з потоком повідомлень.

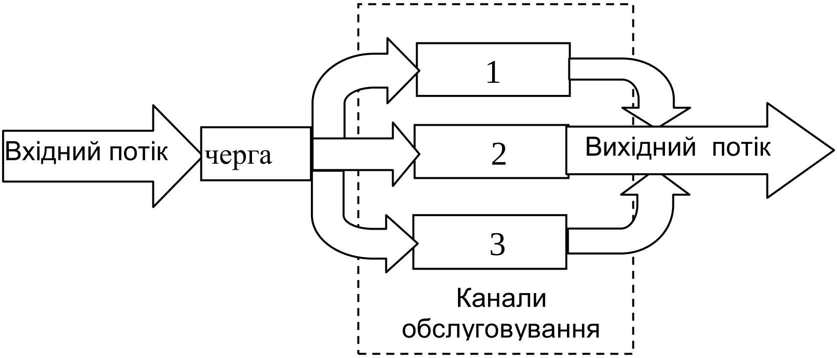


Рис. 4.2. Модель трьохканальної, однофазної СМО

Предметом теорії масового обслуговування є встановлення залежності між:

- характером потоку замовлень на обслуговування;
- продуктивністю окремого каналу;
- числом каналів;
- ефективністю обслуговування.

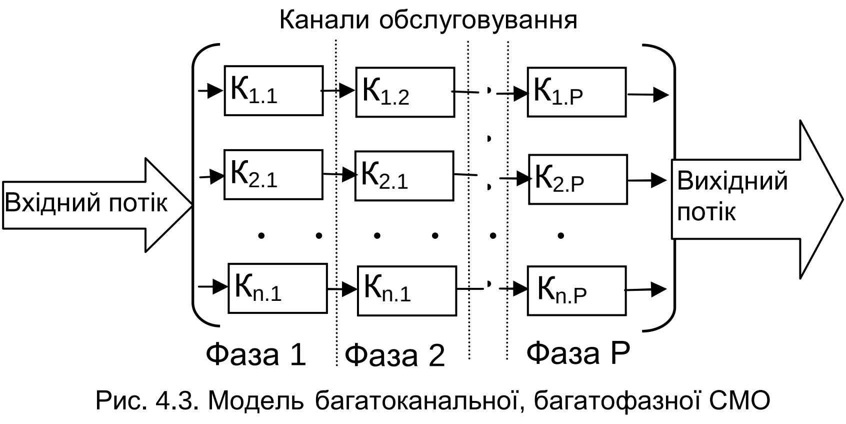


Рис. 4.3. Модель багатоканальної, багатофазної СМО

Принциповий інтерес представляють наступні ***характеристики системи масового обслуговування***:

- довжина черги в різні моменти часу;

- загальна тривалість перебування замовлення у системі обслуговування (очікування у черзі та безпосередньо обслуговування);
- частка часу, протягом якого обслуговуючі прилади були незайняті.

Кожна з цих характеристик описує, з тієї або іншої сторони пропускну здатність СМО. Можна виділити:

- ***абсолютну пропускну здатність системи*** - середнє число замовлень на обслуговування, що система може обробити в одиницю часу;
- ***відносну пропускну здатність*** - середнє відношення числа оброблених замовлень на обслуговування до числа замовлень, що надійшли.

Пропускна здатність (як абсолютна, так і відносна) у загальному випадку залежить не тільки від параметрів системи, але і від характеру потоку замовлень на обслуговування. Якби замовлення на обслуговування надходили регулярно, через точно визначені проміжки часу, і обслуговування кожного замовлення теж мало строго визначену тривалість, розрахунок пропускної здатності системи не представляв би ніяких труднощів. На практиці моменти надходження замовлень на обслуговування і тривалість їхньої обробки випадкові. У зв'язку з цим процес роботи системи протікає нерегулярно, у потоці замовлень на обслуговування утворюється неоднорідність. Підвищена концентрація замовлень на обслуговування може привести або до відмовлень в обслуговуванні, або до утворення черг. Знижена концентрація замовлень на обслуговування може привести до непродуктивних простоїв окремих каналів або системи у цілому. На ці випадки, зв'язані з неоднорідністю потоку замовлень на обслуговування, накладаються ще випадки, що пов'язані із затримками обробки окремих замовлень на обслуговування.

Крім закономірностей, що характеризують потік замовлень на обслуговування, для розрахунку систем масового обслуговування необхідно задатися тією або іншою дисципліною обслуговування. ***Дисципліна обслуговування*** визначає порядок, у якому дана система обробляє замовлення на обслуговування, що надходять. У ***системах з відмовами*** замовлення на обслуговування, що надійшло у момент, коли всі канали обслуговування зайняті, негайно одержує відмову, залишає систему і надалі у процесі обслуговування не приймає участі. У ***системах з очікуванням*** замовлення на обслуговування, що застало всі канали зайнятими, не залишає систему, а стає у чергу й очікує, поки не звільниться який-небудь канал. Зустрічається і змішаний випадок, коли при досягненні деякої довжини черги нові замовлення на обслуговування, що надходять до СМО губляться. Можливий випадок черги, що лякає, коли при збільшенні довжини черги зростає імовірність того, що замовлення на обслуговування, яке надійшло не встане в чергу і буде загублено.

На практиці часто застосовують обробку замовлень на обслуговування у порядку їх надходження. Можливі також випадки, коли замовлення на обслуговування розділяються по визначених пріоритетах, і обслуговування здійснюється в порядку пріоритетності замовлень. Зустрічаються і такі системи, обробка замовлень на обслуговування у яких здійснюється за принципом «останнє замовлення обслуговується першим».

Основним елементом дисципліни обслуговування є порядок, у якому система виділяє пристрої обслуговування. Це питання виникає у кратних системах, коли кожне із замовлень на обслуговування, що надійшли, може бути обслуговано одним з декількох пристроїв, які мають загальне призначення (наприклад, перукарня). У цьому випадку можливі три основні дисципліни обслуговування. По-перше, - ***циклічне обслуговування*,** коли перше замовлення на обслуговування надходить на перший пристрій, друге — на другий і т.д. Цей варіант найбільш простий для математичного аналізу, але відносно рідко зустрічається.

На практиці частіше застосовуються два інших варіанти. У першому із них утворюються різні черги до різних пристроїв, і в момент надходження нового замовлення на обслуговування приймається рішення, до якої з черг його приєднати. У другому варіанті є одна загальна черга, з якої замовлення на обслуговування надходять до першого з пристроїв, що звільнився.

У теорії СМО прийнятий наступний формалізований запис формули надання однофазних СМО: ***M|M|V|S|f.***

Символ на першому місці задає закон розподілу вхідного потоку замовлень на обслуговування (повідомлень).

Символ на другому місці задає закон розподілу часу обслуговування потоку замовлень на обслуговування. Тут використовуються наступні символи:

- ***M*** – при експоненціальному законі;
- ***E* -** при законі Ерланга;
- ***D* –** при детермінованому законі;
- ***G*** – при довільному законі.

Третій символ указує число обслуговуючих приладів, що працюють паралельно.

Четвертий символ указує припустиму довжину черги.

П'ятий символ визначає дисципліну обслуговування замовлень на обслуговування, що надходять. Обов'язковими є перші три символи.

Відмітимо, що до області застосування математичних методів теорії масового обслуговування відносяться багато задач автоматизації виробництва наприклад: робота складального конвеєра. Потоки деталей, що надходять для виконання над ними різних операцій, можуть розглядатися як «потоки замовлень», ритмічність надходження яких порушується за рахунок випадкових причин. Близькими до теорії масового обслуговування виявляються і задачі, що відносяться до надійності технічних пристроїв: такі їхні характеристики, як середній час безвідмовної роботи, необхідна кількість запасних деталей, середній час простою у зв'язку із ремонтом і т.д., визначаються методами, безпосередньо запозиченими з теорії масового обслуговування.

4.2. Випадковий процес із рахованою безліччю станів

Система масового обслуговування являє собою фізичну систему ***дискретного типу з кінцевою*** **(**або рахованою**)** ***безліччю станів*****.** Перехід системи із одного стану у інший відбувається стрибком, у момент, коли відбувається якась подія (прихід нової замовлення на обслуговування, звільнення каналу обслуговування, відправлення замовлення з черги і т.д.). Тому, випадковий процес, що протікає у системі масового обслуговування, полягає у тому, що система у випадкові моменти часу переходить з одного стану у інший: змінюється число зайнятих каналів обслуговування, число замовлень на обслуговування, що очікують у черзі, і т.д.

Розглянемо фізичну систему X з рахованою безліччю станів $x_1, x_2, \ldots x_n \ldots$

У будь-який момент часу **t** система **X** може бути в одному з цих станів x_k з імовірністю $p_k(t)$ $(k = 1,2,\ldots,n)$. Очевидно, для будь-якого моменту часу **t**

$$\sum_{k=1}^{n} p_k(t) = 1 . \tag{4.1}$$

Випадкові процеси із рахованою безліччю станів бувають двох типів: з дискретним і неперервним часом.

Випадкові процеси ***із дискретним часом*** відрізняються тим, що переходи зі стану у стан можуть відбуватися тільки у строго визначені, розділені кінцевими інтервалами моменти часу $t_1, t_2, \dots$

Випадкові процеси ***із неперервним часом*** відрізняються тим, що переходи системи зі стану у стан можливі у будь-який момент часу t.

Як приклад дискретної системи X, у якій протікає випадковий процес із неперервним часом, розглянемо гру у футбол з метою забити n голів у ворота супротивника. Ні момент початку атаки, ні моменти проведення ударів по воротам, заздалегідь невідомі. Різні стани системи відповідають різному числу забитих голів: x_0 — не забито жодного гола, x_1 - забитий рівно один гол, $x_к$ — забито рівно k голів, x_n — забито n голів.

Схема можливих станів системи і можливих переходів зі стану в стан показана на рис. 4.4.

Стрілками показані можливі переходи системи зі стану в стан.

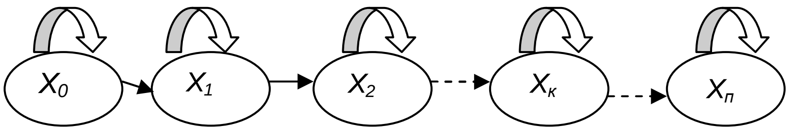


Рис. 4.4. Схема можливих станів системи і переходів СМО.

Закруглена стрілка (петля), спрямована зі стану x_k у нього ж, означає, що система може не тільки перейти в сусідній стан $x_{к+1}$, але і залишитися в колишньому. Для даної системи характерні незворотні переходи (наприклад, забитий гол).

Відзначимо, що на схемі можливих переходів (рис. 4.4.) показані тільки переходи з поточного стану в сусідній стан і не показані «перескоки» через стан.

Випадкові процеси, що протікають у системах масового обслуговування, як правило, являють собою процеси з неперервним часом. Це пов'язано із випадковістю потоку замовлень на обслуговування. Для СМО характерні зворотні переходи: зайнятий канал може звільнитися, черга може «розсмоктатися».

Як приклад розглянемо одноканальну систему масового обслуговування (наприклад, одну телефонну лінію), у якій замовлення на обслуговування, що застала канал зайнятим, не стає у чергу, а залишає систему (одержує «відмову»). Це - дискретна система із безперервним часом і двома можливими станами: x_0 — канал вільний,

x_1 — канал зайнятий. Переходи зі стану у стан - зворотні. Схема можливих переходів показана на рис.4.5.

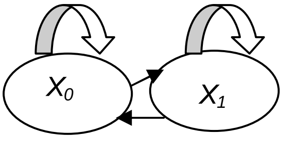


Рис. 4.5. Схема переходів

Розглянемо ще один приклад дискретної системи із неперервним часом: одноканальну СМО, що може знаходитися у чотирьох станах: ***x_0*** — канал справний і вільний, ***x_1***-канал справний і зайнятий, ***x_2*** — канал несправний і чекає ремонту, ***x_3*** — канал несправний і ремонтується. Схема можливих переходів для цього випадку показана на рис. 4.6. Перехід системи зі стану ***x_3*** безпосередньо у стан ***$x_г$****,* минаючи ***x_0****,* можна вважати практично неможливим, тому що для цього потрібно, щоб закінчення ремонту і прихід чергової замовлення відбулися строго одночасно.

Для того щоб описати випадковий процес, що протікає у дискретній системі із неперервним часом, насамперед, потрібно проаналізувати причини, що викликають перехід системи із стану в стан. Для системи масового обслуговування основним фактором, що обумовлює процеси, що протікають у ній, є потік замовлень. Тому математичний опис будь-якої системи масового обслуговування починається з опису потоку замовлень.

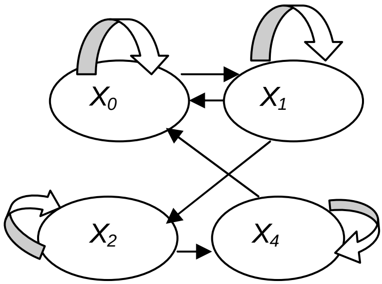


Рис. 4.6. Схема переходів

4.3. Потік подій. Найпростіший потік і його властивості

Під ***потоком подій*** розуміється послідовність подій, що відбуваються одна за одною у якісь моменти часу. Прикладами можуть служити: потік викликів на телефонній станції; потік включень приладів у побутовій електромережі; і т.п.

Події, що утворюють потік, у загальному випадку можуть бути різними, але тут ми будемо розглядати лише потік однорідних подій, що розрізняються тільки моментами появи. Потік замовлень називають ***однорідним*****,** якщо він задовольняє умовам:

- усі замовлення потоку з погляду обслуговування є рівноправними;

➢ замість подій, що по своїй природі можуть бути різними, розглядаються тільки моменти їхньої появи.

Такий потік можна зобразити як послідовність точок $\boldsymbol{t_l, t_2, .., t_k, ...}$ на числовій осі (рис. 4.7), що відповідають моментам появи подій.

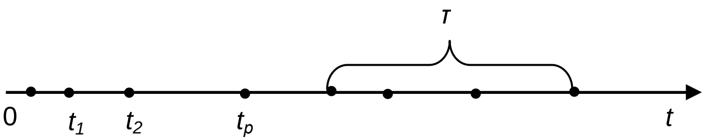


Рис. 4.7. Потік однорідних подій

Потік подій називається ***регулярним***, якщо події виникають одне за іншим через строго визначені проміжки часу. Такий потік порівняно рідко зустрічається в реальних системах, але становить інтерес як граничний випадок.

Типовим для системи масового обслуговування є випадковий потік замовлень.

Розглянемо потоки подій, що володіють деякими простими властивостями. Для цього введемо ряд визначень.

Потік подій називається ***стаціонарним***, якщо імовірність влучення того або іншого числа подій на ділянку часу довжиною $\boldsymbol{\tau}$ (рис. 4.7.) залежить тільки від довжини ділянки і не залежить від того, де саме на осі ***0,t*** розташована ця ділянка, тобто умові стаціонарності задовольняє потік замовлень, імовірності характеристики якого не залежать від часу. Зокрема, для стаціонарного потоку характерна постійна щільність (середнє число замовлень у одиницю часу). На практиці часто зустрічаються потоки замовлень (на обмеженому відрізку часу), що можуть розглядатися як стаціонарні. Наприклад, потік викликів на міській телефонній станції на ділянці часу від 12 до 13 годин може вважатися стаціонарним. Той же потік протягом цілої доби вже не може вважатися стаціонарним (вночі щільність викликів значно менше, ніж вдень). У багатьох задачах теорії масового обслуговування викликає інтерес аналіз роботи системи при постійних умовах, тоді задача вирішується для стаціонарного потоку замовлень.

Потік подій називається ***потоком без післядії***, якщо для будь-яких ділянок часу, що не перекриваються, число подій, які попадають на один з них, не залежить від числа подій, що попадають на інші. Умова відсутності післядії означає, що замовлення надходять у систему незалежно одна від одної. Наприклад, потік пасажирів, що входять на станцію метро, можна вважати потоком без післядії тому, що прихід окремого пасажира саме у той, а не інший момент, як правило, не зв'язаний із приходом інших пасажирів. Однак умова відсутності післядії може бути легко порушена за рахунок появи такої залежності. Наприклад, потік пасажирів, що залишають станцію метро, вже не може вважатися потоком без післядії, тому що моменти виходу пасажирів, які прибули тим самим потягом, залежні між собою. Післядію, що властива вихідному потоку, необхідно враховувати, якщо цей потік є вхідним для якої-небудь іншої СМО, або коли те саме замовлення послідовно переходить із системи у систему.

Відзначимо, що регулярний потік, у якому події відділені друг від друга рівними інтервалами, не є «найпростішим» у нашому сенсі слова, тому що в ньому мається яскраво виражена післядія: моменти появи подій, що слідують одне за одним, зв'язані твердою, функціональною залежністю. Саме через наявність післядії аналіз процесів, що протікають у системі масового обслуговування при регулярному потоці замовлень, набагато складніше, ніж при найпростішому.

Потік подій називається ***ординарним,*** якщо імовірність влучення на елементарну ділянку Δt двох або більше подій досить мала у порівнянні із імовірністю влучення однієї події. Умова ординарності означає, що замовлення на обслуговування приходять поодинці, а не парами, трійками і т.д. Наприклад, потік клієнтів, що входять у перукарню, може вважатися практично ординарним, чого не можна сказати про потік клієнтів, що направляються в ресторан. Якщо в неординарному потоці замовлення надходять тільки парами, тільки трійками і т.д., то неординарний потік легко звести до ординарного, для цього досить замість потоку окремих замовлень розглянути потік пар, трійок і т.д. Складніше буде, якщо кожне замовлення випадковим образом може виявитися подвійним, потрійним і т. д. Тоді вже маємо справу із потоком не однорідних, а різнорідних подій.

Якщо потік подій має всі три властивості (тобто стаціонарний, ординарний і не має післядії), то він називається ***найпростішим потоком*** (або стаціонарним пуассоновським).

Назва «пуассоновський» зв'язана із тим, що при дотриманні зазначених умов число подій, що попадають на будь-який фіксований інтервал часу, буде розподілено за законом Пуассона, тобто випадкова величина X розподілена за законом Пуассона, якщо імовірність того, що вона прийме визначене значення m виражається формулою

$$P_m = (a^m / m!)e^{-a}, \qquad (4.2)$$

де ***a*** – відповідає математичному очікуванню числа подій на ділянці від t_0 до $t_0 + \tau$.

Можна показати, що при підсумовуванні (взаємному накладенні) великого числа ординарних, стаціонарних потоків із практично будь-якою післядією отримаємо потік, як завгодно близький до найпростішого. Умова, що повинна для цього дотримуватися - потоки, що складаються, повинні робити приблизно рівномірно малий вплив на узагальнений потік.

Якщо потік є сумою великого числа ***n*** незалежних ординарних, стаціонарних потоків інтенсивності яких $\lambda_i \ (i = \overline{1,n}\ )$ і жоден з них не є порівнянним по потужності з усім сумарним потоком, то при деяких аналітичних обмеженнях сумарний потік сходиться до найпростішого з інтенсивністю

$$\lambda_{\Sigma} = \sum_{i=1}^{n} \lambda_i .$$

Проілюструємо це положення. Нехай мається ряд незалежних потоків $П_1,\ П_2,\ldots,П_n$. «Підсумовування» потоків полягає у тому, що всі моменти появи подій зносяться на загальну вісь ***0,t***, як показано на рис. 4.8.

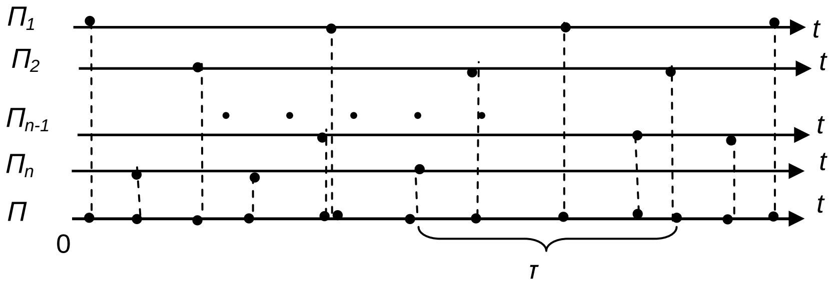


Рис. 4.8. Узагальнений потік однорідних подій

Припустимо, що потоки порівнянні за своїм впливом на сумарний потік (тобто мають щільності одного порядку), крім того ці потоки стаціонарні й ординарні, але кожний з них може мати післядію. Розглянемо сумарний потік на осі ***0,t***

$$П = \sum_{к=1}^{n} П_к .$$

Очевидно, що потік ***П*** повинен бути стаціонарним і ординарним, тому що кожен особистісний потік $П_к$ має цю властивість і вони незалежні. Крім того, досить ясно, що при збільшенні числа доданків $\Pi_к$ післядія в сумарному потоці, навіть якщо вона значна в окремих потоках, повинна поступово зменшуватися. На практиці, виявляється, звичайно досить скласти 4 - 5 потоків, щоб одержати потік, з яким можна оперувати як з найпростішим.

Найпростіший потік грає в теорії масового обслуговування важливу роль. По-перше, найпростіші і близькі до найпростішого потоки замовлень часто зустрічаються на практиці. По-друге, навіть при потоці замовлень, що відрізняється від найпростішого, часто можна одержати задовільні по точності результати, замінивши потік будь-якої структури найпростішим із тією же щільністю.

Розглянемо на осі ***0,t*** найпростіший потік подій ***П*** (Рис. 4.9) як необмежену послідовність випадкових точок.

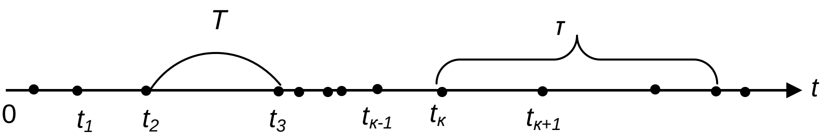


Рис. 4.9. Найпростіший потік подій

У найпростішому потоці виділимо довільну ділянку часу довжиною τ. Число точок, що попадають на ділянку τ, розподілено за законом Пуассона із математичним очікуванням

$$a = \lambda\tau, \qquad (4.3)$$

де λ– інтенсивність потоку замовлень (середнє число подій, що приходиться на одиницю часу або величина, зворотна середньому проміжку часу між замовленнями, що надходять).

Імовірність того, що за час τ відбудеться рівно m подій визначається виразом (4.2), а імовірність того, що ділянка виявиться порожньою (не відбудеться жодної події), буде

$$P_0 = e^{-a}\,. \qquad (4.4).$$

Важливою характеристикою потоку є ***закон розподілу довжини проміжку між сусідніми подіями***. Розглянемо випадкову величину T - проміжок часу між довільними двома сусідніми подіями в найпростішому потоці (рис. 4.9) і знайдемо її функцію розподілу

$$F(t) = P(T < \tau).$$

Перейдемо до імовірності протилежної події

$$1 - F(t) = P(T \geq \tau).$$

Це є імовірність того, що на ділянці часу довжиною τ, яка починається у момент появи одної із подій потоку $t_к$,, не з'явиться жодної з наступних подій. Тому що найпростіший потік не має післядію, то наявність на початку ділянки (у точці t_k) якоїсь події ніяк не впливає на імовірність появи тих або інших подій надалі. Отже знаючи, що

$$P_0 = e^{-\lambda t},$$

отримаємо

$$F(t) = 1 - e^{-\lambda t} \quad (t > 0).$$

Диференціюючи $F(t)$ знайдемо щільність розподілу

$$f(t) = \lambda e^{-\lambda t} \quad (t > 0). \tag{4.5}$$

Закон розподілу із такою щільністю називається показовим законом, а величина τ - його параметром. Графік щільності розподілу $f(t)$ представлений на рис. 4.10.

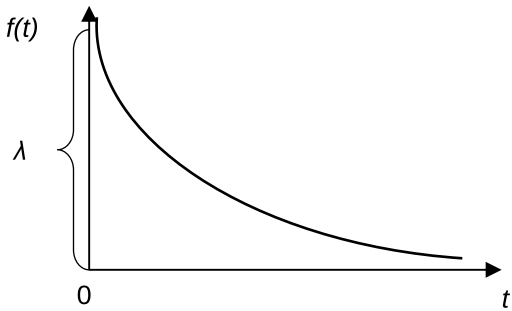


Рис. 4.10. Графік щільності розподілу ***f(t)***

Знайдемо математичне очікування величини T, розподіленої по показовому законі:

$$m_t = M[T] = \int_0^\infty t\,f(t)\,dt = \lambda\int_0^\infty te^{-\lambda t}\,dt, \tag{4.6}$$

або, інтегруючи по частинах

$$m_t = 1/\lambda. \tag{4.7}$$

Дисперсія величини T дорівнює:

$$D_t = D[T] = \int_0^\infty t^2 f(t)\,dt - \frac{1}{\lambda^2} = \int_0^\infty \lambda t^2 e^{-\lambda t}dt - \frac{1}{\lambda^2}. \tag{4.8}$$

Відкіля

$$D_t = 1/\lambda^2, \quad \sigma_t = 1/\lambda. \tag{4.9}$$

Згадаємо, що ***математичним очікуванням*** випадкової величини називається сума добутків усіх можливих значень випадкової величини на імовірності цих значень, а ***дисперсія*** випадкової величини характеризується відхилення окремих значень випадкової величини від математичного очікування тобто це характеристика розсіювання випадкової величини – середньоквадратичне відхилення.

4.4. Нестаціонарний пуассоновський потік

Розглянемо потік однорідних подій, ординарний і без післядії, але не стаціонарний, з змінною щільністю $\lambda(t)$. Такий потік називається ***нестаціонарним пуассоновським потоком***. Це перша ступінь узагальнення у порівнянні із найпростішим потоком (тут величина a залежить не тільки від довжини ділянки τ, але і від її положення на осі $0,t$).

Якщо потік подій є нестаціонарним, то його основною характеристикою є миттєва щільність $\lambda(t)$. ***Миттєвою щільністю*** потоку називається границя відношення середнього числа подій, що приходиться на елементарну ділянку часу $t,\ t + \Delta t$ до довжини цієї ділянки, коли остання прагне до нуля

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{m(t+\Delta t) - m(t)}{\Delta t} = m'(t), \tag{4.10}$$

де $m(t)$ — математичне очікування числа подій на ділянці $(0, t)$.

Знайдемо для нестаціонарного потоку закон розподілу проміжку часу T між сусідніми подіями. Через нестаціонарність потоку цей закон буде залежати від того, де на осі $0,t$ розташована перша із подій. Крім того, він буде залежати від виду функції $\lambda(t)$. Припустимо, що перша з двох сусідніх подій з'явилася у момент t і знайдемо при даній умові закон розподілу часу T між цією подією і наступною:

$$F_{t0}(t) = P(T < t) = 1 - P(T \ge t).$$

Знайдемо $P(T \ge t)$ - імовірність того, що на ділянці від t_0 до $t_0 + t$ не з'явиться жодної події:

$$P(T \ge t) = e^{-\alpha} = e^{-\int_{t_0}^{t_0+t} \lambda(t)\,dt}. \tag{4.11}$$

Відкіля

$$F_{t_0}(T \ge t) = 1 - e^{-\int_{t_0}^{t_0+t} \lambda(t)\,dt}. \tag{4.12}$$

Диференціюючи, знайдемо щільність розподілу

$$f_{t_0}(t) = \lambda(t_0 + t)\, e^{-\int_{t_0}^{t_0+t} \lambda(t)\,dt} \quad (t > 0). \qquad (4.13)$$

Цей закон розподілу вже не буде показовим. Його вид залежить від параметра ***t*** і виду функції ***λ(t)*** Наприклад, при лінійній зміні

$$\lambda(t) = a + bt,$$

щільність розподілу (4.13) має вигляд:

$$f_{t_0}(t) = [a + b(t_0 + t)]\, e^{-at - bt_0 t - \frac{bt^2}{2}}.$$

Графік для цього закону при $a = 0{,}4$, $b = 2$ і $t_0 = 0{,}3$ представлений на рис. 4.11.

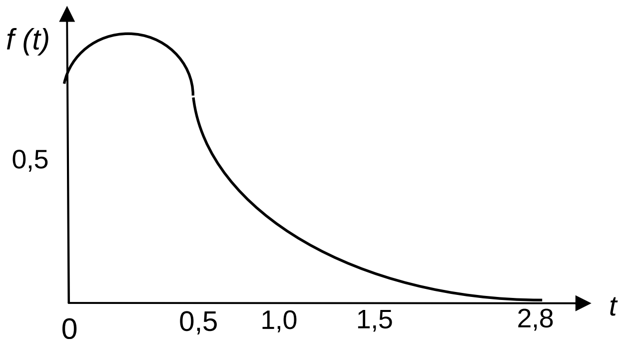


Рис. 4.11. Графік щільності розподілу ***f(t)*** для a = 0,4, b = 2 і t_0 = 0,3

Незважаючи на те, що структура нестаціонарного пуассоновського потоку трохи складніше, ніж найпростішого, він дуже зручний у практичних застосуваннях, головна властивість найпростішого потоку - відсутність післядії - у ньому збережено.

4.5. Потік з обмеженою післядією

Тепер коли ми вже познайомилися з природним узагальненням найпростішого потоку - з нестаціонарним пуассоновським потоком, розглянемо узагальнення найпростішого потоку в іншому напрямку - потік з обмеженою післядією.

Розглянемо ординарний потік однорідних подій (рис. 4,12). Цей потік називається ***потоком з обмеженою післядією*** (або потоком Пальма), коли проміжки часу між послідовними подіями $T_1, T_2, \ldots, T_M$. являють собою незалежні випадкові величини.

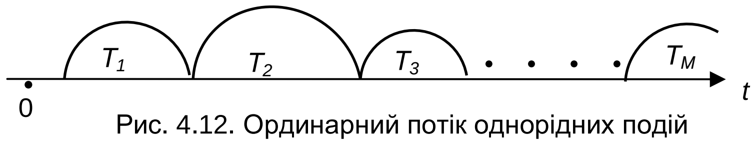


Рис. 4.12. Ординарний потік однорідних подій

Очевидно, найпростіший потік є особистісним випадком потоку Пальма: у ньому відстані $T_1, T_2,\ldots,T_M$ являють собою незалежні випадкові величини, що розподілені по показовому закону.

Що стосується нестаціонарного пуассоновського потоку, то він не є потоком Пальма. Дійсно, розглянемо два сусідніх проміжки T_k і T_{k+1} у нестаціонарному пуассоновському потоці. Як ми бачили раніше, закон розподілу проміжку між подіями у нестаціонарному потоці залежить від того, де цей проміжок починається, а початок проміжку T_{k+1} збігається з кінцем проміжку T_k; тобто виходить, довжини цих проміжків залежні.

Приклад потоку Пальма. Деякий елемент технічного пристрою (наприклад, тиристор) працює безупинно до свого відмовлення (виходу з ладу), після чого він миттєво заміняється новим. Термін безвідмовної роботи елемента випадковий, окремі екземпляри виходять із ладу незалежно друг від друга. При цих умовах потік відмов (або потік «відновлень») являє собою потік Пальма. Якщо, до того ж, термін роботи деталі розподілений по показовому закону, то потік Пальма перетворюється у найпростіший.

Потоки Пальма часто виходять у виді вихідних потоків систем масового обслуговування. Якщо на яку-небудь систему надходить якийсь потік замовлень, то він цією системою розділяється на два потоки замовлень: обслугованих і таких, що необслуговані.

Потік замовлень, що не обслуговані, часто надходить на яку-небудь іншу СМО, тому являє інтерес вивчити його властивості.

У теорії вихідних потоків важливу роль відіграє ***теорема Пальма***, яку ми сформулюємо и приведемо без доказу.

Нехай на систему масового обслуговування надходить потік замовлень типу Пальма, причому замовлення, що застало всі канали зайнятими, одержує відмовлення (не обслуговується), якщо при, цьому час обслуговування має показовий закон розподілу, то потік замовлень,що не обслуговані, є також потоком типу Пальма.

Зокрема, якщо вхідний потік замовлень буде найпростішим, то потік замовлень, що не обслуговані, не будучи найпростішим, буде все-таки мати обмежену післядію.

Цікавим прикладом потоків із обмеженою післядією є так звані ***потоки Ерланга***, які утворюються «просіванням» найпростішого потоку.

Розглянемо найпростіший потік подій (рис. 4.13) і викинемо з нього кожну другу точку (викинуті точки відзначені хрестами). точки, що залишилися, утворять потік, який називається ***потоком Ерланга першого порядку*** (E_1). Очевидно, цей потік є потоком Пальма: оскільки незалежні проміжки між подіями у найпростішому потоці, то незалежні і величини $T_1, T_2, \ldots,$ що отримуються підсумовуванням таких інтервалів по два.

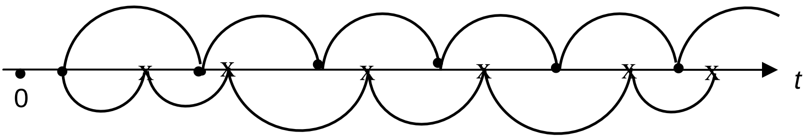


Рис. 4.13. Потік Ерланга першого порядку

Потік Ерланга другого порядку утворюється, якщо зберегти у найпростішому потоці кожну третю точку, а дві проміжні викинути.

Взагалі, ***потоком Ерланга k-го порядку*** ($E_к$) називається потік, який одержано із найпростішого, якщо зберегти кожну ***(к + 1)-*** у точку, а інші викинути. Очевидно, найпростіший потік можна розглядати як потік Ерланга нульового порядку (E_o).

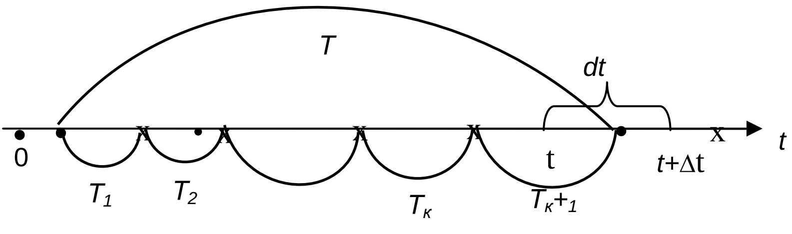


Рис. 4.14. Потік Ерланга ***к***-го порядку

Знайдемо закон розподілу проміжку часу ***T*** між сусідніми подіями в потоці Ерланга ***к***-го порядку. Розглянемо на осі ***0,t*** (рис. 4.14.) найпростіший потік з інтервалами $T_1, T_2, \ldots$. Величина ***T*** являє собою суму ***к + 1*** незалежних випадкових величин

$$T = \sum_{i=1}^{k+1} T_i ,$$

де $T_1, T_2, \ldots, T_{k+1}$ - незалежні випадкові величини, що підпорядковані показовому закону

$$f(t) = \lambda\, e^{-\lambda t} \quad (t > 0). \qquad (4.14).$$

Можна було б знайти закон розподілу величини T як композицію $k+1$ законів (4.14). Однак простіше вивести його елементарними міркуваннями.

Позначимо $f(t)$ щільність розподілу величини T для потоку E_k, тобто $f(t)dt$ - є імовірність того, що величина T прийме своє значення між точками t і $t + dt$ (рис. 4.14). Це значить, що остання точка інтервалу T повинна потрапити на елементарну ділянку $(t, t + dt)$, а попередні k точок найпростішого потоку - на ділянку $(0, t)$. Імовірність першої події дорівнює λdt; імовірність другої, буде

$$P_k(t) = [(\lambda t)^k / k!] e^{-\lambda t} .$$

Перемножуючи ці імовірності, одержимо

$$f_k(t)dt = [\lambda(\lambda t)^k / k!] e^{-\lambda t} dt ,$$

або

$$f_k(t) = [\lambda(\lambda t)^k / k!] e^{-\lambda t} \quad (t > 0). \qquad (4.15).$$

Закон розподілу із щільністю (4.15) називається ***законом Ерланга k-го порядку***. Очевидно, що при $k = 0$ він звертається у показовий

$$f_0(t) = \lambda e^{-\lambda t} \quad (t > 0).$$

Знайдемо характеристики закону Ерланга $f_k(f)$: математичне очікування m_k і дисперсію D_k. Згідно теореми додавання математичних очікувань

$$m_k = \sum_{i=1}^{k+1} m_0 = (k+1)\, m_0 ,$$

де $m_0 = 1/\lambda$ — математичне очікування інтервалу між подіями в найпростішому потоці.

Звідси

$$m_k = (k + 1)/\lambda .$$

Аналогічно, згідно теореми додавання дисперсій

$$D_k = \frac{k+1}{\lambda^2} \quad \text{та} \quad \sigma_k = \sqrt{\frac{k+1}{\lambda}} .$$

Щільність Λ_k, потоку E_k буде зворотна величині m_k тобто

$$\Lambda_k = \lambda/(k + 1).$$

Таким чином, при збільшенні порядку потоку Ерланга збільшуються як математичне очікування, так і дисперсія інтервалу часу між подіями, а щільність потоку зменшується.

З'ясуємо, як буде змінюватися потік Ерланга при $k \to \infty$, якщо його щільність буде зберігатися постійної. Пронормуємо величину T так, щоб її математичне очікування (і, отже, щільність потоку) залишалося незмінним. Для цього змінимо, масштаб по осі часу і замість T розглянемо величину

$$\check{T} = T/(k + 1).$$

Назвемо такий потік ***нормованим потоком Ерланга k-го порядку.*** Закон розподілу інтервалу T між подіями цього потоку буде

$$\tilde{f}_k(t) = \frac{\Lambda_k (\Lambda_k t)^k}{k!} e^{-\Lambda_k t} \quad (t > 0), \quad \Lambda_k = \lambda (k+1),$$

або

$$\tilde{f}_k(t) = \frac{\lambda_k (k+1)^k}{k!} (\lambda (k+1) t)^k e^{-\lambda (k+1) t} \quad (t > 0). \tag{4.16}$$

Математичне очікування величини T, розподіленої за законом (4.16), не залежить від k і дорівнює

$$\tilde{m}_k = m_0 = 1/\lambda,$$

де λ— щільність потоку, що збігає, при будь-якому k, із щільністю вихідного найпростішого потоку.

Дисперсія величини T дорівнює

$$\tilde{D}_k = \frac{D_k}{k+1} = \frac{1}{\lambda^2 (k+1)}, \tag{4.17}$$

і необмежено зменшується із зростанням k.

Таким чином, ми дійдемо наступного висновку: при необмеженому збільшенні k нормований потік Ерланга наближається до регулярного потоку з постійними інтервалами, рівними $1/\lambda$.

Ця властивість потоків Ерланга дає можливість, задавати різні k і одержувати будь-яку ступінь післядії: від повної її відсутності $(k = 0)$ до твердого функціонального зв'язку між моментами появи подій $(k = \infty)$. Таким чином, порядок потоку Ерланга може служити як би «мірою післядії», що має місце у потоці.

З метою спрощення часто буває зручно замінити реальний потік замовлень, що має післядію, нормованим потоком Ерланга з приблизно тими ж характеристиками інтервалу між замовленнями: математичним очікуванням і дисперсією.

Приклад 4.3. У результаті статистичної обробки інтервалів між замовлення у потоці були отримані оцінки для математичного очікування і диспепсії величини $\boldsymbol{T}$ (хв.): $\boldsymbol{m_t = 2; D_t = 0.8.}$

Замінити цей потік нормованим потоком Ерланга із тими ж характеристиками.

Рішення. Маємо $\boldsymbol{\lambda = 1/ m_t = 0.5}$.

З формули (4.17) одержимо $\boldsymbol{k + 1 = (1/D_t\lambda^{2)} = 1/0.2 = 5,}$ звідкіля $\boldsymbol{k = 4.}$

Потік можна приблизно замінити нормованим потоком Ерланга четвертого порядку.

4.6. Час обслуговування

Однієї з найважливіших величин, що зв'язані із системою, є час обслуговування одного замовлення $\boldsymbol{T_{об}}$. Ця величина може бути як невипадкової, так і випадкової. Очевидно, більш загальним є випадковий час обслуговування.

Розглянемо випадкову величину $\boldsymbol{T_{об}}$ і позначимо $\boldsymbol{G(t)}$ її функцію розподілу:

$$G(t) = P(T_{об} < t),$$

а $\boldsymbol{g\ (t)}$ щільність розподілу:

$$g\ (t) = G'(t).$$

Для практики особливий інтерес представляє випадок, коли величина $\boldsymbol{T_{об}}$ має показовий розподіл

$$g\ (t) = \mu e^{-\mu t} \quad (t > 0).$$

де параметр $\boldsymbol{\mu}$. — величина, зворотна середньому часу обслуговування одного замовлення:

$$\mu = 1/ m_{t\ об}, \quad m_{t\ об} = M[T_{об}].$$

Особлива роль, яку грає в теорії масового обслуговування показовий закон розподілу величини $\boldsymbol{T_{об}}$, пов'язана із наступною ***властивістю показового закону***: якщо у якийсь момент $\boldsymbol{t_0}$ відбувається обслуговування замовлення, то закон розподілу часу обслуговування, що залишився, не залежить від того, скільки часу обслуговування вже продовжувалося.

На перший погляд у ряді практичних задач, здається природним припустити час обслуговування або зовсім не випадковим, або розподіленим по нормальному закону. Однак існують умови, у яких час обслуговування дійсно розподіляється за законом, близьким до показового.

Це, насамперед, усі задачі, у яких обслуговування зводиться до ряду «спроб», кожна з яких приводить до необхідного результату з якоюсь імовірністю ***p*.** До такого типу задач можна віднести обслуговування по усуненню несправностей технічних пристроїв, коли пошук несправної деталі або вузла здійснюється послідовністю тестів або перевірок. До такого ж типу можна віднести задачі, де «обслуговування» полягає у виявленні якого-небудь об'єкта радіолокатором, якщо об'єкт із якоюсь імовірністю може бути виявлений при кожному циклі огляду.

Показовим законом добре описуються і ті випадки, коли щільність розподілу часу обслуговування по тим або інших причинах убуває при зростанні аргументу ***t*.** Це буває, коли основна маса замовлень обслуговується дуже швидко, а значні затримки в обслуговуванні спостерігаються рідко.

Приклад 4.4. Розглянемо роботу поштового відділення, де продають марки та конверти, а також приймають поштові відправлення і переклади. Основна маса відвідувачів купує марки або конверти й обслуговується дуже швидко. Рідше зустрічаються замовлення на відправлення замовлених листів, вони обслуговуються трохи довше. Переклади посилаються ще рідше й обслуговуються ще довше. Нарешті, у самих рідких випадках представники організацій відправляють відразу велику кількість листів. Графік розподілу часу обслуговування має вигляд, представлений на рис. 4.16. Тому що щільність розподілу убуває зі зростанням ***t*,** можна без особливої погрішності вирівняти розподіл за допомогою показового закону, підібравши відповідним чином його параметр ***μ*.**

Зрозуміло, що показовий закон не є універсальним законом розподілу часу обслуговування. Часто час обслуговування краще описується, наприклад, законом Ерланга. Однак, пропускна здатність і інші характеристики системи масового обслуговування порівняно мало залежать від виду закону розподілу часу обслуговування, а, головним чином, залежать від його середнього значення $m_{t\ об}$ Тому у теорії масового обслуговування найчастіше користуються допущенням, що час обслуговування розподілений по показовому закону. Ця гіпотеза дозволяє спростити математичний апарат, що застосовується для рішення задач масового обслуговування, і, у ряді випадків, одержати прості аналітичні формули для характеристик пропускної здатності системи.

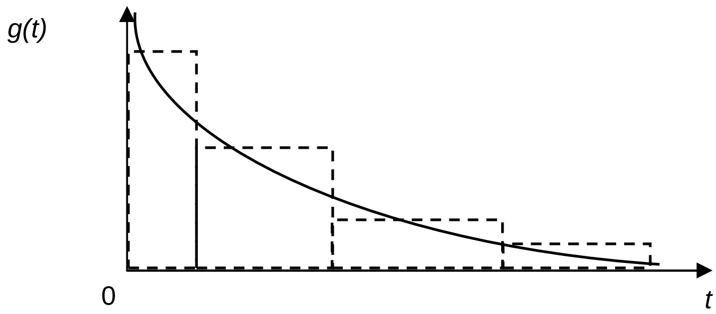


Рис. 4.16. Графік розподілу часу обслуговування

4.7. Марковський випадковий процес

Процес, що протікає у фізичній системі, називається ***марковським*** (або процесом без післядії), якщо для кожного моменту часу імовірність будь-якого стану системи у майбутньому залежить тільки від стану системи у даний момент ***(t_0)*** і не залежить від того, яким чином та коли система прийшла у цей стан.

Розглянемо елементарний приклад марковського випадкового процесу. По осі абсцис ***0,X*** випадковим образом переміщується точка ***X***. У момент часу ***t = 0*** точка ***X*** знаходиться на початку координат ***(X = 0)*** і залишається там протягом однієї секунди. Через секунду кидається монета. Якщо випав герб - точка ***X*** переміщається на одну одиницю довжини вправо, якщо цифра — вліво. Через секунду знову кидається монета і виконується таке ж випадкове переміщення і т.д. Процес зміни положення точки являє собою випадковий процес із дискретним часом ***(t = 0, 1, 2,…)*** і рахованою безліччю станів $x_0 = 0,\ x_1 = 1,\ x_{-1} = -1,\ldots$ Схема можливих переходів для цього процесу показана на рис. 4.17.

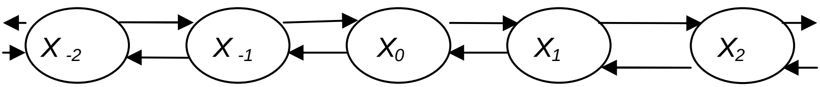


Рис. 4.17. Схема можливих переходів

Розглянемо інший приклад. Мається технічний пристрій ***X***, що складається з елементів типів ***a*** і ***b***, які мають різну довговічність. Ці елементи у випадкові моменти часу і незалежно друг від друга можуть виходити з ладу. Час безвідмовної роботи елементу - випадкова величина, що розподілена по показовому закону. Для елементів типу ***a*** і ***b*** параметри цього закону різні і рівні відповідно λ_a і λ_b. У випадку відмовлення пристрою виявлений несправний елемент негайно замінюється новим. Час, необхідний для ремонту пристрою, розподілено по показовому закону із параметром μ_a (якщо вийшов із ладу елемент типу ***a***) і μ_b (якщо вийшов із ладу елемент типу ***b***).

У даному прикладі випадковий процес, що протікає у системі, є марковським процесом із неперервним часом і кінцевою безліччю станів:

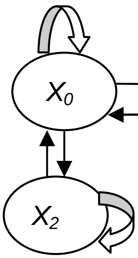


Рис. 4.18. Схема можливих переходів

x_0 - всі елементи справні, система працює;

x_l - несправний елемент типу ***a***, система ремонтується;

x_2 - несправний елемент типу ***b***, система ремонтується.

Схема можливих переходів дана на рис. 4.18.

Дійсно, процес є марковським.

Нехай наприклад, у момент t_0 система знаходиться у стані x_0. Так як час безвідмовної роботи кожного елементу підлягає показовому закону, то момент відмови кожного елементу у майбутньому не залежить від того, скільки часу він вже працював (коли встановлений). Тому імовірність того, що у майбутньому система залишиться в стані x_0 або піде з нього, не залежить від «передісторії» процесу.

Відмітимо, що показовий розподіл часу роботи елементу і показовий розподіл часу ремонту - істотні умови, без яких процес не був би марковським. Дійсно, припустимо, що час справної роботи елементу розподілено не по показовому закону, а по якому-небудь іншому - наприклад, за законом рівномірної щільності на ділянці (t_1,t_2). Це визначає, що кожен елемент гарантовано працює час t_1 а на ділянці від t_1 до t_2 може вийти з ладу в будь-який момент з однаковою щільністю імовірності. Припустимо, що у якийсь момент часу t_0 елемент працює справно. Очевидно, імовірність того, що елемент вийде із ладу на якійсь ділянці часу у майбутньому, залежить від того, наскільки давно поставлений елемент, тобто залежить від передісторії, і процес не буде марковським.

Аналогічно обстоїть справа із часом ремонту T_p; якщо він не показовий й елемент у момент t_0 ремонтується, то час ремонту, що залишився, залежить від того, коли він почався, процес знову не буде марковським.

У випадку, коли процес, що протікає у фізичній системі із рахованою безліччю станів і неперервним часом, є марковським, цей процес можна описати за допомогою звичайних диференціальних рівнянь, у яких невідомими функціями є імовірності станів $p_1(t)$, $p_2(t),\ldots$

Імовірності переходу можна представити двома різними способами. Перший спосіб складається у запису ймовірностей, переходу у вигляді таблиці, що має назву ***матриця переходів*** і позначається через ***P***. Для $n = 3$ матриця має вигляд:.

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$$

Виконання у даному випадку умови (4.1) означає, що сума елементів будь-якого рядка у матриці переходів обов'язково повинна дорівнювати одиниці.

Другий спосіб представлення ймовірностей, переходу складається у побудові діаграми переходів, що представляє собою граф вершини якого відповідають станам системи, а спрямовані дуги вказують можливі переходи від одного стану до іншого. Імовірності переходів відзначаються числами, що приписані до кожної дуги. Відповідно до умови (4.1) сума ймовірностей, для дуг, що виходять із будь-якої вершини графа, повинна дорівнювати одиниці. Приклад діаграми представлений на рис.4.19.

Іноді марковський випадковий процес визначається як однорідний марковський ланцюг. До марковських ланцюгів можуть бути зведені багато процесів із різних областей науки і техніки. У техніці за допомогою марковських ланцюгів описуються деякі процеси передачі інформації, ряд технологічних процесів, процес пошуку несправностей у складних системах.

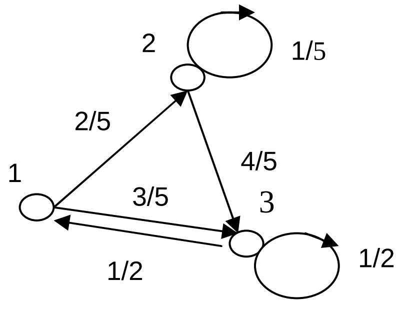


Рис. 4.19. Діаграма переходів

Приклад 4.5. Завод випускає визначений тип комп'ютерів. У залежності чи знаходить дана продукція попит завод може виявитися в двох станах: *1* – попит є; *2* – попиту немає. Згодом попит змінюється так , що має імовірність *4/5* і під кінець року завод залишиться у стані *1*. Якщо завод перейде у стан *2* то будуть прийматися міри щодо модернізації комп'ютерів, що випускаються, такі, що із імовірністю *3/5* до кінця наступного року завод перейде у стан *1*. Отже розвиток виробництва можна представити марковським ланцюгом із наступною матрицею переходів:

$$P = \begin{bmatrix} 4/5 & 1/5 \\ 3/5 & 2/5 \end{bmatrix},$$

та діаграмою переходів, що представлені на рис. 4.20.

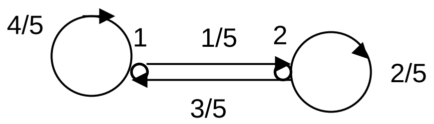


Рис. 4.20. Діаграма переходів

Тепер з'ясуємо як буде змінюватися розподіл ймовірностей, станів p_n коли система робить один крок. Позначимо через A подію, що складається у тому, що на $n + 1$ кроці система буде у стані j. Отже імовірність цієї події буде дорівнює $P(A) = p^{(j)}_{n+1}$.

Позначимо через S_i подію, що складається в тому, що на n - му кроці система знаходиться у стані i, так що $P(S_i) = p^{(i)}_n$. При цьому p_{ij} буде являти собою імовірність того, що система на $(n + 1)$ кроці буде знаходитися у стані j, якщо на n кроці вона була у стані i, тобто одержуємо умовний розподіл ймовірностей, $p_{ij} = p(A|S_i)$. Тоді відповідно до формули повної імовірності

$$P(A) = \sum_i P(A/S_i)\,P(S_i), \qquad (4.18)$$

Знаходимо, що:

$$p^{(j)}_{n+1} = \sum_{i=1}^{N} p^{(i)}_n p_{ij}, \quad (j = 1,..N). \qquad (4.19)$$

Вираз (4.19) дозволяє послідовно крок за кроком визначити зміну розподілу ймовірностей, станів системи, якщо відомий початковий розподіл ймовірностей,.

Якщо у марковського ланцюга існує граничний розподіл ймовірностей,, що відповідає $n \to \infty$ і такий, що не залежить від початкового стану системи, той цей розподіл ймовірностей, визначає граничний або сталий режим системи. У цьому випадку система називається ***статистично стійкою***, а марковський процес у такій системі – ***ергодичним.***

Компоненти $p^{(i)}$ сталого розподілу ймовірностей, ергодичного марковського ланцюга з обліком (4.19) при $n \to \infty$ визначаються як:

$$p^{(j)} = \sum_{i=1}^{N} p^{(i)} p_{ij}, \quad (j = 1,..N).$$

Перехід ергодичного марковського процесу від початкового стану до сталого режиму називається ***перехідним процесом***. Перехідний процес описується послідовністю розподілів ймовірностей, p_n для $n = 1, 2,\ldots$ і при заданому початковому розподілі ймовірностей, p_0 може бути отриманий або шляхом послідовного застосування виразу (4.19), або розглядаючи імовірність l покрокового переходу системи зі стану i у стан j з використанням ***рівняння Чепмена – Колмогорова***:

$$p_{ij}(n,n_1) = \sum_{k=1}^{L} p_{ik}(n,n') \, p_{kj}(n',n_1),$$

при будь-яких $n < n' < n_1$.

4.8. Система масового обслуговування із відмовами.

4.8.1. Рівняння Ерланга

Розглянемо систему масового обслуговування з відмовами як найбільш просту. Нехай є n - канальна система масового обслуговування із відмовами. Розглянемо її як фізичну систему **X** з кінцевою безліччю станів: x_0 — вільні всі канали, x_l — зайнятий рівно один канал, x_k — зайнято рівно $к$ каналів, x_n — зайняті всі n каналів. Схема можливих переходів надана на рис. 4.21.

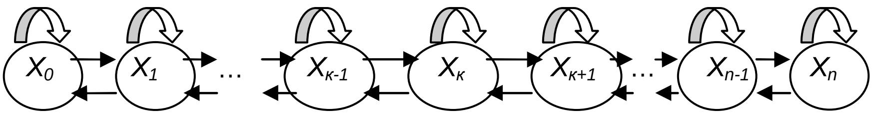


Рис. 4.21. Схема можливих переходів

Поставимо завдання: визначити імовірність станів системи $p_k(t)$ $(k = 1,2,..,n)$ для будь-якого моменту часу t. Задачу розв'язуватимемо при наступних припущеннях:

- потік замовлень — найпростіший, із щільністю λ;
- час обслуговування $T_{об}$ — підлягає показовому закону, із параметром $\mu = 1/m_{tоб}$, і розподілом

$$g(t) = \mu e^{-\mu t} \quad (t > 0).$$

Відмітимо, що параметр μ повністю аналогічний параметру λ показового закону розподілу проміжку T між сусідніми подіями простого потоку:

$$f(t) = \lambda e^{-\lambda t} \quad (t > 0).$$

Параметр λ має сенс «щільності потоку замовлень». Аналогічно, величину μ можна тлумачити як «щільність потоку звільнень» зайнятого каналу.

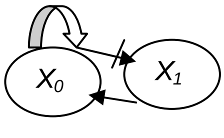


Рис. 4.22. Схема переходів

Оскільки обидва потоки - замовлень і звільнень - найпростіші, процес, що протікає у системі, буде марковським.

Розглянемо можливі стани системи і їх імовірності $p_0(t), p_1(t),..., p_n(t)$, враховуючи для будь-якого моменту часу умову (4.1).

Складемо диференціальні рівняння для всіх ймовірностей,, починаючи з $p_0(t)$. Зафіксуємо момент часу t і знайдемо імовірність $p_0(t+\Delta t)$ того, що у момент $t + \Delta t$ система знаходитиметься у стані x_0 (всі канали вільні). Це може відбутися двома способами (рис. 4.22):

А - у момент t система знаходилася у стані x_0 і за час Δt не перейшла з нього у x_1 (не прийшло жодного замовлення);

В **-** у момент t система знаходилася у стані x_1 і за час Δt канал звільнився, система перейшла в стан x_0 (На рис. 4.22 можливі способи появи стану x_0 у момент $t + \Delta t$ показані стрілками, направленими у x_0; стрілка, направлена із x_0 у x_1, перекреслена на знак того, що система не повинна вийти із стану x_0).

Можливістю «перескоку» системи через стан за малий проміжок часу можна нехтувати, як величиною вищого порядку малості у порівнянні з $P(A)$ і $P(B)$.

Згідно теореми складання ймовірностей, маємо

$$p_0(t + \Delta t) \approx P(A) + P(B).$$

Знайдемо імовірність події A по теоремі множення. Імовірність того, що у момент t система була у стані x_0, дорівнює $p_0(t)$. Імовірність того, що за час Δt не прийде жодної замовлення, дорівнює $e^{-\lambda\Delta t}$. З точністю до величин вищого порядку малості

$$e^{-\lambda\Delta t} \approx 1 - \lambda\Delta t.$$

(у границі, при $\Delta t \to 0$ наближена рівність перейде в точне рівняння).

Отже

$$P(A) \approx p_0(t)(1 - \lambda\Delta t).$$

Знайдемо $P(B)$, Імовірність того, що у момент t система була у стані x_1, яка дорівнює $p_1(t)$. Імовірність того, що за час Δt канал звільниться, точністю до величин вищого порядку малості дорівнює

$$1 - e^{-\mu\Delta t} \approx \mu\Delta t.$$

Отже

$$P(B) \approx p_1(t)\,\mu\Delta t.$$

Звідси

$$p_0(t+\Delta t) \approx p_0(t)\,(1 - \lambda\Delta t) + \mu\, p_1(t)\,\Delta t.$$

Переносячи $p_0(t)$ в ліву частину та ділячи ліву та праву частини на Δt і переходячи до границі при $\Delta t \to 0$ отримаємо диференціальне рівняння для $p_0(t)$:

$$\frac{dp_0(t)}{dt} = -\lambda p_0(t) + \mu p_1(t).$$

Аналогічні диференціальні рівняння можуть бути складені і для інших ймовірностей, станів.

Візьмемо будь-яке k $(0<k<n)$ і знайдемо імовірність $p_k(t+\Delta t)$того, що у момент $t+\Delta t$ система буде у стані x_k (рис. 4.23).

Ця імовірність обчислюється як імовірність суми вже не двох, а трьох подій (по числу стрілок, направлених до стану x_k):

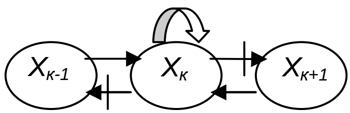


Рис. 4.23. Схема переходів

A - у момент t система була у стані x_k (зайнято k каналів) і за час Δt вона не перейшла з нього ні у x_{k+l}, ні у x_{k-l} стан (жодне замовлення не поступило, жоден канал не звільнився);

B - у момент t система була у стані x_{k-l} (зайнято $k-1$ каналів), а за час Δt перейшла у стан x_k (прийшло одне замовлення);

C - у момент t система була у стані x_{k+l} (зайнято $k+1$ каналів), а за час Δt один із каналів звільнився.

Знайдемо $P(A)$. Обчислимо спочатку імовірність того, що за час Δt не прийде жодне замовлення і не звільниться жоден з k каналів:

$$e^{-\lambda\Delta t}(e^{-\mu\Delta t})^k = e^{-(\lambda + k\mu)\Delta t}.$$

Нехтуючи малими величинами вищих порядків, маємо

$$e^{-(\lambda + k\mu)\Delta t} \approx 1 - (\lambda + k\mu)\,\Delta t.$$

Звідки

$$P(A) \approx p_k(t)[1-(\lambda+\mu k)\,\Delta t].$$

Аналогічно

$$P(B) \approx p_{k-1}(t)\lambda\Delta t \quad \text{і} \quad P(C) \approx p_{k+1}(t)(k+1)\mu\Delta t.$$

Отже:

$$p_k(t+\Delta t) \approx p_k(t)[1-(\lambda+k\mu)\Delta t]+p_{k-1}(t)\lambda\Delta t+p_{k+1}(t)(k+1)\mu\Delta t.$$

Звідси отримуємо диференціальне рівняння для $p_k(t)$ $(0<k<n)$

$$p_k(t)/dt = \lambda\, p_{k-1}(t) - p_k(t)(\lambda+k\mu) + (k+1)\mu\, p_{k+1}(t).$$

Складемо рівняння для останньої імовірності $p_n(t)$ (рис. 4.24). Маємо

$$p_n(t+\Delta t) \approx p_n(t)[1 - n\mu\Delta t) + p_{n-1}(t)\lambda\Delta t ,$$

де $1 - n\mu\Delta t$ **—** імовірність того, що за час Δt не звільниться жоден канал;
$\lambda\Delta t$ **—** імовірність того, що за час Δt прийде одне замовлення .
Отримуємо диференціальне рівняння для $p_n(t)$:

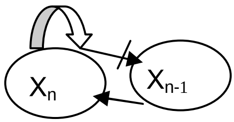


Рис. 4.24. Схема переходів

$$p_n(t)/dt = \lambda p_{n-1}(t) - n\mu\, p_n(t).$$

Таким чином, отримана система диференціальних рівнянь для ймовірностей, $p_l(t), p_2(t),\ldots, p_n(t)$ **виду**:

$$\begin{cases} dp_0(t)/dt = -\lambda p_0(t) + \mu\, p_1(t), \\ \quad . \quad . \quad . \\ dp_k(t)/dt = \lambda\, p_{k-1}(t) - p_k(t)(\lambda+k\mu) + (k+1)\mu\, p_{k+1}(t).\ (0<k<n), \\ \quad . \quad . \quad . \\ dp_n(t)/dt = \lambda p_{n-1}(t) - n\mu\, p_n(t). \end{cases} \qquad (4.20)$$

Рівняння (4.20) називаються ***рівняннями Ерланга***. Інтеграція системи рівнянь (4.20) за наявності початкових умов

$$p_0(0) = 1;\ p_1(0) = \ldots p_n(0) = 0.$$

(у початковий момент всі канали вільні) дає залежність $p_k(t)$ для будь-якого значення **к**. Імовірність $p_k(t)$ характеризує середнє завантаження системи і її зміну з часом. Зокрема, $p_n(t)$ є імовірність того, що замовлення, яке прийшло у момент t, застане всі канали зайнятими (отримує відмову):

$$P_{від} = p_n(t).$$

Величина

$$q(t) = 1 - p_n(t),$$

називається ***відносною пропускною спроможністю*** системи. Для даного моменту t це є відношення середньої кількості замовлень, що обслужені за одиницю часу до середнього числа поданих.

Система лінійних диференціальних рівнянь (4.20) порівняно легко може бути проінтегрована при будь-якому конкретному числі каналів n.

Відмітимо, що при виведенні рівнянь (4.20) не користувалися допущенням про те, що величини μ та λ (щільність потоку замовлень і потоку звільнень) постійні. Тому рівняння (4.20) залишаються справедливими і для залежних від часу $\lambda(t)$, $\mu(t)$, лише б потоки подій, що переводять систему із стану в стан, залишалися пуассоновськими (без цього процес не буде марковським).

4.8.2. Сталий режим обслуговування. Формули Ерланга

Розглянемо n - канальну систему масового обслуговування з відмовами, на вхід якої поступає найпростіший потік замовлень із щільністю λ, час обслуговування має показовий закон розподілу, із параметром. μ. Виникає питання: чи буде стаціонарним випадковий процес, що протікає у системі? Очевидно, що на початку, відразу після включення системи у роботу процес, що протікає в ній ще не буде стаціонарним. У системі масового обслуговування (як і у будь-якій динамічній системі) виникне так званий «перехідний», нестаціонарний процес. Проте, через деякий час, цей перехідний процес затухне, і система перейде на стаціонарний, так званий «сталий» режим, імовірнісні характеристики якого вже не будуть залежати від часу.

У багатьох завданнях практики нас цікавлять саме характеристики граничного сталого режиму обслуговування.

Можна довести, що для будь-якої системи з відмовами такий граничний режим існує, тобто при $t \to \infty$ всі імовірності $p_0(t)$, $p_l(t)$, $p_2(t),\ldots,p_n(t)$. прагнуть до постійних меж p_0, p_l, $p_2,\ldots$, p_n, а всі їх похідні - до нуля.

Щоб знайти граничні імовірності p_0, $p_l,\ldots$, p_n (імовірності станів системи у сталому режимі), замінимо у рівняннях (4.20) всі імовірності $p_k(t)$ $(0 \le k \le n)$ їх межами p_k, а всі похідні покладемо рівними нулю. Отримаємо систему вже не диференціальних, а алгебраїчних рівнянь:

$$\begin{cases} 0 = -\lambda p_0 + \mu p_1, \\ 0 = \lambda p_0 - p_1(\lambda+\mu)+2\mu p_2, \\ \quad . \quad . \quad . \\ 0 = \lambda p_{k-1} - p_k(\lambda+k\mu) + (k+1)\mu p_{k+1} \qquad (0<k<n), \\ \quad . \quad . \quad . \\ 0 = \lambda p_{n-2} - p_{n-1}(\lambda+(n-1)\mu)+n\mu p_n, \\ 0 = \lambda p_{n-1} - n\mu p_n. \end{cases} \tag{4.21}$$

До цих рівнянь необхідно додати умову (4.1).

Вирішимо систему (4.21) щодо невідомих p_0, $p_1,\ldots$, p_n. З першого рівняння маємо

$$p_1 = \lambda p_0/\mu,$$

відповідно з другого

$$p_2 = (1/2\mu)[-\lambda p_0+(\lambda+\mu)p_1]= (1/2\mu)[-\lambda p_0+ \lambda^2 p_0/\mu+ \lambda p_0]= \lambda^2 p_0/2\mu^2,$$

аналогічно з третього

$$p_3 = (1/3\mu)[-\lambda^2 p_0/\mu+ \lambda^3 p_0/2\mu^2 + \lambda^2 2\mu p_0/2\mu^2]= \lambda^3 p_0/3!\mu^3,$$

і взагалі, для будь-якого $k \le n$

$$p_k = (\lambda^k/k!\mu^k)p_0 . \qquad (4.22)$$

Введемо позначення $(\lambda / \mu) = \alpha$, де величина α є ***приведеною щільністю потоку замовлень***. Це є не що інше, як середнє число замовлень, що доводиться на середній час обслуговування одного замовлення. Дійсно $\alpha = \lambda / \mu = \lambda m_{t\,об}$, де $m_{tоб} = M [T_{об}]$ - середній час обслуговування одного замовлення.

У нових позначеннях формула (4.22) приймає вигляд:

$$p_k = (\alpha^k/k!)p_0. \qquad (4.23)$$

Ця формула виражає всі імовірності p_k через p_0. Щоб виразити їх безпосередньо через α та n, скористаємося умовою (4.1). Підставивши в неї (4.23), отримаємо:

$$\sum_{k=0}^{n} p_k = p_0 \sum_{k=0}^{n} \frac{\alpha^k}{k!} = 1 ,$$

звідки

$$p_0 = \frac{1}{\sum_{k=0}^{n} \frac{\alpha^k}{k!}} . \qquad (4.24)$$

Підставляючи (4.24) в (4.23), остаточно отримаємо:

$$p_k = \frac{\frac{\alpha^k}{k!}}{\sum_{k=0}^{n} \frac{\alpha^k}{k!}} \quad (0 \le k \le n) \qquad (4.25)$$

Формули (4.25) називаються ***формулами Ерланга***. Вони дають граничний закон розподілу числа зайнятих каналів залежно від характеристик потоку замовлень і продуктивності системи обслуговування. Вважаючи у формулі (4.25) $k = n$, отримаємо ***імовірність відмови*** (імовірність того, що замовлення , яке поступило, знайде всі канали зайнятими):

$$P_{від} = p_n = \frac{\frac{\alpha^n}{n!}}{\sum_{k=0}^{n} \frac{\alpha^k}{k!}}. \qquad (4.26)$$

Зокрема, для одноканальної системи *(**n = 1**)* маємо

$$P_{віл} = p_1 = \alpha /(1 + \alpha),$$

а відносна пропускна спроможність

$$q = 1 - P_{від} = 1 /(1 + \alpha).$$

Формули Ерланга були отримані для випадку показового розподілу часу обслуговування. Проте дослідження показали, що ці формули залишаються справедливими і при будь-якому законі розподілу часу обслуговування, лише б вхідний потік був найпростішим.

Приклад 4.6. Телефонна станція (ТС) має *4* лінії зв'язку. На станцію поступає простий потік замовлень з щільністю $\lambda = 3$ (викликів за хвилину). Виклик, що поступив у момент часу коли всі лінії зайняті дістає відмову. Середня тривалість розмови *2* хвилини. Визначити:

- імовірність того, що ТС буде взагалі не завантаженою;
- імовірність відмови замовлення.

Маємо: $m_{тоб} = 2$ *(мін),* $\mu = 0.5$ (розмови/мін). Тоді відповідно $\alpha = \lambda/\mu = 6$.

По (4.24) отримаємо

$$p_0 = 1/(1+ \alpha^1/1!+ \alpha^2/2!+ \alpha^3/3!+ \alpha^4/4!) = 0.0087.$$

По (4.26) отримуємо

$$P_{від} = p_4 = (\alpha^4 /4!)/(1+ \alpha^1/1!+ \alpha^2/2!+ \alpha^3/3!+ \alpha^4/4!) = 0.47.$$

Не дивлячись на те, що формули Ерланга у точності справедливі тільки при найпростішому потоці замовлень, ними можна, із відомим наближенням, користуватися і у разі, коли потік замовлень відрізняється від найпростішого.

Розрахунки показують, що заміна довільного стаціонарного потоку, із не дуже великою післядією, простим потоком тієї ж щільності λ, як правило, мало впливає на характеристики пропускної спроможності системи.

Нарешті, можна відмітити, що формулами Ерланга можна наближено користуватися і у разі, коли СМО допускає очікування замовлення у черзі, але коли термін очікування малий у порівнянні із середнім часом обслуговування одного замовлення.

4.9. Система масового обслуговування із очікуванням

4.9.1. Система змішаного типу із обмеженням за часом очікування замовлення у черзі

Система масового обслуговування називається ***системою із очікуванням***, якщо замовлення , що застало всі канали зайнятими, стає у чергу і чекає, поки не звільниться який-небудь канал.

Якщо час очікування замовлення в черзі нічим не обмежено, то система називається ***чистою системою із очікуванням***. Якщо час очікування замовлення у черзі обмежено якимись умовами, то система називається ***системою змішаного типу*.** Це проміжний випадок між чистою системою із відмовами і чистою системою із очікуванням.

Для практики найбільший інтерес представляють саме системи змішаного типу.

Іноді обмеження накладається на час очікування замовлення у черзі якимось терміном $T_{оч}$, що може бути як строго визначеним, так і випадковим. При цьому обмежується тільки термін очікування у черзі, а почате обслуговування доводиться до кінця, незалежно від того, скільки часу продовжувалося очікування (наприклад, якщо клієнт у перукарні сів у крісло, то він звичайно вже не піде до кінця обслуговування).

В інших задачах обмеження накладається не на час очікування в черзі, а на загальний час перебування замовлення у системі (наприклад, об'єкт, що рухається може пробути у зоні захоплення зафіксованої відеокамери лише обмежений час і залишає її незалежно від того, встигли його зняти чи ні).

Нарешті, можна розглянути і таку змішану систему, коли замовлення стає в чергу тільки в тому випадку, якщо довжина черги не занадто велика. Тут обмеження накладається на число замовлень у черзі.

У системах із очікуванням істотну роль грає відзначена раніше «дисципліна черги». Замовлення, що очікують, можуть викликатися на обслуговування як у порядку черги, так і у випадковому, неорганізованому порядку. Існують СМО «із пріоритетами», де деякі замовлення обслуговуються переважно перед іншими (позачергово).

Кожен тип системи із очікуванням має свої особливості і свою математичну теорію. Зупинимося тільки на найпростішому випадку.

Розглянемо змішану систему масового обслуговування ***X з n*** каналами при наступних умовах:

- потік замовлень — найпростіший, із щільністю λ;
- час обслуговування $T_{об}$ — підлягає показовому закону, із параметром $\mu = 1/m_{tоб}$.

Замовлення, що застало всі канали зайнятими, стає у чергу й очікує обслуговування. Час очікування обмежений деяким терміном $T_{оч}$. Якщо до закінчення цього терміну замовлення не буде прийнято до обслуговування, то воно залишає чергу і залишається необслугованим. Термін очікування $T_{оч}$ будемо вважати випадковим і розподіленим по показовому закону

$$h(t) = ve^{-vt},$$

де параметр v - величина, зворотна середньому терміну очікування:

$$v = 1/m_{tоч}, \quad \text{відповідно} \quad m_{tоч} = M[T_{оч}].$$

Параметр v цілком аналогічний параметрам μ і λ потоку замовлень і потоку звільнень. Його можна інтерпретувати, як щільність потоку уходів замовлень, що знаходяться у черзі.

Очевидно, при $v \to \infty$ $(m_{t\,оч} \to 0)$ система змішаного типу перетворюється у чисту систему із відмовленнями. При $v \to 0$ $(m_{t\,оч} \to \infty)$ СМО перетворюється у чисту систему із очікуванням.

Відмітимо, що при показовому законі розподілу терміну очікування пропускна здатність системи не залежить від того, чи обслуговуються замовлення у порядку черги або у випадковому порядку. Для кожного замовлення закон розподілу часу очікування, що залишився, не залежить від того, скільки часу замовлення вже стояло у черзі.

Завдяки допущенню про пуассоновський характер усіх потоків подій, що приводять до змін станів системи, процес, що протікає у ній, буде марковським.

Запишемо рівняння для ймовірностей, станів системи. Для цього, насамперед, перелічимо ці стани. Будемо них нумерувати не за номером зайнятих каналів, а за номером зв'язаних із системою замовлень. Замовлення будемо називати ***зв'язаним із системою***, якщо воно або знаходиться у стані обслуговування, або очікує у черзі.

Можливі стани системи будуть: x_0 - жоден канал не зайнятий, x_l - зайнятий рівно один канал, $x_к$ - зайнято рівно k каналів, x_n - зайняті всі n каналів (черги немає), x_{n+1} - зайняті всі n каналів, одне замовлення знаходиться у черзі, x_{n+s} - зайняті всі n каналів, s замовлень знаходяться у черзі (рис. 4.25).

Число замовлень ***s***, які очікують у черзі, може бути як завгодно великим. Таким чином, система ***X*** має нескінченну (хоча і раховану) безліч станів. Відповідно, число її диференціальних рівнянь, що її описують, теж буде нескінченним.

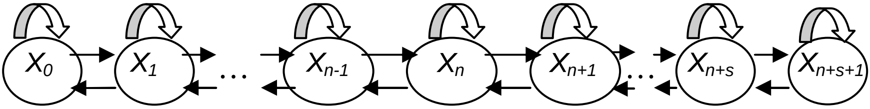


Рис. 4.25. Схема можливих переходів

Очевидно, перші ***n*** диференціальних рівнянь нічим не будуть відрізнятися від відповідних рівнянь Ерланга для СМО з відмовами:

$$\begin{cases} dp_0(t)/dt = -\lambda p_0(t) + \mu\, p_1(t), \\ \quad . \quad . \quad . \\ dp_k(t)/dt = \lambda\, p_{k-1}(t) - p_k(t)(\lambda+k\mu) + (k+1)\mu\, p_{k+1}(t), \\ \quad . \quad . \quad . \\ dp_{n-1}(t)/dt = \lambda\, p_{n-2}(t) - p_{n-1}(t)(\lambda+(n-1)\mu) + n\mu\, p_n(t). \end{cases} \tag{4.27}$$

Відмінність нових рівнянь від рівнянь Ерланга (4.27) розпочнеться при $k = n$. Дійсно, у стан x_n система з відмовами може перейти тільки зі стану x_{n-1}. Що стосується системи із очікуванням, то вона може перейти у стан x_n не тільки з x_{n-1} але і зі стану x_{n+-1} (усі канали зайняті, одне замовлення знаходиться у черзі).

Складемо диференціальне рівняння для $p_n(t)$. Зафіксуємо момент t і знайдемо $p_n(t+\Delta t)$ - імовірність того, що система у момент $t+\Delta t$ буде у стані x_n. Це може здійснитися трьома способами:

- у момент t система вже була у стані x_n і за час Δt не вийшла з нього (не прийшло жодне замовлення і жоден з каналів не звільнився);
- у момент t система була у стані x_{n-1} а за час Δt перейшла у стан x_n (прийшло одне замовлення);
- у момент t система була у стані x_{n+1} (усі канали зайняті, одне замовлення знаходиться у черзі) і за час Δt перейшла у стан x_n (або звільнився один канал і замовлення, що знаходиться у черзі зайняло його, або замовлення, що знаходиться у черзі, пішло із черги у зв'язку із закінченням терміну очікування).

Тоді маємо:

$$p_n(t+\Delta t) \approx (1 - \lambda\,\Delta t - n\mu\,\Delta t)\, p_n(t) + p_{n-1}(t)\lambda\,\Delta t + (n\mu + \nu)\,\Delta t\, p_{n+1}(t).$$

Відкіля маємо:

$$dp_n(t)/dt = -(\lambda + n\mu)\, p_n(t) + \lambda\, p_{n-1}(t) + (n\mu + v)\, p_{n+1}(t).$$

Обчислимо тепер імовірність того, що у момент $t + \Delta t$ усі n каналів будуть зайняті і рівно s замовлень будуть стояти в черзі $p_{n+s}(t + \Delta t)$, при будь-якому $s > 0$. Ця подія знову може здійснитися трьома способами:

- у момент t система вже була у стані x_{n+s} і за час Δt цей стан не змінився (жодне замовлення не прийшло, жоден канал не звільнився і жодне із s замовлень, що очікують у черзі, не пішло);
- у момент t система була у стані x_{n+s-1} і за час Δt перейшла у стан x_{n+s} (тобто прийшло одне замовлення);
- у момент t система була у стані x_{n+s+1} і за час Δt перейшла у стан x_{n+s} (для цього або один із каналів повинний звільнитися, і тоді одне із $s + 1$ замовлень, що очікують у черзі, займе його, або одне з замовлень, що очікують у черзі, повинно піти у зв'язку із закінченням терміну очікування).

Отже:

$$p_{n+s}(t + \Delta t) \approx (1 - \lambda\,\Delta t - n\mu\,\Delta t - sv\,\Delta t))\, p_{n+s}(t) + p_{n+s-1}(t)\lambda\,\Delta t + \\ + (n\mu + (s+1)v)\,\Delta t p_{n+s+1}(t).$$

Відкіля маємо:

$$dp_{n+s}(t)/dt = -(\lambda + n\mu + sv)\, p_{n+s}(t) + \lambda\, p_{n+s-1}(t) + (n\mu + (s+1)v)\, p_{n+s+1}(t).$$

Таким чином, ми одержали для ймовірностей, станів систему нескінченного числа диференціальних рівнянь:

$$\begin{cases} dp_0(t)/dt = -\lambda p_0(t) + \mu\, p_1(t), \\ dp_1(t)/dt = \lambda p_0(t) - p_1(t)(\lambda + \mu) + 2\mu\, p_2(t), \\ \quad . \quad . \quad . \\ dp_k(t)/dt = \lambda\, p_{k-1}(t) - p_k(t)(\lambda + k\mu) + (k+1)\mu\, p_{k+1}(t)\ (0 \le k \le n-1), \\ \quad . \quad . \quad . \\ dp_n(t)/dt = \lambda p_{n-1}(t) - (\lambda + n\mu)\, p_n(t) + (n\mu + v)\, p_{n+1}(t), \\ \quad . \quad . \quad . \\ dp_{n+s}(t)/dt = \lambda p_{n+s-1}(t) - (\lambda + n\mu + sv)\, p_{n+s}(t) + (n\mu + (s+1)v)\, p_{n+s+1}(t). \end{cases} \quad (4.28)$$

Рівняння (4.28) є природним узагальненням рівнянь Ерланга на випадок системи змішаного типу із обмеженням за часом очікування. Параметри λ, μ, v у цих рівнняннях можуть бути як постійними, так і змінними.

При інтегруванні системи (4.28) потрібно враховувати, що хоча теоретично число можливих станів системи нескінченно, але на практиці імовірності $p_{n+s}(t)$ при зростанні s стають дуже малими, і відповідні рівняння можуть бути відкинуті.

Виведемо формули, аналогічні формулам Ерланга, для ймовірностей, станів системи при сталому режимі обслуговування (при $t \to \infty$). З рівнянь (4.28), вважаючи всі p_k $(k = 0, 1, ..., n.)$ постійними, а всі похідні - рівними нулю, одержимо систему алгебраїчних рівнянь:

$$\begin{cases} 0 = -\lambda p_0 + \mu\, p_1, \\ 0 = \lambda p_0 - p_1(\lambda + \mu) + 2\mu\, p_2\,, \\ \quad . \quad . \quad . \\ 0 = \lambda\, p_{k-1} - p_k(\lambda + k\mu) + (k+1)\mu\, p_{k+1} \quad (0 \le k \le n-1), \\ \quad . \quad . \quad . \\ 0 = \lambda p_{n-1} - (\lambda + n\mu)\, p_n + (n\mu + v)\, p_{n+1}\,, \\ \quad . \quad . \quad . \\ 0 = \lambda p_{n+s-1} - (\lambda + n\mu + sv)\, p_{n+s} + (n\mu + (s+1)v)\, p_{n+s+1}. \end{cases} \qquad (4.29)$$

До них потрібно приєднати умову (4.1).

Знайдемо рішення системи (4.29). Для цього застосуємо той же прийом, яким ми користувалися у випадку системи з відмовами: розв'яжемо перше рівняння відносно p_1 і підставимо у друге, і т.д. Для будь-якого $k \le n$, як і у випадку системи з відмовами, одержимо:

$$p_k = (\lambda^k / k!\, \mu^k) p_0. \qquad (4.30)$$

Перейдемо до рівнянь для $k > n$ $(k = n + s,\ s \ge 1)$. Тим же способом одержимо:

$$p_{n+1} = \lambda^{n+1} p_0 / n! \mu^n (n\mu + v),$$

$$p_{n+2} = \lambda^{n+2} p_0 / n! \mu^n (n\mu + v)(n\mu + 2v),$$

і взагалі при будь-якому $s \ge 1$ маємо:

$$p_{n+s} = \frac{\lambda^{n+s} p_0}{n! \mu^n \prod_{m=1}^{s} (n\mu + mv)}\,. \qquad (4.31)$$

В обох формул (4.30) та (4.31) як співмножник входить імовірність p_0. Визначимо її з умови

$$\sum_{k=0}^{n} \sum_{s=1}^{\infty} p_{n+s} + p_k = 1\,.$$

Підставляючи у нього вираження (4.30) та (4.31) для $k \le n$ і $s \ge 1$, одержимо:

$$p_0\left\{\sum_{k=0}^{n}\frac{\lambda^k}{k!\mu^k}+\sum_{s=1}^{\infty}\frac{\lambda^{n+s}}{n!\mu^k\prod\limits_{m=1}^{s}(n\mu+m\nu)}\right\}=1\,.$$

Відкіля:

$$p_0=\frac{1}{\sum\limits_{k=0}^{n}\frac{\lambda^k}{k!\mu^k}+\sum\limits_{s=1}^{\infty}\frac{\lambda^{n+s}}{n!\mu^k\prod\limits_{m=1}^{s}(n\mu+m\nu)}}\,. \qquad (4.32)$$

Перетворимо вирази (4.30), (4.31) і (4.32), вводячи в них замість щільностей λ і ν «приведені» щільності:

$$\lambda/\mu = \lambda m_{\ t\ про} = \alpha,$$
$$\nu/\mu = \nu m_{\ t\ про} = \beta.$$

Параметри α і β виражають відповідно середнє число замовлень і середнє число відходів замовлень, що знаходяться у черзі, які приходяться на середній час обслуговування одного замовлення.

У нових позначеннях формули (4.30), (4.31) і (4.32) приймуть вид:

$$p_k=(\alpha^k/k!)p_0 \quad (0<k\le n), \qquad (4.33)$$

$$p_{n+s}=\frac{\frac{\alpha^{n+s}}{n!}p_0}{\prod\limits_{m=1}^{s}(n+m\beta)}. \quad (s\ge 1), \qquad (4.34)$$

$$p_0=\frac{1}{\sum\limits_{k=0}^{n}\frac{\alpha^k}{k!}+\frac{\alpha^n}{n!}\sum\limits_{s=1}^{\infty}\frac{\alpha^s}{\prod\limits_{m=1}^{s}(n+m\beta)}}\,. \qquad (4.35)$$

Підставляючи (4.33) у (4.34) і (4.35), одержимо остаточні вирази для ймовірностей, станів системи:

$$p_k=\frac{\frac{\alpha^k}{k!}}{\sum\limits_{k=0}^{n}\frac{\alpha^k}{k!}+\frac{\alpha^n}{n!}\sum\limits_{s=1}^{\infty}\frac{\alpha^s}{\prod\limits_{m=1}^{s}(n+m\beta)}} \qquad (0\le k\le n), \qquad (4.36)$$

$$p_{n+s} = \frac{\dfrac{\alpha^k}{n!} \dfrac{\alpha^s}{\prod_{m=1}^{s}(n+m\beta)}}{\sum_{k=0}^{n} \dfrac{\alpha^k}{k!} + \dfrac{\alpha^n}{n!} \sum_{s=1}^{\infty} \dfrac{\alpha^s}{\prod_{m=1}^{s}(n+m\beta)}} \quad (s \geq 1) . \tag{4.37}$$

Знаючи імовірності всіх станів системи, можна легко визначити інші характеристики, що цікаві для нас, зокрема, імовірність того, що замовлення залишить систему необслугованим $p_н$. Визначимо її з наступних міркувань: при сталому режимі імовірність того, що замовлення залишить систему необслугованим, є не що інше, як відношення середнього числа замовлень, що ідуть з черги у одиницю часу, до середнього числа замовлень, що надходять у систему в одиницю часу

$$P_н = vm_s / \lambda.$$

Обчислимо математичне очікування m_s числа замовлень, що знаходяться в черзі:

$$m_s = M[s] = \sum_{s=1}^{\infty} s p_{n+s} \cdot \tag{4.38}$$

Щоб одержати $p_н$, потрібно m_s помножити на середню «щільність відходів» одного замовлення v і розділити на середню щільність замовлень λ тобто помножити на коефіцієнт

$$v/\lambda = (v/\mu)/(\lambda/\mu) = \beta / \alpha .$$

Тоді маємо, що

$$p_н = \beta\, m_s / \alpha. \tag{4.39}$$

Відносна пропускна здатність системи характеризується імовірністю того, що замовлення, що потрапило у систему, буде обслуговано:

$$q = 1 - p_н.$$

Очевидно, що пропускна здатність системи із очікуванням, при тих же λ і μ буде завжди вище, ніж пропускна здатність системи із відмовами, так як у випадку наявності очікування необслугованими ідуть не всі замовлення, що застали n каналів зайнятими, а тільки деякі. Пропускна здатність збільшується при збільшенні середнього часу очікування

$$m_{t\,оч} = 1 / v .$$

Безпосереднє користування формулами (4.36), (4.37) і (4.39) трохи утруднене тим, що в них входять нескінченні суми. Однак члени цих сум швидко убувають.

Подивимося, у що перетворяться формули (4.36) і (4.37) при $\beta\to\infty$. Очевидно, що при $\beta\to\infty$ система із очікуванням повинна перетворитися у систему із відмовами (замовлення миттєво іде з черги).

Інший крайній випадок - це чиста система із очікуванням ($\beta \to 0$). У такій системі замовлення взагалі не ідуть з черги, і тому $p_н = 0$ (кожне замовлення рано або пізно дочекається обслуговування).

Зате у чистій системі із очікуванням не завжди мається граничний стаціонарний режим при $t\to\infty$. Можна довести, що такий режим існує тільки при $\alpha < n$ тобто коли середнє число замовлень, що приходиться на час обслуговування одного замовлення, не виходить за межі можливостей *i* - канальної системи. Якщо ж $\alpha > n$ число замовлень, що очікують у черзі, буде із часом необмежено зростати.

Припустимо, що $\alpha<n$ і знайдемо граничні імовірності $p_k\,(0\le k \le n)$ для чистої системи із очікуванням. Для цього покладемо у формулах (4.35), (4.36) і (4.37) $\beta \to 0$. Одержимо:

$$p_0 = \frac{1}{\sum_{k=0}^{n}\frac{\alpha^k}{k!}+\frac{\alpha^n}{n!}\sum_{s=1}^{\infty}\frac{\alpha^s}{n^s}}.$$

Звідси, користуючись формулами (4.33) і (4.34), знайдемо

$$p_k = \frac{\frac{\alpha^k}{k!}}{\sum_{k=0}^{n}\frac{\alpha^k}{k!}+\frac{\alpha^n}{n!}\sum_{s=1}^{\infty}\frac{\alpha^s}{n^s}}\quad (0 \le k \le n), \tag{4.40}$$

й аналогічно для $k = n + s \quad (s \ge 0)$

$$p_{n+s} = \frac{\frac{\alpha^{n+s}}{n!n^s}}{\sum_{k=0}^{n}\frac{\alpha^k}{k!}+\frac{\alpha^n}{n!}\sum_{s=1}^{\infty}\frac{\alpha^s}{n^s}}.$$

Середнє число замовлень, що знаходяться у черзі, визначається з формули (4.38) при $\beta \to 0$

$$m_s = \frac{\dfrac{\alpha^{n+1}}{nn!\,(1 - \dfrac{\alpha}{n})^2}}{\sum\limits_{k=0}^{n} \dfrac{\alpha^k}{k!} + \dfrac{\alpha^n}{n!} \sum\limits_{s=1}^{\infty} \dfrac{\alpha^s}{n^s}} . \qquad (4.41)$$

Приклад 4.7. Нехай у систему, що має три канали обслуговування, із необмеженим часом очікування надходить найпростіший потік замовлень із щільністю $\lambda = 4$ (замовлень/година). Середній час обслуговування одного замовлення – $m_{t\ об} = 30$ ***хв***. Визначити, чи існує сталий режим обслуговування, якщо так,то знайти p_0, p_1, p_2, p_3. імовірність наявності черги і середню довжину черги.

Рішення. Маємо $\mu = 1/ m_{t\ об},\ \alpha = \lambda / \mu = 2.$ Тому, що $\alpha < n$ - сталий режим існує. Використовуючи (4.40) знаходимо

$$p_0 = 1/9 \approx 0.111,\ p_1 = 2/9 \approx 0.222,\ p_2 = 2/9,\ p_3 = 8/54 \approx 0.148.$$

імовірність наявності черги

$$p_{чер} = 1 - (p_0 + p_1 + p_2 + p_3) = 0.297.$$

А середня довжина черги за (4.41) буде $m_s \approx 0.89$ (замовлення).

4.9.2. Система масового обслуговування змішаного типу із обмеженням по довжині черги

Розглянемо систему змішаного типу із обмеженням по числу замовлень, що очікують у черзі. Припустимо, що замовлення, яке застала всі канали зайнятими, стає у чергу, тільки якщо у ній знаходиться менш m замовлень. Якщо ж число замовлень у черзі дорівнює m (більше m воно бути не може), то останнє замовлення, що надійшло у чергу не стає і залишає систему необслугованим. Інші допущення - про найпростіший потік замовлень і про показовий розподіл часу обслуговування залишаються, як і у попередньому випадку.

Отже, є n - канальна система з очікуванням, у якій кількість замовлень, що очікують у черзі, обмежено числом m. Складемо диференціальні рівняння для ймовірностей, станів системи. Відмітимо, що у даному випадку число станів системи буде $n + m$ (n - замовлення, що обслуговуються і m – замовлення, що стоять у черзі). Перелічимо стани системи: x_0 - усі канали вільні, черги немає; x_1 - зайнятий один канал; x_k - зайнято k каналів; x_{n-1} - зайнято $n - 1$ каналів; x_n - зайняті всі n каналів; x_{n+1} - зайняті всі n каналів і одне замовлення стоїть у черзі; $\boldsymbol{x_{n+m}}$ - зайнято всі $\boldsymbol{n}$ каналів, $\boldsymbol{m}$ замовлень стоїть у черзі (рис. 4.26). Параметр $\boldsymbol{v}$ у даному випадку не фігурує.

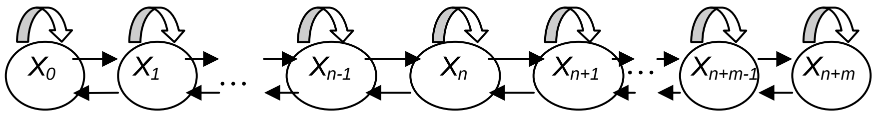


Рис. 4.26. Схема можливих переходів

Очевидно, перші **n** рівнянь для ймовірностей, $p_0(t), \ldots, p_{n-1}(t)$ будуть збігатися із рівняннями Ерланга, що були розглянуті раніше. Виведемо інші рівняння. Маємо:

$$p_n(t+\Delta t) = p_{n+1}(t)n\mu\Delta t + p_n(t)(1 - \lambda\Delta t - n\mu\Delta t)+p_{n-1}(t)\lambda\Delta t\ .$$

Відкіля

$$dp_n(t)/dt = \lambda p_{n-1}(t) - (\lambda + n\mu)\, p_n(t) + (n\mu)\, p_{n+1}(t),$$

Далі виведемо рівняння для $p_{n+s}(t)$ $(1\le s \le m)$

$$p_{n+s}(t+\Delta t) = (1 - \lambda\, \Delta t - n\mu\, \Delta t\,)\, p_{n+s}(t) + p_{n+s-1}(t)\lambda\, \Delta t + n\mu\, \Delta t\, p_{n+s+1}(t).$$

Відкіля

$$dp_{n+s}(t)/dt = \lambda p_{n+s-1}(t) - (\lambda + n\mu\,)\, p_{n+s}(t) + n\mu\, p_{n+s+1}(t).$$

Останнє рівняння буде

$$dp_{n+m}(t)/dt = \lambda p_{n+m-1}(t) - n\mu\, p_{n+m}(t).$$

Таким чином, отримана систему диференціальних рівнянь:

$$\begin{cases} dp_0(t)/dt = -\lambda p_0(t) + \mu\, p_1(t)\ , \\ dp_1(t)/dt = \lambda p_0(t) - p_1(t)(\lambda+\mu)+2\mu\, p_2(t)\ , \\ \quad . \quad . \quad . \\ dp_k(t)/dt = \lambda\, p_{k-1}(t) - p_k(t)(\lambda+k\mu) + (k+1)\mu\, p_{k+1}(t)\ (0 \le k \le n-1)\ , \\ \quad . \quad . \quad . \\ dp_n(t)/dt = \lambda p_{n-1}(t) - (\lambda+n\mu)\, p_n(t) + n\mu\, p_{n+1}(t)\ , \\ \quad . \quad . \quad . \\ dp_{n+s}(t)/dt = \lambda p_{n+s-1}(t) - (\lambda+n\mu)\, p_{n+s}(t) + n\mu\, p_{n+s+1}(t), \quad (0 \le s \le m)\ , \\ \quad . \quad . \quad . \\ dp_{n+m}(t)/dt = \lambda p_{n+m-1}(t) - n\mu\, p_{n+m}(t)\ ., \end{cases} \qquad (4.42)$$

Розглянемо граничний випадок при $t\to\infty$. Дорівнюючи всі похідні нулю, а всі імовірності вважаючи постійними, одержимо систему алгебраїчних рівнянь:

$$
\begin{cases}
0 = -\lambda p_0 + \mu p_1, \\
0 = \lambda p_0 - p_1(\lambda+\mu)+2\mu p_2, \\
\quad . \quad . \quad . \\
0 = \lambda p_{k-1} - p_k(\lambda+k\mu) + (k+1)\mu p_{k+1} \qquad (0 \le k \le n-1), \\
\quad . \quad . \quad . \\
0 = \lambda p_{n-1} - (\lambda+n\mu) p_n + n\mu p_{n+1}, \\
\quad . \quad . \quad . \\
0 = \lambda p_{n+s-1} - (\lambda+n\mu) p_{n+s} + n\mu p_{n+s+1}, \quad (0 \le s \le m), \\
\quad . \quad . \quad . \\
0 = \lambda p_{n+m-1} - n\mu p_{n+m}.
\end{cases} \qquad (4.43)
$$

Рівняння (4.43) можуть бути вирішені так само, як ми вирішили аналогічні алгебраїчні рівняння. Не зупиняючись на цьому рішенні, приведемо тільки остаточні формули:

$$
p_k = \frac{\dfrac{\alpha^k}{k!}}{\sum\limits_{k=0}^{n} \dfrac{\alpha^k}{k!} + \dfrac{\alpha^n}{n!} \sum\limits_{s=1}^{m} \left(\dfrac{\alpha}{n}\right)^s} \quad (0 \le k \le n) \qquad (4.44)
$$

$$
p_{n+s} = \frac{\dfrac{\alpha^n}{n!} \left(\dfrac{\alpha}{n}\right)^s}{\sum\limits_{k=0}^{n} \dfrac{\alpha^k}{k!} + \dfrac{\alpha^n}{n!} \sum\limits_{s=1}^{m} \left(\dfrac{\alpha}{n}\right)^s} \quad (0 \le s \le m) \qquad (4.45)
$$

Імовірність того, що замовлення залишить систему необслугованим, дорівнює імовірності того, що у черзі вже очікують m замовлень p_{n+m}.

Неважко помітити, що формули (4.45) і (4.44) виходять із формул (4.36), (4.37), якщо покласти в них $\beta = 0$ і обмежити підсумовування по s верхньою границею m.

Приклад 4.8. На станцію ремонту автомашин надходить найпростіший потік замовлень із щільністю $\lambda = 0{,}5$ (машини за год.). Мається одне приміщення для ремонту. У дворі станції можуть одночасно знаходитися, очікуючи у черзі, не більше **3** машин. Середній час ремонту однієї машини $m_{тоб} = 1/\mu = 2$ (години).

Визначити: пропускну здатність системи; середній час простою станції; наскільки зміняться ці характеристики, якщо обладнати друге приміщення для ремонту.

Рішення. Маємо $\lambda = 0.5$, $\mu = 0.5$, $\alpha = 1$, $m = 3$.

По формулі (4.45), вважаючи $n = 1$, знаходимо імовірність того, що замовлення, яке прийшло, залишить систему необслугованим:

$$p_H = p_{1+3} = 1/(1+1+3) = 0.20.$$

Абсолютна пропускна здатність:

$Q = \lambda q = 0.4$ (машини за годину).

Середню частку часу, що система буде простоювати, знайдемо по формулі (4.44):

$$p_0 = 1/5 = 0.2.$$

Вважаючи $n = 2$, знайдемо:

$$P_н = p_{2+3} = (1/16)/(1+1+1/2+1/4+1/8+1/16) = 1/47 = 0.021.$$

Відносна пропускна здатність системи

$$q = 1 - p_H = 0.979.$$

тобто задовольнятися буде близько 98% усіх замовлень).

Абсолютна пропускна здатність:

$Q = \lambda q = 0.49$ (машини за годину).

Відносний час простою:

$$p_0 = (16/47) = 0.34,$$

тобто устаткування буде простоювати близько **34%** усього часу.

4.10. Методи рішення задачі розрахунку СМО

Усі закони розподілу, що характеризують потік замовлень, можуть задаватися або на основі апріорних допущень загального характеру, або у результаті спеціального експериментального дослідження відповідного потоку.

Прикладом, що характеризує перший випадок, може служити телефонна станція, що обслуговує велику кількість індивідуальних абонентів. Апріорним, тут є природне припущення про незалежність телефонних дзвінків різних абонентів і повної випадковості моменту, коли такий дзвінок буде. З цих припущень випливає, що розглянутий потік буде пуассоновським.

На частку експериментального дослідження залишається лише визначення щільності λ цього потоку. Для стаціонарного потоку ця величина може бути отримана у результаті визначення числа N дзвінків протягом досить великого терміну часу T:

$$\lambda \approx N / T.$$

Якщо апріорні припущення, що дозволяють визначити вигляд закону розподілу, відсутні, або маються підстави сумніватися у їхній справедливості, то необхідно більш детальне експериментальне дослідження розглянутого потоку. Для цієї мети протягом досить тривалого проміжку часу фіксуються моменти надходження замовлень і фактична тривалість обслуговування. На підставі отриманих даних будуються експериментальні графіки розподілу, які потім апроксимуються.

У загальному випадку задача розрахунку СМО полягає в тому, щоб визначити закони розподілу і середні значення різних (випадкових) величин, зв'язаних з цією системою:

- довжина черги;
- час очікування обслуговування;
- час зайнятості обслуговуючих приладів і т.п.

Розглянемо основні методи рішення задачі розрахунку СМО.

Аналітичні методи. Дозволяють представити опис процесів функціонування елементів реального об'єкта у вигляді математичних співвідношень (алгебраїчних, інтегральних, диференціальних, логічних і т.д.) і функцій , що зв'язують характеристики, що цікавлять нас, із вхідними змінними, при наявності системи обмежень.

При цьому передбачається наявність однозначної обчислювальної процедури одержання точного рішення рівнянь.

Недолік даних методів міститься у тому, що прості аналітичні (формальні) вирази для такого роду функцій вдається одержувати лише у найпростіших випадках.

Аналітичні методи можуть бути корисні лише при розгляді порівняно простих СМО. Тому при аналізі реальних систем, особливо у випадках, коли характер вхідного потоку замовлень і часу обслуговування підкоряється іншим законам розподілу, ніж ті, що розглянуті, прибігають до ***методів математичного моделювання*** процесу масового обслуговування. Ці методи припускають встановлення відповідності реальної СМО до деякої математичної моделі й дослідження цієї моделі, що дозволяє одержати характеристики реальної системи.

Застосування математичного моделювання дозволяє досліджувати об'єкти, реальні експерименти над якими утруднені або неможливі (великі затрати, небезпечно для здоров'я, однократні процеси, які неможливі через фізичні або часові обмеження, процеси, що відбуваються далеко, процеси які ще, або вже не існують і т.п.).

Методи математичного моделювання можна класифікувати наступним чином (рис. 4.27).

Комп'ютерне - математичне моделювання –є різновидом математичного моделювання, що включає:

- чисельне моделювання;
- імітаційне моделювання;
- статистичне моделювання.

Комп'ютерне - математичне моделювання формулюється у вигляді алгоритму (програми для ЕОМ), яка визначає зв'язок між вихідними, вхідними й внутрішніми параметрами, що дозволяє проводити над моделлю, яка досліджується обчислювальні експерименти.

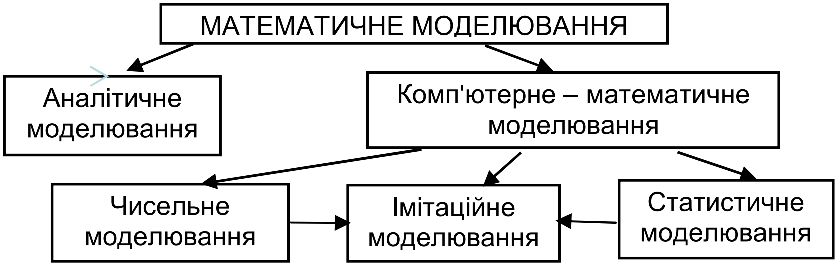


Рис. 4.27. Класифікація методів математичного моделювання

Сутність методу полягає у тому, що програма, яка установлена на ЕОМ:

- повторює крок за кроком весь алгоритм функціонування системи масового обслуговування, яка досліджується;
- відповідно до заданого закону розподілу, визначаються:
  - моменти появи замовлень;
  - тривалості обробки замовлень;
  - порядок виділення обслуговуючих приладів;
  - постановку у чергу замовлень і т.п.

При цьому спеціальна частина програми:

- увесь час веде підрахунок наступних величин:
  - довжини черг;
  - простоїв обслуговуючих приладів;
  - кількість замовлень, що не були обслуговані і т.п.
- будує експериментальні графіки розподілу ймовірностей, по яким обчислюються:
  - середні значення, дисперсії;
  - середньоквадратичні відхилення;
  - інші характеристики експериментальних розподілів.

Методом математичного моделювання в принципі можуть бути вирішені будь-які задачі із розрахунку систем масового обслуговування, у тому числі і складних систем зі змінною (залежною від стану системи) дисципліною обслуговування.

Недолік даного методу - великі витрати машинного часу, оскільки для одержання стійких статистичних характеристик системи, що досліджується доводиться робити велике число експериментів. Але при цьому витрати на дослідження СМО скорочуються у середньому в 10 - 100 разів.

Приклад 4.9. Нехай було здійснено сто обчислень довжини черги. Відкладаючи по горизонталі різні значення довжини, а по вертикалі. кількість випадків, протягом яких довжина черги приймала задані значення, одержимо деяку гістограму (рис. 4.28). Із цієї гістограми видно, що довжина черги, яка дорівнює нулю, зустрічалася 10 разів, а довжина черги, яка дорівнює трьом - 25 разів, що відповідає експериментальним ймовірностям

$$P_0=10/100 = 0.1, \qquad P_3 =25\backslash 100 = 0.25.$$

Середнє значення довжини черги

$$n_{cp} =0P_0+1P_1+2P_2+3P_3+4P_4+5P_5 = 0+0,2+0,6+0,75+0,4+0,25=2,2.$$

Найбільш ймовірна довжина черги дорівнює ***2***.

Дисперсія визначається, як

$$d=\sum_{i=0}^{5}(i-n_{cp})^2 P_i =$$

$$=2,2^2\,0,1+1,2^2\,0,2+0,2^2\,0,3+0,8^2\,0,25+1,8^2\,0,1+2,8^2\,0,05=1,307.$$

А середньоквадратичне відхилення $\sigma = d^{-2} =1,14$.

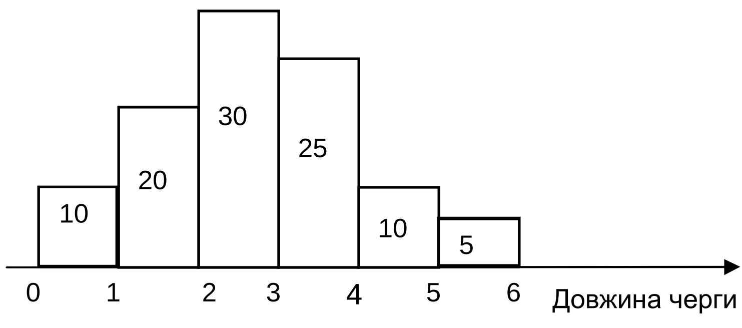


Рис. 4.28. Гістограма зміни довжини черги

Технологію комп'ютерного математичного моделювання можна надати у вигляді алгоритму наданого на рис.4.29.

Результат перших двох етапів визначає вибір математичного методу, який покладеного у основу моделювання (і навіть навпаки - їхній результат визначає вибір методу). А от реалізація інших п'яти етапів істотно розрізняється для кожного із двох основних підходів до побудови моделі - аналітичного і комп'ютерного математичного моделювання.

Рис. 4.29.Технологія комп'ютерного математичного моделювання

Розглянемо більш детально класифікацію математичних моделей (табл. 4.1.)

Таблиця 4.1

Класифікація математичних моделей

Ознаки класифікації	Види математичних моделей
1. Приналежність до ієрархічного рівня	Моделі мікрорівня Моделі макрорівня Моделі метарівня
2. Характер властивостей об'єкта	Структурні моделі Функціональні моделі
3. Спосіб подання властивостей об'єкта	Аналітичні моделі Алгоритмічні моделі Імітаційні моделі
4. Спосіб одержання моделі	Теоретичні моделі Емпіричні моделі
5. Особливості поводження об'єкта	Детерміновані моделі Імовірнісні моделі

Відповідно приналежності до ієрархічного рівня математичні моделі, що описують виробничий процес можна класифікувати, як:

- ***моделі мікрорівня*** – ті, що відбивають фізичні процеси, які протікають, наприклад, при різанні металів (описують найпростіші процеси на рівні переходу);
- ***моделі макрорівня*** – ті, що описують технологічні процеси;
- ***моделі метарівня*** – ті, що описують технологічні системи (ділянки, цехи, підприємство в цілому).

За характером властивостей об'єкту виділяють:

- ***структурні математичні моделі*** - ті, що призначені для відображення структурних властивостей об'єктів, наприклад, у САПР для подання структури технологічного процесу використаються структурно - логічні моделі;
- ***функціональні математичні моделі*** - ті, що призначені для відображення інформаційних, фізичних і часових процесів, що протікають у працюючому устаткуванні, у ході виконання технологічних процесів і т.д.

За способом подання властивостей об'єкту відрізняють:

- ***аналітичні математичні моделі*** - являють собою явні математичні вираження вихідних параметрів як функцій від параметрів вхідних і внутрішніх;
- ***алгоритмічні математичні моделі*** - виражають зв'язок між вихідними параметрами й параметрами вхідними й внутрішніми, у вигляді алгоритму;
- ***імітаційні математичні моделі*** - це алгоритмічні моделі, що відбивають розвиток процесу (поводження досліджуваного об'єкта) у часі, при завданні зовнішніх впливів на процес (об'єкт), наприклад, це моделі систем масового обслуговування, що задані у алгоритмічній формі.

За способом одержання моделі розрізняють:

- ***теоретичні математичні моделі*** - створюються у результаті дослідження об'єктів (процесів) на теоретичному рівні (не завжди прийнятні для практичного використання, тому що дуже громіздкі й не зовсім адаптовані до реальних процесів);
- ***емпіричні математичні моделі*** - створюються у результаті проведення експериментів (вивчення зовнішніх проявів властивостей об'єкта за допомогою виміру його параметрів на вході й виході) і обробки їхніх результатів методами математичної статистики.

Моделі, що визначаються особливостями поводження об'єкту:

- ***детерміновані математичні моделі*** - описують поводження об'єкта з позицій повної визначеності в теперішній час та у майбутньому, приклади таких моделей - формули фізичних законів, технологічні процеси обробки деталей і т.д.;

- ***імовірнісні математичні моделі*** - враховують вплив випадкових факторів на поводження об'єкту, тобто оцінюють його майбутнє з позицій імовірності тих або інших подій, приклади таких моделей: опис очікуваних довжин черг у системах масового обслуговування, очікуваних обсягів випуску надпланової продукції виробничою ділянкою, точності розмірів у партії деталей із урахуванням явища розсіювання й т.д.

Відповідно до класифікації, що надана на рис. 4.27 розглянемо більш детально основні методи математичного моделювання.

Сутність методу ***імітаційного моделювання*** міститься у відтворенні на ЕОМ (імітації) процесу функціонування системи, що досліджується, дотримуючись логічної й часової послідовності протікання процесів. Це дозволяє визначити дані про стан СМО або окремих її елементів у певні моменти часу. При цьому припускається, що математична модель, яка використовується, відтворює алгоритм (логіку) функціонування СМО у часі при різних комбінаціях значень параметрів системи, що досліджується й зовнішнього середовища.

Істотною характеристикою таких моделей є структурна відповідність об'єкта й моделі. Це означає, що відповідно до завдання дослідження кожному істотному елементу СМО ставиться у відповідність елемент її моделі. При побудові імітаційної моделі описуються закони функціонування кожного елемента об'єкта й зв'язки між ними. Робота з імітаційною моделлю полягає у проведенні імітаційного експерименту. Процес, що протікає у моделі в ході експерименту, подібний до процесу у реальному об'єкті. Тому дослідження об'єкту по його імітаційній моделі зводиться до вивчення характеристик процесу, що протікає в ході експерименту.

Основою створення всякої імітаційної моделі є:

- розробка моделі системи, що досліджується, на основі особистісних імітаційних моделей підсистем, які об'єднані у єдине ціле своїми взаємодіями;
- вибір інформативних характеристик об'єкту та способів їхнього одержання й аналізу;
- побудова моделі впливу зовнішнього середовища на систему у вигляді сукупності імітаційних моделей зовнішніх факторів, що впливають;
- вибір способу дослідження імітаційної моделі відповідно до методів планування імітаційних експериментів.

Імітаційну модель можна представити у вигляді структурної схеми, що реалізується програмно або апаратно (рис.4.30).

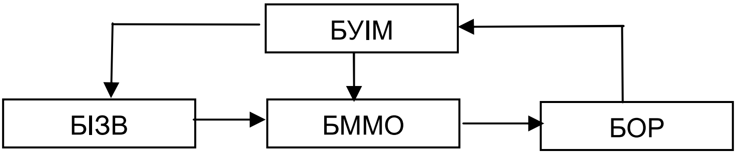


Рис. 4.30 Імітаційна модель

БІЗВ - блок імітації зовнішніх впливів - формує реалізації випадкових або детермінованих процесів, що імітують впливи зовнішнього середовища на об'єкт.

БОР - блок обробки результатів - призначений для одержання інформативних характеристик об'єкту, що досліджується.

БММО - блок математичної моделі об'єкту.

БУІМ - блок управління - реалізує спосіб дослідження імітаційної моделі, основне його призначення - автоматизація процесу проведення імітаційних експериментів.

Ще один спосіб імітаційного моделювання систем заснований на використанні поняття транзакту.

Транзакт - це деяке повідомлення (замовлення на обслуговування), що надходить ззовні на вхід СМО й підлягає обробці. У деяких випадках, наприклад при моделюванні АСУ, більш зручно простежити функціонування системи саме щодо алгоритму обробки транзакту. У рамках однієї імітаційної моделі можуть розглядатися транзакти декількох типів. Кожний транзакт характеризується відповідним алгоритмом обробки й необхідними для його реалізації ресурсами системи. Вважаючи це, проходження транзакту по системі можна у деяких випадках розглядати як послідовну активізацію процесів, що реалізують його обробку (обслуговування замовлення).

Статистичне моделювання СМО - передбачає обробку даних про систему (модель) з метою одержання статистичних характеристик СМО.

В основі статистичного моделювання лежить імітація процесу масового обслуговування, при якій моменти надходження замовлень і тривалості їх обслуговування знаходять по таблицях випадкових чисел або від спеціальних датчиків випадкових чисел за допомогою ЕОМ.

Оскільки як датчики, так і таблиці випадкових чисел маються лише для деяких найбільш розповсюджених видів функцій розподілу ймовірностей,, то для одержання випадкових чисел з довільною функцією розподілу ймовірностей, використовують спеціальні методи одержання випадкової вибірки.

Нехай ***x*** — випадкова величина зі щільністю розподілу ймовірностей, ***f(x)*** і з функцією розподілу ймовірностей, ***F(x)**,* яка зв'язана із ***f(x)*** співвідношенням

$$f(x)dx = dF(x).$$

Нехай також ***y = F(x)*** є випадковою величиною із рівномірним розподілом ймовірностей, від *0* до *1*. Тоді імовірність того, що випадкова величина **x** із щільністю розподілу ймовірностей, ***f(x)*** попадає на інтервал від ***x*** до ***x + dx*** дорівнює імовірності того, що випадкова величина ***y = F(x)*** із рівномірним розподілом ймовірностей, від *0* до *1* попадає на інтервал від ***y*** до ***y + dy*** (рис. 4.31).

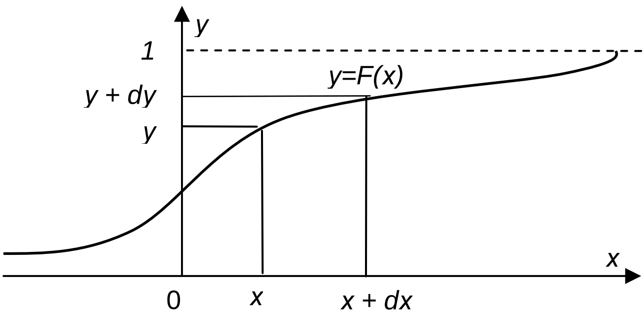


Рис. 4.31. Одержання випадкової вибірки із заданою функцією розподілу.

Звідси можна отримати алгоритм одержання випадкової вибірки із заданою щільністю розподілу ***f(x)*.**

1. Для даної щільності розподілу ***f(x)*** будується графік або знаходиться аналітичне вираження функції розподілу ймовірностей, ***y = F(x)***.

2. По таблиці випадкових чисел із рівномірним розподілом ймовірностей, від *0* до *1* знаходиться вибірка випадкових чисел величини ***y = F(x)***.

3. Для кожного числа із вибірки значень ***y*** по залежності ***y = F(x)*** знаходиться відповідне значення ***x***.

4. Сукупність отриманих значень ***x*** дає випадкову вибірку з щільністю розподілу ймовірностей, ***f(x).***

***Чисельне моделювання* -** припускає використання, для побудови моделі СМО, методів обчислювальної математики (відрізняється від чисельного аналітичного тім, що можливо завдання різних параметрів моделі).

Аналітичне моделювання - допускає використання математичної моделі реального об'єкта у формі алгебраїчних, диференціальних, інтегральних і інших рівнянь і функцій, що зв'язують вихідні змінні із вхідними змінними, з урахуванням системи обмежень. При цьому аналітичні математичні моделі являють собою явні математичні вирази вихідних параметрів як функцій від параметрів вхідних і внутрішніх.

Аналітичне моделювання засноване на непрямому описі об'єкту, що моделюється, за допомогою набору математичних формул. Мова аналітичного опису містить наступні основні групи семантичних елементів:

- критерій (критерії);
- невідомі;
- дані;
- математичні операції;
- обмеження.

Аналітичні моделі є ефективним інструментом для рішення завдань оптимізації й обчислення характеристик технологічних систем. Вони завжди являють собою конструкцію, яку можна проаналізувати й вирішити математичними засобами. Але аналітична модель не є структурно подібною до об'єкту моделювання. Під ***структурною подобою*** розуміється однозначна відповідність елементів і зв'язків моделі елементам і зв'язкам об'єкту, що моделюється.

До аналітичних моделей відносяться ті, що побудовані на основі апарату математичного програмування, кореляційного, регресійного аналізу.

Якщо використається апарат математичного програмування, то модель в основі своєї складається із :

- цільової функції;
- змінних;
- системи обмежень на змінні.

Цільова функція, як правило, виражає ту характеристику СМО, яку потрібно обчислити або оптимізувати (продуктивність системи).

Змінні - виражають технічні характеристики СМО.

Обмеження – виражають допустимі граничні значення технічних характеристик .

Для реальних технологічних систем (автоматичних ліній, гнучких виробничих систем) розмірність їхніх аналітичних моделей дуже велика й вимагає значних обчислювальних витрат. Тому для підвищення їх обчислювальної ефективності використають різні засоби. Наприклад розбиття завдання великої розмірності на задачі меншої розмірності так, щоб автономні рішення таких особистісних задач у певній послідовності давали рішення основного завдання. Або зменшення точності обчислень, за рахунок чого вдається скоротити час рішення завдання.

Найпростішою аналітичною моделлю є рівняння прямолінійного рівномірного руху.

4.11. Контрольні питання

1. Дайте визначення предмету теорії масового обслуговування.
2. Дайте визначення випадковому процесу із рахованою множиною станів.
3. Дайте визначення поняттю найпростіший потік і перерахуйте його властивості.
4. Дайте визначення поняттю нестаціонарний пуассоновський потік.
5. Дайте визначення поняттю потік з обмеженою післядією.
6. Дайте визначення поняттю час обслуговування.
7. Дайте визначення поняттю марківський випадковий процес.
8. Дайте визначення поняттю система масового обслуговування з відмовами.
9. Дайте визначення поняттю система масового обслуговування з очікуванням і отримайте для неї рівняння Ерланга.
10. Дайте визначення поняттю система змішаного типу з обмеженням по довжині черги і отримайте для неї рівняння Ерланга.
11. Дайте визначення поняттю система змішаного типу з обмеженням по довжині черги і отримайте для неї рівняння Ерланга.
12. Дайте визначення поняттю сталий режим обслуговування.
13. Отримайте формули Ерланга для СМО з відмовами.
14. Охарактеризуйте основні методи рішення задачі розрахунку СМО.
15. Дайте визначення поняттю комп'ютерне - математичне моделювання СМО.
16. Дайте визначення поняттю імітаційне моделювання СМО.
17. Дайте визначення поняттю статистичне моделювання СМО.
18. Дайте визначення поняттю аналітичне моделювання СМО.
19. Дайте визначення поняттю чисельне моделювання СМО.
20. Охарактеризуйте аналітичні методи вирішення задачі розрахунку СМО.
21. Наведіть технологію комп'ютерного математичного моделювання.

Розділ 5

Застосування елементів теорії ухвалення рішень у системах управління

Розглянуто основні поняття і визначення теорії ухвалення рішень, загальну схему процесу ухвалення рішень. Надана класифікація задач ухвалення рішень. Також розкриті переваги особи, що ухвалює рішення. Наведено контрольні питання.

5.1. Загальна схема процесу ухвалення рішень

Протягом всього свого життя людина вимушена ухвалювати ті або інші рішення: куди піти вчитися, як краще витратити зайві (або останні) гроші і т.д.

Якщо певна ситуація, що вимагає ухвалення рішення, повторюється достатньо часто, то рішення приходить «само собою», немовби автоматично. Якщо ж ситуація недостатньо знайома, або людина не має в своєму розпорядженні всієї необхідної інформації, то ухвалення рішення істотно ускладнюється. У таких випадках людина вимушена, як правило, порівнювати між собою декілька можливих варіантів і вибирати той, який здається їй найбільш привабливим (або найменше небезпечним).

Наприклад, деякий службовець спізнюється на роботу; до зупинки, яка знаходиться на протилежній стороні вулиці, підходить автобус, а світлофор сяє червоним світлом. Є два варіанти дій — перетнути вулицю з ризиком для життя на червоне світло або поїхати на наступному автобусі. Для ухвалення рішення він повинен знати, коли з'явиться наступний автобус, а також співвіднести втрати від запізнення і можливого невдалого переходу вулиці.

У приведеному вище прикладі службовець сам ухвалює рішення, і сам же запроваджує його у життя. Значно важливіші наслідки мають так звані рішення, що управляють. Вони характеризуються тим, що вибір і реалізація рішення покладаються на різні елементи єдиної і, як правило, достатньо складною, системи. Ухвалюється рішення органом, що управляє, а реалізується — виконавчим органом. Система, засобами якої формується і реалізується рішення, може бути організаційною, технічною, або змішаною (комбінованою). Прикладом організаційної системи може служити будь-який навчальний заклад. У ньому органом, що управляє, є ректорат, а виконавчим — професорсько-викладацький і інженерно-технічний склад. Найбільш поширеними у даний час є змішані системи, які іноді називають ще «людино-машинними» системами або АСУ. До таких систем відноситься, зокрема, будь-яке сучасне промислове підприємство.

Одним з найважливіших атрибутів складної системи є наявність цілеспрямованої поведінки. У процесі досягнення мети система, так або інакше, взаємодіє із зовнішнім середовищем, яке може бути або «доброзичливим», або «ворожим», або «нейтральним». Очевидно, чим складніше система і чим складніше її взаємодія із середовищем, тим більше існує різних варіантів руху спрямованому на досягнення цілі, що поставлена. Одні варіанти можуть бути краще, інші гірше, треті взагалі можуть привести до руйнування системи. А може відшукатися один, самий найкращий, оптимальний, який задовольняє усіх.

Отже, щоб будь-яка система ефективно функціонувала, необхідно уміти:

- по-перше, оцінювати якість всіх можливих способів досягнення мети;
- по-друге, вибирати з них якнайкращий з погляду інтересів системи.

Для вирішення вказаних завдань розроблена спеціальна теорія, яка так і називається, — теорія ухвалення рішень.

Теорія ухвалення рішень – це область досліджень, що використовує поняття й методи математики, статистики, економіки, менеджменту й психології з метою вивчення закономірностей вибору людьми шляхів рішення різного роду завдань, а також способів пошуку найбільш вигідних (оптимальних) із можливих рішень.

У основі ухвалення рішення лежить дослідження операції. Під ***операцією*** у даному випадку розуміється процес досягнення цілі, що стоїть перед системою (з урахуванням її взаємодії із зовнішнім середовищем). ***Дослідження операції*** полягає в оцінці і порівнянні можливих способів її проведення з урахуванням наявних обмежень. Обмеження, як правило, пов'язані із часовими, матеріальними, людськими або іншими видами ресурсів, які знаходяться у розпорядженні сторони, що виконує операцію (суб'єкта операції). Таким чином, спосіб проведення операції визначається стратегією використання наявних ресурсів. Тому частіше використовують термін стратегія. Стратегії, що задовольняють накладеним обмеженням, називаються ***допустимими***. Поняття «допустима стратегія» є відносним. Безліч допустимих стратегій змінюються, якщо змінюються обмеження (або наявні ресурси).

Реалізація тієї або іншої допустимої стратегії приводить до різних результатів операції. Якість проведення операції, її «успішність» оцінюється з позицій ***особи, що ухвалює рішення*** (ОУР). Під цим терміном у теорії ухвалення рішень розуміється будь-який орган, що управляє, персональний або колегіальний, такий, що має біологічне або технічне втілення. У вказаному сенсі оцінка якості проведення операції завжди є суб'єктивною. Проте, для отримання такої оцінки повинні використовуватися об'єктивні методи.

Мірою ефективності проведення операції служить показник ефективності. У загальному випадку ***показник ефективності*** (ПЕ) відображає результат проведення операції, який, у свою чергу, є функцією трьох чинників:

- корисного ефекту операції ***(q)***;
- витрат ресурсів на проведення операції ***(c)***;
- витрат часу на проведення операції ***(t)***.

Значення ***q, c*** і ***t*** залежать від стратегії проведення операції ***(о)***. У формальному вигляді сказане можна записати так:

$$Y_{оп} = Y(q(o), c(o), t(o)).$$

Залежно від того, які сторони планованої операції цікавлять ОУР, список аргументів у вказаному виразі може змінюватися. Наприклад, якщо ефективність операції не залежить від її тривалості, то чинник часу ***t(о)*** може бути опущений. І навпаки, чинники, найбільш істотні з погляду ОУР, повинні бути деталізовані. Зокрема, витрати на проведення операції ***c(о)*** можуть бути представлені у вигляді вектора $\{c_1(o), c_2(o), , c_3(o),...\}$, кожна компонента якого відповідає певному типу ресурсів.

Необхідно відзначити, що чинники ***q, c*** і ***t*** можуть носити не тільки кількісний, але і якісний характер. Причому, форма їх опису залежить як від сфери діяльності, до якої відноситься дана операція (або, як то кажуть, «предметної області»), так і від можливості і вимог до точності їх оцінки. Зрозуміло, кількісні оцінки у багатьох випадках є об'єктивнішими, проте при рішенні деяких задач вони або просто не потрібні, або їх отримання є дуже трудомістким, а спрощення, що вводяться, спотворюють суть задачі, що вирішується.

Як ілюстрацію до викладеного розглянемо операцію по порятунку рибаків, що відносяться вітром від берега на відірваних крижинах. Очевидно, що кожна така операція може бути проведена або більш, або менш успішно залежно від вибраної стратегії пошуку потерпілих, сил і засобів, що використовуються і т.д. При кількісній оцінці результату операції чинники у показаному вище виразі отримують наступну інтерпретацію:

- ***q(o)*** - число врятованих рибаків;
- ***c(o)*** - вартість рятувальних робіт;
- ***t(o)*** - час на проведення операції.

Маючи можливість розрахувати вказані величини, можна отримати достатньо об'єктивну оцінку ефективності вибраної стратегії і проведення рятувальної операції. Проте не менш реалістичну оцінку можна отримати і при використанні якісних значень тих же величин:

- ***q(o)*** - зберігаючи той же сенс, приймає тільки одне з двох значень - вдалося врятувати всіх чи ні;
- ***c(o)*** - також має два можливі значення — перевищений кошторис на проведення робіт чи ні;
- ***t(o)*** - вдасться закінчити операцію до настання темряви чи ні.

Не дивлячись на певні переваги якісного підходу до оцінки результату операції, порівняння таких оцінок пов'язане із деякими труднощами. Тому все подальше викладення матеріалу буде відноситися до таких завдань пошуку рішення, в яких показник ефективності має кількісний вираз.

Отже, показник ефективності дозволяє оцінити (точніше, описати) результат операції, який отримано при використанні конкретної стратегії. Проте навіть якщо такі оцінки будуть отримані для всієї безлічі допустимих стратегій, цього ще не достатньо, щоб вибрати одну з них, ту, яка буде реалізована.

Наприклад, при оцінці завантаженості обчислювальної мережі виявилось, що коефіцієнт її використання дорівнює 0.7. Добре це або погано? Щоб відповісти на подібне питання, необхідно сформулювати правило, що дозволяє ОУР порівнювати між собою стратегії, які характеризуються різними значеннями ПЕ. У одних випадках правило порівняння може бути дуже простим, у інших же його взагалі не вдається знайти і доводиться змінювати (уточнювати) показник ефективності. Скажімо, якщо автомобіль однієї і тієї ж марки у двох різних автосалонах продається за різними цінами (за інших рівних умов), то правило вибору салону напрошується само собою. Зовсім інша справа, коли автомобілі розрізняються вартістю, фірмою, що їх виготовляє, дизайном, організацією гарантійного обслуговування і т.д. У такій ситуації покупець повинен спочатку визначити правило вибору і лише після цього порівнювати між собою різні варіанти.

У теорії ухвалення рішень правило, на підставі якого проводиться вибір стратегії, що відповідає інтересам ОУР, називається ***критерієм ефективності***.

Таким чином, показник ефективності і критерій ефективності у сукупності відображають цілі, які переслідує ОУР при проведенні даної операції, а також найбільш переважний для нього спосіб досягнення цієї мети.

Відома трагедія буриданова віслюка, який загинув з голоду, так і не вибравши одну з двох в'язанок сіна. Чому це відбулося? Та тому, що не було у нього ні показника, ні критерію ефективності, що відображають його переваги.

Необхідно відзначити, що на практиці переваги ОУР непостійні і можуть змінюватися навіть у одній і тій же ситуації вибору. У зв'язку з цим важливе значення має поняття концепції раціональної поведінки ОУР. Та лінія поведінки (концепція), якою дотримується ОУР, і визначає вибір правила, на основі якого порівнюватимуться стратегії.

Згідно теорії ухвалення рішень, ОУР може використовувати одну з трьох концепцій раціональної поведінки:

- придатності;
- оптимальності;
- адаптивності.

При використанні ***концепції придатності*** прийнятною вважається будь-яка стратегія, що забезпечує значення ПЕ не гірше заданого.

Концепція оптимальності вимагає, щоб зі всієї безлічі допустимих стратегій була вибрана тільки та, яка приводить до якнайкращого («екстремального») значення ПЕ.

Концепція адаптивної поведінки припускає, що правило вибору, може змінюватися відповідно до характеристик даної ситуації, що змінюються.

Приклад 5.1. Пояснимо відмінність у використанні різних концепцій раціональної поведінки, скориставшись приведеним раніше прикладом (оцінка завантаженості мережі). Нехай є три альтернативні стратегії організації роботи мережі. Перша забезпечує значення коефіцієнту завантаження, рівне *0.6*, друга — *0.7*, третя — *0.9*.

При використанні концепції придатності повинно бути задано мінімально допустиме значення коефіцієнта завантаження. Якщо воно дорівнює *0.7*, а оцінка стратегій проводилася у порядку їхньої нумерації, то як рішення, що управляє, буде вибрана друга (хоча при зміні порядку оцінки на її місці може опинитися і третя).

При використанні концепції оптимальності вибір у даній ситуації буде завжди однозначний — як стратегія управління роботи мережею прийматиметься тільки третя.

Навіть на тлі цього простого прикладу можна дати коротку порівняльну оцінку перших двох концепцій: концепція придатності вимагає, як правило, менших витрат часу на пошук рішення і володіє певною гнучкістю, зате концепція оптимальності гарантує вибір якнайкращого рішення з числа допустимих.

Ще більшою гнучкістю володіє концепція адаптивності. Для приведеного прикладу вона може бути реалізована так: допустимий рівень завантаження мережі змінюватиметься при зміні параметрів мережі (зокрема, її конфігурації), а разом з ним змінюватиметься і стратегія управління мережею.

І ось тепер знову повертаємося до поняття результату операції і пов'язаного з ним показника ефективності. Щоб порівнювати між собою різні стратегії, необхідно мати у своєму розпорядженні їх кількісні оцінки (тобто відповідні значення ПЕ). Яким чином вони можуть бути отримані? Найнадійніший спосіб — це вимірювання результату операції після її реального проведення (при цьому під вимірюванням може розумітися підрахунок, хронометраж, зважування і т. д.). Очевидно, що такий підхід пов'язаний із цілою низкою проблем.

По-перше, далеко не завжди можна повторити операцію у одних і тих же умовах (погода змінилася, виконавці втомилися або вийшли з ладу і т. п.), що, природно, не дозволяє говорити про об'єктивність вибору.

По-друге, багато операцій просто неможливо провести повторно, використовуючи іншу стратегію (наприклад, якусь історичну битву).

По-третє, реальне втілення системи, що використовується при проведенні операції, як правило, є вельми дорогою і трудомісткою справою, а якщо йдеться про порівняння проектних або конструкторських рішень, то витрати засобів і часу зростають пропорційно числу варіантів, що порівнюються.

Список проблем можна було б продовжити, але і приведених цілком достатньо, щоб зробити висновок: методу вимірювань повинна існувати якась альтернатива. І ось тут на перше місце виходить моделювання. Поки обмежимося достатньо загальним трактуванням даного поняття.

Моделювання — це заміщення об'єкту, що досліджується (оригіналу) його умовним образом або іншим об'єктом (моделлю) і вивчення властивостей оригіналу шляхом дослідження властивостей моделі.

Очевидно, що дійсна користь від моделювання може бути отримана тільки при дотриманні двох умов:

- модель забезпечує коректне (або, як то кажуть, адекватне) відображення властивостей оригіналу, істотних з погляду операції, що досліджується;
- модель дозволяє усунути перераховані вище проблеми, властиві проведенню вимірювань на реальних об'єктах.

Як відомо, залежно від способу реалізації, всі моделі можна розділити на два великі класи: фізичні і математичні. Згадаємо, що ***фізичні моделі*** припускають, як правило, реальне втілення тих фізичних властивостей оригіналу, які цікавлять ОУР. Наприклад, при проектуванні нового літака створюється його макет, що володіє тими ж аеродинамічними властивостями; при плануванні забудови архітектори виготовляють макет, що відображає просторове розташування її елементів. У зв'язку з цим фізичне моделювання називають також макетуванням.

Математична модель є формалізованим описом системи (або операції) на деякій абстрактній мові, наприклад у вигляді сукупності математичних співвідношень або схеми алгоритму. За великим рахунком, будь-який математичний вираз, у якому фігурують фізичні величини, можна розглядати як математичну модель того або іншого процесу або явища. Саме математичні моделі ми і розглядатимемо надалі як основний інструмент оцінки ефективності альтернативних стратегій. Викладене в даному розділі

Процес ухвалення рішення можна представити у вигляді загальної схеми наданої на рис. 5.1. Найбільш примітним у даній схемі є те, що процес пошуку (вибору) рішення носить циклічний характер. Мається на увазі, що будь-який з етапів може повторюватися неодноразово до тих пір, поки не буде знайдено рішення, що задовольняє вимогам ОУР (або не закінчиться час, відпущений на ухвалення рішення). При цьому можуть уточнюватися цілі і умови проведення операції, коректуватися модель переваг ОУР і модель самої операції.

Розглянемо загальна структура процесу ухвалення рішення. Ухвалення рішення - складний психологічний процес. Серед відомих видів мислення із процесом ухвалення рішення найбільше тісно зв'язане оперативне мислення, у ході якого формується "суб'єктивна модель передбачуваної сукупності дій", спрямованої на рішення поставленого завдання.

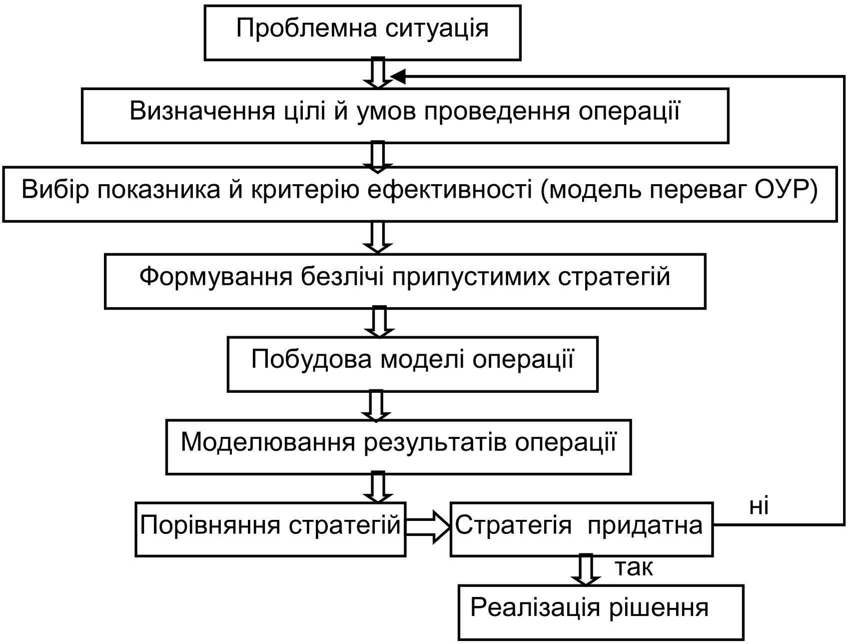


Рис. 5.1. Загальна схема процесу ухвалення рішення

У загальному випадку процес ухвалення рішення оператором АСУ доцільно аналізувати з наступних точок зору:

- логіко-психологічної;
- операційної;
- функціонально-динамічної;
- описової.

Логіко-психологічне дослідження дозволяє відбити послідовність ухвалення рішення, що містить у собі:

- постановку завдання;
- пошук, накопичення й регулювання інформації, необхідної для ухвалення рішення;
- виявлення й оцінку поточної ситуації з урахуванням проблемної орієнтації діяльності операторів;

- висування сукупності гіпотез;
- вибір рішення;
- реалізацію рішення.

Операційний опис дозволяє розглядати процес ухвалення рішення у вигляді композиції трьох множин:

$$H = H_1 + H_2 + H_3,$$

де: безліч H_1 характеризує сукупність операцій інформаційної підготовки процесу ухвалення рішення;

безліч H_2 характеризує етап вибору рішення;

безліч H_3 характеризує дії для реалізації рішення.

Інформаційна підготовка ухвалення рішення пов'язана із селекцією інформації про об'єкт, що управляється та середовище, що дозволяє досягти максимальної ефективності рішення. Важливо відмітити, що інформаційна підготовка ухвалення рішення складається із:

- зовнішнього інформаційного забезпечення;
- внутрішнього інформаційного забезпечення.

Зовнішнє інформаційне забезпечення виконується при апріорній підготовці ухвалення рішення і вирішує питання відбору необхідної інформації й вибору способів її оптимального подання.

Внутрішня інформаційна підготовка виконується при рішенні конкретних оперативних завдань і містить у собі процедури класифікації й узагальнення інформації про поточні ситуації, побудову оперативних моделей діяльності.

Безпосередньо вибір рішення складається із наступних етапів:

- формування робочих гіпотез і зіставлення їх із концептуальними моделями поточних ситуацій;
- коректування сформованих моделей;
- оцінки співвідношення гіпотез і результатів, що досягаються;
- вибору найкращої гіпотези й послідовності дій для реалізації рішення відповідно до обраної гіпотези.

Дії оператора по вибору рішення (H_2) принципово не можливо формалізувати. На цьому етапі оператор використає три основні форми розумової діяльності:

- ***діалектичне мислення*** - є вищою формою психічних процесів людини, невіддільно від неї й служить джерелом багатопрограмності, адаптації, вибору рішення по неповній і неточній інформації;
- ***емпіричне мислення*** - базується на узагальненні успіхів попереднього досвіду;
- ***аксіоматичне мислення*** - пов'язане із вихідним знанням правил рішення завдань.

Функціонально-динамічні дослідження - спрямовані на вивчення комплексу психологічних механізмів ухвалення рішення і принципів інформаційно-евристичного пошуку.

Застосування перерахованих напрямків досліджень процесу ухвалення рішення дозволяє, на основі вивчення закономірностей людського мислення, імітувати у технічних засобах ряд функцій мислення, які можна формалізувати.

5.2. Класифікація завдань ухвалення рішень

Одною із найбільш важливих умов скорочення витрат часу і сил під час пошуку оптимального рішення — це уміння правильно вибрати метод пошуку.

Завданням ухвалення рішення називають кортеж (сукупність):

$$\acute{\Omega} = (X,O),$$

де – ***X*** - безліч варіантів рішення задачі;

O – принцип оптимальності, що дає уявлення про якість варіантів, у простому випадку – це правило їх переваги одного перед іншим.

Рішенням задачі ухвалення рішень називається множина $X_{опт}$, яка є підмножиною множини ***X***, що отримана на основі принципу оптимальності.

Завдання ухвалення рішень класифікуються по наявності інформації про множини ***X*** та ***O*.** Можна виділити три види таких завдань:

1. Множини ***X*** і ***O*** – невідомі. Це загальне завдання ухвалення рішень. Дані для отримання $X_{опт}$ визначають у даному завданні в процесі її рішення.
2. Безліч ***X*** – невідома, безліч ***O*** – відома (це завдання пошуку варіантів).
3. Безліч ***X*** і безліч ***O*** – відомі (це завдання оптимізації).

У загальному випадку завдання ухвалення рішення вирішується в два етапи:

1 етап - завдання формалізується, тобто будується його математична модель, в якій конкретні фізичні, технічні, технологічні, економічні умови і вимоги до об'єкту втілюються у вигляді завдання оптимізації із певною цільовою функцією і допустимою безліччю варіантів.

2 етап - рішення задачі оптимізації із використанням відомих методів.

Практично будь-яка ситуація, що вимагає ухвалення рішення, може бути віднесена до того або іншого відомого класу, і дослідникові залишається тільки «впізнати» її. Для цього потрібно, щонайменше, мати уявлення про характерні ознаки різних класів. У даний час відсутня єдина універсальна класифікаційна схема завдань ухвалення рішень, проте практично у всіх виданнях, присвячених цим питанням, фігурують наступні класифікаційні ознаки (рис. 5.2):

- число осіб, що ухвалюють рішення;
- вид показника ефективності;
- ступінь визначеності інформації про проблемну ситуацію;
- залежність характеристик проблемної ситуації від часу.

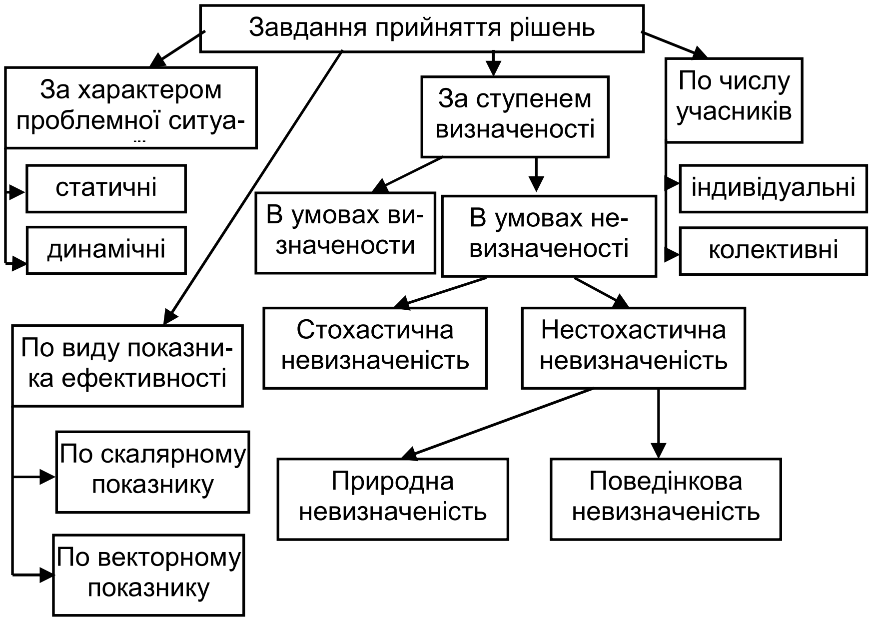


Рис. 5.2. Класифікація завдань ухвалення рішень

За ознакою числа ОУР розрізняють завдання індивідуального і групового ухвалення рішень. При груповому ухваленні рішень визначальну роль грає проблема узгодження індивідуальних переваг членів групи.

По вигляду показника ефективності завдання ухвалення рішення підрозділяють на завдання зі скалярним або векторним показником.

При використанні скалярного ПЕ передбачається, що ОУР цікавить тільки одна зі складових результату операції, наприклад, її тривалість. Це найбільш простій випадок при виборі стратегії. Але простота не означає, що визначити значення скалярного ПЕ теж легко. Практично всі методи математичного програмування призначені для пошуку рішень саме по скалярному показнику. При використанні цих методів у ролі показника ефективності виступає цільова функція.

При порівнянні стратегій по векторному ПЕ можуть бути використані спеціальні методи, що дозволяють звести векторний показник до скалярного.

За ступенем визначеності інформації про проблемну ситуацію розрізняють завдання ухвалення рішень в умовах визначеності і завдання ухвалення рішень в умовах невизначеності.

Завдання ухвалення рішень в умовах визначеності характеризуються наявністю повної і достовірної інформації про проблемну ситуацію, цілі, обмеження і наслідки рішень, що ухвалюються. У таких завданнях наперед, до початку операції, відомо, до якого результату приведе застосування кожної із стратегій. Це, зокрема, означає, що всі зовнішні чинники відомі, враховані, і вони не можуть яким-небудь непередбаченим чином вплинути на результат операції.

Характерна особливість всіх завдань ухвалення рішень, в умовах невизначеності, полягає в тому, що результат операції залежить не тільки від стратегій ОУР і фіксованих чинників, але і від невизначених чинників, які не контролюються ОУР і не відомих йому у момент ухвалення рішення (або недостовірно відомих). У результаті кожна стратегія виявляється пов'язаною із безліччю можливих результатів операції, що істотно ускладнює процес вироблення рішення.

Завдання ухвалення рішень в умовах невизначеності підрозділяють на завдання стохастичної і нестохастичної невизначеності. У разі стохастичної невизначеності кожній стратегії відповідає деяка кінцева безліч результатів, причому досліднику відомі їх імовірнісні характеристики. Але навіть якщо він орієнтується на найбільш імовірний результат, це не означає, що операція розвиватиметься саме за означеним сценарієм. Тому завдання такого типу називають також ухваленням рішень в умовах ризику. Вони мають місце у тих випадках, коли на результат операції можуть вплинути ті або інші випадкові чинники. Наприклад, якщо зустрічаються дві приблизно рівні по силі футбольні команди, то можливі три різні результати матчу (або перемагає перша команда, або друга, або зустріч закінчується внічию). Імовірність кожного із результатів може бути відома, але який саме буде реалізований, залишається загадкою практично до закінчення гри.

Завдання ухвалення рішень в умовах нестохастичної невизначеності підрозділяються, у свою чергу, на завдання ухвалення рішень в умовах природної або поведінкової невизначеності.

Такі завдання виникають у тих випадках, коли ОУР не має у своєму розпорядженні імовірнісних характеристик можливих результатів операції, або вони взагалі не є випадковими. У кращому разі відомі лише діапазони їх значень.

Якщо обмеженість інформації обумовлена недостатнім рівнем вивчення природи явищ, що досліджуються, то говорять про завдання з «природною» невизначеністю. Якщо ж недолік інформації обумовлений впливом на хід операції інших суб'єктів крім ОУР, то має місце завдання «поведінкової» невизначеності. Для вирішення завдань із поведінковою невизначеністю використовуються методи теорії ігор.

По характеру залежності проблемної ситуації від часу розрізняють статичні і динамічні завдання. У динамічних завданнях параметри (характеристики) проблемної ситуації змінюються у часі.

Розглянута класифікація завдань ухвалення рішень за перерахованими ознаками передбачає їх різноманітні комбінації. Наприклад, вибір варіанту атаки льотчиком винищувача може бути класифікований як динамічне, скалярне завдання індивідуального ухвалення рішення в умовах поведінкової невизначеності, оскільки у даному випадку як ПЕ використовується імовірність знищення цілі і для будь-якої вибраної стратегії результати випадкові, що обумовлене наявністю чинника невизначеності — поведінкою супротивника.

Для скорочення часу й підвищення якості прийняття рішень при управлінні складними технічними системами реального часу необхідна розробка та впровадження відповідних програмно-апаратних засобів, що забезпечують (підтримують) діяльність ОУР та створюють умови для ефективного планування діяльності по управлінню роботою підприємства. Такі програмно-апаратні комплекси отримали назву системи підтримки ухвалення рішень (СПУР). Вони є людино-машинними інформаційними системами, які дозволяють ОУР, використовувати фактологічну інформацію, накопичені знання, об'єктивні й суб'єктивні моделі для аналізу й рішення слабко структурованих або неструктурованих проблем.

Системи підтримки ухвалення рішень призначені для виконання інформаційної підтримки процесу ухвалення рішення людиною оператором і повинні забезпечувати наочність форм подання інформації, швидкість одержання нових видів звітності, можливість аналізу поточних і історичних даних.

Застосування систем підтримки ухвалення рішень є необхідною умовою, що дозволяє підприємствам «виживати» в умовах жорсткої конкурентної боротьби.

При виборі альтернативних варіантів рішень, що ухвалюються вибирається те, яке найбільше повно відповідає поставленій цілі, з урахуванням великої кількості обмежень, суперечливих вимог і багатьох критеріїв.

У загальному випадку можна виділити три основні класи невизначеностей, які необхідно враховувати у процесі ухвалення рішень:

- невизначеності, що пов'язані із неповнотою наших знань про проблему по який приймається рішення;
- невизначеність, що виникає у зв'язку із непередбачуваністю реакції навколишнього середовища на наші дії;
- невизначеність, що виникає у зв'язку із неточним розумінням цілі безпосередньо самою особою, що ухвалює рішення.

Визначимо фактори, що утрудняють процес ухвалення рішень. По-перше, не можна звести всі завдання із невизначеністю до формалізованих, тому треба робити правки на суб'єктивність рішення експерта. По-друге, в умовах стрімкого розвитку і впровадження у промисловості новітніх технологій кількість факторів, які підлягають обліку незупинно зростає, а час на їх аналіз знижується.

Також різноманітні невизначеності викликають помилкові рішення наслідок від яких принципово розрізняються у частині того, що не отриманий виграш іноді несе менше шкоди ніж реалізований програш, тобто у першому випадку це не реалізована вигода, у другому - це прямі суттєві втрати. Наприклад Для генерала почати воєнну операцію, що буде програна, набагато гірше, ніж упустити ситуацію, коли можна було провести успішну операцію.

Узагальнюючи розглянутий матеріал можна вважати, що процес підтримки ухвалення рішення полягає у наступному:

- надання інформаційної допомоги ОУР, необхідної для аналізу об'єктивної складової проблеми, яка підлягає вирішенню;
- виявлення і аналіз переваг ОУР;

- облік невизначеностей, які необхідно враховувати у процесі ухвалення рішень ОУР;
- генерація набору можливих рішень проблеми, яка існує;
- оцінка можливих рішень проблеми, виходячи із переваг ОУР і обмежень, що накладаються і впливу зовнішнього середовища;
- аналіз можливих наслідків від реалізації варіантів рішень, які підлягають ухваленню;
- вибір найкращого варіанту, з погляду особи, що ухвалює рішення.

5.3. Опис переваг особи, що ухвалює рішення

Як було показано раніше, модель переваг ОУР служить основою для ухвалення остаточного рішення. Слово «модель» у даному випадку означає формалізований опис відповідних категорій, який забезпечує повторення процедури вибору у однотипних ситуаціях при різних початкових даних. Крім того, модель переваг ОУР може бути використана для автоматизації процесу пошуку оптимального рішення.

У свою чергу, переваги ОУР можуть бути виражені за допомогою вибору показника і критерію ефективності. Найбільш простий варіант передбачає, що для кожної стратегії поведінки де якого об’єкту, яка ухвалюється, відповідно до поточної ситуації, може бути поставлено у відповідність значення скалярного ПЕ (тут застосовуються методи математичного програмування). Складніша (і поширеніша на практиці) ситуація виникає, коли кожна з допустимих стратегій характеризується векторним ПЕ. Труднощі вибору ще більше зростають в умовах невизначеності, за відсутності однозначної відповідності між стратегіями і їх векторними оцінками.

Необхідно відзначити, що модель переваг ОУР може бути побудована, якщо така система переваг володіє властивостями повноти і спрямованості.

Але у будь-якому випадку для успішного вирішення завдання вибору оптимальної стратегії поведінки необхідно виявити і виміряти переваги ОУР.

Зазвичай під виявленням переваг ОУР розуміють процес отримання інформації про його думки щодо можливих результатів операції і найбільш раціональних шляхів її реалізації. Існує два підходи до виявлення переваг ОУР:

- на основі аналізу інформації про раніше ухвалені рішення (при багатократному повторенні вибору у незмінних умовах);
- за допомогою спеціальної процедури опитування (до ухвалення рішення ОУР).

Саме у результаті виявлення переваг ОУР виявляється можливим визначити, чи володіють ці переваги властивостями повноти і спрямованості. Дамо визначення цим властивостям.

Система переваг ОУР володіє ***властивістю повноти*** на безлічі елементів вибору, якщо вона дозволяє порівняти між собою будь-які два елементи і винести одну з трьох альтернативних думок:

- c_1 переважно c_2;
- c_1 і c_2 рівноцінні;
- c_2 переважно c_1.

Властивість спрямованості означає наступне. Якщо, наприклад, при порівнянні елементів c_1 і c_2 ОУР виносить думку: «c_1, переважно c_2», а при порівнянні елементів c_2 і c_3 — «c_2 переважно c_3», то при порівнянні елементів c_1 і c_3 його висновок повинен бути однозначний: «c_1 переважно c_3».

Вимірювання переваг є відображенням альтернативних варіантів рішень на числову вісь.

У разі опису стратегій поведінки об’єкта, що управляється, за допомогою скалярного ПЕ вимірювання переваг не викликає труднощів. Іноді може бути отримана функціональна залежність між особистісними ПЕ і результатом операції, що дозволяє представити результат у вигляді скалярної величини.

У складніших випадках використовуються інші способи вимірювання переваг із застосуванням так званих вимірювальних шкал.

Шкала вимірювань — це система позначень, що дозволяє поставити у відповідність об'єкту деяку ознаку і використовувати її надалі для порівняння об'єктів між собою. Найбільшого поширення набули метричні, порядкові і номінальні шкали. Розглянемо їх в порядку зростання можливостей.

Номінальна шкала (або шкала найменувань) - це, по суті, якісна шкала. Її застосовують для позначення приналежності об'єктів до певних класів. Вона дозволяє описати відношення еквівалентності і відмінності між об'єктами. Проте перевага між об'єктами і між класами не встановлюється. Числа у цій шкалі використовуються тільки для позначення класу об'єктів. Приклад - номер цеху або виробничої ділянки на підприємстві.

Порядкова (рангова) шкала — застосовується для вимірювання впорядкованості об'єктів за однією ознакою або по їх сукупності. Числа у ній задають тільки порядок проходження об'єктів, але не дозволяють визначити, наскільки один об'єкт переважніше іншого. Приклад: розподіл місць у естафетній гонці (коли інтерес представляє тільки порядковий номер на фініші, а не різниця в часі); призначення пріоритетів заявкам, що поступають на обслуговування у СМО, і т.д.

Метричні шкали є найбільш здійсненими. На практиці застосовуються наступні їх види.

Шкала інтервалів - використовується для опису відмінності властивостей об'єктів у вигляді різниці. Вимірювання по даній шкалі дозволяють визначити, наскільки один об'єкт кращий від інших. При цьому обов'язково задаються масштаб вимірювань і початок відліку. У шкалі інтервалів вимірюються, наприклад, терміни виконання різних робіт, гарантійні терміни служби пристроїв, об'єм витрат на проведення операції.

Шкала відносин - є окремим випадком шкали інтервалів, вона дозволяє визначати не тільки різницю, але і відношення між значеннями ПЕ. Показники, що вимірюються у шкалі відносин, найбільш поширені в техніці і математиці. До них відносяться, наприклад, довжина, маса, напруга і т.д.

Абсолютну шкалу прийнято вважати найбільш досконалою. У цій шкалі використовується нульова точка відліку і одиничний масштаб. Це означає, що вимірювання у ній можуть бути проведені єдиним способом (у відмінність, наприклад, від шкали відносин). Абсолютною є, наприклад, шкала температур по Кельвіну, шкала значень імовірності події і т.п.

5.4. Контрольні питання

1. Охарактеризуйте предметну область теорії ухвалення рішень.
2. Дайте визначення поняттю допустима стратегія.
3. Дайте визначення поняттю шкала інтервалів.
4. Дайте визначення поняттю абсолютна шкала.
5. З яких етапів складається процес вибору рішення.
6. Дайте визначення поняттю шкала відносин.
7. Дайте визначення поняттю порядкова шкала.
8. Охарактеризуйте властивості спрямованості переваг ОУР.
9. Охарактеризуйте властивості повноти переваг ОУР.
10. Наведіть і поясніть класифікацію завдань ухвалення рішень.

11. Дайте визначення поняттю номінальна шкала.
12. Наведіть і поясніть загальну схему процесу ухвалення рішення.
13. Дайте визначення поняттю порядкова шкала.
14. Дайте визначення поняттю операційний опис процесу ухвалення рішення.
15. Дайте визначення поняттю концепція оптимальності.
16. Дайте визначення поняттю концепція придатності.
17. Дайте визначення поняттю концепція адаптивності.
18. Які чинники визначають результат проведення операції.
19. Охарактеризуйте основні підходи до виявлення переваг ОУР.
20. У чому полягає процес підтримки ухвалення рішення.
21. Дайте визначення поняттю діалектичне мислення.
22. Дайте визначення поняттю аксіоматичне мислення.
23. Дайте визначення поняттю емпіричне мислення.
24. Дайте визначення і охарактеризуйте основні класи невизначеностей, які необхідно враховувати у процесі ухвалення рішень.
25. Які фактори утрудняють процес ухвалення рішень.

Розділ 6

Застосування елементів теорії ігор у процесі ухвалення рішення

Розглянуто основні поняття і визначення теорії ігор, рішення гри у змішаних стратегіях, основна теорема теорії матричних ігор. Також розкрито принципи максиміна й мінімакса, формальний опис рішення гри, гра як модель конфліктної ситуації. Надано приклади застосування теорії ігор. Наведено контрольні питання.

6.1. Основні поняття та визначення теорії ігор. Гра як модель конфліктної ситуації

Розглянемо методи теорії ігор, які застосовуються для аналізу й ухвалення рішень у конфліктних ситуаціях, коли є дві сторони, що переслідують протилежні цілі. У загальному випадку теорія ігор це - математична теорія вирішення конфліктів. Уперше математичні аспекти й додатки теорії ігор були викладені у 1944 року в класичній книзі Джона фон Неймана й Оскара Моргенштерна "Теорія ігор і економічне поводження". Перші концепції теорії ігор аналізували антагоністичні ігри, коли є гравці, що програли і такі, що виграли за їхній рахунок. Але практичний досвід показав, що класичний підхід до конкуренції, який описав А. Сміт, коли кожний сам за себе, виявився неоптимальним варіантом поведінки у конфліктних ситуаціях. Більш оптимальними виявилися дії гравців, коли кожний намагається зробити краще для себе, роблячи краще для інших.

Дамо визначення поняттю ***конфлікт*** - це така ситуація (збіг обставин), у якій зіштовхуються інтереси сторін, відбувається боротьба інтересів. Тобто кожний з учасників конфлікту переслідує особисті інтереси, які не збігаються з інтересами інших.

Стосовно АСУ, під ***конфліктною ситуацією*** будемо розуміти ситуацію, що виникає у процесі управління складною технічною системою при неузгодженості дійсного й необхідного станів системи, яка потребує від особи, що ухвалює рішення конкретної альтернативи управління при наявності інформації про стан об'єкта й системи управління, критеріїв і вирішальних правил та власної системи переваг. Найбільш типовим прикладом конфліктних ситуацій є гра у шахи, або конкурентна боротьба у економічних системах за міжнародний ринок.

Причини, що породжують конфліктні ситуації, можна умовно розділити на п'ять груп:

- ситуації першої групи виникають внаслідок ненадійності елементів системи;
- ситуації другої групи є наслідком недосконалості самого процесу управління, що обумовлено неповнотою й неточністю інформації про об'єкт управління, недосконалістю методів і алгоритмів управління, недоліками й помилками оперативного персоналу й т.д.;
- ситуації третьої групи виникають у зв'язку із обмеженими можливостями системи управління, обмеженнями на ресурси й т.д.;
- ситуації четвертої групи пов'язані із подоланням багатозначності, що виникає у процесі управління;
- ситуації п'ятої групи потрібують ухвалення рішення, коли система управління стає нездатною до рішення виникаючих завдань.

Для АСУ реального часу характерно проявлення усіх ***джерел невизначеності*** таких, як:

- невідомість ситуації, що склалася;
- неповнота інформації;
- невірогідність інформації;
- випадковість подій, що трапляються;
- неточність оцінок;
- багатозначність факторів, які підлягають оцінюванню.

У процесі функціонування системи ці невизначеності повинні долатися оперативним персоналом АСУ на основі:

- знання конструкції й принципів функціонування об'єктів та підсистем АСУ;
- постійного аналізу й передбачення спрямованості процесу управління;
- особистого досвіду, інтуїції й високої професійної підготовки.

Усунення означених невизначеностей і безпосереднє ухвалення рішення є результатом мислення оператора, який виступає в якості ОУР із обліком його:

- суб'єктивних поглядів, суджень і емоцій;

- суб'єктивної моделі дій, що передбачувані, які спрямовані на рішення завдання, що поставлено.

Теорію ігор можна використовувати для опису й моделювання поводження людських популяцій - гравці повинні вибирати поводження, яке максимізує їхню сумарну вигоду. Однак у ході багатьох досліджень було виявлено, що люди не додержуються рівноважних стратегій на практиці (нераціональність, особисті мотиви гравців, включаючи альтруїзм). Тому теорія ігор розглядається не як інструмент передбачення поведінки деякого біологічного об’єкту, а як інструмент аналізу ситуації, що склалася, з метою виявлення найкращого поводження гравця.

Для того, щоб зробити можливим математичний аналіз конфліктної ситуації, її необхідно спростити, урахувавши тільки основні фактори і представити конфлікт у ігровій формі, указавши:

- дії, які можливі для учасників;
- до якого результату приведе гра, якщо кожний із гравців вибере певне поводження.

Спрощена формалізована модель конфліктної ситуації називається ***грою***, а сторони, що конфліктують - ***гравцями***. Інакше кажучи під грою розуміється процес, у якому беруть участь дві й більше сторони, що ведуть боротьбу за реалізацію своїх інтересів. Кожна зі сторін має свою ціль й вживає деякі дії, які можуть вести до виграшу або програшу - залежно від поводження супротивників, з урахуванням їх ресурсів і їхніх можливих вчинків.

Гра являє собою сукупність правил, які описують поведінку гравців. Але не кожний конфлікт, що зустрічається на практиці, протікає за правилами. Гра охоплює декілька періодів, протягом яких гравці здійснюють послідовні або одночасні дії, які є елементами гри і позначаються терміном ***хід***. Дії можуть бути пов'язані із цінами, обсягами продаж, витратами на наукові дослідження й розробки й т.д. Періоди, протягом яких гравці роблять свої ходи, називаються ***етапами гри***. Правила гри указують, яка повинна бути послідовність ходів і указують характер кожного ходу. Обрані на кожному етапі ходи у остаточному підсумку визначають "***платежі***" (виграш або збиток) кожного гравця, які можуть виражатися в матеріальних цінностях або грошах. Для спрощення, обмежимося розглядом ігор, у яких є тільки дві конфліктуючі сторони.

Узагальнюючи розглянутий матеріал можна виділити наступні ознаки, що характеризують, гру як математичну модель конфліктної ситуації:

- наявність декількох учасників;

- невизначеність поводження учасників, яка пов'язана із наявністю у кожного з них декількох варіантів дій;
- розходження (розбіжність) інтересів учасників конфлікту;
- взаємозв'язок поводження учасників конфлікту, оскільки результат, одержуваний кожним з них, залежить від поводження всіх учасників;
- наявність правил поведінки, що відомі всім учасникам гри.

Наслідок конфлікту передбачити у точності не можливо, тому що він залежить від випадку (влучення або невлучення у ціль при стрільбі). Отже, замість "результату гри" потрібно говорити про результат, що доводиться у середньому на одну партію гри, якщо буде зіграна досить велика кількість партій. У одній партії може випадково "поталанити" будь-якому гравцю. Якщо ж партій буде багато, то у середньому виграє той, хто поводиться розсудливо. Отже ***основну ідею теорії ігор*** можна сформулювати так: як повинен поводитися розумний гравець у конфлікті з розумним супротивником, щоб забезпечити собі у середньому найбільший можливий виграш.

Результат гри, навіть у тому випадку, коли він не має прямої кількісної оцінки, характеризується деяким значенням, наприклад, виграш *+ 1*, програш *- 1*, нічия *0*. Гра може бути парною або. множинною (з багатьма учасниками).

Розвиток гри відбувається у результаті послідовного виконання тих або інших ходів. Хід називається ***особистим***, якщо він виконується одним із гравців у результаті свідомого аналізу ситуацій.

На противагу особистим, ***випадкові*** ходи виникають не у результаті свідомого рішення, а у результаті того або іншого випадкового процесу (наприклад, роздавання карт).

Сукупність правил, що визначають вибір варіанту дій при кожному ході, залежно від ситуації, що склалася у ході гри називається - ***стратегія гравця***. Інакше кажучи, стратегія гравця - це можливі дії, що дозволяють гравцю на кожному етапі гри вибирати із певної кількості альтернативних варіантів такий хід, що представляється йому "кращою відповіддю" на дії інших гравців. Гра називається ***кінцевою***, якщо у кожного гравця є лише кінцеве число стратегій. Якщо таких стратегій нескінченно багато (нехай навіть тільки у одного гравця), то гра називається ***нескінченною***. Кожна стратегія гри повинна мати свою концепцію. ***Концепція стратегії*** припускає, що гравець визначає свої дії не тільки для етапів, яких фактично досягла конкретна гра, але й для всіх ситуацій, включаючи й ті, які можуть і не виникнути у ході даної гри.

Основні типи ігор, з погляду на них, як на конфліктні ситуації можна класифікувати наступним чином (рис. 6.1.).

У ***екстенсивній формі*** гра представляється у вигляді орієнтованого дерева, де кожна вершина відповідає ситуації, яка породжується при виборі гравцем своєї стратегії. Це ігри, що передбачають наявність більш ніж двох гравців виконання послідовних ходів.

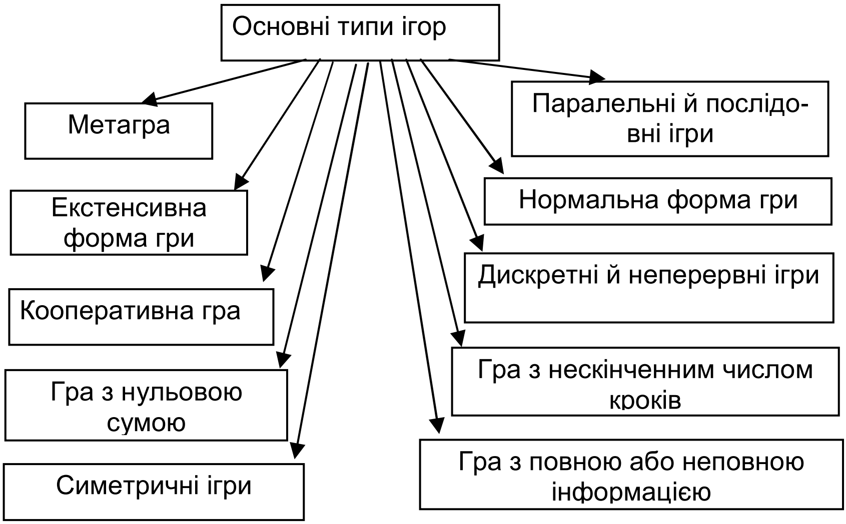


Рис. 6.1. Класифікація основних типів ігор

Наприклад, гравець ***1*** ходить першим і вибирає стратегію ***F*** або стратегію ***U***. Аналіз можливих наслідків гри підказує гравцю, що це скоріше буде стратегія ***U*** (рис. 6.2). Гравець ***2*** аналізує свою позицію й вирішує вибрати стратегію ***A*** або ***R***. Скоріше це буде стратегія ***A*.** Для кожного із гравців це будуть оптимальні стратегії.

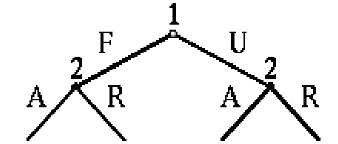


Рис. 6.2. Граф ходів

Гра буде ***симетричною*** тоді, коли відповідні стратегії у гравців будуть рівні, тобто матимуть однакові платежі та якщо гравці можуть помінятися місцями й при цьому їхні виграші за ті самі ходи не зміняться. Багато ігор для двох гравців є симетричними.

У ***нормальній формі*** гра описується платіжною матрицею (табл. 6.1.). Рядки визначають стратегії першого гравця, а стовпці -

другого. На перетинанні двох стратегій можна побачити виграші, які одержать гравці.

Якщо гравець 1 вибирає першу стратегію, а другий гравець - другу стратегію, то на перетинанні ми бачимо значення *(-1, -1),* тобто у результаті ходу обидва гравця втратили по одному очку. Чому?

Таблиця 6.1.

Стратегії гравців	Гравець 2 стратегія 1	Гравець 2 стратегія 2
Гравець 1 стратегія 1	4, 3	–1, –1
Гравець 1 стратегія 2	0, 0	3, 4

Гравці вибирали стратегії з максимальним для себе можливим результатом, але програли, через незнання ходу іншого гравця. Таким чином була отримана гра з неповною інформацією.

Гра називається ***кооперативною,*** або ***коаліційною***, якщо гравці можуть поєднуватися в групи, беручи на себе деякі зобов'язання перед іншими гравцями й координуючи свої дії. У некооперативних іграх - кожний зобов'язаний грати сам за себе.

У кооперативних іграх із ***трансферабельною*** корисністю, тобто можливістю передачі засобів від одного гравця до іншого, використають так звану ***характеристичну функцію***, що визначає виграш кожної коаліції гравців. Якщо у грі із двома сторонами утвориться коаліція ***C***, то проти її виступає коаліція ***B***. Утвориться як би гра для двох гравців. Але тому що варіантів можливих коаліцій багато (а саме $\boldsymbol{2^N}$, де ***N*** - кількість гравців), то виграш для ***C*** буде деякою характеристичною величиною, що залежить від складу коаліції.

Ігри ***із нульовою сумою*** є особливим різновидом ігор з постійною сумою, коли гравці не можуть збільшити або зменшити наявні ресурси, або фонд гри. У даному випадку сума всіх виграшів дорівнює сумі всіх програшів при будь-якому ході. Прикладами таких ігор може служити покер, де один виграє всі ставки інших. Грою із не нульовою сумою є торгівля, де кожний учасник має деяку користь. Або відомі ігри, коли сума виграшу зменшується, прикладом є війна.

Надалі, для спрощення сприйняття матеріалу, ми будемо розглядати парні ігри із нульовою сумою - при яких одна сторона виграє те, що програє інша.

Нехай у грі беруть участь два гравці ***A*** і ***B***. Гравці виконують ряд послідовних ходів, тобто роблять ряд послідовних дій, передбачених правилами гри.

Завдання пари стратегій гравців ***A*** і ***B*** у грі двох осіб повністю визначає її результат, тобто виграш одного гравця й програш іншого.

Як було визначено раніше, результати кінцевої парної гри із нульовою сумою можна задавати платіжною матрицею, рядки й стовпці якої відповідають різним стратегіям, а її елементи є відповідно виграшами однієї сторони (які дорівнюють програшам іншої). При цьому зручно програш першої сторони розглядати як її негативний виграш, а виграш - як її негативний програш.

Розглянемо кінцеву гру, у якій гравець ***A*** має ***m*** стратегій $(A_1, A_2, ..., A_m)$, а гравець ***B*** - ***n*** стратегій $(B_1, B_2, ..., B_n)$. Така гра називається грою ***m*** *x* ***n***. Якщо гравці ***A*** і ***B*** використають тільки особисті ходи, то вибір стратегій ***A*** і ***B*** однозначно визначає ***результат гри*** a_{ij}, тобто число, що характеризує виграш гравця ***A*** та програш гравця ***B***. Причому a_{ij} може бути як позитивним, так і негативним. Будемо вважати, що при $a_{ij}>0$ гравець ***A*** виграє, а гравець ***B*** програє величину a_{ij}. Якщо $a_{ij}<0$, то, навпаки, виграє гравець ***B*** і програє (має негативний виграш) гравець ***A***.

Якщо у грі використовуються випадкові ходи, то виграш при двох стратегіях A_i і B_j є випадковим. У цьому випадку за оцінку виграшу, що передбачається береться його математичне очікування.

Припустимо, що нам відомі всі значення a_{ij} у грі ***m*** *x* ***n***. Ці значення записуємо у вигляді таблиці платіжної матриці (табл. 6.2.), де рядки відповідають стратегіям A_i, а стовпці - стратегіям B_j.

Таблиця 6.2.

B A	B_1	B_2	…	B_j	…	B_n	β_j
A_1	a_{11}	a_{12}	…	a_{1j}	…	a_{1n}	
A_2	a_{21}	a_{22}	…	a_{2j}	…	a_{2n}	
…	…	…	…	…	…	…	
A_i	a_{i1}	a_{i2}	…	a_{ij}	…	a_{in}	
…	…	…	…	…	…	…	
A_m	a_{m1}	a_{m2}	…	a_{mj}	…	a_{mn}	
α_i							α β

Підхід до гри, коли для виграшу доводиться використовувати випадковий вибір стратегій, називається підходом із застосуванням змішаних стратегій. Отже, ***змішані стратегії*** отримуються шляхом випадкового вибору окремих чистих стратегій (одних єдиних) і для кожної стратегії, що змішуються, вказується імовірність її вибору.

Припустимо спочатку, що гра складена таким чином, що існує її рішення у чистих стратегіях. Спочатку визначимо найкращу зі стратегій гравця ***A***, тобто найкращу з $A_1, A_2, ..., A_m$ із урахуванням того, що на будь-яку стратегію A_i гравець ***B*** відповість стратегією B_j, для якої виграш гравця ***A*** виявиться мінімальним. Щоб знайти цю стратегію B_j, треба у рядку платіжної матриці, що відповідає стратегії A_i (рядку з номером ***i***), знайти мінімальне із чисел a_{ij}. Позначимо його як α_i,

$$\alpha_i = \min_j \alpha_{ij}, \quad j = \overline{1,n},$$

де мінімум визначається шляхом перебору всіх номерів стовпців.

При зміні стратегій гравця **А** відповідно до кожної із цих стратегій число α_i теж буде мінятися. Природно, що гравцю ***A*** вигідніше всього зупинитися на такій стратегії A_i, для якої значення α_i буде максимальним. Позначимо це максимальне значення α, тобто

$$\alpha = \max_i \alpha_i, \quad i = \overline{1,m},$$

або, враховуючи вираження для α_i, одержимо

$$\alpha = \max_i \min_j \alpha_{ij}.$$

Величину α прийнято називати ***нижньою ціною гри*** або ***максимінним виграшем*** (скорочено максиміном). Стратегію гравця ***A***, що відповідає максиміну, назвемо максимінною стратегією.

Якщо гравець ***A*** буде дотримуватися максимінної стратегії, то йому при будь-якому поводженні гравця ***B*** гарантовано виграш, у всякому випадку не менше, ніж α. Тому величину α називають нижньою ціною гри**,** тобто це той гарантований мінімум, що одержить гравець ***A*** в даній грі.

Аналогічно можна визначити найкращу зі стратегій гравця ***B***. Він прагне звернути виграш гравця ***A*** у мінімум. Для цього гравець ***B*** намагається для кожної своєї стратегії B_j одержати максимальне значення виграшу при будь-якій стратегії гравця ***A***, тобто він шукає значення β_j таке, що

$$\beta_j = \max_i \alpha_{ij}, \quad i = \overline{1,m}.$$

Однак гравець ***B*** не може розраховувати на те, що гравець ***A*** дозволить йому одержати любий з виграшів β_j**.** Єдине, на що може розраховувати гравець ***B****,* так це на те, що одержить виграш, що буде не менше, ніж величина β, яка обумовлена виразом

$$\beta = \min_j \max_i \alpha_{ij}.$$

Ця величина β називається ***верхньою ціною гри***, або ***мінімаксним виграшем*** (мінімаксом). Відповідна мінімаксу стратегія гравця ***B*** називається мінімаксною стратегією. Це більш обережна стратегія гравця ***B***, що забезпечує йому у любому випадку програш не більше β і, відповідно, виграш гравцю ***A*** також не більше β.

У теорії ігор принцип обережності, що рекомендує гравцям дотримуватися максимінної і мінімаксної стратегії, називається ***принципом мінімакса***. Він випливає із припущення про обережність гравців або із бажання вирішити конфліктну ситуацію найкращим чином для всіх сторін, що беруть участь у ній.

Якщо верхня ціна гри збігається з нижньою ціною, то їх загальне значення

$$\beta = \alpha = v$$

називається ***чистою ціною гри v***. Мінімаксні стратегії, що відповідають чистій ціні гри, є оптимальними стратегіями, а їхня сукупність - ***оптимальним рішенням.*** Пара чистих стратегій дає оптимальне рішення тоді й тільки тоді, коли відповідний їй елемент a_{ij} є одночасно найбільшим у своєму стовпці й найменшим у своєму рядку. У цьому випадку говорять, що гра має ***сідлову точку***. Стратегії, які змішуються для одержання оптимальної стратегії, будемо називати ***корисними***.

6.2. Основна теорема теорії матричних ігор. Рішення гри у змішаних стратегіях.

Змішані стратегії являють собою математичну модель мінливої й гнучкої тактики гравця, при якій гравець, що протистоїть йому, не може довідатися заздалегідь про ті обставини, з якими йому доведеться зустрітися у грі, тому що перед кожною партією здійснюється випадковий вибір однієї із чистих стратегій за допомогою деякого механізму, що робить цей вибір зі заздалегідь визначеними й заданими ймовірностями. Покажемо, як визначити ці ймовірності, виходячи з вимоги одержання максиміну, тобто оптимального рішення гри.

Розглянемо гру $m \times n$. Введемо позначення для змішаних стратегій гравців ***A*** і ***B***:

$$S_A=S_A(p_1, p_2, \dots, p_m),$$
$$S_B=S_B(q_1, q_2, \dots, q_n),$$

де $p_1, p_2, \ldots, p_m$ – імовірності використання чистих стратегій $A_1, A_2, \ldots, A_m$;

$q_1, q_2, \ldots, q_n$ – імовірності використання чистих стратегій $B_1, B_2, \ldots, B_n$.

Оскільки ці стратегії, за умовою гри, повністю вичерпують можливі ходи гравців ***A*** і ***B***, то вони утворять повну групу подій. Тому для ймовірностей p_i справедлива умова

$$\sum_{i=1}^{m} p_i = 1 \cdot$$

Аналогічна умова має місце для ймовірностей q_j:

$$\sum_{j=1}^{n} q_j = 1 \cdot$$

Очевидно, що кожна чиста стратегія є особистісним випадком змішаної стратегії, коли всі стратегії, крім обраної чистої стратегії, мають імовірності, що дорівнюють нулю, а ця обрана стратегія - імовірність, що дорівнює одиниці.

Виявляється, якщо допустити можливість використання змішаних стратегій, то для кожної гри можна знайти оптимальне рішення, тобто пару стійких оптимальних стратегій гравців.

У теорії ігор доводиться ***основна теорема матричних ігор***: будь-яка матрична гра має рішення у змішаних стратегіях.

Із цієї теореми витікає, що кожна матрична гра має ціну v, причому ця ціна завжди лежить між нижньою α й верхньої β ціною гри:

$$\alpha \le v \le \beta \cdot$$

Дійсно, α є мінімальним гарантованим виграшем, що може забезпечити собі гравець ***A***, застосовуючи тільки свої чисті стратегії. Застосовуючи змішані стратегії, гравець ***A*** у всякому разі не зменшить свій виграш, тобто $\alpha \le v$. Для гравця ***B*** аналогічні міркування приведуть до висновку, що $v \le \beta$. Тому у загальному випадку:

$$\alpha \le v \le \beta \cdot$$

Якщо у деякій грі $m \times n$ знайдене оптимальне рішення, що міститься у використанні двох оптимальних змішаних стратегій

$$S^*_A(p_1, p_2, \ldots, p_m),$$
$$S^*_B(q_1, q_2, \ldots, q_n),$$

то чисті стратегії $A_1, A_2, ..., A_m$ і $B_1, B_2, ..., B_n$, для яких імовірності p_i й q_j виявилися відмінними від нуля, називаються ***активними стратегіями*** гравців A и B.

Також у теорії ігор доводиться ***теорема про активні стратегії***: якщо один із гравців дотримується своєї оптимальної змішаної стратегії, то його виграш не менше ціни гри, незалежно від того, що робить інший гравець.

Зміст цієї теореми полягає в тому, що якщо один із гравців почне послідовно дотримуватися своєї оптимальної змішаної стратегії, то конфронтуючий йому гравець вже не зможе змінити результат гри.

6.3. Формальний опис рішення гри.

Рішення гри володіє однією важливою властивістю: якщо один із гравців використовує свою оптимальну стратегію, а інший змішує свої корисні стратегії, то середній виграш продовжує залишатися рівним ціні гри.

При цьому, щоправда, як і при будь-яких відступах від оптимальної стратегії, відповідна зміна стратегії супротивником може привести до збільшення його середнього виграшу. Доведено, що для гри m x n число корисних стратегій з кожної сторони не перевершує мінімального із чисел m і n.

Рішення всякої парної кінцевої гри із нульовою сумою може бути отримано методами лінійного програмування. Однак, перед тим, як використати ці методи, на практиці звичайно проводять попереднє спрощення завдання, виключаючи стратегії, які можуть бути безрезультатними.

Приклад 6.1. Нехай, є гра **3** x **3** із платіжною матрицею, що надана у табл. 6.3.

Таблиця. 6.3.

	B_1	B_2	B_3
A_1	1	0	2
A_2	0	2	1
A_3	-1	2	-2

Всі елементи третього рядка матриці не більше, ніж відповідні елементи другого рядка. Тому, застосовуючи другу стратегію, перший гравець одержить при будь-яких стратегіях супротивника не менший виграш, ніж при застосуванні другої стратегії. Виходить, що третя стратегія свідомо не може бути корисної, і її варто виключити з розгляду.

У матриці **2** x **3**, що залишилася, всі елементи третього стовпця більше, ніж відповідні елементи першого стовпця. Отже, якщо другий гравець застосує третю стратегію, то виграш його супротивника (тобто його власний програш) буде більше, ніж при застосуванні першої стратегії. Тому при пошуку рішення стратегія B_3 повинна бути свідомо виключена з розгляду.

У такий спосіб у результаті виключення свідомо безперспективних стратегій гра **3** x **3** звелася до гри **2** x **2** з матрицею, що надана у табл. 6.4.

Наступним кроком знаходимо нижню й верхню ціну гри:

Таблиця. 6.4.

	B_1	B_2
A_1	1	0
A_2	0	2

$$\alpha = \max_i \min_j \alpha_{ij} = \max_i (0,0) = 0,$$

$$\beta = \min_j \max_i \alpha_{ij} = \min_j (1,2) = 1 .$$

Оскільки максимін (нульовий елемент) зустрічається як у першому, так і у другому рядку, то обидві стратегій першого гравця є максимінними. Мінімаксною же стратегією другого гравця є лише стратегія B_1.

У нашому випадку $\alpha \neq \beta$, отже, рішення потрібно шукати серед змішаних стратегій.

Приклад 6.2. Нехай гра **2** x **3** із матрицею, що надана у табл. 6.5.

Знаходимо нижню й верхню ціну гри:

Таблиця. 6.5.

	B_1	B_2	B_3
A_1	0	3	-1
A_2	3	2	1

$$\alpha = \max_i \min_j \alpha_{ij} = \max_i (-1,1) = 1,$$

$$\beta = \min_j \max_i \alpha_{ij} = \min_j (3,3,1) = 1 .$$

У цьому випадку $\alpha = \beta = 1$. Оскільки $\beta < \gamma < \beta$ то ціна гри також дорівнює *1*. Рішення ж гри задається парою чистих стратегій (A_2, B_3), тому що максимін (він же мінімакс) розташований у другому рядку й третьому стовпці.

У області чистих стратегій рішення може бути отримане безпосередньо. Якщо ж рішення потрібно шукати у області змішаних стратегій (при $\alpha \neq \beta$), то у загальному випадку ***m*** x ***n*** матриці застосовується наступний прийом.

Вважаючи всі ***m*** стратегій першого гравця корисними, визначимо імовірності їхнього застосування у змішаній оптимальній стратегії (якщо якась стратегія в дійсності не корисна, то відповідна ймовірність звернеться до нуля). Позначимо ці імовірності $p_1, p_2, \ldots p_n$, а ціну гри (поки невідому) – γ.

Оскільки при оптимальній стратегії середній виграш першого гравця не менше ***m*** x ***n*** при будь-якій стратегії супротивника, то запишемо ***n*** нерівностей:

$$\begin{cases} p_1a_{11} + p_2a_{21} + \ldots + p_ma_{m1} \geq \gamma, \\ p_1a_{12} + p_2a_{22} + \ldots + p_ma_{m2} \geq \gamma, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ p_1a_{1n} + p_2a_{2n} + \ldots + p_ma_{mn} \geq \gamma. \end{cases} \quad (6.1.)$$

Введемо нові невідомі

$$x_1 = p_1 / \gamma.; \ldots, x_m = p_m / \gamma.$$

Щоб виключити можливість ділення на нуль, можна завжди дотримуватися умови щоб $\gamma > 0$. Для цього досить до всіх елементів матриці $\|a_{ij}\|$ додати те саме позитивне число C й всі її елементи зробити позитивними. Легко зрозуміти, що ця операція збільшує ціну гри на C, але не міняє оптимальних стратегій, що шукаються.

Оскільки $p_1 + p_2 + \ldots + p_m = 1$, то $x_1 + x_2 + \ldots + x_m = p_m / \gamma$.

Розділивши ліві й праві частини нерівностей (6.1.) на γ, одержимо нову систему нерівностей:

$$\begin{cases} x_1a_{11} + x_2a_{21} + \ldots + x_ma_{m1} \geq 1, \\ x_1a_{12} + x_2a_{22} + \ldots + x_ma_{m2} \geq 1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_1a_{1n} + x_2a_{2n} + \ldots + x_ma_{mn} \geq 1. \end{cases} \quad (6.2.)$$

При зроблених припущеннях всі x_i ненегативні. Оскільки ціль оптимальної стратегії - максимізація виграшу, то при її досягненні лінійна функція

$$f = x_1 + x_2 + \ldots + x_m = 1 / \gamma,$$

повинна звернутися у мінімум. Отже, оптимальна стратегія першого гравця (тобто набір ймовірностей $p_i = \gamma\, x_i$) знаходиться у результаті мінімізації функції

$$f = x_1 + x_2 + \ldots + x_m,$$

при ненегативних x_i, що задовольняють системі нерівностей (6.2.) це звичайне завдання лінійного програмування у результаті рішення якого знаходяться не тільки оптимальна стратегія першого гравця, але й ціна гри γ.

Знаючи ціну гри, оптимальну стратегію другого гравця звичайно можна знаходити вже без рішення завдання лінійного програмування. Для цього вибирається $n - 1$ корисна стратегія першого гравця (маючи можливість міняти місцями гравців, можна вважати, що $m \geq n$) і для кожної з них записується середній виграш, який при цьому повинен бути обов'язково дорівнювати ціні гри γ. Наприклад, якщо для першого гравця корисна стратегія A_i, то їй відповідає рівняння:

$$a_{i1}q_1 + a_{i2}q_2 + \ldots + a_{in}q_n = \gamma.$$

Крім того, є ще одне рівняння

$$q_1+q_2+\ldots+q_n= 1.$$

Усього маємо n рівнянь для n величин $q_1,q_2,\ldots,q_n$.

Для ігор 2×2 оптимальна стратегія обох гравців може бути, отримана також без рішення завдання лінійного програмування.

При $\alpha = \beta$ це очевидно, оскільки рішенням є чисті стратегії. При $\alpha \neq \beta$ обидві стратегії першого гравця й обидві стратегії другого обов'язково будуть корисними (інакше існувало б рішення у чистих стратегія). Тому нерівності (6.1.) перетворюються у цьому випадку в рівності:

$$\begin{cases} p_1a_{11} + p_2a_{21} =\gamma, \\ p_1a_{12} + p_2a_{22} =\gamma . \end{cases} \qquad (6.3.)$$

Оскільки $p_1 + p_2 = 1,$ то є достатнє число рівнянь для визначення значень всіх трьох невідомих p_1, p_2,γ.

Приклад 6.3. Застосуємо цей прийом для рішення першого з прикладів, що було розглянуто раніше (табл. 6.4.). Тоді рівняння (6.3.) запишемо у вигляді:

$$p_1 =\gamma, \quad 2p_2 =\gamma.$$

Приєднуючи до них рівняння

$$p_1+ p_2=1,$$

одержуємо рішення $p_1=2/3;\ \ p_2=1/3;\ \gamma=2/3.$

Отже, при випадковому чергуванні першої й другої стратегії з відносними частотами $2/3$ і $1/3$ - першому гравцеві забезпечений середній виграш у розмірі $2/3$.

Відзначимо, що цей висновок іде врозріз із інтуїцією, що підказує більш часто використати другу стратегію, оскільки вона має багатообіцяючий рядок виграшів: $(0,2)$ у порівнянні з (1.0) у першій стратегії.

Для знаходження оптимальної стратегії другого гравця достаточно (оскільки вже, відома ціна гри) скласти рівняння для його середнього програшу при кожній із двох корисних стратегій першого гравця, наприклад, стратегії A_1. Величина середнього програшу (дорівнює ціні гри) представиться при цьому у вигляді $1q_1+0q_2$.

Дорівнюючи її ціні гри $\gamma=2/3$, одержуємо $q_1=2/3$ і $q_2=1/3$. Отже, другий гравець повинен змішувати стратегії B_1 і B_2 із частотами $2/3$ і $1/3$ відповідно.

Описаний метод простого рішення гри може бути узагальнений також на ігри $2 \times n$.

Теоретико - ігрові методи можуть застосовуватися для вивчення ситуацій, які не є у точному значенні слова конфліктними. Мова йде про ситуації, де другим гравцем є навколишнє середовище.

Приклад 6.4. Є ділянка землі й дві стратегії її використання. Перша стратегія A_1 полягає у тому, щоб засівати ділянку ярицею, а друга A_2 - озимої (з можливим пересіванням навесні). У природи також є дві стратегії B_1 і B_2: перша, коли сильні морози наступають після випадіння снігу, а друга - до випадіння (що приводить до загибелі озимих). У цій грі ціною буде грошовий прибуток, що отриманий від реалізації врожаю.

Задамо деяку матрицю гри (табл. 6.6.), яка на практиці визначається з досвіду.

Визначимо нижню й верхню ціни гри:

$$\alpha = \max_i \min_j \alpha_{ij} = \max_i (5,2) = 5,$$

$$\beta = \min_j \max_i \alpha_{ij} = \min_j (10,6) = 6.$$

Таблиця. 6.6.

	B_1	B_2
A_1	5	6
A_2	10	2

Так як $\alpha \neq \beta$, то оптимальною буде змішана стратегія. Рівняння (6.3.) для нашого випадку дають:

$$\begin{cases} p_1 + 10p_2 = \gamma, \\ 6p_1 + 2p_2 = \gamma. \end{cases}$$

З урахуванням рівняння $p_1 + p_2 = 1$ отримаємо рішення:

$$p_1 = 8/9; \quad p_2 = 1/9; \quad \gamma = 50/9.$$

Визначимо оптимальну стратегію другої сторони:

$$5q_1 + 6q_2 = 50/9. \quad q_1 + q_2 = 1, \quad \text{звідки} \quad q_1 = 4/9,\ q_2 = 5/9.$$

Найбільш підступною стратегією з боку природи у цьому випадку буде така, для якої у середньому п'ять зим із дев'яти призводять до вимерзання озимих. Якби природа була б свідомим супротивником і дотримувалася б саме цієї стратегії, то кращою стратегією для нас було б висівання озимих лише один раз протягом дев'яти років. При цьому мали б середній річний прибуток у розмірі ***50/9*** тис. грн.

Якщо ми не маємо у своєму розпорядженні ніяку статистику зим, то, найдоцільніше припускати гірший випадок, тобто найбільш підступну стратегію з боку природи. Однак відмінність розглянутої ситуації від справді конфліктної складається у тому, що природа не є свідомим супротивником і її стратегія не є обов'язково оптимальною.

Припустимо, що ця реальна. (змішана) стратегія задається ймовірностями $q_1 = 0,.9$; $q_2 = 0,1$ (які можуть бути отримані у результаті спостережень), причому у чергуванні "поганих" і "добрих" зим немає ніякої іншої закономірності, крім цих ймовірностей. Завдання полягає у тому, щоб виробити оптимальну стратегію нашого поводження (імовірності p_1 і p_2), що дозволяє забезпечити максимальний середній виграш у цих умовах.

Математичне очікування виграшу виразиться сумою

$$\delta = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} p_i q_j ,$$

тобто

$$\delta = (5q_1+6q_2)p_1+(10q_1+2q_2)p_2 = 5{,}1p_1+9{,}2p_2.$$

Необхідно знайти максимум цієї величини при. умовах $p_1 \geq 0$, $p_2 \geq 0$, $p_1 + p_2 = 1$ отже $p_2 = 1-p_1$, звідки $\delta = 9{,}2 — 4{,}1p_1$. Ясно, що максимум цієї величини, що дорівнює $9{,}2$, може бути отриманий при $p_1=0;\ p_2=1$. У цих умовах правильною стратегією є висівання озимих у кожному році.

Більш ретельний аналіз ситуації показує, що оптимальна стратегія супротивника (у цьому випадку природи) є свого роду перевальною точкою: якщо ймовірність поганої зими перевищує обчислену вище оптимальну ймовірність $q_2=5/9$, то найбільш сприятливою стратегією буде щорічне висівання ярових. Якщо ця ймовірність менше ніж $5/9$, то раціонально висівати щороку озимі. Виграш у обох цих випадках буде більше, ніж ціна гри, що дорівнює $50/9$.

На ідеї пошуку оптимальної відповіді на кожну (не обов'язково оптимальну) стратегію іншої сторони ґрунтуються наближені методи рішення ігор, які застосовуються при великих платіжних матрицях і не близьких одна до одної α і β.

Для одержання наближення вживається наступний ітераційний метод. Гравець A вибирає деяку зі своїх стратегій, наприклад A_{i1}. Гравець B вибирає як відповідь ту стратегію B_{j1}, що найбільш невигідна для гравця A при стратегії A_{i1}. Гравець A відповідає на це кращою (проти стратегії B_{j1}) стратегією A_{i2}. Гравець B шукає стратегію B_{j2}, найкращу проти рівноімовірної суміші у рівних пропорціях стратегій A_{i1} і A_{i2}. Стратегія A_{i3} повинна бути найкращою проти суміші стратегій B_{j1} і B_{j2}, стратегія B_{j3} повинна бути кращою проти рівноімовірної суміші стратегій $A_{i1}, A_{i2},\ldots, A_{i3}$ і т.д.

Якщо продовжити цей процес досить довго, то частоти, з якими зустрічаються різні стратегії, прагнуть до імовірностей їхнього застосування у парі оптимальних стратегій, а середній виграш до ціни гри.

6.4. Рішення гри методами лінійного програмування

Покажемо, що рішення будь-якої кінцевої гри може бути зведено до завдання лінійного програмування.

Будемо вважати, що платіжна матриця гри a_{ij} нам відома (табл. 6.2.). Знайдемо рішення гри, тобто дві оптимальні змішані стратегії гравців ***A*** и ***B***:

$$S^*_A(p_1, p_2, \ldots, p_m);$$
$$S^*_B(q_1, q_2, \ldots, q_n).$$

Знайдемо спочатку оптимальну стратегію S^*_A. Ця стратегія повинна забезпечувати гравцю ***A*** виграш не менше v при будь-якому поводженні гравця ***B*** и виграш, що дорівнює v, при оптимальній поведінці гравця ***B***.

Припустимо, що ціна гри $v > 0$. Дійсно, для виконання умови $v > 0$ достатньо, щоб всі елементи платіжної матриці a_{ij} були ненегативними. Цього можна домогтися у любій грі, додаючи до всіх елементів a_{ij} одну і ту саму досить велику величину M. При цьому ціна гри збільшиться на M, а саме рішення (вибір стратегій) не зміниться.

Далі припустимо, що гравець ***A*** застосовує свою оптимальну змішану стратегію S^*_A (поки нам невідому), а гравець ***B*** - тільки свою будь-яку чисту стратегію B_j. Тоді середній виграш гравця ***A*** складе:

$$a_j = p_1 a_{1j} + p_2 a_{2j} + \ldots + p_m a_{mj}\,.$$

Але кожне із чисел a_j не може бути менше v. Звідси одержуємо у загальному випадку n нерівностей:

$$\begin{cases} p_1 a_{11} + p_2 a_{21} + \ldots + p_m a_{m1} \geq v, \\ p_1 a_{12} + p_2 a_{22} + \ldots + p_m a_{m2} \geq v, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ p_1 a_{1n} + p_2 a_{2n} + \ldots + p_m a_{mn} \geq v. \end{cases}$$

Введемо нові позначення:

$$x_1 = \frac{p_1}{v},\ x_2 = \frac{p_2}{v}, \ldots,\ x_m = \frac{p_m}{v}.$$

Тоді після ділення всіх нерівностей на **v** одержимо:

$$\begin{cases} x_1 a_{11} + x_2 a_{12} + \ldots + x_m a_{1m} \geq 1, \\ x_1 a_{21} + x_2 a_{22} + \ldots + x_m a_{2m} \geq 1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_1 a_{n1} + x_2 a_{n2} + \ldots + x_m a_{nm} \geq 1. \end{cases}$$

де за умовою всі x_i - ненегативні змінні, що задовольняють умові:

$$x_1 + x_2 + \ldots + x_m = 1/v.$$

Гравець **А** прагне зробити свій гарантований виграш v максимальним. Це значить, що величина $1/v$ повинна прийняти мінімальне значення. Таким чином, завдання рішення гри звелося до наступного завдання лінійного програмування: визначити ненегативні значення змінних:

$$L(\overline{x}) = x_1 + x_2 + \ldots + x_m = \min,$$

за умови, що виконуються зазначені обмеження.

Міркуючи аналогічно, тобто послідовно фіксуючи чисті стратегії A_i гравця **А**, оптимальну стратегію гравця **В** сформулюємо, вирішуючи наступне завдання лінійного програмування: визначити такі ненегативні значення змінних $y_1, y_2, \ldots, y_n$, щоб звернулася у максимум лінійна функція:

$$L(\overline{y}) = y_1 + y_2 + \ldots + y_n = \max,$$

за умови, що виконуються обмеження:

$$\begin{cases} y_1 a_{11} + y_2 a_{12} + \ldots + y_n a_{1n} \le 1, \\ y_1 a_{21} + y_2 a_{22} + \ldots + y_n a_{2n} \le 1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_1 a_{m1} + y_2 a_{m2} + \ldots + y_n a_{mn} \le 1. \end{cases}$$

Тут введені наступні позначення:

$$y_1 = \frac{q_1}{v},\ y_2 = \frac{q_2}{v}, \ldots,\ y_n = \frac{q_n}{v}.$$

Замість вимоги мінімізації функції:

$$L(\overline{x}) = \frac{1}{v},$$

тепер використовується вимога максимізації функції:

$$L(\overline{y}) = \frac{1}{v},$$

тому що гравець **В** у відмінність від гравця **А** прагне мінімізувати v (програти як найменше).

Отже, рішення гри $m \times n$ звелося до рішення двоякої пари завдань лінійного програмування. Відповідно до теорем теорії ігор, це рішення завжди існує.

Приклад 6.5. Нехай підприємство випускає три види продукції $A_i,\ i=\overline{1,3}$ одержуючи при цьому прибуток, що залежить від попиту, який може перебувати у станах $B_j,\ j=\overline{1,3}$. Елементи платіжної матриці (табл. 6.7) характеризують прибуток, що одержить підприємство при випуску ***i*** - ої продукції у ***j*** - ому стані попиту.

Таблиця 6.7.

A_i / B_j	B_1	B_2	B_3	β_j
A_1	3	6	8	8
A_2	9	4	2	9
A_3	7	5	4	7
α_i	3	4	2	$\beta=7$ $\alpha=4$

Визначити оптимальні пропорції продукції, що випускається, яка гарантує середню величину прибутку у будь-якому стані попиту, вважаючи його невизначеним.

Рішення. Із таблиці 6.7 одержуємо матрицю коефіцієнтів a_{ij} розмірності ***3х3***.

$$P=\begin{pmatrix} 3\ 6\ 8 \\ 9\ 4\ 2 \\ 7\ 5\ 4 \end{pmatrix}.$$

Знайдемо нижню й верхню ціну гри:

$$\alpha = \max_i \min_j \alpha_{ij} = 4;$$

$$\beta = \min_j \max_i \alpha_{ij} = 7.$$

Тому що $\alpha \neq \beta$, то сідлова точка відсутня, тобто рішення у чистих стратегіях не існує. Будемо шукати рішення завдання у змішаних стратегіях:

$$S^*_A(p_1,\ p_2,\ p_3);$$
$$S^*_B(q_1,\ q_2,\ q_3).$$

де $p_1,\ p_2,\ p_3$ - імовірності того, що підприємство випустить продукцію виду $A_1,\ A_2,\ A_3$;

$q_1,\ q_2,\ q_3$ - імовірності того, що попит на продукцію буде перебувати у станах $B_1,\ B_2,\ B_3$.

Введемо змінні:

$$x_i = \frac{p_i}{v}, \; i = \overline{1,3}, \; y_j = \frac{q_j}{v}, \; j = \overline{1,3}.$$

Математична модель цього завдання має такий вигляд:

Завдання 1

$$L(\overline{x}) = x_1 + x_2 + x_3 = min,$$

$$\begin{cases} 3x_1 + 9x_2 + 7x_3 \geq 1, \\ 6x_1 + 4x_2 + 5x_3 \geq 1, \\ 8x_1 + 2x_2 + 4x_3 \geq 1. \end{cases}$$

$$x_i \geq 0, \; i = \overline{1,3}.$$

Завдання 2

$$L(\overline{y}) = y_1 + y_2 + y_3 = max,$$

$$\begin{cases} 3y_1 + 6y_2 + 8y_n \leq 1, \\ 9y_1 + 4y_2 + 2y_n \leq 1, \\ 7y_1 + 5y_2 + 4y_n \leq 1. \end{cases}$$

$$y_j \geq 0, \; j = \overline{1,3}.$$

Ці два взаємоподвійні завдання лінійного програмування можна вирішити, використовуючи симплексний метод.

Оптимальне базисне рішення завдання 1 симплекс методом наступне: ***(0,111**; **0**; **0,074)***, причому

$$L(\overline{x}) = L(\overline{y}) = 0.1852.$$

Знайдемо ціну гри

$$v = \frac{1}{minL(\overline{x})} = \frac{1}{maxL(\overline{y})} = \frac{1}{0,1852} = 5,4.$$

Оптимальну стратегію $S^*_A(p_1, p_2, p_3)$ знаходимо зі співвідношення:

$$p_i^* = x_i v, \quad i = \overline{1,3}.$$

Після проведення розрахунків отримаємо:

$$p_1^* = 5,4 \cdot 0,111 = 0,4; \quad p_2^* = 5,4 \cdot 0 = 0; \quad p_3^* = 5,4 \cdot 0,074 = 0,6.$$

Одержуємо $S^*_A(0,4;\ 0;\ 0,6)$. Отже, підприємство повинне випускати ***40%*** продукції A_1 і ***60%*** продукції A_3, а продукцію A_2 не випускати.

Оптимальна стратегія попиту визначається аналогічно:

$$S^*_B(0,2;\ 0,8;\ 0),$$

тобто оптимальний попит у ***20%*** перебуває в стані B_1 і у ***80%*** - у стані B_2.

6.5. Контрольні питання

1. Дайте визначення поняттю конфлікт.
2. Дайте визначення поняттю конфліктна ситуація.
3. Назвіть причини, що породжують конфліктні ситуації.
4. Назвіть джерела невизначеності, які характерні для АСУ реального часу.
5. Дайте визначення поняттю стратегія гравця.
6. Дайте визначення поняттю особистий хід гравця.
7. Дайте визначення поняттю кінцева гра з нульовою сумою.
8. Дайте визначення поняттю змішані стратегії.
9. Перерахуйте ознаки, що характеризують, гру як математичну модель конфліктної ситуації.
10. Дайте визначення поняттю випадковий хід гравця.
11. Наведіть класифікацію основних типів ігор.
12. Дайте визначення поняттю максимінний виграш.
13. Дайте визначення поняттю мінімаксний виграш.
14. Дайте визначення поняттю активні стратегії.
15. Сформулюйте основну теорему матричних ігор.
16. Сформулюйте теорему про активні стратегії.
17. Дайте визначення поняттю нескінченна гра.
18. Дайте визначення поняттю оптимальні стратегії.
19. Знайти нижню й верхню ціну гри для варіантів: а), б) і с):

а) $$P = \begin{pmatrix} 0,3 & 0,6 & 0,8 \\ 0,9 & 0,4 & 0,2 \\ 0,7 & 0,5 & 0,4 \end{pmatrix}.$$ Відповідь: $\alpha=0,4$, $\beta=0,6$.

б) $$P = \begin{pmatrix} 4 & 5 & 3 \\ 6 & 4 & 7 \\ 5 & 2 & 3 \end{pmatrix}.$$ Відповідь: $\alpha=4$, $\beta=5$.

в) $$P = \begin{pmatrix} 4 & 5 & 3 \\ 6 & 7 & 4 \\ 5 & 2 & 3 \end{pmatrix}.$$ Відповідь: $\alpha=4$, $\beta=4$.

20. Завод випускає продукцію, що швидко псується і може її:

A_1 - відразу відправити споживачу;

A_2 - відправити на склад для зберігання;

A_3 - піддати додатковій обробці для тривалого зберігання.

Споживач, у свою чергу, може придбати цю продукцію:

$\boldsymbol{B_1}$ - негайно;

$\boldsymbol{B_2}$ - протягом невеликого періоду;

$\boldsymbol{B_3}$ - після тривалого періоду часу.

У випадку застосування стратегій $\boldsymbol{A_2}$ і $\boldsymbol{A_3}$ завод несе додаткові витрати на зберігання й обробку продукції, які не потрібні д ля при застосуванні стратегії $\boldsymbol{A_1}$. Однак при стратегії $\boldsymbol{A_2}$ варто врахувати можливі збитки через псування продукції, якщо споживач вибере стратегії $\boldsymbol{B_2}$ або $\boldsymbol{B_3}$.

Визначити оптимальні пропорції продукції для застосування стратегій $\boldsymbol{A_1}$, $\boldsymbol{A_2}$, $\boldsymbol{A_3}$, керуючись мінімальним критерієм при матриці витрат, що надана у табл. 6.8:

Таблиця 6.8.

B_j / A_i	B_1	B_2	B_3
A_1	2	5	8
A_2	7	6	10
A_3	12	10	8

Відповідь:

$$S_A^* = \left(0; \frac{1}{3}; \frac{2}{3}\right), \quad v = 8\frac{2}{3}.$$

21. Торговельна фірма розробила кілька варіантів плану продажу товарів на майбутньому ярмарку з урахуванням мінливої кон'юнктури ринку й попиту покупців. Показники доходу надані у табл. 6.9. Визначити оптимальний план продажу.

Таблиця 6.9:

K_j / $П_i$	K_1	K_2	K_3	K_4
$П_1$	8	4	9	2
$П_2$	2	8	3	4
$П_3$	1	2	3	8

Відповідь:

$$S_A^* = (0,44; 0,24; 0,31), \quad S_B^* = (0,31; 0,24; 0,44), \quad v = 4,36 \text{ у.е.}$$

Розділ 7

Використання елементів теорії розкладів для вирішення задачі оптимізації

Розглянуто загальну постановку і методи рішення задачі оптимізації із застосуванням теорії розкладів, комбінаторні задачі на складання розкладу, задачу комівояжера та приклади задач, що зводяться до задачі комівояжера. Наведено контрольні питання.

7.1. Постановка задачі оптимізації із застосуванням теорії розкладів

Теорія розкладів, як і деякі розділи прикладної математики широко використовується для вирішення задач оптимізації АСУ. Теорія розкладів оперує з двома основними поняттями: процес, що протікає (операція, задача, робота) та ресурси, що використовуються при цьому (машини, люди, обладнання). Наприклад, задача досягнення мінімальних витрат часу на виконання сукупності операцій за рахунок оптимального їх закріплення за ресурсами системи. Самі ж операції можуть бути залежними, або не залежними.

Теорія розкладів надана сукупністю типових задач і методів їх рішення. Важливе практичне значення має ціла група задач, що пов'язані із виконанням комплексу робіт, що мають назву «операції», «замовлення» і т. п., за наявності обмежень на порядок їх виконання, тривалість кожної роботи, наявних ресурсів і т.п. З такими задачами в АСУ доводиться стикатися як при рішенні питань календарного планування, та і у оперативному управлінні.

Задачі подібного роду досить різноманітні і їх дослідження залежно від характерних ознак проводиться у рамках різних наукових дисциплін. У теорії мережевого планування основна увага надається визначенню порядку виконання складних розробок, що включають велике число взаємозв'язаних робіт, які вимагають багато виконавців і значних матеріальних витрат. У теорії упорядкування розглядаються питання встановлення порядку виконання великого комплексу однорідних робіт за допомогою одного або декількох спеціалізованих пристроїв, машин, приладів. У теорії масового обслуговування розглядаються задачі призначення пріоритетів в обслуговуванні замовлень, що поступають, деякими пристроями.

У загальному випадку цілеспрямовану діяльність можна розглядати як деякий процес, що протікає у часі та полягає у реалізації (виконанні) певної сукупності робіт

$$S = \{P_1, \ldots, P_s\},$$

яка має назву - ***програма***.

Виконання програми має цілий ряд обмежень і умов, які можна розбити на дві групи.

Обмеження, що відносяться до першої групи описують взаємну залежність робіт, що виконуються. Кожній роботі $P_i \in S$ можна поставити у відповідність деяку безліч $\Pi_i \subset S$ безпосередньо попередніх робіт, без виконання яких не можна починати виконання роботи P_i, а також ряд робіт, що не входять у множину S, які можуть розглядатися як зовнішні чинники, необхідні для виконання роботи P_i.

Для опису деяких інших умов, що входять у обмеження, які відносяться до першої групи зручно хід виконання робіт розглядати у дискретному часі, прийнявши деяку величину Δt за одиничний інтервал часу. Цей інтервал можна назвати «умовний день», «умовний рік» або ***кроком***.

Деякі роботи можуть бути виконані цілком за один крок. Інші можуть зажадати декілька кроків. У цьому випадку у обмеження, що відносяться до першої групи повинна входити умова допустимості або недопустимості переривання робіт.

Припустимо, що $\bar{t}$ - номер даного кроку ($\bar{t}$ =1, 2...). Позначимо через $y_i(\bar{t})$ частку роботи P_i, що виконується на кроці з номером $\bar{t}$, яку назвемо ***одиничним об'ємом роботи*** (денний, річний і т. п.) P_i. Через $S(\bar{t})$ позначимо перелік робіт, що виконується на кроці з номером $\bar{t}$

$$S(\bar{t}) = \sum_{i=1}^{n} y_i(\bar{t}).$$

Припустимо, що робота P_i складається з m одиничних об'ємів y_i, і переривання цієї роботи недопустимо. Умова недопустимості переривання роботи полягає у тому, що якщо ця робота починається на кроці з номером $\bar{t}$, то одиничні об'єми $\mathbf{y_i}$ повинні входити до переліку робіт $S(\bar{t}), S(\bar{t}+1),...,S(\bar{t}+m-1)$.

Обмеження, що відносяться до другої групи пов'язані із обмеженим об'ємом ресурсу, який може бути виділений на виконання програми S.

Позначимо через $F(\bar{t})$ вектор ресурсу, який може бути виділений на виконання робіт на кроці з номером $\bar{t}$. Компонентами вектора $F(\bar{t})$ можуть бути величини: F_1 - сировина, матеріали; F_2 - устаткування; F_3 - робоча сила і т.п. Позначимо також через $f[y_i(\bar{t})]$ кількість ресурсу, яку необхідна витратити для виконання роботи $y_i(\bar{t})$. У цих позначеннях обмеження, що відносяться до другої групи запишуться у вигляді

$$\sum_{i=1}^{s} q\left[y_i\left(\bar{t}\right)\right] \le M\left(\bar{t}\right), \bar{t} = 1,2,...,M,$$

де M - загальне число кроків.

У даній задачі часто вводиться припущення, що одиничний об'єм роботи $y_i(\bar{t})$ або виконується повністю, якщо для неї виділений весь необхідний ресурс, або не виконується зовсім, якщо ресурс виділений не повністю. Загальніша, проте, і складніша задача має місце тоді, коли частка виконаної роботи залежить від кількості виділеного ресурсу.

Задача на складання розкладу зводиться до того, щоб для кожного кроку $\bar{t}$ назвати перелік робіт

$$S(\bar{t}) = \left\{P_{\bar{t}_1},...,P_{\bar{t}_k}\right\},$$

які повинні виконуватися протягом цього кроку, а також частку $y_i(\bar{t})$ кожної з цих робіт з урахуванням всіх обмежень, що накладаються на порядок виконання робіт і на ресурси, що виділяються.

Якщо кількість робіт велика, то кількість варіантів розкладу може бути величезною. Вибір конкретного варіанту розкладу пов'язаний з поняттям директивного терміну. Якщо перелік робіт і ресурс на їх виконання вже задані, то якнайкращим (оптимальним) варіантом розкладу слід вважати такий варіант, при якому термін виконання всього переліку буде мінімальним. Цей термін і прийнято називати ***директивним терміном***.

Припустимо, що є складна програма робіт $S=\{P_1, ..., P_s\}$, у якій діють як обмеження першого типу, що визначають послідовність виконання окремих робіт, так і обмеження на ресурси, що виділяються. На практиці звичайно буває так, що обмеження зв'язують між собою не всі роботи, а діють лише усередині деяких незалежних груп робіт. Самі ж такі групи між собою жорсткими обмеженнями не зв'язані.

У таких випадках програму робіт S можна представити у вигляді деякого розбиття $M = \{R_1, ..., R_m\}$, елементи якого $R \subset M$ є деякими самостійними підпрограмами, які володіють тією властивістю, що між ними відсутні жорсткі обмеження послідовності, так що кожна підпрограма $R \in M$ може виконуватися майже незалежно від виконання інших підпрограм. Проте усередині кожної підпрограми

$$R_i = \left\{P_{i_1}, ..., P_{i_r}\right\},$$

порядок виконання робіт підлягає жорстким обмеженням, що відносяться до першої групи.

Таким чином, задачу на складання розкладу можна розбити на дві самостійні задачі.

Задачі першого типу передбачають визначення порядку виконання робіт усередині кожної підпрограми з урахуванням обмежень на черговість виконання робіт і розподіл ресурсу по етапах виконання підпрограми. Оскільки жорсткі обмеження першого типу у таких задачах в значній мірі зумовлюють порядок виконання окремих робіт то кількість можливих варіантів розкладу тут не дуже велика. Проте варіанти завжди можливі і серед них можуть знайтися такі, при яких виявиться мінімальним час виконання підпрограми, найраціональніший розподіл ресурсів і т.п. Для вирішення задач цього типу широке застосування одержали методи мережевого планування і управління (МПУ).

У задачах другого типу обмеження на послідовність виконання окремих робіт незначні. Проте це не означає, що порядок виконання робіт тут може бути довільним. Від порядку виконання робіт залежить розподіл ресурсу по кроках. На перехід від однієї роботи до іншої витрачається час, можуть бути потрібними перекваліфікація персоналу, переналагодження або заміна устаткування і т.п. Тому і тут стоїть задача складання такого розкладу робіт, який у певному значенні був би якнайкращим.

Проте відсутність обмежень у задачах цього типу приводить до необхідності розглядати дуже велике число різних комбінацій порядку виконання робіт і давати оцінку кожній комбінації. Тому такі задачі одержали назву ***комбінаторних задач*** на складання розкладу або ***задач впорядкування***.

7.2. Комбінаторні задачі на складання розкладу

Сутність комбінаторних задач проілюструємо на прикладі визначення оптимальної послідовності обробки деталей.

Припустимо, що є два верстати, наприклад токарний і фрезерний, і n різних деталей, кожна з яких повинна бути послідовно оброблена на кожному з цих верстатів. Треба визначити, в якій послідовності потрібно обробляти ці деталі.

Дана задача має характерні ознаки задач на складання розкладу з обмеженнями типу α на послідовність операцій (спочатку токарна обробка, а потім фрезерування) і з обмеженнями типу β на ресурси, що використовуються (є тільки один токарний і один фрезерний верстати). Проте дані обмеження не є надзвичайно жорсткими і допускають велику свободу вибору варіантів обробки деталей. Почати обробку можна з будь-якою з n деталей, другою можна обробляти будь-яку з $n - 1$ деталей, що залишилися, і т.д. Загальне число варіантів $n(n-1)... 1 = n!$. Не всі ці варіанти рівноцінні. Від вибору порядку обробки деталей залежать простої устаткування (фрезерний верстат вже звільнився, а токарна обробка чергової деталі не закінчена) і перерви у обробці деталей (закінчена токарна обробка чергової деталі, але фрезерний верстат ще зайнятий).

Очевидно, зі всіх можливих варіантів обробки слід вибрати такий, який дозволить закінчити обробку всієї партії деталей за найкоротший час. Рішення даної задачі оптимізації надзвичайно утруднено наявністю величезного числа варіантів обробки, що і примушує віднести її до задач комбінаторного типу.

Розглянемо простіший варіант задачі про визначення оптимальної послідовності обробки деталей. Припустимо, що n різних деталей, занумерованих від 1 до n, треба обробити на одному верстаті. При переході від обробки деталі i до деталі k потрібен час d_{ik} на переналагодження верстата, різний при різних i та k. Значення d_{ik} задані у вигляді табл. 7.1.

Таблиця 7.1

i	k			
	1	2	…	n
1	x	d_{12}	…	d_{1n}
2	d_{21}	x	…	d_{2n}
…	…	…	…	…
n	d_{n1}	d_{n2}	…	x

Потрібно визначити такий порядок обробки деталей, при якому загальний час на обробку буде мінімальним. Оскільки можливо ***n*** варіантів обробки першої деталі, ***n -1*** - другої і т.д., то загальне число варіантів обробки буде рівне ***n!***. Отже, ця задача відноситься до задач комбінаторного типу.

Оскільки час, що витрачається безпосередньо на обробку всіх деталей, буде одним і тим же при будь-якому порядку обробки, то мінімізація загального часу обробки зводиться до мінімізації часу на переналагодження верстатів. Дані для вирішення цієї задачі цілком містяться у табл. 7.1.

У математиці задача подібного вигляду одержала назву ***задачі комівояжера***. Комівояжер (агент, що рекламує товар своєї фірми), що виходить з якого-небудь міста, повинен відвідати ***n -1*** інших міст і повернутися до початкового пункту. Відомі відстані між містами - d_{ik}, задані у вигляді вже розглянутої табл. 7.1. Треба встановити, у якому порядку комівояжер повинен відвідувати міста, щоб загальна пройдена відстань була мінімальною.

Якщо дещо змінити формулювання задачі і не вимагати повернення до початкового пункту, то одержимо незамкнуту задачу комівояжера. Якщо ж не заданий і початковий пункт, то одержуємо повну аналогію із задачею про вибір оптимальної послідовності обробки деталей.

Дуже наочно задачу комівояжера можна представити у вигляді графа ***G=(X, U)***, вершини якого відповідають містам, а дуги - дорогам, що сполучають міста між собою. Кожній дузі, що сполучає міста ***i, k,*** приписується число d_{ik}, яке називається довжиною дуги і дорівнюється відстані між містами ***i*** та ***k***. Якщо з кожного міста є безпосередній шлях у будь-яке інше місто без заходу в проміжні міста, то кожні дві вершини графа будуть сполучені дугами в обох напрямах. Граф, що виходить при цьому, виявляється ***повним***.

У практичних задачах повні графи зустрічаються рідко. Якщо граф не повний, то його можна вважати повним, відновивши дуги, яких бракує і приписавши їм довжину $d = \infty$. У таблиці відстаней відповідні клітки зручно залишати незаповненими.

Елементарний шлях у графі, що проходить через всі вершини, називається ***гамільтоновим шляхом***. Замкнутий елементарний шлях, тобто елементарний контур, що проходить через всі вершини графа, називається ***гамільтоновим контуром***. Ці назви пов'язані з задачею Гамільтона, яка буде розглянута пізніше. Як бачимо, задача комівояжера зводиться до знаходження найкоротшого гамільтонова контуру або найкоротшого гамільтонова шляху.

Надалі основним видом задачі комівояжера вважатимемо замкнутий варіант, коли задана вершина, в якій починається і закінчується маршрут комівояжера. Проте незамкнуті варіанти легко приводяться до замкнутих.

Припустимо, що є граф $\boldsymbol{G=(X, U)}$ і початкова точка $x_0 \in X$, але не вказана кінцева точка. Замінимо довжини всіх дуг, що закінчуються у x_0, нулями. Тоді будь-який гамільтонів контур містить дугу нульової довжини, що закінчується у x_0. Якщо знайти найкоротший контур і виключити з нього дугу нульової довжини, що закінчується у x_0, то одержимо найкоротший гамільтонів шлях.

Якщо не вказані ні початкова, ні кінцева вершини графа, то вводимо фіктивну вершину z, яка з'єднується дугами нульової довжини зі всіма вершинами графа у обох напрямах. Цю вершину приймають за початкову і знаходять найкоротший гамільтонів контур. Виключивши з нього дугу нульової довжини, що виходить із z, і дугу нульової довжини, що заходить у z, одержимо найкоротший гамільтонів шлях.

Розглянемо приклади задач, що зводяться до задачі комівояжера.

Задача Гамільтона. На рис. 7.1 зображений додекаедр, тобто правильний багатокутник з 12 п'ятикутними гранями і 20 вершинами.

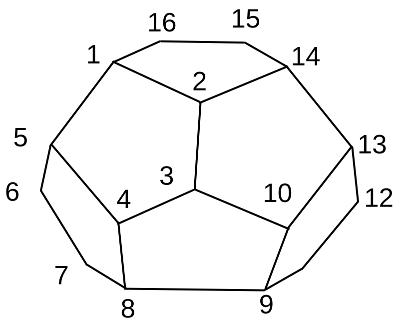


Рис. 7.1. Додекаедр

Потрібно, рухаючись по ребрах додекаедра, обійти усі вершини, заходивши у кожну з них тільки один раз, і повернутися у початкову вершину. Кожний такий шлях, якщо він існує, називається гамільтоновим циклом або гамільтоновим контуром.

Для того, щоб звести цю задачу до задачі комівояжера, вважатимемо, що відстань d_{ij} між двома вершинами дорівнює одиниці, якщо вершини сполучені ребром, і дорівнює нескінченності у протилежному випадку. За цими даними можна скласти таблицю взаємних відстаней між вершинами по типу табл. 7.1 і далі розв'язувати задачу комівояжера, яка склалася.

Транспортні задачі. При проектуванні різноманітних технологічних процесів виникає величезне число задач, пов'язаних із об'їздом ряду пунктів і розвезення сировини, заготівок, напівфабрикатів деталей, вузлів і т.п. Оскільки транспортна мережа у багатьох випадках є схемою організації доставки чого-небудь (інформації, вантажів), то рішення транспортної задачі дозволяє визначити найбільш раціональний план перевезень, тобто такий розподіл маршрутів, який забезпечує: по-перше, мінімальні витрати, а по-друге мінімальний час на доставку чого-небудь до споживача.

Оптимізація процесу програмування для ЕОМ. При виконанні складних обчислень на ЕОМ доводиться виконувати різні операції, які здійснюються підпрограмами $B_1, \ldots, B_n$. При попередньому програмуванні встановлюється певний порядок виконання операцій. Проте при розв’язанні конкретних задач порядок виконання операцій іноді доводиться змінювати.

У подібних випадках програмування зводиться до введення команд умовного переходу, що порушують природний хід виконання програми, так що за операцією B_i може слідувати не операція B_{i+1}, а деяка інша операція, наприклад, операція B_k (рис. 7.2).

$$B_1 \ldots B_i \;\; B_{i+1} \ldots B_i \ldots B_k \ldots \mathrm{B_n}$$

Рис. 7.2. Введення команд умовного переходу

Кожна команда умовного переходу вимагає витрати додаткового часу і ускладнює як програмування, так і роботу ЕОМ. Тому доцільно так скласти попередню програму, щоб при розв'язуванні заданого набору задач використовувати мінімальну кількість команд умовного переходу.

Припустимо, що є набір з N задач, призначених для вирішення на ЕОМ. Позначимо через p_k імовірність (частоту) рішення k-ї задачі. Введемо також числа:

$$a_{ij}^{k} = \begin{cases} 1, & \text{якщо у } k\text{-й задачі за операцією } B_i \text{ слідує операція } B_j; \\ 0, & \text{інакше.} \end{cases}$$

При цьому величина

$$d_{ij} = \sum_{k=1}^{N} p_k a_{ij}^{(k)} \quad (i, j = \overline{1,n}),$$

чисельно виражатиме математичне очікування того, що у розглянутих умовах за операцією B_i - слідуватиме операція B_j. Величину d_{ij} можна умовно прийняти за відстань між операціями B_i і B_j, а самі операції розглядати як вершини графа. Тоді оптимізація програми, що мінімізує середнє число команд умовного переходу, зводиться до знаходження найкоротшого гамільтонова шляху у одержаному графі. Відповідно до послідовності операцій на цьому шляху і слід здійснювати попереднє програмування.

Визначення оптимальної послідовності обробки деталей на двох верстатах. Повернемося до розглянутої вище задачі про обробку n деталей на двох верстатах. Позначимо через a_k і b_k час обробки k-ї деталі на першому і другому верстатах відповідно. Треба знайти таку послідовність обробки деталей, при якій загальний час обробки всіх деталей буде якнайменшим.

Наголосимо, перш за все, на особливостях роботи кожного з верстатів. Ясно, що перший верстат може почати обробляти наступну деталь зразу ж, як тільки закінчить обробку попередньої і простоїв у нього не буде. Другий верстат може почати обробку чергової деталі тільки в тому випадку, якщо:

- закінчилася обробка цієї деталі на першому верстаті;
- закінчилася обробка попередньої деталі на другому верстаті.

Розглянемо спочатку простий випадок обробки всього двох деталей.

Якщо спочатку обробляється перша деталь, то момент закінчення її обробки на першому верстаті буде a_i. Моменти a_1+b_1 та a_1+a_2 відповідають закінченню обробки першої деталі на другому верстаті і другої деталі на першому верстаті. Другу деталь на другому верстаті можна почати обробляти, коли обидві ці операції закінчено, тобто в момент

$$\max(a_1 + b_1, a_1 + a_2) = a_1 + \max(a_2, b_1).$$

Обробка обох деталей закінчиться у момент

$$t' = a_1 + \max(a_2, b_1) + b_2.$$

Аналогічно, якщо спочатку обробляється друга деталь, то момент закінчення обробки обох деталей буде дорівнювати

$$t'' = a_2 + \max(a_1, b_2) + b_1.$$

Обробку вигідніше починати з першої деталі, якщо $t' \le t''$, тобто якщо

$$a_1 + \max(a_2, b_1) + b_2 \le a_2 + \max(a_1, b_2) + b_1. \quad (7.1)$$

Легко побачити справедливість відносин

$$a_1 + b_2 = \max(a_1, b_2) + \min(a_1, b_2);$$

$$a_2 + b_1 = \max(a_2, b_1) + \min(a_2, b_1),$$

з урахуванням яких (7.1) запишеться у вигляді

$$\min(a_1, b_2) \le \min(a_2, b_1).$$

Це і є умова, коли обробку вигідніше починати з першої деталі.

Розгляд випадку обробки n деталей складніший, але можна показати, що з двох деталей i та j раніше потрібно починати обробку деталі i, якщо виконується умова

$$\min(a_i, b_j) \le \min(a_j, b_i). \quad (7.2)$$

Якщо тепер поставити у відповідність кожній деталі вершину графа і вважати, що відстань d_{ij} між вершинами i та j рівна одиниці, при виконанні умови (7.2), і дорівнює нескінченності у іншому випадку, то приходимо до матриці відстаней, що визначає відповідну задачу комівояжера. Найкоротший гамільтонів шлях, що визначається цією матрицею, і дає оптимальну послідовність обробки деталей.

Задача комівояжера розв'язується методами Монте - Карло, динамічного та лінійного програмування і методом меж і гілок.

7.3. Методи рішення задачі оптимізації

7.3.1. Застосування методу Монте-Карло

Методом Монте-Карло називають будь-яку статистичну процедуру, що включає прийоми статистичної вибірки. Розглянемо їх застосування для вирішення задачі комівояжера.

Вершину 1 приймають за початкову і закладають в урну жетони із номерами від 2 до n. Ретельно перемішавши жетони, витягують їх поодинці і записують номери, наприклад, $i_2, ..., i_n$. При цьому отримують гамільтонів контур $1, i_2, ..., i_n, 1.$ Підраховують довжину цього контура і запам'ятовують її. Після цього процедуру повторюють, і якщо новий маршрут виявиться гіршим, його негайно забувають, а якщо він виявляється кращим, то забувають попередній, а новий запам'ятовують. Проводячи таку процедуру багато разів, можна з високим ступенем вірогідності розраховувати на те, що вдасться знайти якщо і не якнайкращий, то достатньо хороший маршрут.

Для проведення таких випробувань зазвичай використовують ЕОМ, забезпечені датчиками випадкових чисел, що дозволяє за короткий час проглянути значне число маршрутів.

7.3.2. Зведення задачі оптимізації до рішення завдання цілочисельного лінійного програмування

У графі $G=(X, U)$, що відповідає завданню комівояжера, розглянемо деякий гамільтонів контур $1, i_2, ..., i_n, 1$. Сукупність дуг, що входять у гамільтонів контур, можна описати у вигляді матриці:

$$Y = \left[y_{ij} \right] = \begin{bmatrix} y_{11} \cdots y_{1n} \\ \cdots\cdots\cdots \\ y_{n1} \cdots y_{nn} \end{bmatrix},$$

де

$$y_{ij} = \begin{cases} 1, \text{ якщо дуга } (i,j) \text{ входить у контур, що розглядається;} \\ 0 - \text{ в протилежному випадку.} \end{cases}$$

Оскільки кожна вершина зустрічається у гамільтоновому контурі один і лише один раз, то у кожну вершину обов'язково входить і з кожної вершини обов'язково виходить одна і лише одна дуга. Це означає, що у матриці в кожному рядку і в кожному стовпці повинна стояти одна і лише одна одиниця, що математично запишеться у вигляді

$$\sum_{i=1}^{n} y_{ij} = 1, \qquad \sum_{j=1}^{n} y_{ij} = 1. \tag{7.3}$$

З обліком (7.3) умову, щоб всі y_{ij} були рівні 0 або 1, можна замінити вимогою, щоб всі y_{ij} були ненегативними цілими числами:

$$y_{ij} \geq 0, \quad (i, j = \overline{1,n}), \tag{7.4}$$

$$E(y_{ij}) = y_{ij}, \tag{7.5}$$

де - $E(y_{ij})$ означає операцію взяття цілої частини числа y_{ij}.

Як бачимо, будь-який гамільтонів контур може бути описаний матрицею, елементи якої y_{ij}, задовольняють співвідношенням (7.3) - (7.5). Довжина гамільтонова контуру, дорівнює сумі довжин дуг, що його складають знайдеться так:

$$L = \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} y_{ij}. \tag{7.6}$$

Знаходження гамільтонова контура мінімальної довжини зводиться, таким чином, до знаходження значень змінних y_{ij}, що задовольняють лінійним рівнянням і нерівностям (7.3) і (7.4) та умові цілочисельності (7.5) і що обертають у мінімум лінійну форму (7.6), тобто до рішення задачі цілочисельного лінійного програмування.

Слідує, проте, врахувати, що умови (7.3) - (7.5) визначають, що у кожну вершину входить і з кожної вершини виходить тільки одна дуга, але не обов'язково визначають наявність гамільтонова контуру. Цим умовам відповідатиме і контур, що проходить не через всі вершини, якщо у вершинах, що залишилися, утворюються петлі. Це відповідає зверненню у одиницю деяких елементів матриці, що стоять на головній діагоналі, і може бути виключено накладенням на y_{ij}, додаткових обмежень виду

$$y_{ij} = 0 \text{ коли } i = j, \quad j = \overline{1,n}.$$

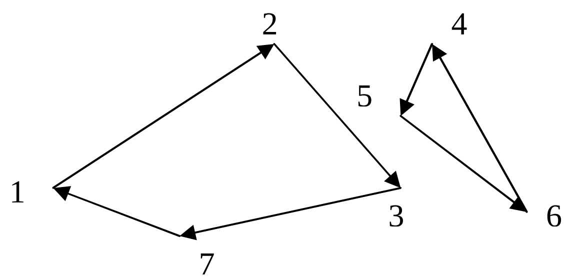


Рис. 7.3. Утворення двох незв'язаних контурів.

Можливі також випадки, коли виходить не один гамільтонів контур, а два або більше замкнутих контурів, кожний з яких охоплює тільки частину вершин, як показано, наприклад, на рис. 7.3 для $n = 7$. Якщо вийшло саме таке рішення, то вводять додаткове обмеження виду

$$y_{12} + y_{23} + y_{37} + y_{71} \leq 3,$$

яке забороняє виникнення контуру ***(1, 2, 3, 7, 1)*** і вирішують задачу наново спільно з цим обмеженням.

7.3.3. Метод гілок і меж

Нехай U — безліч варіантів рішення деякої задачі. Якщо ці варіанти перенумерувати від 1 до n, то під U можна розуміти безліч номерів варіантів рішення, тобто покласти, що $U = \{1, ..., m\}$.

Кожний (*i*-й) варіант рішення задачі будемо характеризувати деяким показником (критерієм) якості q_i. Причому оптимальним буде такий варіант рішення задачі, якому відповідає мінімальне значення показника якості

$$q^* = \min_{i \in U} q_i .$$

Величина q^* називається ***точною нижньою межею*** (найменшим значенням) множини показників якості $Q = \{ q_1, q_2, ..., q_m \}$, що відповідає множині варіантів рішення задачі U.

При знаходженні оптимального рішення задачі методом гілок і меж широко використовується поняття оцінки знизу для всієї множини варіантів рішення задачі U та для її складових частин. ***Оцінкою знизу*** (оптимістичною оцінкою знизу) $\hat{q}$ для множини U називається таке значення показника якості, яке задовольняє співвідношенням

$$\hat{q} < q^* \text{ або } \hat{q} \le q^* .$$

У останньому випадку оцінка називається ***досяжною***.

Якщо розбити множину варіантів рішення задачі U на дві підмножини U_A і U_B з точними нижніми межами (мінімальними значеннями) показника якості q^*_A і q^*_B відповідно, то оптимальне значення показника якості множини U співпадає з мінімальним з q^*_A і q^*_B

$$q^* = \min\{ q^*_A, q^*_B \}.$$

Оскільки множини **A** і **B** мають менше число елементів, чим безліч U, то є можливість отримати оцінки q_A, q_B, які ближчі до значень q^*_A, q^*_B, чим у первинній безлічі U. Це означає, що оцінка $\min(\hat{q}_A, \hat{q}_B)$ буде ближча до q^*, чим оцінка $\hat{q}$. Тобто чим менше множина варіантів рішення задачі, тим простіше і точніше можна отримати оцінку знизу. Більш того, імовірно, що оптимальний варіант рішення задачі $q = q^*$, буде знаходитися у тій підмножині з U_A і U_B, для якої оцінка знизу буде меншою. При цьому імовірність знаходження оптимального варіанта рішення у підмножині варіантів рішення, для якої оцінка знизу менша, буде тим більша, чим більша різниця між оцінками знизу для отриманих у результаті розбиття підмножин.

Якщо виявилось, що $\hat{q}_A > \hat{q}_B$, тобто є підстави підозрювати, що оптимальний варіант входить у множину B і цю множину треба досліджувати детальніше. Для цього розбиваємо його у свою чергу на дві підмножини ***C*** і ***D*** так, щоб оцінки $\hat{q}_D$, $\hat{q}_C$ розрізнялися як можливо більше. Якщо виявилось, що $\hat{q}_D < \hat{q}_C$, то подібному ж розбиттю піддаємо безліч ***D*** і т. д., до тих пір, поки не прийдемо до підмножини, що складається всього з одного елементу із значенням параметра $q = q_o$-Всю процедуру розбиття зручно представити у вигляді дерева розбиття, приведеного на рис. 7.4.

Проте знайдений елемент із значенням $q = q_o$ може лише підозрюватися як оптимальний. Необхідно перевірити, чи немає серед нерозглянутих множин ***A, C*** і т.д.. елементу із значенням $q < q_o$. При цьому множини, у яких $\hat{q} > q_o$, можуть не розглядатися, оскільки елементу з $q < q_o$ у них бути не може.

Розбиття вищезгаданих множин або приведе до елементу із $q < q_o$, тоді цей елемент і повинен бути прийнятий за оптимальний, або приведе до оцінок знизу $\hat{q} > q_o$. У цьому випадку елемент із значенням параметра q_o і буде варіантом рішення задачі, що шукається.

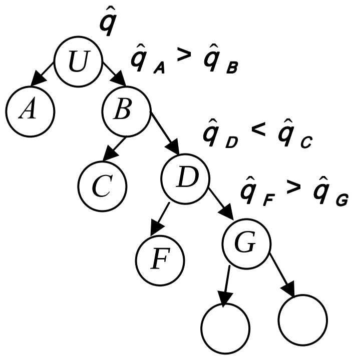


Рис. 7.4. Побудова дерева розбиття

На цьому заснований метод гілок і меж, який зводиться до поступового розбиття множини варіантів рішення задачі на підмножини та знаходження оцінок знизу для кожної з підмножин. Розбиття продовжується, доки не отримують підмножину, що відповідає варіанту рішення задачі, який має мінімальну оцінку знизу з усіх отриманих у ході розбиття підмножин. При цьому процес розбиття відображають у вигляді дерева розбиття, що починається із множини варіантів рішення задачі ***U*** і поділяється на гілки, які відображають отримані на кожному етапі підмножини і їх оцінки знизу.

7.3.4. Динамічне програмування

На практиці, типовою для автоматизованих систем управління є задача знаходження найкоротшого шляху доведення інформації до декількох споживачів за умови обслуговування кожного споживача лише один раз та отримання вихідним джерелом інформації контрольного сигналу.

Такі задачі зводяться до класичного завдання про комівояжера, суть її полягає в наступному. Є ***n*+1** місто $A_1, A_2, ... A_n \ (n \geq 1)$ із заданими між ними відстанями $d_{ij} \ (i, j = 0,1,...,n)$. Потрібно, відправляючись з A_1, вибрати такий маршрут пересування $A_1, A_2, ... A_n, A_1$, при якому комівояжер, побувавши у кожному місті по одному разу, повернувся б в початковий пункт A_1, пройшовши мінімально можливий сумарний шлях.

Ця задача може бути у принципі вирішена методом простого перебору всіх можливих маршрутів. Легко бачити, проте, що загальне число таких маршрутів дорівнює $1 Ч 2 Ч 3 Ч ... Ч (n-1) Ч n$, а ця величина вже при n, вимірюваному декількома десятками, така величезна, що практично виключає можливість прямих обчислень навіть при використанні ЕОМ. Проблема, отже, полягає у тому, щоб різко скоротити перебір, відкидаючи наперед свідомо непридатну безліч варіантів. Методи вирішення цієї проблеми складають суть так званого динамічного програмування.

Основний прийом динамічного програмування — знаходження правил домінування, які дозволяють проводити порівняння варіантів розвитку послідовностей і завчасне відсіювання безперспективних варіантів. У ряді випадків у завданнях динамічного програмування вдається отримати такі сильні правила домінування, що вони визначають елементи оптимальної послідовності однозначно один за іншим. У цьому випадку правила домінування називають зазвичай правилами, що вирішують.

Вирішуючи правила (у разі їх існування) зазвичай виводяться за допомогою ***принципу оптимальності Беллмана***. Суть принципу оптимальності полягає у наступному. Нехай критерій F (що задається формулою або алгоритмом), який дає числову оцінку якості варіанту (послідовності) $A_n = a_{i1}, a_{i2}, ..., a_{in}$, можна застосовувати не тільки до всієї послідовності, але і до будь-якого її початкового відрізку $A_k = a_{i1}, a_{i2}, ..., a_{ik}$. Послідовність, A_n якої відповідає екстремальне значення критерію, F називається оптимальною. Якщо будь-який початковий відрізок оптимальної послідовності також оптимальний (у класі всіх послідовностей, складених з тих же елементів і, можливо ще маючий початок і кінець, що і даний відрізок), то говорять, що для відповідного завдання справедливий принцип оптимальності.

Принцип оптимальності, як неважко зрозуміти, завжди виконується для адитивних критеріїв ***F***, які характеризуються тією властивістю, що для будь-якої послідовності ***A***, складеною з двох підпослідовностей ***B*** і ***C (A=BC)***:

$$F(A) = F(BC) = F(B) + F(C).$$

Число тих траєкторій, що підлягають розгляду може бути ще більш зменшено, якщо до правил домінування, що витікають із принципу оптимальності, пред'являються додаткові обмеження, що витікають з конкретного сенсу завдання (наприклад, відсутність різких зламів траєкторії, її обов'язкове проходження через задані проміжні пункти і т.п.).

Розглянемо основні принципи динамічного програмування на прикладі завдання про комівояжера. Нехай є всього п'ять пунктів: A_1, A_2, A_3, A_4, A_5 (рис. 7.5) із наступною таблицею відстаней (табл. 7.2).

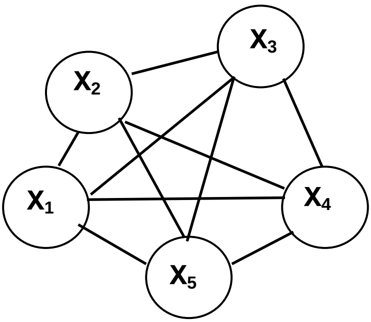


Рис. 7.5. Схема маршрутів

Таблиця 7.2.

i	*j*				
	1	2	3	4	5
1	x	5	4	7	3
2	10	x	2	1	5
3	3	1	x	2	1
4	7	2	4	x	4
5	5	5	1	7	x

У кожний момент часу комівояжер знаходяться в одному з об’єктів A_i ($i = 1, 2, ..., 5$), і приймає рішення у який наступний об’єкт A_j необхідно рухатися, щоб показник якості (сумарний шлях) був мінімальним. Область допустимих напрямків руху на кожному кроці зменшується, оскільки з неї виключаються всі об’єкти, в яких вже був комівояжер.

Почнемо з перебору послідовних комбінацій $A_1A_{i2}A_{i3}A_{i4}$, що поєднують чотири об’єкта (три переходи), і групують їх по останньому об’єкту A_{i4}. Вибір такої довжини початкової послідовної комбінації обумовлений тим, що такі комбінації, з одного боку, охоплюють велику кількість можливих шляхів, а з іншого боку, є достатньо нескладними для їх повного перебору.

Всі можливі послідовні комбінації $A_1 A_{i2} A_{i3} A_{i4}$ записують у таблицю, у другий стовбець якої записують довжину сумарного шляху, що відповідає цим комбінаціям (суму довжин дуг (A_1, A_{i2}), (A_{i2}, A_{i3}), (A_{i3}, A_{i4}). Далі проводять аналіз отриманих послідовних комбінацій. З усіх послідовних комбінацій, що закінчуються в тому самому об’єкті A_{i4} і проходять крізь однакові об’єкти, вибирають послідовну комбінацію, якій відповідає шлях мінімальної довжини, і записують напроти неї у третьому стовпці "так", розуміючи, що саме її доцільно надалі розглядати для побудови оптимального маршруту. Напроти інших послідовних комбінацій, що закінчуються у цьому ж об’єкті A_{i4} і проходять крізь однакові об’єкти і яким не відповідає шлях мінімальної довжини записують "ні" і виключають їх з подальшого розгляду як свідомо неперспективні варіанти рішення.

На наступному кроці у нову таблицю записують лише ті послідовні комбінації, що позначені у попередній таблиці позначкою "так", і додають до них ще по одному об’єкту A_{i5} (одному переходу), такому, якого не було у послідовній комбінації $A_1 A_{i2} A_{i3} A_{i4}$, до якої додається об’єкт. Знов записують довжини шляхів, що задаються отриманими послідовними комбінаціями $A_1 A_{i2} A_{i3} A_{i4} A_{i5}$. З усіх послідовних комбінацій, що закінчуються в тому самому об’єкті A_{i5} і проходять крізь однакові об’єкти, вибирають для подальшого розгляду послідовну комбінацію, якій відповідає шлях мінімальної довжини, і записують напроти неї у третьому стовпці "так". Напроти інших послідовних комбінацій, що закінчуються у цьому ж об’єкті A_{i5} і проходять крізь однакові об’єкти, записують "ні" і виключають їх з подальшого розгляду як свідомо неперспективні варіанти рішення.

Кількість таких кроків визначається кількістю об’єктів, крізь які повинний пройти комівояжер. Коли отримують послідовні комбінації, що проходять крізь всі об’єкти, переходять до останнього етапу, на якому до отриманих послідовних комбінацій додають перехід у об’єкт A_1, з якого починався рух. Розраховують довжини отриманих контурів і як оптимальний шлях вибирають той, що має мінімальну довжину.

Приклад. Знайдемо замкнений шлях мінімальної довжини, що проходить крізь п’ять об’єктів, для таблиці 7.2.

Таблиця 7.3.

Варіант маршруту	Довжина маршруту,км	Перспективно чи ні
$A_1 A_3 A_4 A_2$	8	так
$A_1 A_4 A_3 A_2$	12	ні
$A_1 A_3 A_5 A_2$	10	ні
$A_1 A_5 A_3 A_2$	5	так
$A_1 A_4 A_5 A_2$	16	ні
$A_1 A_5 A_4 A_2$	12	так
$A_1 A_2 A_4 A_3$	10	так
$A_1 A_4 A_2 A_3$	11	ні
$A_1 A_2 A_5 A_3$	11	ні
$A_1 A_5 A_2 A_3$	10	так
$A_1 A_4 A_5 A_3$	12	так
$A_1 A_5 A_4 A_3$	14	ні
$A_1 A_2 A_3 A_4$	9	ні
$A_1 A_3 A_2 A_4$	6	так
$A_1 A_2 A_5 A_4$	17	ні
$A_1 A_5 A_2 A_4$	9	так
$A_1 A_3 A_5 A_4$	12	ні
$A_1 A_5 A_3 A_4$	6	так
$A_1 A_2 A_3 A_5$	8	так
$A_1 A_3 A_2 A_5$	10	ні
$A_1 A_2 A_4 A_5$	10	так
$A_1 A_4 A_2 A_5$	14	ні
$A_1 A_3 A_4 A_5$	10	так
$A_1 A_4 A_3 A_5$	12	ні

Перший крок. Складемо всі можливі послідовні комбінації $A_1 A_{i2} A_{i3} A_{i4}$, що поєднують чотири об'єкта, і запишемо їх у табл. 7.3, групуючи їх по останньому об'єкту A_{i4}. Розрахуємо довжини шляхів, що відповідають цим послідовним комбінаціям, і запишемо їх у другий стовбець табл. 7.3.

Далі проведемо аналіз отриманих послідовних комбінацій. З послідовних комбінацій $A_1 A_3 A_4 A_2$ і $A_1 A_4 A_3 A_2$, що закінчуються в одному об'єкті A_2 і проходять крізь ті ж самі об'єкти A_3, A_4, напроти послідовної комбінації $A_1A_3A_4A_2$ пишемо "так", оскільки їй відповідає шлях меншої довжини (8 км) в порівнянні з послідовною комбінацією $A_1A_4A_3A_2$ (12 км), напроти якої пишемо "ні".

Аналогічно з послідовних комбінацій $A_1A_3A_5A_2$ і $A_1A_5A_3A_2$, що закінчуються в одному об'єкті A_2 і проходять крізь ті ж самі об'єкти A_3, A_5, напроти послідовної комбінації $A_1A_5A_3A_2$ пишемо "так", оскільки їй відповідає шлях меншої довжини (5 км) в порівнянні з послідовною комбінацією $A_1A_3A_5A_2$ (10 км), напроти якої пишемо "ні". Таким чином заповнюємо весь третій стовбець табл. 7.3.

Другий крок. У наступну таблицю (табл.7.4) запишемо лише ті послідовні комбінації з табл.7.3, напроти яких написано "так", і додамо до цих послідовних комбінацій ще по одному об'єкту A_{i5}, якого ще не має в послідовній комбінації. Розрахуємо довжини отриманих шляхів, що відповідають цим послідовним комбінаціям, і запишемо їх у другий стовбець табл. 7.4.

Таблиця 7.4.

Варіант маршруту	Довжина маршруту,км	Перспективно чи ні
$A_1\ A_3A_4\ A_2\ A_5$	13	ні
$A_1\ A_5A_3\ A_2\ A_4$	6	так
$A_1\ A_5A_4\ A_2\ A_3$	14	ні
$A_1\ A_2A_4\ A_3\ A_5$	11	ні
$A_1\ A_5A_2\ A_3\ A_4$	12	ні
$A_1\ A_4A_5\ A_3\ A_2$	13	ні
$A_1\ A_3A_2\ A_4\ A_5$	10	так
$A_1A_5A_2\ A_4\ A_3$	13	ні
$A_1\ A_5A_3\ A_4\ A_2$	8	так
$A_1\ A_2A_3\ A_5\ A_4$	15	ні
$A_1\ A_2A_4\ A_5\ A_3$	11	так
$A_1\ A_3A_4\ A_5\ A_2$	15	ні

Проведемо аналіз отриманих послідовних комбінацій. З послідовних комбінацій $A_1A_3A_4A_2A_5$, $A_1A_3A_4A_2A_5$ і $A_1A_3A_2A_4A_5$, що закінчуються в одному об'єкті A_5 і проходять крізь ті ж самі об'єкти A_2, A_3, A_4 напроти послідовної комбінації $A_1A_3A_2A_4A_5$ пишемо "так", оскільки їй відповідає шлях меншої довжини (10 км) у порівнянні з іншими послідовними комбінаціями $A_1A_3A_4A_2A_5$ (13 км) і $A_1A_3A_4A_2A_5$ (11 км), напроти яких пишемо "ні".

Аналогічно із послідовних комбінацій $A_1A_5A_3A_2A_4$, $A_1A_5A_2A_3A_4$ і $A_1A_5A_2A_3A_4$, що закінчуються в одному об'єкті A_4 і проходять крізь ті ж самі об'єкти A_2, A_3, A_5, напроти послідовної комбінації $A_1A_5A_3A_2A_4$ пишемо "так", оскільки їй відповідає шлях меншої довжини (6 км) у порівнянні із іншими послідовними комбінаціями $A_1A_5A_2A_3A_4$ (12 км) і $A_1A_5A_2A_3A_4$ (15 км), напроти яких пишемо "ні". Таким чином заповнюємо весь третій стовбець табл.7.4.

Третій крок. Оскільки отримані на другому кроці послідовні комбінації проходять крізь усі об'єкти, перейдемо до останнього кроку. У наступну таблицю (табл. 7.5) запишемо лише ті послідовні комбінації з табл. 7.4, напроти яких написано "так", і замкнемо отримані шляхи, додавши до цих послідовних комбінацій перехід до об'єкту A_1. Розрахуємо довжини отриманих шляхів, що відповідають цим послідовним комбінаціям, і запишемо їх у другий стовбець.

З аналізу замкнених шляхів, що записані в табл. 7.5 випливає, що у задачі, що розглядається, оптимальним є шлях $A_1A_5A_3A_2A_4A_1$, який має мінімальну довжину, що складає 13 км.

Таблиця 7.5.

Варіант маршруту	Довжина маршруту,км	Перспективно чи ні
A_1 A_3A_2 A_4 A_5 A_1	15	ні
A_1 A_5A_3 A_2 A_4 A_1	13	так
A_1 A_2A_4 A_5 A_3 A_1	14	ні
A_1 A_5A_3 A_4 A_2 A_1	18	ні

У описаній процедурі у загальному випадку піддається аналізу

$$n(n-1)(n-2)+\frac{n(n-1)(n-2)(n-3)}{2}+\frac{n(n-1)(n-2)(n-3)\,(n-4)}{2Ч3}+\ldots$$

неповних варіантів (за винятком останніх). Якби процедура відсіву починалася із особистісних варіантів, що включають одну ділянку (а не ***3***, як це має місце в даному випадку), то загальне число варіантів, які досліджуються виразилося б сумою

$$c_n^1+c_n^2+\ldots+c_n^{n-1}=2^n-2\ ,$$

що при великих n значно менше, ніж загальне число всіх варіантів, що дорівнює, як було вже відмічено вище $n!$.

7.4. Контрольні питання

1. Дайте визначення поняттю гамільтонів контур.
2. Сформулюйте обмеження, що накладаються на програму виконання робіт.
3. Сформулюйте задачу на складання розкладу.
4. У чому полягає зведення задачі оптимізації до рішення завдання цілочисельного лінійного програмування?
5. У чому полягає сутність динамічного програмування.
6. Дайте визначення поняттю комбінаторні задачі.
7. У чому полягає сутність методу гілок і меж?
8. Сформулюйте принцип оптимальності Беллмана.
9. Дайте визначення поняттю одиничний об’єм роботи.
10. Дайте визначення поняттю директивний термін виконання всього переліку робіт.
11. У чому полягає сутність статистичної процедури методу Монте-Карло?
12. Дайте визначення поняттю гамільтонів шлях.

Розділ 8

Структурна організація АСУ ТП

Розглянуто основні різновиди промислових АСУ, склад і функції АСУ ТП, основні операції управління технологічним процесом. Також розкриті варіанти структурної організації та методи автоматизованого управління. Наведено контрольні питання.

8.1. Основні різновиди промислових АСУ

Як вже відзначалося, основним інструментом для вирішення сучасних проблем управління матеріальним виробництвом є АСУ, у яких центральна роль і творчі здібності людини сполучаються із широким застосуванням сучасних математичних методів і засобів автоматизації, включаючи обчислювальну техніку.

Щоб одержати уявлення про особливості й характер функціонування сучасних систем управління, розглянемо загальну структурну схему управління технологічним об'єктом, яка надана на рис. 8.1.

Інформація про процеси, які протікають у об'єкті, що управляється надходить у систему управління, де порівнюється із завданням системі. Результати порівняння аналізуються, після чого готуються й приймаються відповідні рішення на вироблення управляючих впливів. На рис. 8.1. показана також можливість прийняття рішень на підставі результатів лише обробки або аналізу інформації.

Оскільки термін "управління" розуміється у досить широкому сенсі, АСУ можуть сильно відрізнятися за типом об'єктів управління, характеру й обсягу завдань, що розв'язуються й по ряду інших ознак.

Розглянемо основні три види АСУ, що зустрічаються на більшості промислових підприємств:

- автоматизована система управління підприємством;
- автоматизована система управління технологічним процесом;
- інтегровані автоматизовані системи управління.

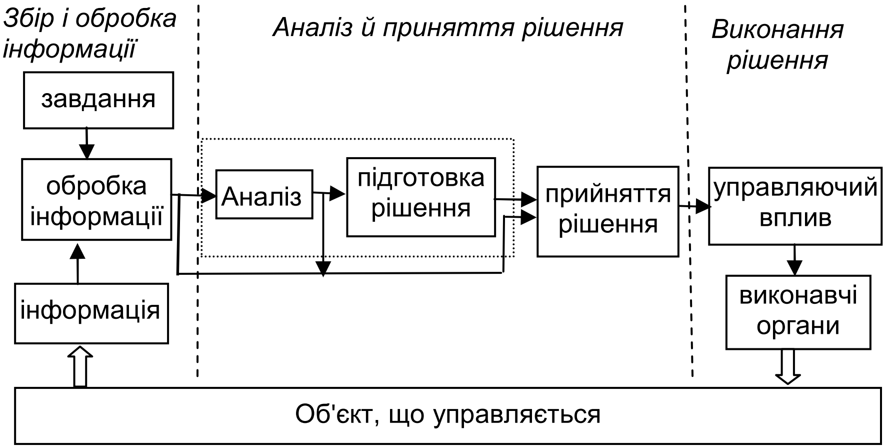


Рис. 8.1. Структура системи управління технологічним об'єктом

Автоматизована система управління підприємством (АСУ-П) призначена для рішення основних завдань управління виробничо-господарською діяльністю промислового підприємства у цілому та його самостійних частин на основі застосування економіко - математичних методів і засобів обчислювальної техніки. Обсяг і ступінь доступу до технологічної інформації залежать і від типу програмного забезпечення, що використовується в управлінських структурах підприємства, та від категорії співробітників-споживачів даної інформації.

Відзначимо основні особливості АСУ П, що визначають специфічні труднощі їхнього створення й використання:

- переважне значення в АСУ П мають економічні завдання управління, так як ритмічне функціонування підприємства можливо лише при наявності безперервних зв'язків між виробництвом і постачанням, виробництвом і фінансовими засобами, виробництвом і реалізацією готової продукції;
- визначальними в управлінні підприємством можуть бути не технологічні обмеження, а директивні вказівки у вигляді плану, що мають силу закону й обов'язкові до виконання;
- істотними є постійний взаємозв'язок з безліччю інших підприємств (організацій) і наявність внаслідок цього таких специфічних завдань, як управління постачанням, збутому, фінансової діяльністю, складання статистичної звітності, економіко-статистичні розрахунки й т.д.;

- важливу роль грають різноманітні завдання управління людьми й трудовими ресурсами (підготовка наказів і розпоряджень, контроль за прийомом і звільненням, розрахунок заробітної плати й т.д.);
- в АСУ П використаються специфічні форми зберігання й руху інформації - документообіг, пов'язаний із участю у рішенні загального завдання управління великого колективу людей.

Внаслідок сильного взаємозв'язку різних показників роботи підприємства основним критерієм управління для АСУ П є прибуток підприємства за планований період (наприклад, за 1 рік). Максимізація цього критерію при обліку інших показників у вигляді відповідних обмежень часто може вважатися формалізованою ціллю роботи підприємства.

Автоматизована система управління технологічним процесом (АСУ ТП) - це АСУ, яка призначена для вироблення й реалізації управляючих впливів на технологічний об'єкт управління відповідно до прийнятого критерію управління. АСУ ТП є одним із різновидів АСУ тому їй властиві наступні ознаки, загальні для всіх АСУ:

- це людино-машинна система, у якій людина відіграє найважливішу роль, приймаючи у більшості випадків змістовну участь у ухваленні рішень по управлінню;
- істотне місце в АСУ ТП займають автоматичні пристрої (у тому числі засобу обчислювальної техніки), що виконують трудомісткі операції по збору, обробці й переробці інформації;
- ціль функціонування АСУ ТП - оптимізація роботи об'єкта шляхом відповідного вибору управляючих впливів.

Слід зазначити, що АСУ може бути віднесена до класу АСУ ТП тільки в тому випадку, якщо вона:

- здійснює вплив на об'єкт у темпі протікання у ньому технологічних процесів;
- забезпечує управління технологічним об'єктом у цілому;
- технічні засоби АСУ ТП беруть участь у виробленні рішень по управлінню об'єктом.

Останніми двома обставинами АСУ ТП якісно відрізняється від традиційних систем автоматизації й різноманітних локальних систем автоматики, які власно кажучи являють собою технічні засоби для автоматизації дій людини на той або іншій дільниці процесу. На відміну від цього в АСУ ТП реалізується автоматизований процес прийняття рішень по управлінню технологічним об'єктом як єдиним цілим. Для цього в АСУ ТП застосовуються різноманітні, "інтелектуальні" автоматичні пристрої обробки інформації, і насамперед - сучасні засоби обчислювальної техніки.

Призначення будь-якої АСУ, її функціональні можливості, бажані технічні характеристики й інші особливості визначаються тим об'єктом, для якого створюється дана система.

Для АСУ ТП об'єктом, що управляється є ***технологічний об'єкт управління*** (ТОУ), що представляє собою сукупність технологічного устаткування й реалізованого на ньому, технологічного процесу виробництва цільового продукту. У якості ТОУ можуть розглядатися:

- технологічні агрегати й установки;
- окремі виробництва, що реалізують самостійний, закінчений технологічний цикл;
- виробничий процес усього промислового підприємства, якщо управління їм полягає у виборі й узгодженні раціональних режимів роботи взаємозалежних агрегатів, ділянок і виробництв.

Таким чином призначення АСУ ТП можна визначити як цілеспрямоване ведення технологічного процесу й забезпечення суміжних і вищестоящих систем управління необхідною інформацією.

Функціонування АСУ ТП припускає виконання сукупності дій системи (заздалегідь визначених і описаних у експлуатаційній документації послідовності операцій і процедур, що виконуються складовими системи), спрямованих на досягнення особистісної цілі управління. Функціонування кожної АСУ ТП спрямоване на одержання цілком певних техніко-економічних результатів (зниження собівартості продукції, зменшення втрат, підвищення продуктивності праці, і т.п.).

АСУ ТП відносяться до класу організаційно-технічних систем, що складаються із персоналу й комплексу засобів автоматизації (КЗА) його діяльності. Управлінські рішення формуються шляхом аналізу й обробки інформації. У даному випадку інформаційні технології в АСУ реалізуються у вигляді певної логічної послідовності інформаційно-зв'язаних функцій, завдань або процедур, що виконуються в автоматизованому (інтерактивному) або автоматичному режимах.

Розглянемо, у якості прикладу, ієрархію рівнів АСУ ТП цеху підприємства (рис. 8.2.). При вирішенні завдань автоматизації виробництва діють наступні принципи:

- незалежність рівнів;
- незалежна попередня обробка інформації на кожному рівні;
- прийняття рішень на робочому місці;

- час обміну інформацією між рівнями повинен бути мінімальним;
- кожний рівень зв'язаний інформаційними потоками, по можливості, тільки із прилеглими рівнями;
- збереження даних на місці їхнього виникнення.

Рівні автоматизації

Управління виробничими ресурсами
Рівень 4

⇕

Оперативне управління виробництвом
Рівень 3

⇕

Системи візуалізації й управління обладнанням (робочі місця)
Рівень 2

⇕

Контролери
Рівень 1

⇕

Виконуючи пристрої, датчики й механізми
Рівень 0

Рис. 8.2. Ієрархія рівнів АСУ ТП цеху

На рівні 0 реалізуються функції самого нижнього рівня, такі як вплив на параметри технологічного процесу й облік параметрів стану системи.

Рівень 1 передбачає реалізацію функцій, що пов'язані із регулюванням, обслуговуванням, захистом обладнання, контролем процесу функціонування обладнання, управлінням (станками із ЧПУ, виконанням різних поточних завдань), обробкою величин, що підлягали вимірюванню.

На рівні 2 вирішуються завдання, що пов'язані із контролем параметрів виробничого процесу, виконанням технологічних операцій, введення технологічних параметрів і контролем результатів (виконує оператор).

Рівень 3 передбачає управління усім виробничим процесом та координація роботи всіх ділянок цеху. Основними завданнями, що вирішуються на даному рівні є: складський облік, виробниче планування й управління, оптимізація технологічного процесу.

На рівні 4 здійснюється управління виробничими ресурсами, яке включає координуючі й адміністративні функції всього підприємства: матеріально-технічне постачання, статистику, планування потужностей, одержання інформації про поточний стан справ, проведення коректування довгострокових планів.

Формалізовано природну форму опису функції технічної системи можна представити такою, що складається з трьох компонентів

$$F = (D, G, H),$$

де **D** - дії, що виконуються даною системою, для досягнення бажаного результату;

G - визначення об'єкту (об'єктів), на який направлена дія;

H - визначення особливих умов і обмежень, при яких виконується дія.

Будь-яка сформульована функція відповідає деякому класу технічних систем, що реалізовують цю функцію або що можуть її реалізувати. При конкретизації описів компонентів **D** і **G**, при збільшенні числа умов до обмежень **H** і при більшій їх конкретизації, опис функції відповідатиме різним більш вузьким класам технічних систем. У цілому опис функції необхідно формулювати конкретніше, коротше і простіше.

Функція технічної системи може бути простою чи складною. ***Складна функція*** представляє сукупність двох або більше простих функцій. ***Проста функція*** - це функція, яку не можна розкласти на інші функції системи. Прості функції поєднуються за спільністю цілі, ролі у процесі управління, використання інформації й інших ознак. Сукупність всіх функцій АСУ можна розглядати як одну складену функцію.

Розрізняють інформаційні, управляючі й допоміжні функції АСУ ТП.

До ***інформаційних*** відносяться такі функції АСУ ТП, результатом виконання яких є надання оператору системи або зовнішньому споживачу інформацію про хід процесу, що управляється. Інформаційна функція АСУ F^I включає: одержання інформації $F^I_{од}$, обробку $F^I_{об}$; відображення $F^I_{від}$ та передачу інформації $F^I_{пд}$ про стан об'єкта управління й зовнішнього середовища:

$$F^I = \left\{F^I_{од}, F^I_{об}, F^I_{від}, F^I_{пд}\right\}.$$

Характерними прикладами інформаційних функцій АСУ ТП є:

- безперервна перевірка відповідності параметрів процесу припустимим значенням і негайне інформування персоналу при виникненні невідповідностей;
- вимір або реєстрація, по команді оператора, технологічних параметрів, що підлягають контролю;
- інформування оператора (по його запиту) про ситуації на тій або іншій ділянці об'єкта управління у поточний момент часу;
- фіксація часу відхилення параметрів процесу за допустимі межі;
- обчислення комплексних показників, що не піддаються безпосередньому виміру, що характеризують якість продукції або інших важливих показників технологічного процесу;
- періодична реєстрація параметрів, що вимірюються і показників, що обчислюються;
- виявлення й сигналізація про настання аварійних ситуацій.

Виконуючи ці інформаційні функції АСУ ТП вчасно забезпечує оператора або вищестоящу систему відомостями про стан й будь-які відхилення від нормального протікання технологічного процесу.

Управляючі функції АСУ ТП містять у собі дії по виробленню й реалізації раціональних управляючих впливів на об'єкт. Тут під виробленням розуміється визначення (на підставі отриманої інформації) раціональних впливів, а під реалізацією - дії, що забезпечують здійснення прийнятих рішень. Управляюча функція АСУ $F^{у}$ включає: приведення у готовність АСУ $F^{у}_{г}$; одержання інформації про стан об'єкта управління й зовнішнього середовища $F^{у}_{оі}$, оцінку інформації $F^{у}_{оц}$, вибір управляючих впливів $F^{у}_{в}$ і їхню реалізацію $F^{у}_{р}$:

$$F^{у} = \left\{F^{у}_{р}, F^{у}_{оі}, F^{у}_{оц}, F^{у}_{в}, F^{у}_{г}\right\}.$$

До основних управляючих функцій відносяться:

- стабілізація змінних технологічного процесу на постійних значеннях, обумовлених регламентом виробництва;
- програмна зміна режиму процесу, що протікає по заздалегідь заданим вимогам;
- захист устаткування від аварій;
- формування й реалізація управляючих впливів, що забезпечують досягнення або дотримання режиму, що є оптимальним по деякому технологічному або техніко-економічному критерію;
- розподіл матеріальних потоків і навантажень між технологічними агрегатами;
- управління пусками й зупинками агрегатів і ін.

Допоміжна фікція F^{∂} АСУ включає: збір $F^{\partial}_{зб}$ й обробку даних $F^{\partial}_{об}$ про стан АСУ і або подання $F^{\partial}_{ні}$ цієї інформації персоналу системи або здійснення $F^{\partial}_{ув}$ управляючих впливів на відповідні технічні й програмні засоби АСУ. Крім того, допоміжна функція АСУ включає наступні особистісні функції: документування F^{∂}_{∂} інформації, що циркулює у системі управління, синхронізації різних сигналів $F^{\partial}_{сн}$, тренування персоналу АСУ, обліку й планування різних ресурсів $F^{\partial}_{рс}$ і інші $F^{в}_{інш}$, тому складна допоміжна функція автоматизованої системи може бути представлена наступною підмножиною особистісних функцій:

$$F^{\partial} = \left\{F^{\partial}_{зб}, F^{\partial}_{пі}, F^{\partial}_{ув}, F^{\partial}_{\partial}, F^{\partial}_{сн}, F^{\partial}_{рс}, F^{\partial}_{інш}, F^{\partial}_{об}\right\}.$$

Безлічі $F^{i}, F^{у}, F^{\partial}$ мають близькі за змістом елементи. Тому вони можуть бути об'єднані в окремі підмножини, планування $F_{п}$**,** обліку $F_{обл}$**,** контролю $F_{к}$, аналізу $F_{а}$ і безпосереднього оперативного управління $F_{оу}$ в узагальненому функціональному графі АСУ, тобто у ***F***- структурі, яка є основою й вихідними даними для розробки ***О, І, А, Р, Т*** і ***Д*** структур АСУ.

Інтегровані АСУ є органічним об'єднання декількох АСУ ТП або АСУ П між собою, або із галузевими АСУ, яке здійснюється з ціллю підвищення загальної технічної й економічної ефективності їхнього функціонування, підвищення якості продукції, що, випускається, а також забезпечення нової якості управління за рахунок створення єдиного інформаційного простору підприємства.

При впровадженні й експлуатації на одному підприємстві АСУ ТП і АСУ П розглядаються як взаємозалежні, але окремі системи, між якими існують відносини ієрархічної співпідпорядкованості як молодшого до старшого, а не як частини до цілого.

АСУ агрегатами, АСУ П та галузеві АСУ (ГАСУ) не "вкладені" одна в іншу, а утворюють багаторівневу ієрархію автоматизованих систем управління промисловими об'єктами (рис. 8.3.).

Інтегровані АСУ особливо ефективні у тих випадках, коли у них реалізується взаємопов'язане, погоджене управління, як технологією, так і організацією виробництва у масштабі всього підприємства. Комплексна інтеграція: припускає створення у рамках підприємства єдиного банку даних про продукцію, технологічні процеси, даних допоміжних виробництв, що знижує ступінь дублювання інформації й забезпечує стандартизацію всієї діяльності підприємства. АСУ ТП одержує від відповідних підсистем АСУ П завдання й обмеження (обсяг виробництва, задані значення техніко-економічних показників і ін.) і забезпечує підготовку й передачу підсистем АСУ П необхідної техніко-економічної інформації (про виконання завдань, стан устаткування, хід технологічного процесу й ін).

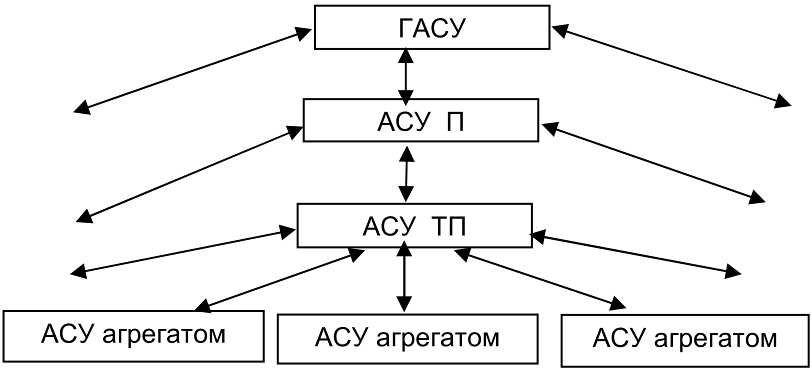


Рис. 8.3. Ієрархія АСУ промисловими об'єктами

Інтеграція систем автоматизації повинна здійснюватися у двох напрямках: горизонтальному й вертикальному.

Горизонтальна інтеграція це об'єднання між собою всіх автономних систем автоматизації технологічних і виробничих процесів, а також адміністративних відділень цехового рівня в єдину інформаційну мережу реального часу. Горизонтальна інтеграція припускає організацію обліку на кожному етапі виробничого процесу від поставки вихідної сировини до складу готової продукції.

Вертикальна інтеграція ґрунтується на організації потоків інформації від нижнього рівня (датчиків і контролерів технологічного устаткування) у внутрішні й зовнішні комп'ютерні мережі підприємства й через них у адміністративні системи управління. Основна ціль вертикальної інтеграції об'єднання промислових і адміністративних мереж та усунення перешкод на шляху інформаційних потоків між рівнями АСУ П і АСУ ТП з ціллю оперативного обміну даними.

8.2. Склад АСУ ТП

Прийнято вважати, що до складу будь-якої АСУ ТП повинні входити наступні основні компоненти (рис. 8.4): оперативний персонал, інформаційне, організаційне, правове, лінгвістичне, метрологічне, програмне і технічне забезпечення. Розробка програмного забезпечення проводиться на основі математичного забезпечення, яке до складу АСУ не входить. Під ***компонентом АСУ*** розуміється виділена за певною ознакою або сукупністю ознак частина АСУ, що розглядається як єдине ціле.

Процес функціонування АСУ за суттю є процесом цілеспрямованого перетворення інформації, що включає: збір обробку і аналіз вхідної інформації, прийняття рішення по управлінню і його реалізацію. Ці дії виконуються спільно двома компонентами: оперативним персоналом і комплексом засобів автоматизації.

У загальному випадку ***комплекс засобів автоматизації АСУ*** являє собою сукупність всіх компонентів АСУ, за винятком людей і об'єкта управління.

У вузькому сенсі КЗА АСУ є сукупність взаємозалежних компонентів і комплексів програмного, технічного й інформаційного забезпечення яка розробляється і виготовляється як продукція науково-технічного й виробничо-технічного призначення, із необхідною експлуатаційною документацією.

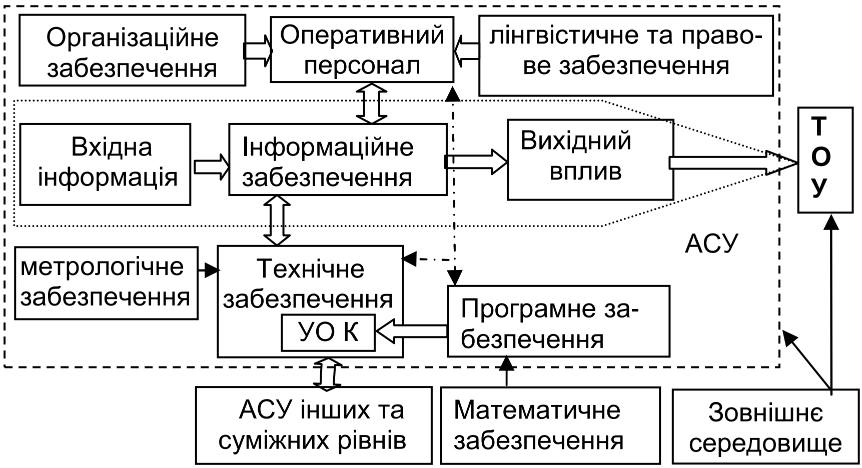


Рис. 8.4. Схема взаємодії основних компонентів АСУ ТП

При функціонуванні КЗА використаються наступні засоби: інформаційні, програмні, технічні, організаційні, методичні, лінгвістичні, математичні, метрологічні, ергономічні, правові й інші, склад яких визначається у процесі створення системи.

Між всіма компонентами АСУ у процесі її функціонування відбувається інтенсивна взаємодія. Організаційне і програмне забезпечення визначають поведінку і взаємодію оперативного персоналу і обчислювального комплексу. Крім того, персонал активно взаємодіє із технічним забезпеченням і при необхідності корегує програмне забезпечення. Всі ці взаємодії усередині системи, а також її взаємодія із зовнішнім середовищем носять, у основному, інформаційний характер, оскільки зводяться до передачі і прийому інформації у вигляді різних сигналів, даних, повідомлень і т.п.

Безліч прийнятих форм масивів даних, переліків і шкал сигналів й кодів, що використовуються і правил їх розшифровки, утворює основний компонент АСУ, який має назву ***інформаційне забезпечення***. Важлива роль цього компоненту полягає у тому, що саме через нього, а точніше - з його допомогою здійснюються всі процеси обміну інформацією як всередині АСУ, так і із зовнішнім середовищем.

До складу інформаційного забезпечення АСУ входять класифікатори техніко-економічної інформації, нормативно-довідкова інформація, форми уявлення і організації даних у системі, у тому числі форми документів, відеограми, масиви інформації і логічні інтерфейси (протоколи обміну даними).

Оперативний персонал АСУ складається з технологів - операторів (диспетчерів), що здійснюють контроль і управління об'єктом, і експлуатаційного персоналу, що забезпечує правильність функціонування всіх технічних і програмних засобів системи.

Склад оперативного персоналу конкретної АСУ і встановлені взаємини між її працівниками визначають ***організаційну структуру системи***. Елементами такої структури є окремі посадовці - виробничі або адміністративні працівники, що здійснюють у той чи іншій мірі управління даним технологічним об'єктом. Основні зв'язки між елементами організаційної структури відповідають відносинам оперативної підпорядкованості вказаних працівників. При необхідності на схемі організаційної структури відображають також територіальне розміщення оперативного персоналу АСУ і його взаємодії з персоналом інших систем і (або) рівнів управління.

До складу ***математичного забезпечення*** АСУ входять методи розв'язання задач управління, моделі і алгоритми.

У функціонуючій системі математичне забезпечення реалізується у складі програмного забезпечення.

До складу ***програмного забезпечення*** АСУ входять програми (у тому числі програмні засоби) з програмною документацією на них, необхідні для реалізації всіх функцій АСУ в обсязі, передбаченому у технічному завданні на створення АСУ.

До складу ***технічного забезпечення*** АСУ входять технічні засоби, необхідні для реалізацій функцій АСУ. У загальному випадку воно включає засоби отримання, введення, підготовки, обробки, зберігання (накопичення), реєстрації, виведення, відображення, використання, передачі інформації і засоби реалізації управляючих впливів.

До складу ***організаційного забезпечення*** АСУ входять документи, що визначають функції підрозділів управління, дії і взаємодії персоналу АСУ.

До складу ***метрологічного забезпечення*** АСУ входять метрологічні засоби і інструкції по їхньому застосуванню.

До складу ***правового забезпечення*** АСУ входять нормативні документи, що визначають правовий статус АСУ, персоналу АСУ, правил функціонування АСУ і нормативи на документи, що автоматично формуються, у тому числі на машинних носіях інформації.

Правове забезпечення АСУ у складі функціонуючої системи реалізується у вигляді документів організаційного забезпечення АСУ.

До складу ***лінгвістичного забезпечення*** АСУ входять тезауруси і мови опису і маніпулювання даними. Лінгвістичне забезпечення функціонуючої АСУ може бути присутнім у ній самостійно або у вигляді рішень в інформаційному забезпеченні АСУ і у документах організаційного забезпечення АСУ.

8.3. Структурна організація АСУ ТП

8.3.1. Варіанти структурної організації АСУ ТП

Розглянемо основні (типові) варіанти структурної організації АСУ ТП. Для нескладних об'єктів, з малою кількістю технологічних агрегатів, які до того ще і зосереджені на невеликій території, звичайно використовують однорівневу централізовану структуру (рис. 8.5).

У даному випадку управління усіма технологічними агрегатами (ТА) і обслуговування системи достатньо надійно і зручно здійснюється за допомогою одного (загального) управляючого пристрою (УП). Однак, така структура стає мало ефективною при збільшенні кількості агрегатів і відстані між ними, бо при цьому, перш за все, ускладнюється обслуговування системи (налагодження агрегатів, пошук несправностей), значно збільшуються витрати на реалізацію з'єднань УП з об'єктами управління (витрати на канали зв'язку), а також виникає можливість виходу із ладу всієї системи при відмовах УП.

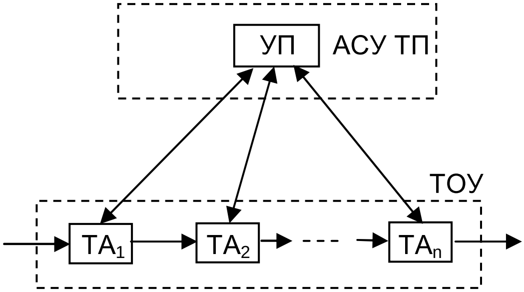


Рис. 8.5. Однорівнева централізована структура

Щоб уникнути вказаних вище недоліків, застосовують однорівневі децентралізовані (розподілені) структури АСУ ТП, один із варіантів якої наведений на рис. 8.6.

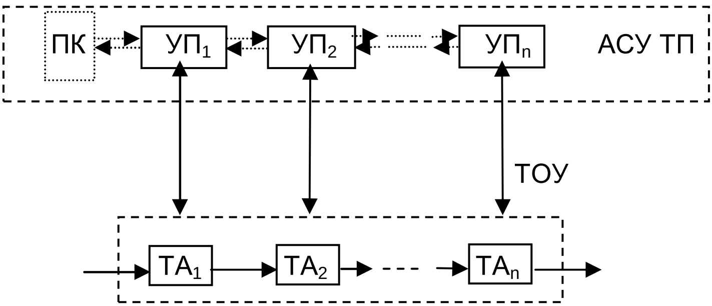


Рис. 8.6. Однорівнева децентралізована структура

При застосуванні децентралізованої структури управління кожним агрегатом здійснюється від окремого УП, що забезпечує підвищення надійності функціонування АСУ ТП, бо при відмові одного з УП інші УП продовжують працювати і ТОУ у цілому, хоча і не у повному обсязі, залишається у робочому стані.

Слід відзначити, що, хоча управління кожним агрегатом здійснюється в основному незалежно, деякі зв'язки між УП залишаються (показані пунктиром) - це, як правило, - сигнали блокування між суміжними агрегатами. Крім того, у сучасних розподілених системах безпосереднє управління агрегатами здійснюється саме окремими

УП, але для введення планових показників, виведення інформації про стан агрегатів, про кількість зробленої продукції використовується окремий пристрій координації (ПК), який виконує функції координації роботи УП і який зв'язаний з ними за допомогою локальної мережі.

Більшість сучасних АСУ ТП мають багаторівневу ієрархічну структуру, приклад якої наведений на рис. 8.7.

На 1-му рівні ієрархії здійснюється локальне (незалежне) управління технологічними агрегатами, для кожного з яких застосовується окремий ПК.

ПК, що встановлені на 2-му рівні ієрархії, забезпечують координацію сумісної роботи агрегатів у межах, наприклад, дільниці технологічної лінії. ПК цього рівня забезпечують також схемну реалізацію та контроль виконання системи блокувань між суміжними агрегатами, контроль станів агрегатів ("працює" – "не працює"), контроль вхідних (на початку дільниці) і вихідних (на виході з дільниці) параметрів технологічного процесу та інші функції.

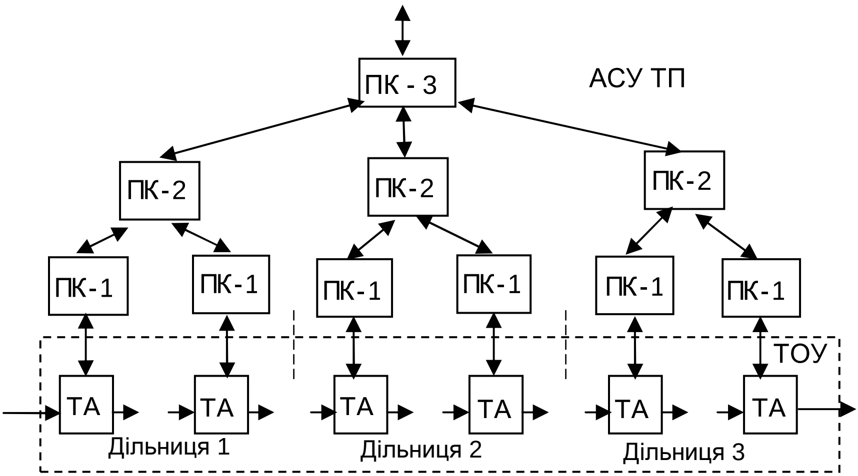


Рис. 8.7. Багаторівнева ієрархічна структура

ПК, що встановлюється на 3-му рівні ієрархії, забезпечує формування планових завдань та режимів роботи дільниць технологічної лінії та передачу цієї інформації до ПК 2-го рівня, накопичення та візуалізацію даних про хід технологічного процесу, формування підсумкових відомостей про роботу лінії за зміну, добу, тиждень, а також може виконувати інші функції.

8.3.2. Різновиди структур АСУ ТП

Сучасні АСУ ТП дуже різноманітні і можуть відрізнятися одна від одної по функціональному складу, ступеню автоматизації управління об'єктом, технічним засобам, що використовуються та багатьом іншим ознакам і характеристикам. Проте, щоб отримати уявлення про те, якими бувають АСУ ТП, доцільно розглянути характерні різновиди структур таких систем, що розрізняються способами виконання основних інформаційних та управляючих функцій.

АСУ ТП, що функціонують без обчислювального комплексу. Подібні людино-машинні системи зазвичай застосовуються для управління окремими відносно простими технологічними агрегатами, установками або групами апаратів. У загальній структурі управління виробництвом такі системи займають самий нижній ступінь ієрархії і тому характеризуються тісним зв'язком із об'єктом, деякою автономністю "поведінки", найбільшою оперативністю контролю і управління. Основні функції таких систем управління наступні:

- контроль параметрів технологічного процесу;
- стабілізація технологічного процесу на заданому постійному режимі, що визначається регламентом виробництва;
- програмне управління (включаючи пуск та зупинку устаткування);
- захист устаткування від аварій;
- оперативний зв'язок з вищестоящими рівнями управління.

Завдання управління на нижньому рівні мають свої специфічні особливості. Такі системи управління історично були першою областю, у якій почали застосовуватися автоматичні пристрої. Якщо завдання координації (середній рівень управління) і планування (верхній рівень управління), у більшості випадків вирішуються людиною без застосування безпосередньо пов'язаних із об'єктом технічних засобів автоматизації, то завдання автоматичного контролю і стабілізації технологічних параметрів, і у ряді випадків завдання оптимізації, вже давно успішно вирішуються за допомогою автономних спеціалізованих технічних пристроїв.

Практично вся інформація про стан об'єкту вводиться у таку систему автоматично від датчиків, а дії, що управляють, поступають від неї безпосередньо на регулюючі органи. При цьому збір інформації і формування дій, що управляють, зазвичай проводяться або безперервно, або із достатньо високою частотою, що визначається темпом технологічного процесу, який управляється.

Унаслідок різноманіття форм зв'язку з об'єктом технічні засоби, що використовуються на нижньому ступені управління, більш різнотипні і численні, чим засоби, що використовуються на верхніх рівнях ієрархії. До складу навіть простих систем нижнього рівня входять різноманітні вимірювачі, регулюючі, логічні й інші спеціалізовані пристрої. Крім того, у розпорядженні оператора системи є різні автоматичні пристрої контролю і управління, що дозволяють розвантажити його від виконання численних одноманітних дій по спостереженню за станом устаткування і управлінню їм, зосередити увагу на головних технологічних параметрах і операціях. Один із варіантів таких систем управління зображений на рис.8.8, де показані її основні функціонально - апаратурні частини (підсистеми).

Призначення підсистеми дистанційного управління полягає у передачі дій оператора на виконавчі механізми, віддалені від центрального пункту управління.

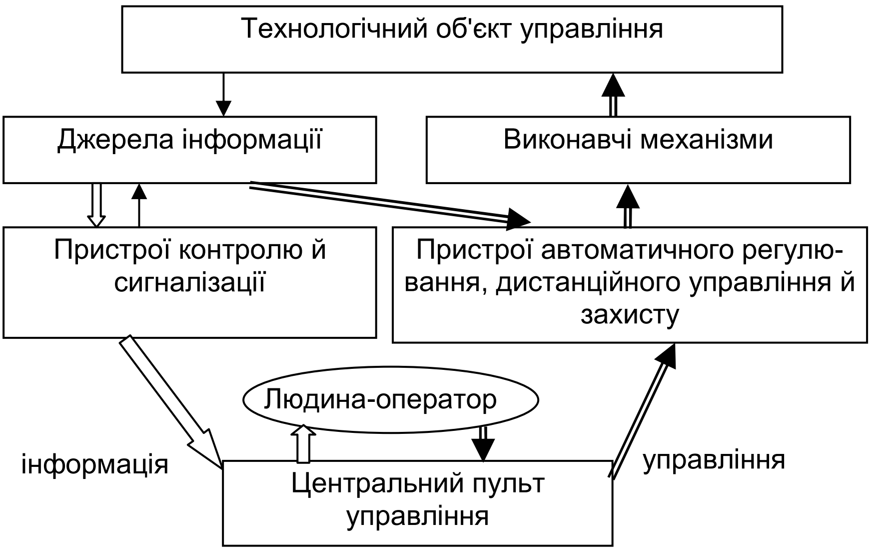


Рис.8.8. АСУ ТП без обчислювального комплексу

Підсистема автоматичного регулювання повинна підтримувати задану продуктивність промислових установок і стабілізувати технологічні параметри на заданому рівні.

Пристрої захисту покликані запобігати виникненню і розвитку аварій і оберігати установки від пошкоджень й руйнувань при виході із ладу окремих елементів устаткування, відмовах або при помилкових діях оператора.

Призначення пристроїв сигналізації полягає у фіксації моментів перевищення контрольованими параметрами гранично допустимих значень, інформування про цей факт оператора і реєстрації причин виникнення, ходу і розвитку аварій.

Пристрої контролю необхідні для спостереження за відповідальними параметрами, перевірки поточних значень великої кількості однотипних допоміжних параметрів.

Вимірювальні прилади, перемикачі і інші технічні засоби всіх підсистем розміщуються на центральних щитах і пультах управління, які встановлюються у спеціально відведених для них приміщеннях і обслуговуються оператором. При цьому частина другорядних технологічних параметрів контролюється за допомогою приладів, встановлених на так званих місцевих щитах управління, що розташовані поблизу агрегатів, що діють.

Описана система, хоча і не містить у своєму складі обчислювального комплексу, є людино-машинною системою, тобто простим, але достатньо представницьким і поширеним різновидом АСУ ТП.

АСУ ТП з обчислювальним комплексом, що виконує інформаційні функції. Системи цього вигляду (рис.8.9.) містять всі функціональні і апаратурні елементи, властиві попередній системі, але відрізняються від неї наявністю обчислювального комплексу (ОК), який виконує функції централізованого контролю процесу, що протікає, обчислення комплексних технічних і техніко - економічних показників, а також контроль роботи і стану устаткування.

Обчислювальний комплекс отримує всю необхідну інформацію про протікання процесу і стан об'єктів, зокрема про величини, що регулюються і управляються. Характерною особливістю даного виду системи є те, що завдання аналізу інформації, що поступає, ухвалення рішень, а також здійснення дій, що управляють, як і у системах попереднього вигляду, покладаються на оператора.

Дані про об'єкт, що отримані за допомогою обчислювального комплексу, що управляє (УОК), надаються на засоби відображення інформації і можуть або передаватися у вищестоящі пункти управління (ВПУ) для подальшої обробки, або виводиться на зовнішні накопичувачі інформації. Ціллю збору даних може бути також вивчення технологічного процесу за різних умов. У результаті накопичується інформація, що дозволяє побудувати і уточнити математичну модель процесу, яким потрібно управляти.

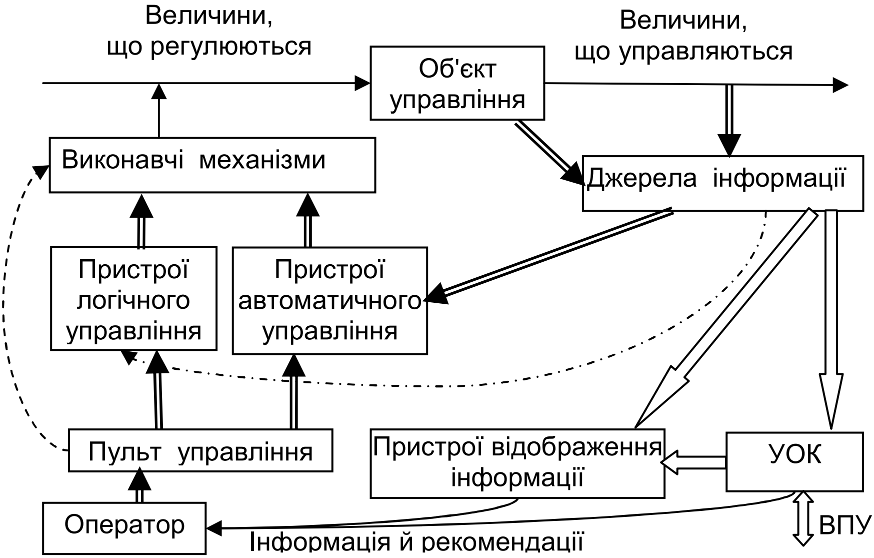


Рис.8.9. АСУ ТП з обчислювальним комплексом, що виконує інформаційні функції

АСУ ТП із обчислювальним комплексом, що виконує управляючі функції у режимі "порадника". Структурна схема такої системи аналогічна наведеної раніше АСУ ТП з обчислювальним комплексом, що виконує інформаційні функції. Але крім функцій, що виконуються УОК у попередньому варіанті системи, на нього покладають завдання аналізу інформації, що поступає й пошуку оптимальних рішень із видачею рекомендацій із управління (порад) оператору-технологу (рис.8.10).

У цій системі, на засобах відображення, людині пред'являється не тільки інформаційна модель ситуації, що виникла на об'єкті, але й обраний системою один із альтернативних варіантів рішення по управлінню, у вигляді інформаційної моделі рішення. Оператору необхідно враховувати, що моделі функціонування об'єкту, які закладені у УОК для вироблення альтернативних варіантів рішень, не повністю адекватні поводженню об'єкта, так як багато факторів не включені у ці моделі або із причин важкого їхнього обліку, або із причин їхнього рідкого виникнення. Людина, що веде управління, порівнюючи рішення,що запропоноване системою, із рішенням, що він вибрав сам, може або погодитися із рішенням системи (дозволити системі його реалізацію), або ввести у систему своє рішення. Остаточний вибір і здійснення впливів, що управляють, як і раніше залишаються за оператором.

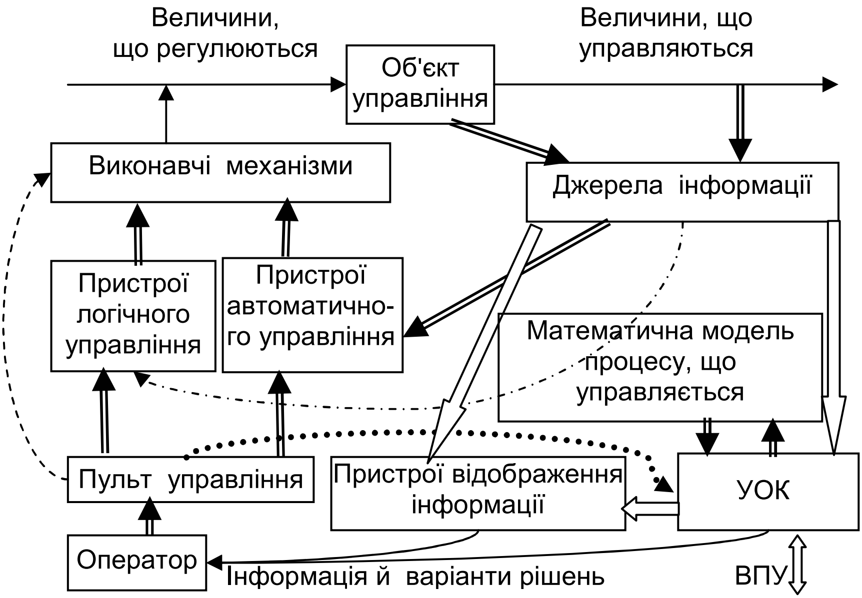


Рис. 8.10. АСУ ТП з обчислювальним комплексом, що виконує управляючі функції у режимі "порадника"

Така АСУ функціонує у такий спосіб. Через задані проміжки часу (залежно від конкретних умов) отримані УВК дані про стан об'єкту й комплексні, технічні й техніко - економічні показники аналізуються за допомогою математичної моделі процесу, що управляється. Шляхом обчислень по моделі визначаються впливи, необхідні для наближення процесу до оптимуму. Результати надаються оператору, який здійснює управління процесом шляхом зміни уставок регуляторів або виконуючи інші дії відповідно до рекомендацій, що вироблені УВК.

Регулятори у такій системі є засобами не тільки стабілізації, але й програмної зміни технологічних параметрів процесу, а оператор відіграє роль ланки, що стежить і координує й вносить зміни відповідно до порад УОК, що безупинно допомагає оператору у його зусиллях оптимізувати технологічний процес.

Число вхідних змінних параметрів у системі, що працює у режимі порадника оператора може перебувати у межах від 10 до 100, але УОК може, якщо це економічно доцільно, обробляти й більше число змінних. Число змінних, що управляються порівняно невелике, тому що оператору самому доводиться змінювати завдання. Намагатися виконувати регулювання в системі кожні 10-:-15 хвилин можливо, але, звичайно, недоцільно. Робота у такому темпі не може виконуватися безпомилково всю робочу зміну. Тому одним із недоліків режиму управління, що розглядається є наявність обмежень, пов'язаних із участю у системі людини.

Однак управління цього типу має й переваги. Воно задовольняє вимозі обережного підходу до нових способів управління. Застосування УОК у режимі порадника забезпечує також гарні можливості для перевірки нових моделей процесу. У якості оператора при цьому може виступати інженер-технолог, що звичайно краще оператора відчуває технологічний процес. Крім того, УОК у такій системі може стежити за виникненням усіх можливих аварійних ситуацій, що дозволяє оператору звільнитися від виконання цієї стомлюючої роботи. Розглянута АСУ може бути побудована на принципах самонавчання.

АСУ ТП із обчислювальним комплексом, що виконує функції центрального управляючого пристрою (супервізорне управління). Характерна риса таких систем управління полягає у тому, що в них УОК включається у замкнутий контур автоматичного управління й виробляє управляючі впливи, що надходять безпосередньо на вхід системи автоматичного регулювання (рис. 8.11). Цей режим роботи УОК істотно відрізняється від режиму порадника, при якому всі зміни в управління вносить тільки оператор. Основне завдання супервізорного управління - автоматична підтримка технологічного процесу поблизу оптимальної робочої точки шляхом оперативного впливу на нього. У цьому одна із головних переваг даного виду систем.

Робота обчислювального комплексу по збору й переробці інформації у системі супервізорного управління мало відрізняється від роботи, що описана для режиму порадника оператора. Обчислення по визначенню впливів, що управляють теж аналогічні. Однак якщо після виконання розрахунків у попередньому випадку знайдені нові значення уставок перетворювалися у форму, зручну для сприйняття оператором, то у даному випадку вони перетворяться у сигнали, які можна використовувати для зміни завдання й настроювання регуляторів.

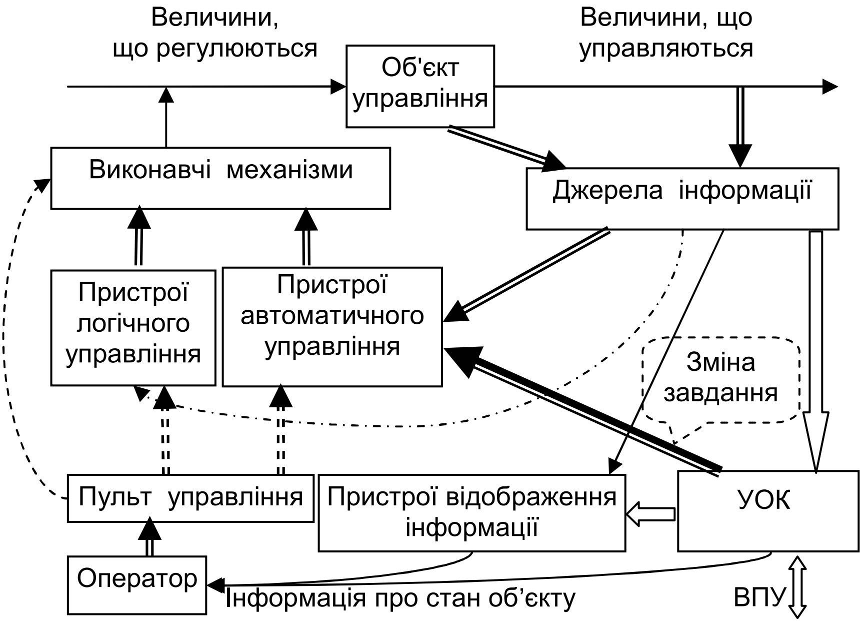


Рис. 8.11. АСУ ТП з обчислювальним комплексом, що здійснює супервізорне управління

Оскільки у таких системах контур управління є замкнутим через УОК, то функції оператора зводяться до загального спостереження за ходом технологічного процесу. Втручання оператора потрібно лише при виникненні якихось непередбачених (наприклад, аварійних) ситуацій. Залишається також необхідність вносити корективи в управління процесом при змінах, наприклад, сировини або складу продукції, що виробляється. Це часто вимагає визначення нових значень коефіцієнтів рівнянь, що описують технологічний об'єкт управління. Відповідні розрахунки можуть виконуватися зовнішньою обчислювальною машиною, здатною вирішувати задачі великої розмірності, або обчислювальним комплексом самої АСУ. У останньому випадку необхідно забезпечити розподіл часу обчислювального комплексу АСУ між завданням управління ТП і додатковими обчисленнями по оптимізації ТП.

Важлива перевага систем супервізорного управління полягає у тому, що у них УОК не тільки безупинно контролює процес, але й автоматично управляє їм поблизу оптимального значення. Це дозволяє виключити флуктуації, що пов'язані із якістю роботи різних операторів, "почерк" кожного з яких обов'язково впливає на процес налагодження регуляторів. Оскільки обчислення виконуються із більшими швидкостями у моделі процесу можна відбити значно більшу кількість змінних.

АСУ ТП із обчислювальним комплексом, що виконує функції безпосереднього (прямого) цифрового управління. Головні розбіжності систем управління, що були розглянуті раніше, міститься у принципах використання обчислювального комплексу. У системі УОК якої виконує роль "порадника" оператора, не виконується пряме управління процесом. Завдання по управлінню вводяться оператором. При супервізорному управлінні уставки регуляторів задаються від ЕОМ, але команди на пристрої управління об'єкта надходять від регуляторів.

У АСУ ТП, що надана на рис. 8.12 обчислювальний комплекс працює у режимі безпосереднього цифрового управління (БЦУ), сигнали, що використовуються для приведення у дію виконавчих механізмів, надходять безпосередньо від УОК, а відповідні регулятори взагалі виключаються із системи (або використаються як резерв).

Концепція БЦУ дозволяє замінити сукупність регуляторів безпосередньо обчислювальним комплексом. Замість того, щоб розраховувати уставки, необхідні для оптимальної роботи технологічного обладнання, як при супервізорному управлінні, у даному випадку УОК розраховує необхідні значення впливів, що управляють і передає відповідні сигнали безпосередньо на виконавчі механізми засобів регулювання. Це робиться для кожного контуру управління. Число контурів може становити від одиниць до декількох сотень, залежно від типу процесу й потужності УОК.

Для більш глибокого розуміння принципу дії АСУ ТП із УОК, що працює в режимі БЦУ, розглянемо рис. 8.12, на якому показаний один контур управління. Сигнал від датчика надходить у УОК, де після перетворення його у цифрову форму, обчислюється помилка (відхилення величини, що регулюється, від її заданого значення), яка використається у алгоритмі управління даним контуром. Результати обчислень, що надані у цифровій формі й, у свою чергу, перетворюються вихідним пристроєм у сигнал, який впливає на виконавчі механізми засобів регулювання. Таким чином, контур управління контролюється й регулюється безпосередньо УОК, який звертається до кожного з контурів по черзі, із частотою, що обумовлена характеристиками технологічного процесу. Уставки для контурів вводяться у

УОК оператором або зовнішнім ОК, що виконує розрахунки по оптимізації технологічного процесу.

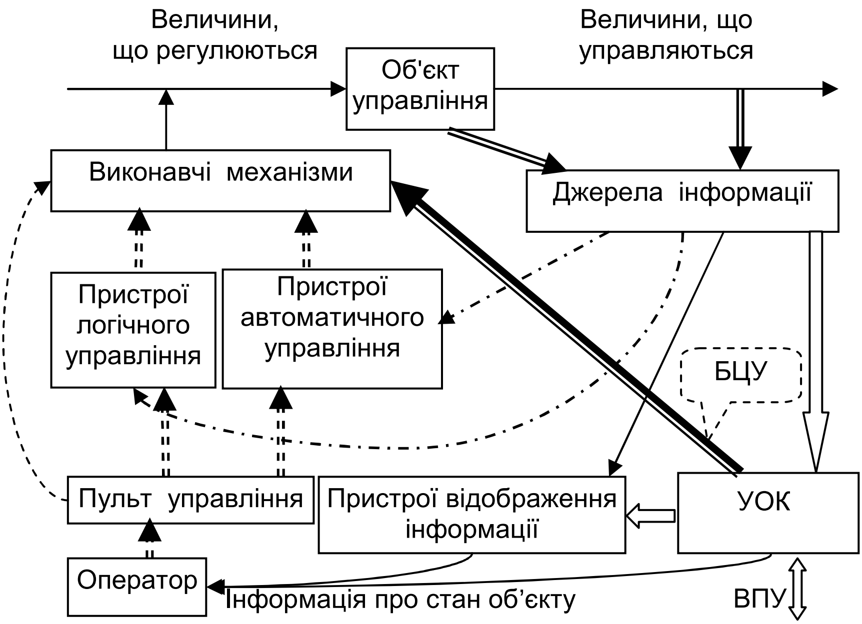


Рис. 8.12. АСУ ТП із обчислювальним комплексом, що виконує функції безпосереднього цифрового управління

При наявності системи БЦУ оператор повинен мати можливість змінювати уставки, контролювати деякі обрані змінні, варіювати діапазони припустимого відхилення змінних, що вимірюються, змінювати параметри налагодження й взагалі мати доступ до програми, що управляє. Для забезпечення всіх цих функцій необхідно мати повне й функціонально розвинене сполучення людини й машини (інтелектуальний інтерфейс).

Одна із головних переваг застосування АСУ із УОК, що функціонує у режимі БЦУ полягає у можливості зміни алгоритмів управління для любого з контурів простим внесенням змін у програмне забезпечення ЕОМ. Однак такі зміни повинні бути ретельно підготовлені, тому що нова програма повинна бути повністю перевірена до її реального використанням. Гнучкість такої системи у принципі нічим не обмежена. Деякі із існуючих АСУ є комбінацією систем БЦУ й супервізорного управління.

Пряме управління від УОК дозволяє реалізувати не тільки функції, що оптимізують технологічний процес, але й операції перемикання виконавчих механізмів основного й допоміжного устаткування.

Найбільш очевидний недолік систем із БЦУ проявляється при відмові УОК. Незважаючи на те, що надійність всіх засобів системи може бути винятково високою, відмови УОВК проте можливі, що у системі із БЦУ може привести до повної втрати управляємості об'єктом. Тому при створенні будь-якої системи БЦУ необхідно ретельно враховувати цю обставину.

Впливи, що управляють, які формуються АСУ безперервними технологічними процесами у автоматичному режимі, повинні забезпечувати підтримку заздалегідь заданих значень технологічних змінних або досягнення таких їхніх значень, які будуть обчислені як оптимальні. Якщо необхідні значення змінних задаються раніше, то АСУ виконує тільки функції регулювання. При створенні схем регулювання у АСУ ТП, у загальних випадках, застосовують два принципи регулювання: по відхиленню й по збурюванню.

Регулювання по відхиленню (рис. 8.13. а), тобто із використанням принципу зворотного зв'язку, у супервізорному режимі, як було відзначено раніше, виконується аналоговими регуляторами, які одержують значення уставки від УОК, а у режимі безпосереднього цифрового управління - програмним шляхом.

При регулюванні по збурюванню модель об'єкта зберігається у УОК, і по ній обчислюється вплив, що управляє, який компенсує збурювання, що діє на контрольований об'єкт. У сучасних АСУ складними технологічними процесами найчастіше застосовують комбіноване регулювання, що враховує відхилення й збурювання (рис. 8.13. б). Використання УОК дозволяє також зручно будувати програмним шляхом системи каскадного й багатозв’язного регулювання, що враховують взаємозв'язки між окремими ділянками об'єкту управління.

Інші різновиди АСУ ТП. Чітких границь між розглянутими різновидами АСУ провести не можна, кожна з них має окремі ознаки, що притаманні іншим АСУ. На різних рівнях управління ці риси зустрічаються у різних модифікаціях або зовсім відсутні. Однак при плануванні, проведенні й узагальненні розробок АСУ ТП бажано мати чітку й обґрунтовану класифікацію АСУ, тобто правила об’єднання АСУ ТП у підмножини, в межах яких вони близькі, схожі у тому або іншім відношенні.

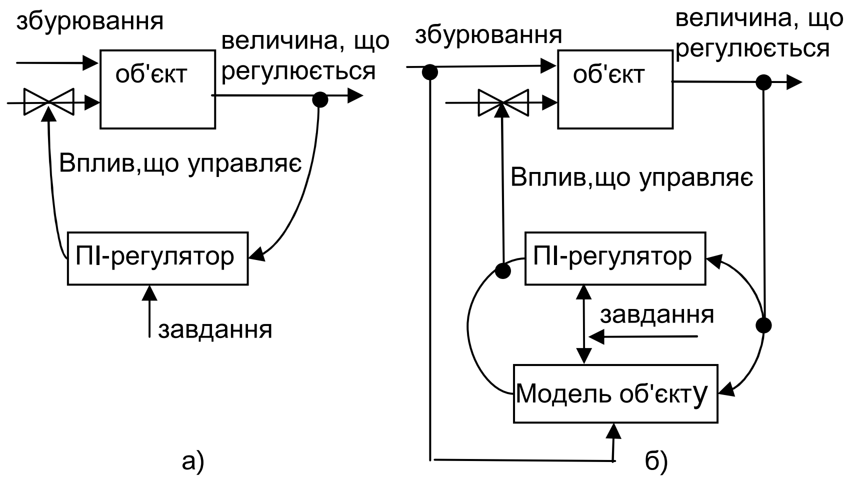


Рис. 8.13. Принципи регулювання: а) по відхиленню; б) комбіноване регулювання

На теперішній час загальноприйнятої класифікації АСУ ТП поки не існує. Проте, при виборі систем - аналогів АСУ ТП доцільно враховувати наступні умовні ознаки:

- рівень, що займає система у організаційно - виробничій ієрархії;
- характер протікання технологічного процесу, що управляється у часі;
- показник умовної "інформаційної потужності" об'єкта управління;
- ступінь функціональної розвиненості АСУ.

За рівнем, займаним у організаційно - виробничій ієрархії, розрізняють АСУ ТП нижнього (першого) і верхнього (другого) рівнів, а також, багаторівневі системи. Як було розглянуто, ієрархічна система управління - це централізована система, підсистема ухвалення рішень якої розподілена по декількох підлеглих рівнях, кожний з яких виконує певну частину функції ухвалення рішення.

До систем ***нижнього*** (першого) ***рівня*** відносять АСУ, що управляють агрегатами, установками, ділянками виробництва й не мають у своєму складі інших АСУ ТП. До АСУ ***другого рівня*** відносять автоматизовані системи, що управляють групами установок, цехами, виробництвами, у яких окремі ділянки (агрегати, установки) оснаще-

оснащені своїми системами управління, у тому числі, можливо, АСУ ТП першого рівня. Відзначимо, що АСУ із дворівневою організаційною структурою, яка поєднує у своєму складі АСУ ТП першого й другого рівнів і реалізує погоджене управління, як окремими технологічними установками, так і їхньою сукупністю (цехом, виробництвом), можна вважати інтегрованими АСУ ТП.

За характером протікання у часі технологічні процеси, що управляються можна розділити що найменше на три групи: дискретні, неперервні й неперервно-дискретні, що розрізняються за відносним часом перебування їх у перехідному і сталому режимах.

Дискретний характер має, наприклад, більшість процесів у машинобудуванні й приладобудуванні. Їхньою характерною ознакою є велика кількість виробів, інформація про які формується у більшості випадків вручну за допомогою документів (накладних, нарядів) або різних пристроїв ручного введення інформації, а також може вводитися від автоматичних датчиків. Впливи, що управляють звичайно передаються оперативно - виробничому персоналу у вигляді графіків випуску деталей, виробів, а також у вигляді команд і розпоряджень.

До виробництв, що мають ***неперервний характер*** технологічних процесів, відноситься вироблення й розподіл енергії, видобуток і транспортування нафти й газу, виробництво нафтохімічних продуктів і т.п. Для виробництв неперервного типу характерне одержання інформації про хід процесу головним чином за допомогою автоматичних датчиків із неперервним або дискретним вихідним сигналом. Управління технологічними процесами виконується шляхом впливу на різні види виконавчих механізмів (клапани, перемикачі й т.п.).

До процесів ***неперервно-дискретного типу*** відносять, наприклад, багато процесів у електронній, цементній і іншій галузях промисловості, у яких вони характеризуються наявністю циклів і сполученням особливостей дискретних й неперервних процесів.

Умовна інформаційна потужність технологічного об'єкту управління і його АСУ найчастіше характеризується числом технологічних змінних, що вимірюються або контролюються даною системою.

Ступінь функціональної розвиненості АСУ ТП можна характеризувати двома найбільш складними функціями, інформаційною й управляючою, що реалізовані у даній системі за допомогою засобів автоматичної обробки інформації.

8.4. Контрольні питання

1. Дайте визначення поняттю АСУ підприємством.
2. Дайте визначення поняттю АСУ технологічним процесом.
3. Дайте визначення поняттю технологічний об'єкт управління.
4. Дайте визначення поняттю інтегрована АСУ.
5. Дайте визначення поняттю технічне забезпечення.
6. Дайте визначення поняттю програмне забезпечення.
7. Дайте визначення поняттю інформаційне забезпечення.
8. Дайте визначення поняттю організаційне забезпечення.
9. Дайте визначення поняттю оперативний персонал.
10. Дайте визначення поняттю проста функція.
11. Дайте визначення поняттю складна функція.
12. Дайте визначення поняттю інформаційна функція.
13. Дайте визначення поняттю функція управління.
14. Наведіть структуру АСУ ТП, що функціонують без обчислювального комплексу.
15. Наведіть структуру АСУ ТП з обчислювальним комплексом, що виконує інформаційні функції.
16. Наведіть однорівневу централізовану структуру АСУ ТП.
17. Наведіть однорівневу децентралізовану структуру АСУ ТП.
18. Наведіть схему взаємодії основних компонентів АСУ ТП.
19. Наведіть багаторівневу ієрархічну структуру АСУ ТП.
20. Наведіть структуру АСУ ТП із обчислювальним комплексом, що виконує управляючі функції у режимі "порадника".
21. Наведіть структуру АСУ ТП із обчислювальним комплексом, що виконує функції центрального управляючого пристрою (супервізорне управління).
22. Наведіть структуру АСУ ТП із обчислювальним комплексом, що виконує функції безпосереднього (прямого) цифрового управління.
23. Наведіть схему регулювання по відхиленню.
24. Наведіть схему комбінованого регулювання.
25. Дайте визначення поняттю неперервний технологічний процес.
26. Дайте визначення поняттю дискретний технологічний процес.
27. Дайте визначення поняттю неперервно-дискретний технологічний процес.

Розділ 9

Технічні засоби АСУ ТП

Розглянуто універсальні та спеціалізовані засоби обробки, зберігання, візуалізації інформації, технічні засоби регулювання АСУ ТП. Також розглянуто питання надійності системи управління. Наведено контрольні питання.

9.1. Універсальні та спеціалізовані засоби обробки, зберігання та візуалізації інформації

До складу технічного забезпечення АСУ ТП входять усі технічні засоби, необхідні для реалізацій функцій АСУ ТП. У загальному випадку воно включає засоби отримання, введення, підготовки, обробки, зберігання (накопичення), реєстрації, виведення, відображення, використання, передачі інформації і засоби реалізації управляючих впливів.

Усі засоби обробки, зберігання та візуалізації інформації, що застосовуються в АСУ ТП, умовно можна поділити на такі групи:

- однокристальні мікроконтролери (МК);
- програмовані логічні контролери (ПЛК);
- індустріальні комп'ютери (ІК);
- універсальні комп'ютери.

Однокристальні мікроконтролери. Мікроконтролер являє собою спеціалізований мікропроцесорний пристрій, що управляє, який виконаний на одному кристалі і що програмується на мові Асемблер. Головною перевагою МК є те, що усі його функції реалізуються на одному кристалі, у результаті чого маємо високу відмовостійкість і дуже малу вартість.

Мікроконтролери застосовуються у багатьох галузях промисловості, у побутовій техніці, пристроях передачі та прийому інформації, особливо широко у пристроях і системах управління окремими технологічними агрегатами на нижньому рівні ієрархії АСУ ТП.

Програмовані логічні контролери. Програмований логічний контролер являє собою спеціалізований мікропроцесорний пристрій, що управляє, який пристосований до використання безпосередньо у виробничих умовах і що програмується за допомогою спрощених мов, які доступні непрофесійним користувачам.

Більшість сучасних ПЛК побудовані за блочно-модульним принципом, що дозволяє одержувати необхідну для конкретної задачі управління конфігурацію контролера шляхом зміни складу базового комплекту. До складу базового комплекту звичайно входять: уніфікована касета (каркас, або об'єднуюча панель), модуль процесора, модуль пам'яті, джерело живлення, модулі входів-виходів і спеціальні модулі, номенклатуру і кількість яких визначає користувач. У модулях введення-виведення ПЛК передбачені так звані гальванічні розв'язки внутрішніх (низьковольтних і слабкострумових) електричних кіл контролера від впливу зовнішніх (високовольтних і потужнострумових) електричних кіл об'єкта, що управляється. Зовнішні електричні кола є джерелами потужних електричних і електромагнітних перешкод, що порушують стійке функціонування контролера.

Основною особливістю функціонування ПЛК є те, що він працює циклічно. Одноразове опитування усіх входів ПЛК називається циклом сканування або ***робочим циклом***, тривалість якого $T_ц$, характеризує швидкодію ПЛК. Для нормального функціонування ПЛК повинна виконуватися умова

$$T_ц < T_{спр.їм},$$

де $T_{спр.їм}$ - час спрацьовування виконавчих механізмів.

Виконання цієї умови виключає можливість видачі з виходів ПЛК хибних команд управління при наявності на входах контролера короткочасних імпульсних перешкод.

Різноманітність моделей, функціональних можливостей і технічних характеристик ПЛК дозволяє розглядати їх як універсальні засоби, за допомогою яких можна вирішувати практично будь-які задачі промислової автоматизації на нижніх рівнях ієрархії інтегрованих систем автоматизації.

Програмовані логічні контролери застосовуються не тільки для управління окремими технологічним агрегатами, а і роботизованими технологічними модулями, автоматизованими технологічними лініями та їх дільницями. У цих випадках ПЛК вже виконують не тільки функції безпосереднього управління механізмами, але і функції програмної координації роботи окремих технологічних агрегатів та їх груп, а також ефективно вирішують задачі автоматичного діагностування елементів та пристроїв систем управління та об’єктів управління.

Індустріальні (промислові) комп'ютери. До індустріальних комп'ютерів у загальному випадку відносяться всі комп'ютери, що пристосовані до роботи безпосередньо у виробничих умовах. На даний час ІК об’єднують наступні комп'ютерні засоби промислової автоматизації:

- індустріальні персональні комп'ютери (IPC);
- індустріальні одноплатні комп'ютери (PCM);
- модульні індустріальні комп'ютери (MIC);
- пристрої розподіленого і віддаленого збору даних і управління;
- панельні персональні комп'ютери (PPC);
- промислові робочі станції (AWS).

До числа провідних фірм, що випускають усю гаму перерахованих IBM PC сумісних пристроїв промислової автоматизації, відносяться фірми: Advantech (Тайвань), Octagon systems (США), Siemens (Німеччина) та ін.

Індустріальні персональні комп'ютери орієнтовані у основному на роботу у якості персонального комп'ютера на верхніх рівнях ієрархії інтегрованих систем автоматизації і не призначені для безпосереднього управління технологічними агрегатами й іншими промисловими установками. Вони відносяться до класу так званих відмовостійких промислових комп'ютерів, що оснащені системою виявлення несправності шасі (корпусу) і видачі сигналу тривоги. До числа параметрів шасі, що контролюються відносяться: справність джерела живлення, вентиляторів і температура усередині шасі.

Головна відміна індустріальних персональних комп'ютерів від звичайних універсальних (офісних) полягає в тому, що комп'ютери типу IPC розроблені для використання у жорстких (промислових) умовах експлуатації (наявність пилу, значних механічних і кліматичних впливів, а також інтенсивного електромагнітного випромінювання).

Для захисту від пилу і підвищеної температури навколишнього повітря ІК забезпечуються високошвидкісними охолоджуючими вентиляторами і повітряними фільтрами, що забезпечують припливно-витяжну вентиляцію усередині блоку.

Для захисту від вібрації й ударів більшість ІК має фіксуючи пристрої для плат з ударостійкими відсіками для дискових накопичувачів. Захист від електромагнітного випромінювання забезпечується за рахунок того, що корпуса ІК виготовляються з високоякісної сталі зі спеціальним покриттям. Джерела живлення, що використовуються у ІК не чутливі до кидків і коливанням напруги у мережі.

Вбудовані у плати світлодіоди самодіагностики полегшують процес обслуговування і ремонту. Передбачено також самотестування плат при включенні електроживлення.

Комп'ютери типу IPC можуть комплектуватися змінними модулями дуже широкої номенклатури.

Індустріальні одноплатні комп'ютери призначені для використання на нижніх рівнях ієрархії інтегрованих систем автоматизації і являють собою універсальний комп'ютер, що виконаний на одній платі із можливістю її монтажу безпосередньо у об'єкті управління. Вони пристосовані до роботи у промислових умовах при температурі від 0 до 60^{o}С.

Комп'ютери тип PCM можуть застосовуватися і як вбудовані персональні комп'ютери і як вбудовані модулі, що управляють, але у загальному випадку потрібне їхнє дообладнання спеціальними вузлами (модулями) гальванічної розв'язки. Ці комп'ютери просто і ефективно стикуються з усіма іншими комп'ютерними засобами інтегрованих систем автоматизації.

Модульні індустріальні комп'ютери є досить універсальними засобами промислової автоматизації, де оптимальним чином сполучаються властивості ПЛК та ІК. Вони застосовуються:

- у розподілених системах збору даних і управління;
- у системах, що вбудовуються;
- у системах управління технологічними агрегатами і роботизованими комплексами;
- у системах автоматизації контрольних випробувань.

Конструктивне виконання модульних індустріальних комп'ютерів аналогічно конструкції комп'ютерів типу IPC і являє собою шасі касетного типу, що пристосовано до роботи безпосередньо у особливо жорстких промислових умовах, у режимі реального часу.

У комп'ютерах MIC забезпечується безпосередній доступ до інтерфейсів введення-виведення за рахунок використання з'єднань, що рознімаються, які розташовані на передніх панелях модулів. Усі модулі мають однакове конструктивне виконання, аналогічне конструкції модулів ПЛК.

Пристрої розподіленого і віддаленого збору даних і управління (наприклад пристрої ADAM-5000 і ADAM 4000 фірми Advantech) призначені для вирішення різних задач у багатьох предметних галузях.

Серія ADAM-5000/485 розроблена для великих розподілених систем, у яких не потрібно миттєвої реакції на події, що відбуваються на об'єкті, що управляється, а серія ADAM-5000/CAN орієнтована на застосування у системах "жорсткого" реального часу з малим часом реакції на зміну станів об'єкту, що управляється.

Серія ADAM-5000 являє собою апаратно-програмний комплекс, призначений для збору інформації про територіально розподілений об'єкт, що контролюється, первинної обробки даних шляхом фільтрації і нормалізації аналогових і дискретних сигналів, видачі впливів, що управляють на об'єкт, що контролюється, а також для обміну даними із центральною обчислювальною або управляючою системою, за допомогою об'єднання пристроїв у багатоточкову мережу.

Для обміну даними можуть використовуватися різні лінії зв'язку: симетрична кручена пара, волоконно-оптична лінія зв'язку і радіоканал. Вбудовані програмні засоби дозволяють настроювати діапазони вхідних сигналів і встановлювати умови видачі впливів, що управляють.

У пристроях серії ADAM-5000 широко використовуються апаратно-програмні засоби самодіагностики. Всі пристрої оснащені трьохрівневою гальванічною розв'язкою, у тому числі: по колах живлення і для модулів введення-виведення з напругою ізоляції 3000 В, а також для портів послідовного зв'язку з напругою ізоляції 2500 В.

Наявність гальванічної розв'язки дозволяє знизити вплив електромагнітних перешкод, усунути гальванічний зв'язок з електроустаткуванням об'єкту, що контролюється, а також запобігти несправностей, що можуть бути викликані випадковими кидками напруги живлення і перехідними процесами при комутації силового устаткування. Конструктивно вироби цієї серії являють собою закриті пилозахищені модулі.

Стандартна конфігурація пристрою ADAM-5000 містить у собі два компоненти: блок процесора і до 4-х модулів введення-виведення (до 64 каналів дискретного введення-виведення або до 32 каналів аналогового введення). Мається можливість гнучкого конфігурування системи у залежності від кількості і виду контрольованих параметрів, а також від розташування об'єктів, що контролюються.

У модулях ADAM-4000 реалізована функція віддаленого програмного настроювання типів і діапазонів прийнятих аналогових сигналів, що забезпечує можливість їхнього сполучення із різними датчиками і перетворювачами.

Пристрої ADAM-4000 за допомогою стандартного послідовного інтерфейсу RC-485 можуть об'єднуватися у багатоточкові мережі. Тип і діапазон вхідного сигналу встановлюється шляхом передачі на адресу модуля послідовним каналом зв'язку відповідної команди від пристрою управління. Таким чином, для рішення різних вимірювальних задач можливе застосування модулів одного і того ж типу.

Панельні персональні комп'ютери, що називаються також пультовими комп'ютерами, являють собою сукупність персонального комп'ютера і дисплея і можуть застосовуватися на різних рівнях ієрархії інтегрованих систем автоматизації в усіх випадках, де є необхідність оперативного відображення якості протікання технологічного процесу і активного використання інтерфейсу "людина – машина". Вони можуть вбудовуватися у лицьові панелі приладів і технологічного устаткування, а також використовуватися як інтелектуальні пульти управління промисловим устаткуванням.

Ці комп'ютери є повнофункціональними IBM-сумісними індустріальними комп'ютерами на основі високопродуктивних процесорів і пристосовані до роботи у жорстких промислових умовах. Вони оснащені автоматичною системою зниження потужності, що споживається.

Промислові робочі станції, представляють собою високопродуктивні пульти диспетчера (оператора) на основі PPC, у яких крім функцій панелі оператора реалізуються функції об'єднання і координації роботи окремих ПЛК, PC і IPC, універсальних комп'ютерів, контролерних та комп'ютерних мереж, реєстрації параметрів та інш.

У порівнянні з традиційними пультами оператора (диспетчера) AWS мають суттєво більш високі показники продуктивності, виконують значно більший набір функцій, а, головне являють собою більш гнучкий, легко адаптуємий програмований управляючий пристрій.

AWS забезпечують:

- автоматичну передачу даних до оператора і від оператора;
- збереження й аналіз отриманої інформації;
- надання користувачу інформації у зручному для нього форматі;
- управління технологічним процесом з ініціативи оператора;
- управління технологічним процесом у автоматичному режимі.

Застосування AWS дозволяє знизити витрати на проектування систем, що управляють, підвищити продуктивність праці оператора і підвищити якість управління технологічним процесом, а також підвищити гнучкість систем управління.

Протікання технологічного процесу за допомогою AWS відображається у графічному виді, що дозволяє миттєво оцінити загальний стан процесу. Автоматична подача сигналів про аварійну ситуацію сприяє забезпеченню безпечної експлуатації технологічних агрегатів і установок.

AWS використовується також для обробки інформації про хід технологічного процесу у реальному часі, що скорочує до мінімуму простої технологічного устаткування при відхиленнях процесу від норми.

Конструктивно AWS виконані у корпусах, що забезпечують повний захист від пилу, дощу, бризок, прямих потоків води, граду і зледеніння.

Вибір між ПЛК і ІК найчастіше залежить не тільки від характеристик об'єкту, що обслуговується та граничних умов рішення задачі. Вирішальну роль тут часто відіграють особисті прихильності і досвід користувача. Найбільш важливе значення при виборі засобів автоматизації завжди мають працездатність системи у реальному часі і її надійність.

Разом із тим, можливість роботи у реальному часі не є єдиним чинником при виборі між ПЛК і ІК. Такі критерії, як можливість підключення системи до інформаційної мережі, функції обробки даних і візуалізації, а також якість графічного інтерфейсу грають також важливу роль.

Для рішення задач безпосереднього управління промисловим устаткуванням у режимі "жорсткого" реального часу перевага віддається ПЛК.

У випадках, коли інформаційні функції починають суттєво превалювати над функціями управління технологічним процесом і потрібно використання всього спектра "інтелектуальних" можливостей, перевага повинна віддаватися персональним комп'ютерам.

Універсальні комп'ютери широко застосовуються на верхніх рівнях усіх інтегрованих системах автоматизації. За їх допомогою вирішується велика гамма завдань, до яких немає вимоги реалізації у режимі "жорсткого" реального часу.

До таких завдань відносяться: планування, облік та контроль виконання планових завдань, оптимізація розподілення ресурсів, необхідних для реалізації технологічних процесів, вибір і оптимізація режимів функціонування технологічних комплексів і їх дільниць, збирання, обробка, аналіз і реєстрація інформації про результати їх функціонування, видача рекомендацій з корегування планових завдань та режимів роботи технологічного устаткування.

9.2. Універсальні технічні засоби автоматизації

Необхідно відзначити, що АСУ реалізує деяку сукупність взаємопов'язаних процесів автоматичного регулювання технологічних параметрів, що здійснюються системами автоматичного регулювання, які будуються шляхом інтеграції, із дотриманням певних правил, із окремих уніфікованих блоків, модулів, приладів, пристроїв і виконавчих механізмів.

Блок – конструктивний збірний пристрій, що виконує одну або декілька функцій, який змонтований у одному корпусі.

Модуль – проста структурна одиниця - уніфікований вузол, що виконує одну елементарну операцію й входить до складу блоку або приладу. Модуль складається з елементів, якими є постійні або змінні резистори, ємкості, дроселі, транзистори, операційні підсилювачі, мікросхеми й інш. Елементи можуть бути не тільки електричними, але й пневматичними та гідравлічними.

Виконавчі механізми – пристрої для перетворення управляючого впливу у механічне переміщення регулювального органу, які розвивають достатні для цього зусилля та потужність.

У загальному випадку структуру системи автоматичного регулювання можна представити наступною схемою (рис. 9.1), яка складається з наступних елементів.

Задавач (ЗД), який призначений для введення необхідного значення величини, що регулюється сигналом $x_{зд}(t)$.

Давач (Д) призначений для перетворення технологічних величин у інформаційні сигнали $x_{д}(t)$.

Вимірювальний блок (ВБ) - сприймає та порівнює сигнали давача та задавача, формує сигнал неузгодження $x(t) = x_{зд}(t) - x_{д}(t)$ і перетворює його у більш зручну для обробки форму.

Формуючий блок (ФБ) призначений для формування необхідного закону регулювання або рівняння регулятора $y(t) = f_p(x(t))$. Пов'язує сигнали $x(t)$ і $y(t)$, що визначають переміщення регулювального органу.

Перемикач режиму роботи системи регулювання (ПРР) і модуль ручного (дистанційного) управління складають блок управління. При автоматичному режимі роботи вплив на об'єкт здійснюється відповідно до закону управління. При ручному управлінні – оператором, за допомогою кнопок дистанційного управління.

Виконавчий підсилювач (ВП) — підсилює по потужності командний сигнал від ФБ або сигнал ручного управління ($z_A(t)$ – сигнал автоматичного регулювання, $z_P(t)$ – сигнал ручного регулювання), не змінюючи його інформації.

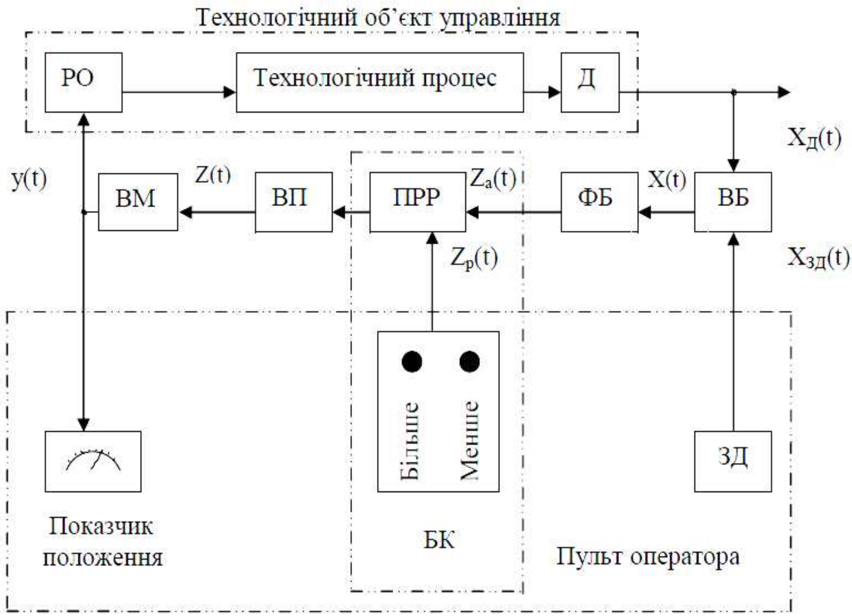


Рис. 9.1. Структурна схема системи автоматичного регулювання

Виконавчий механізм (ВМ) — реалізує командний сигнал ***z(t)***, виробляє регулювальний вплив на регулювальний орган (РО), тобто здійснює його переміщення.

На ***пульті оператора*** знаходяться органи оперативного управління.

Розглянемо особливості і класифікацію універсальних регуляторів стосовно їх основних конструктивних елементів і функцій, що виконуються.

За характером формування впливу, що управляє на регулювальний орган виділяють:

- регулятори прямої дії;
- регулятори непрямої дії.

Регулятори прямої дії виробляють вплив, що управляє за рахунок енергії, яку містить вимірювальний пристрій.

Регулятори непрямої дії використовують для своєї роботи і переміщення регулювального органу джерело додаткової енергії. За допомогою регуляторів непрямої дії вдається досягти більш якісного регулювання.

За характером підтримки заданого значення розрізняють:

- стабілізуючі регулятори;
- програмні регулятори;
- екстремальні регулятори, що підтримують значення регульованої величини на найбільшому або найменшому рівні;
- системи, що стежать, які забезпечують контроль деякого вхідного сигналу, який змінюється у часі по заздалегідь невідомій функції.

За родом енергії, що використовується розрізняють:

- гідравлічні регулятори;
- пневматичні регулятори;
- електричні регулятори;
- комбіновані регулятори.

Електричні регулятори – характеризуються:

- високою точністю та зручністю при відтворенні законів регулювання;
- можливістю будувати розподілені системи регулювання з великими відстанями між елементами та великою швидкістю передачі сигналів.

Електричні виконавчі пристрої мають нелінійні характеристики, що ускладнює точну реалізацію лінійних законів регулювання.

Пневматичні регулятори – характеризуються:

- пожежо та вибухобезпекою;
- гіршою реалізацією законів регулювання, їм властиві менша точність і діапазони встановлення параметрів регулювання;
- передача сигналів здійснюється повільніше;
- ускладнена взаємодія з електронними засобами, що стоять на більш високих рівнях ієрархії.

У той же час пневматичні виконавчі механізми швидкі й надійні, мають лінійні характеристики.

Гідравлічні регулятори –розміщуються безпосередньо у зоні об’єкта і мають:

- найменшу гнучкість, вузький діапазон застосування;
- реалізують лише найпростіші алгоритми й функції;
- характеризуються винятковою надійністю, виконавчі механізми найкращі (високі потужності, зусилля й швидкість при обмежених розмірах).

Вимоги щодо регульованих параметрів: універсальні регулятори повинні мати здатність стабілізувати різні технологічні змінні, тобто мають допускати підключення до них різних давачів (або давачі повинні мати стандартизований уніфікований вихід, або входи регулятора мають сприймати сигнали від різних давачів).

Найбільш розповсюджені наступні уніфіковані сигнали, що використовуються у розглянутих регуляторах:

- сигнали постійного струму: 0...5 мА; -5...0...+5 мА; 0...20 мА; 4...20 мА; 0...100 мА.
- сигнали напруги постійного струму: 0...10 В; -10...0...+10 В; 0...24 В; 0...48 В; 0...110 В; 0...220 В.
- сигнали напруги змінного струму: 0...1 В; 0...2 В; 0...36 В; 0...127 В; 0...220 В.
- стандартні пневматичні сигнали: 0,02...0,1 МПа (0,2...1,0 атм).
- стандартні гідравлічні сигнали: 0,01...0,1 МПа.

Універсальні регулятори мають реалізувати будь-який стандартний закон регулювання. При цьому до них пред'являються такі вимоги:

- точність відтворення закону регулювання;
- можливість змінювати параметри регулятора у широкому діапазоні.

В універсальних регуляторах повинна розвиватися необхідна потужність для переміщення регулювальних органів. Це досягається включенням до складу технічних засобів автоматизації проміжних підсилювачів.

Технічні засоби автоматизації повинні дозволяти будувати систему регулювання із розміщенням її елементів на різній відстані один від одного так, щоб задовольнити інтереси обслуговуючого персоналу. Виконавчі механізми і давачі розташовуються на об'єкті регулювання. Регулятори – на щиті регулювання. Органи оперативного управління, задавачі, покажчики положення розташовуються на пульті або щиті оператора.

Можливість дистанційного розподілу легше за все здійснюється у електричних системах регулювання. Менш пристосовані для цього пневматичні засоби. Практично неможливо побудувати розподілену систему за допомогою гідравлічних пристроїв.

Більшість універсальних регуляторів має потребу у джерелах додаткової енергії, тому вони повинні бути розраховані на стандартні види енергії.

Як було визначено раніше виконавчий механізм здійснює перетворення сигналу, що управляє у переміщення регулювального органу, призначення якого полягає у зміні витрати речовини або енергії у об'єкті управління.

Можна виділити три типи виконавчих механізмів:

- електричний;
- пневматичний;
- гідравлічний.

Пневматичний виконавчий механізм (рис. 9.2) призначений для побудови САР технологічних об'єктів управління із підвищеними вимогами пожежо та вибухобезпеки і являє собою пропорційну ланку. Переміщення вихідного штоку *1*, з'єднаного з регулювальним органом *2*, в одну сторону здійснюється силою, що створюється тиском $\boldsymbol{P_{BX}}$, в іншу – зусиллям від пружини *3*. Сигнал $\boldsymbol{P_{BX}}$ надходить у герметичну мембранну головку, у якій перебуває мембрана *4* із твердим центром. Знизу на мембрану давить пружина *3*. Статичні характеристики більшості механізмів близькі до лінійних, у широкому діапазоні переміщень.

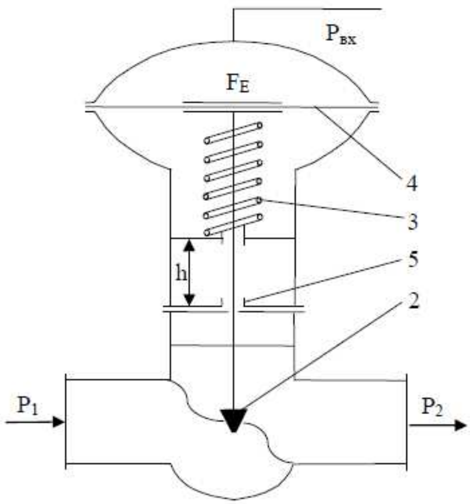


Рис. 9.2. Пневматичний мембранний виконавчий механізм

Пневматичні виконавчі механізми застосовуються при підвищеній загазованості атмосфери, на нафтохімічних виробництвах, газоперекачувальних станціях, хімічних підприємствах та ін. Носієм інформації у пневматичних засобах регулювання є тиск стисненого повітря. Робочий діапазон інформаційних сигналів і сигналів управління складає: 0,02...0,1 МПа (0,2...1,0 атм). Вхідна величина такого виконавчого механізму – тиск, вихідна – переміщення штоку і клапану.

Універсальна система елементів промислової пневмоавтоматики (УСЕППА) являє собою функціонально повний набір окремих елементів, кожний з яких може виконувати лише найпростішу функцію перетворення сигналу у загальній схемі всього пристрою. До складі УСЕППА входять пневматичні ємності, опори, дросельні суматори, підсилювачі, реле та ін.

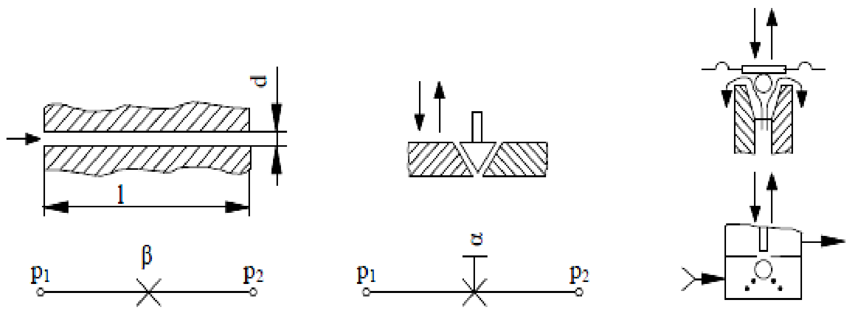
Пневмоємність — елемент, у якому накопичення потенціалу здійснюється за рахунок зміни тиску при сталому об’ємі. Залежно від функціонального призначення розрізняють постійні й змінні пневмоємності.

Пневматичні опори (дросель) — призначені для створення опору течії повітря (дроселювання потоку газу). Залежно від функціонального призначення розрізняють постійні, змінні пневматичні опори й пневматичні опори, що управляються (рис. 9.3).

Постійні (нерегульовані) пневматичні опори, що управляються (рис. 9.3.а) являють собою капіляр довжиною 20 мм з отвором діаметром 0,32 або 0,18 мм.

Змінні (регульовані) пневматичні опори (рис. 9.3.б) - величина яких може змінюватися у певних межах. У їх якості найчастіше використовується конструкція, що містить робочу пару «конус-конус».

Пневматичні опори, що управляються (рис. 9.3.в) – це опори, величина яких може змінюватися під дією якого-небудь параметра автоматично. Найпоширеніші опори - типу «сопло-заслінка», «сопло-кулька» й ін.



а) б) в)

Рис. 9.3. Пневматичні опори

Дросельні суматори реалізують операцію додавання із попереднім масштабуванням двох пневматичних сигналів (тисків), що надходять через постійний (ДП) або регульований (ДР) дроселі. Існує два типи дросельних суматорів, що відрізняються один від одного тим, що у одному з них регульований дросель забезпечений шкалою.

Підсилювач потужності (підсилювач-повторювач) призначений для повторення неперервного вхідного пневматичного сигналу й посилення його по потужності (витраті повітря). Підсилювачі застосовуються як вихідні елементи у більшості приладів, побудованих з елементів УСЕППА. Модифікації підсилювачів потужності розрізняються принциповою схемою й технологічними характеристиками.

***Реле перемикання*.** Використовується для підключення по черзі двох пневматичних каналів до третього, наприклад, для зміни режиму роботи регулятора з автоматичного на ручний та навпаки.

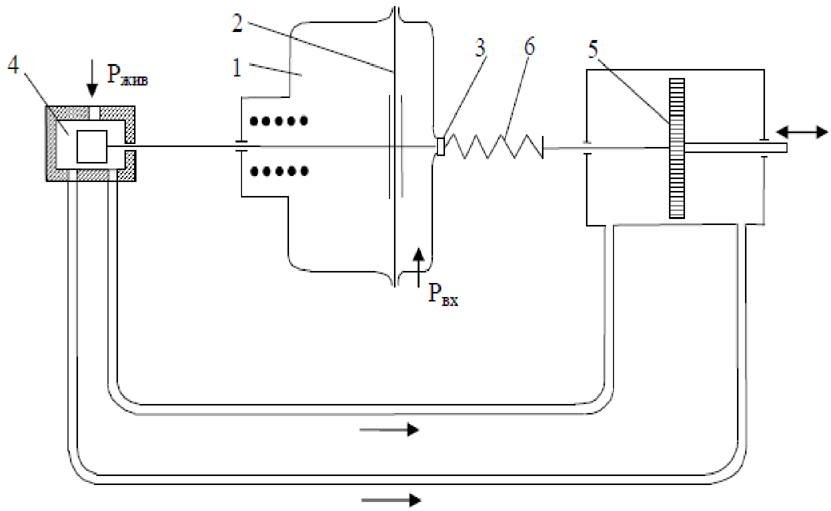


Рис. 9.4. Схема поршневого приводу, що стежить

Розглянемо ***поршневий привід, що стежить*** (рис. 9.4), який складається із виконавчого механізму *1*, що має більшу *2* і меншу *3* мембрани, золотника 4 із трьома клапанами, поршневого механізму *5* із пружиною від’ємного зворотного зв'язку *6.* При збільшенні $P_{вх}$ золотник зміщується ліворуч й тиск живлення $P_{жив}$ надходить до лівої порожнини циліндра *5*, переміщуючи поршень праворуч й збільшуючи натяг пружини зворотного зв'язку до встановлення нового стану рівноваги. Зона нечутливості у поршневому приводі, що стежить, дорівнює 0,5%, повний час переміщення 4с (при відсутності навантаження). У області частот $0 \le \omega \le 0,6$ рад/с поршневий привід, що стежить, можна розглядати як пропорційну ланку.

Гідравлічний виконавчий механізм (рис. 9.5) являє собою інтегральну ланку. Вхідна величина гідравлічного виконавчого механізму – різниця тисків, вихідна — переміщення поршня. Такі виконавчі механізми найчастіше застосовуються у системах, де потрібні висока надійність і великі зусилля і не потрібні складні закони регулювання. Гідравлічні засоби автоматичного регулювання одержали поширення на агрегатах у металургійній і хімічній промисловості, у енергетиці й машинобудуванні.

Тому що носієм інформації у них є зміна тиску рідини, є можливість використати давачі із механічним переміщенням на виході. Регульовані величини: тиск, розрідження, перепад тисків, витрата, рівень, кількість обертів й ін.

Сигналом завдання звичайно служить переміщення або зусилля. Застосовуються пружинні задавачі із регульованим ступенем стиску.

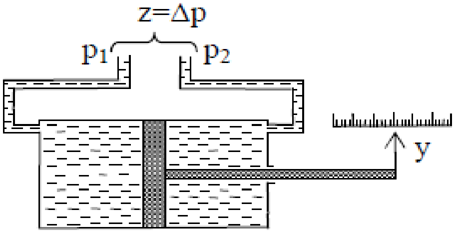


Рис. 9.5. Гідравлічний виконавчий механізм

Формування законів регулювання у гідравлічних засобах автоматичного регулювання здійснюється за рахунок застосування двох видів зворотного зв'язку: жорсткого зворотного зв'язку (система важелів) і гнучкого або ізодромного зворотного зв'язку, що використовується для формування прапорційно - інтегрального закону регулювання.

У якості органів настроювання у гідравлічних засобах автоматичного регулювання використовуються дроселі, недоліками яких є нелінійні характеристики й неширокий діапазон зміни регульованих параметрів.

Властивості гідравлічних підсилювачів - задовільні: можна одержати досить високий коефіцієнт підсилення (це необхідно для точності формування закону регулювання), при невеликих габаритах досягається значне посилення сигналу по потужності (важливо при управлінні потужними виконавчими механізмами).

Гідравлічні виконавчі механізми мають найбільший питомий момент і питому потужність серед всіх виконавчих механізмів (розвивають великі зусилля при достатній швидкодії).

У ролі гідравлічного носія сигналів звичайно використовують масло, іноді воду. Для роботи гідравлічної системи автоматичного регулювання необхідне спеціальне джерело енергії (маслонасосна станція), пристрої, що забезпечують задане значення живильного тиску (редуктор) і система фільтрів.

Швидкість передачі інформації визначається швидкістю передачі тиску у рідині. Ця обставина обмежує застосування гідравлічних засобів регулювання, тому що не дозволяє розташовувати елементи системи регулювання далеко один від одного. Довжина ліній зв'язку між елементами системи не повинна перевищувати: по горизонталі - 100...150 м, по вертикалі - 20...30 м.

На структурних схемах гідравлічних систем регулювання однією лінією показують механічні зв'язки, двома - гідравлічні, хвилястою лінією позначають електричні зв'язки.

Виконавчі механізми бувають двох видів: поршневі й кривошипно-шатунні.

Елемент порівняння найчастіше виконується на основі мембранної конструкції.

Електричний виконавчий механізм будується на базі асинхронного електродвигуна, що живиться змінним струмом. До основних переваг даного виконавчого механізму можна віднести надійність і зручність, використання електричної енергії для живлення. Недоліками електричного ВМ є його нелінійність і стала швидкість обертання електродвигуна.

Управління виконавчим механізмом сталої швидкості здійснюють імпульсним методом (електричними імпульсами змінного струму). Таким чином, електричний ВМ наближають до лінійної (інтегральної) ланки, використовуючи як корисну інформацію шпаруватість – відношення часу імпульсу до суми часу імпульсу і паузи:

$$r\ (t\ ) = \frac{t_i}{t_i + t_п}.$$

Змінюючи відстань між імпульсами і їхню тривалість, одержуємо різну середню швидкість виконавчого механізму. Реверсування двигуна здійснюється за рахунок зміни фаз.

Існують два способи формування сигналу управління виконавчим механізмом ***z(t)***, які призводять до зміни величини шпаруватості.

Широтно-імпульсна модуляція (ШІМ). У цьому випадку фіксується сума $t_i + t_п$, для зміни шпаруватості змінюють t_i.

Часово-імпульсна модуляція (ЧІМ). При цьому способі задаються постійною тривалістю імпульсу t_i, змінюючи $t_п$.

Виконавчі механізми релейно-імпульсних регуляторів повинні забезпечувати переміщення органу, що регулює із постійною швидкістю протягом дії імпульсів управління, які надходять від регулювального блоку або від оператора. У загальному випадку до складу електричних виконавчих механізмів входять наступні елементи (рис. 9.6):

- асинхронний двигун (АД);
- фрикційна муфта і електромагнітне гальмо (ЕМГ);
- планетарний або черв’ячний редуктор (Р);
- механізм ручного управління (МРУ);
- кінцеві та шляхові вимикачі (КВіШВ);
- давачі положення (ДП).

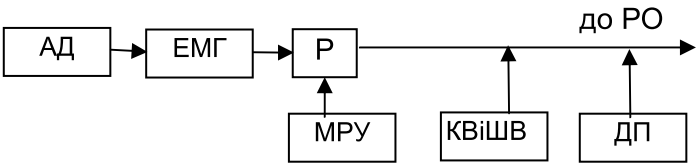


Рис. 9.6. Структура електричного ВМ

Електродвигун із редуктором служить для перетворення електричної енергії у механічну, достатню для подолання опору середовища з боку регулювального органа. Кінцеві вимикачі використовуються для відключення пускового пристрою при досягненні регулювальним органом крайніх положень, тобто виконують захисні функції, а шляхові - для сигналізації або обмеження діапазону переміщення регулювального органа у автоматичному режимі.

Давачі положення формують сигнал, пропорційний куту повороту вихідного вала виконавчого механізму. Цей сигнал подається на індикатор положення, що розташований на пульті оператора, а також може використовуватися у системі управління як сигнал зворотного зв'язку по положенню регулювального органа й ін.

Гальмові пристрої встановлюються у виконавчих механізмах для зменшення вибігу вихідного валу у процесі зупинки електродвигуна (після припинення дії імпульсів управління).

Ручний привід передбачається у виконавчому механізмі для забезпечення можливості переміщення регулювального органа при виході із ладу пускового пристрою.

Параметри сигналів зв'язку регулювальних блоків, пускових пристроїв і виконавчих механізмів стандартизовані. Тому той самий тип виконавчих пристроїв може використовуватися із різними системами регулювальних блоків.

Для управління потужними ВМ із трифазними двигунами використовується контактний пристрій - пускач електромагнітний реверсивний з гальмом (наприклад ПМРТ 69). Реверс електродвигуна ВМ забезпечується зміною послідовності двох фаз, які комутуються контактами $\boldsymbol{P_{б1}}$, $\boldsymbol{P_{б3}}$, $\boldsymbol{P_{м1}}$, $\boldsymbol{P_{м3}}$ (рис. 9.7).

Якщо надходить команда у напрямку «Більше», то через відповідні контакти на ВМ надходить напруга з послідовністю фаз ***A, B, C,*** що приводить до обертання електродвигуна по годинній стрілці.

Якщо надходить команда у напрямку «Менше», то фази ***A*** і ***C*** міняються місцями й напрямок обертання змінюється на протилежний.

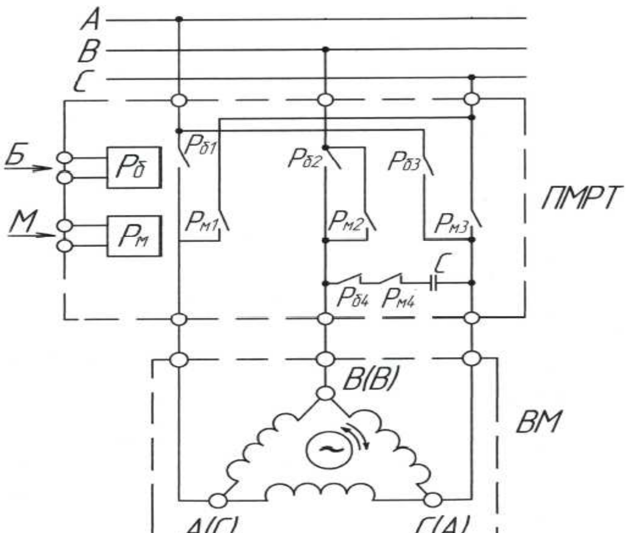


Рис. 9.7. Схема включення пускача магнітного ПМРТ 69

Коли імпульс управління закінчується, двигун повинен зупинитися якнайшвидше. Велика інерційність виконавчого механізму не дозволяє йому зупинитися миттєво. Конденсатор С і нормально-розімкнені контакти $\boldsymbol{P_{б4}}$ і $\boldsymbol{P_{м4}}$ виконують роль електромеханічного гальма. По цьому колу протікає струм, викликаний обертанням ротора по інерції, а це у свою чергу, створює гальмуючий момент, що перешкоджає вибігу ВМ.

У пускачах ПМРТ ефективне блокування здійснюється від спрацювання двох груп силових контактів.

9.3. Надійність АСУ ТП

Надійність системи управління прийнято визначати як здатність цієї системи виконувати задані функції, зберігаючи у часі значення встановлених експлуатаційних показників у заданих межах, що відповідають заданим режимам, умовам використання, а також і технічного обслуговування.

Відомо, що надійність є комплексною характеристикою, яка у залежності від призначення технологічного об'єкта управління може оцінюватися різними методами та показниками.

Надійність у загальному випадку може характеризувати безвідмовність, довговічність, ремонтопридатність і зберігаємість окремо чи визначати сполучення цих параметрів, як для системи у цілому, так і для її окремих частин.

АСУ ТП, як правило, представляє собою систему багаторазового використання і відноситься до класу так званих відновлюваних систем, що оцінюються такими показниками надійності:

1. ***Середній час безвідмовної роботи*** (напрацювання на відмову) ***T*** - середнє значення безвідмовної роботи системи, що обчислюється як середньостатистичне значення часу безвідмовної роботи

$$T = \frac{\sum_{i=1}^{n} t_i}{n},$$

де t_i - час безвідмовної роботи на i-му інтервалі часу;

n - кількість відмов.

2. ***Інтенсивність*** (частота) ***відмов*** λ - характеризує кількість відмов у одиницю часу

$$\lambda = \frac{1}{T}, \quad \text{або} \quad T = \frac{1}{\lambda}.$$

3. ***Середній час відновлення $T_в$*** - середньостатистичний час, затрачений на усунення однієї відмови

$$T_в = \frac{\sum_{к=1}^{n} t_к}{n},$$

де t_k - час, затрачений на усунення k – ї відмови;

n - кількість відмов.

4. ***Коефіцієнт готовності $К_Г$*** (ймовірність справної роботи) – характеризує ймовірність перебування системи у справному стані у довільний момент часу

$$К_Г = \frac{T}{T + T_в}.$$

Аналіз наведеної формули показує, що високе значення $К_Г$ може бути досягнуто як за рахунок підвищення напрацювання на відмову T, так і за рахунок зниження середнього часу відновлення $T_в$, яке у свою чергу може бути досягнуто наступними шляхами:

- застосування модульного принципу побудови управляючих пристроїв та систем (що дає змогу швидкої заміни несправного модуля, часто навіть без відключення електроживлення);
- застосування високоефективних апаратних та програмних засобів діагностування передвідмовних станів та відмов управляючих пристроїв.

Для об'єктивної оцінки показників надійності АСУ ТП використовується система поправочних коефіцієнтів.

Коефіцієнт важливості (цінності, ваги) $К_в$ - оцінює важливість функціонування i-ої функції у роботі системи, тобто показує, у якій мірі знижується ефективність функціонування АСУ при відмові окремих пристроїв і неможливості реалізації (повної або часткової) i - ої функції системи. Чисельно визначається різними методами, але у основному – методом експертних оцінок.

Коефіцієнт зайнятості $К_З$ - враховує середню частку часу використання i-ї функції у загальному часі роботи системи

$$К_з = \frac{t_i}{t_c},$$

де t_i - час використання ***i***-ої функції у загальному часі роботи системи;

t_c - тривалість роботи системи.

Коефіцієнт навантаження K_H **-** враховує співвідношення між фактичним значенням параметра та його номінальним значенням. Коефіцієнт навантаження для резисторів, транзисторів, діодів, мікросхем, ПЛІС і модулів на їхній основі обчислюється за формулою

$$K_н = \frac{P_ф}{P_{ном}},$$

де $P_ф$ - потужність, що фактично розсіюється;

$P_{ном}$ - номінальна потужність.

Для конденсаторів $K_н$ обчислюється за формулою

$$K_н = \frac{U_ф}{U_{ном}}.$$

АСУ ТП є складною, багатофункціональною системою, причому важливість (цінність) реалізації різних функцій (наприклад, - управляючих та інформаційних) різниться інколи у десятки разів. Тому при оцінці надійності АСУ ТП практично ніколи на наводять показники T, $T_в$ та $K_Г$ на усю систему, а виконують розрахунки і дають оцінки надійності за окремими функціями і навіть у межах однієї функції дають диференційовані оцінки її окремих складових.

Загальна методика розрахунку параметрів надійності АСУ ТП як складної системи передбачає такі етапи:

- розрахунок надійності функціональних вузлів;
- розрахунок надійності управляючих пристроїв;
- розрахунок надійності систем управління окремими технологічними агрегатами (установками);
- розрахунок надійності реалізації функцій АСУ ТП.

У зв'язку з тим, що у сучасних АСУ ТП, як правило, застосовуються уніфіковані та стандартні технічні засоби, параметри надійності яких відомі, то більшість розрахунків надійності стосується систем програмного управління окремими технологічними агрегатами та функцій АСУ ТП.

Як приклад розглянемо методику розрахунку параметрів надійності системи програмного управління (СПУ) на основі мікропроцесорного контролера.

Методика передбачає послідовне виконання таких етапів.

1 Визначення поправочних коефіцієнтів $K_в$, $K_з$ і $K_н$.

2 Обчислення сумарного значення інтенсивності відмов $\boldsymbol{\lambda}$ на основі довідкових даних про $\boldsymbol{\lambda}$- характеристики окремих елементів та пристроїв.

3 Обчислення $\boldsymbol{T}$.

4 Визначення або призначення $\boldsymbol{T_в}$.

5 Обчислення $\boldsymbol{K_Г}$.

До заходів, що можуть забезпечити подальше підвищення надійності АСУ ТП, відносяться :

- скорочення часу виявлення відмов за рахунок автоматичної програмної діагностики технічних засобів та програмного забезпечення;
- скорочення часу усунення відмов за рахунок наявності резервних (запасних) модулів, а також наявності універсального та спеціального інструмента для швидкого і якісного монтажу, демонтажу та ремонту модулів системи.

Одним із новітніх та перспективних підходів до підвищення надійності та безпеки функціонування АСУ ТП є застосування ***багатоверсійності*** реалізації технічних засобів та програмного забезпечення. У цьому випадку замість звичайного резервування каналів обробки інформації із застосуванням ідентичних технічних засобів їх реалізації використовують два або декілька паралельних каналів, котрі мають різну технічну реалізацію (наприклад застосовуються контролери або комп'ютери різних фірм). Прикладом багатоверсійності програмного забезпечення є дві або декілька програм, що розроблені різними програмістами для одного і того ж алгоритму.

9.4. Контрольні питання

1. Дайте визначення поняттю однокристальний мікроконтролер.
2. Дайте визначення поняттю програмований логічний контролер.
3. Дайте визначення поняттю індустріальний комп'ютер.
4. Дайте визначення поняттю універсальний комп'ютер.
5. Дайте визначення поняттю індустріальний персональний комп'ютер.
6. Дайте визначення поняттю індустріальний одноплатний комп'ютер.
7. Дайте визначення поняттю модульний індустріальний комп'ютер.
8. Дайте визначення поняттю промислові робочі станції.

9. Дайте визначення поняттю панельний персональний комп'ютер.

10. Дайте визначення поняттю пристрій розподіленого і віддаленого збору даних і управління.

11. Наведіть структурну схему системи автоматичного регулювання.

12. Дайте визначення поняттю регулятор прямої дії.

13. Дайте визначення поняттю регулятори непрямої дії.

14. Дайте визначення поняттю електричний виконавчий механізм.

15. Дайте визначення поняттю пневматичний виконавчий механізм.

16. Дайте визначення поняттю гідравлічний виконавчий механізм.

17. Дайте визначення поняттю надійність системи управління.

18. Дайте визначення поняттю середній час безвідмовної роботи.

19. Дайте визначення поняттю інтенсивність (частота) відмов.

20. Дайте визначення поняттю середній час відновлення.

21. Дайте визначення поняттю коефіцієнт готовності.

22. Дайте визначення поняттю коефіцієнт важливості.

23. Дайте визначення поняттю коефіцієнт зайнятості.

24. Дайте визначення поняттю коефіцієнт навантаження.

25. У чому полягає загальна методика розрахунку параметрів надійності АСУ ТП.

26. У чому полягає підвищення надійності та безпеки функціонування АСУ ТП за рахунок застосування багатоверсійності.

Література

1. Артюх С.Ф., Дуэль М.А., Шелепов И.Г. «Автоматизированные системы управления технологическими процессами в энергетике». Харьков: УИПА. - 2001. - 392 с.
2. Барсов В.І., Краснобаєв В.А., Деренько М.С., Авдєєв І.В. Основи побудови АСУ. Підручник. – Харків: УІПА, 2010. – 400 с.
3. Барсов В.І., Краснобаєв В.А., Фурман І.А., Маліновский М.Л. Системи обробки інформації та управління АСУ ТП на основі застосування кодів модулярної арифметики: Монографія. – Х.: МОН, УІПА, 2009. 145-с.
4. Бусленко Н.П., Калашников В.В., Коваленко И.Н. Лекции по теории сложных систем.- М.: Сов. радио. 1973. - 439 с.
5. Вентцель Е. С. Исследование операций. - М.: Сов. радио, 1972. - 514 с.
6. Глушков В.М. «Введение в АСУ», Киев: «Техніка», 1974. - 320 с.
7. Коршунов Ю.М. Математические основы кибернетики. – М.: Энергия, 1980. – 424 с.
8. Краснобаєв В. А., Харченко В. С., Ілюшко В. М. та ін. Основи надійності цифрових систем: Підручник. - Харків: Нац. Аерокосмічний ун-т, 2005. - 560 с.
9. Краснобаев В.А., Богомолов А.В., Черепков С.Т. Теория технических систем. – Харьков: ХГТУСХ, 2000.-78с.
10. Месарович М., Мако Д., Такахара И. Теория иерархических многоуровневых систем. - М.: Мир, 1973. - 344 с.
11. Общесистемное проектирование АСУ реального времени / Володин С.В., Макаров А.Н., Умрихин Ю.Д., Фараджев В.А./ Под ред. Шабалина В.А.. М.: Радио и связь, 1984. - 232 с.
12. Стефани Е.П. Основы построения АСУ ТП. М.: ЭАИ, 1974. – 318 с.
13. Справочник проектировщика АСУ ТП / Г. Л. Смилянский, Л. 3. Амлинский, В. Я. Баранов и др. / Под ред. Г. Л. Смилянского. - М.: Машиностроение, 1983. - 527 с.
14. Хетагуров Я.А., Древе Ю.Г. Проектирование информационно-вычислительных комплексов. - М.: ВШ, 1987. - 280 с.
15. Архангельский В.И., Богаенко И.Н., Рюмшин Н.А. Интегрированные АСУ в промышленности. — К.: НПК "Киевский институт автоматики", 1995. - 316 с.

Навчальне видання

Барсов Валерій Ігорович
Краснобаєв Віктор Анатолійович
Тиртишніков Олексій Іванович
Орищенко Сергій Анатолійович
Слюсарь Ігор Іванович

Математичні методи та технічні засоби АСУ

підручник

Відповідальний за випуск – Барсов В. І.
Комп’ютерна верстка та технічне редагування – Барсова З. В.

Підписано до друку 24.04.2013 р.
Формат 60х84 1/16 папір офсетний. Гарнітура Times New Roman.
Офсетний друк. Умов. друк. арк. .
Тираж 500 примірників.

Поліграфічний центр
Полтавського національного технічного університету
імені Юрія Кондратюка
36001, Полтава, першотравневий проспект, 24