

Міністерство освіти і науки України
Вінницький національний технічний університет

О. В. Бісікало

**ФОРМАЛЬНІ МЕТОДИ
ОБРАЗНОГО АНАЛІЗУ ТА СИНТЕЗУ
ПРИРОДНО-МОВНИХ КОНСТРУКЦІЙ**

Монографія

Вінниця
ВНТУ
2013

УДК 004.93:159.95

ББК 32.97

Б65

Рекомендовано до друку Вченою радою Вінницького національного технічного університету Міністерства освіти і науки України (протокол № 7 від 6 березня 2013 р.)

Рецензенти:

В. А. Широков, доктор технічних наук, професор

А. М. Петух, доктор технічних наук, професор

Бісікало, О. В.

Б65 Формальні методи образного аналізу та синтезу природно-мовних конструкцій : монографія / О. В. Бісікало. – Вінниця : ВНТУ, 2013. – 316 с.

ISBN 978-966-641-528-1

В монографії розглянуто теоретичні основи образного аналізу текстової інформації у відповідності до ідеї формалізації поняття образного сенсу через визначення його властивості та параметра. Запропоновано методи синтезу структурно-функціональних моделей системи образної обробки природно-мовного контенту. У межах функцій єдиної онтогенетичної системи з властивістю до самовдосконалення бази загальних знань образного сенсу отримано корисні моделі на рівні алгебраїчних операцій з мовними образами. Розроблено інформаційну технологію образного аналізу та синтезу природно-мовних конструкцій, що дозволило отримати нові розв'язки семантико-залежних задач.

УДК 004.93:159.95

ББК 32.97

ISBN 978-966-641-528-1

© О. Бісікало, 2013

ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ І ПОЗНАЧЕНЬ.....	6
ПЕРЕДМОВА	8
РОЗДІЛ 1 АНАЛІЗ МЕТОДІВ МОДЕЛЮВАННЯ ПРОЦЕСІВ ОБРОБКИ ПРИРОДНО-МОВНОЇ ІНФОРМАЦІЇ ТА ОБҐРУНТУВАННЯ СТРУКТУРНО-ФУНКЦІОНАЛЬНОГО ПІДХОДУ.....	11
1.1 Аналіз існуючих методів моделювання процесів обробки природно-мовної інформації та побудови баз знань	11
1.2 Мультидисциплінарні основи моделювання когнітивної сфери людини	19
1.3 Нейропсихологічні основи структурно-функціонального підходу.....	26
1.4 Аналіз лінгвістичних основ моделювання мовленнєвої діяльності	33
1.4.1 Загальнолінгвістичне представлення процесів мовлення	33
1.4.2 Формування мовних висловлювань.....	41
1.4.3 Розуміння мовних висловлювань	48
1.5 Вибір напрямку, мети та постановка завдань дослідження.....	57
РОЗДІЛ 2 ОСНОВИ ТЕОРІЇ ОБРАЗНОГО АНАЛІЗУ ТЕКСТОВОЇ ІНФОРМАЦІЇ ТА ФОРМАЛІЗАЦІЯ ПОНЯТТЯ ІНФОЛОГІЧНОЇ СИСТЕМИ	64
2.1 Концептуальні поняття інфологічної системи та онтогенетичного принципу її побудови	64
2.2 Концепція визначення образного сенсу природно-мовних конструкцій	72
2.3 Формалізація комутативної напівгрупи образних конструкцій на основі прикладної теорії першого порядку	80
2.4 Метод побудови нечіткого відношення образного сенсу	89
2.5 Дослідження простору образного сенсу з нечіткою мірою.....	95
2.6 Підхід до формалізації механізму функціонування інфологічної системи	102
2.7 Структурно-функціональна модель образної обробки природно-мовного контенту	107

РОЗДІЛ 3 МЕТОДИ МОДЕЛЮВАННЯ ПРОЦЕСІВ ОБРАЗНОЇ ОБРОБКИ ПРИРОДНО-МОВНОГО КОНТЕНТУ	120
3.1 Формалізація асоціативної мережі образів за допомогою графів	120
3.2 Інтерпретація простору образного сенсу на основі булеану	129
3.3 Поняття «піраміди сенсу»	134
3.4 Булева алгебра сенсу та формалізація концептів теорії.....	138
3.5 Операції, предикати та відношення БАС	141
3.6 Організація бази знань інфологічної системи.....	146
РОЗДІЛ 4 СИНТЕЗ ФУНКЦІЙ ОБРАЗНОГО ПОШУКУ ТА ГЕНЕРАЦІЇ ЗНАНЬ СИСТЕМИ ОБРОБКИ ПРИРОДНО- МОВНОГО КОНТЕНТУ	157
4.1 Функціональна модель системи обробки природно-мовного контенту на основі класифікації можливих типів образного пошуку.....	157
4.2 Розробка алгоритмів асоціативного та інсайтного пошуку.....	162
4.3 Алгоритм визначення ланцюга образів у зваженому графі.....	170
4.4 Алгоритм пошуку найвагомшого шляху в орієнтованому графі.....	179
4.5 Метод моделювання механізму оперативної пам'яті СОПМК....	182
4.5.1 Визначення основних понять та загальних принципів моделювання.....	184
4.5.2 Формалізація образного механізму оперативної пам'яті СОПМК.....	187
4.5.3 Алгебраїчна модель орієнтувального рефлексу	190
4.6 Метод самовдосконалення бази знань системи на основі моделювання складових парадигматичного устрою мови	196
РОЗДІЛ 5 РЕАЛІЗАЦІЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ ТА РОЗВ'ЯЗАННЯ СЕМАНТИКО-ЗАЛЕЖНИХ ЗАДАЧ	204
5.1 Побудова інформаційної технології на основі СОПМК.....	204
5.2 Формалізація результатів пізнавальної діяльності	212
5.3 Конструювання образу розв'язування проблемної ситуації	218

5.4 Генерація повідомлень щодо стану та потреб системи.....	223
5.5 Побудова відповіді на питання в процесі діалогу	229
5.6 Програмна реалізація інформаційної технології	232
5.7 Аналіз результатів впровадження інформаційної технології.....	242
ПІСЛЯМОВА.....	255
ЛІТЕРАТУРА	259
Додаток А Приклади застосування формальної теорії Th для російськомовних речень	284
Додаток Б Наскрізний тестовий приклад даних з тематики «WEB- технології: стандарти Semantic WEB»	287
Додаток В Підтримка базових функцій пошуку та генерації знань СОПМК: тексти програм і результати тестування	297
Додаток Д Діючий прототип інформаційної технології: результати тестування	309

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ І ПОЗНАЧЕНЬ

АВМ – асоціативно-вербальна мережа

АМО – асоціативна мережа образів

АО – ансамбль образів

БАС – булева алгебра сенсу

ВЕ – вектор емоцій

ЕК – електронний контент

ІС – інфологічна система

КМО – конструкція мовних образів

МО – мовний образ

ОК – образна конструкція

ПМК – природно-мовна конструкція

ППЗ – програмно-педагогічний засіб

СОПМК – система обробки природно-мовного контенту

Концепти мовних образів:

I (Image) – образ

N (Notion) – поняття

O (Object) – об'єкт

M (Method) – метод

Q – якість

ON – поняття об'єкта

QN – поняття якості

MN – поняття методу

OQ (Object-Quality) – якість об'єкта

MQ (Method-Quality) – якість методу

H – власне обставина (відповідь на питання як?)

T – обставина часу (відповідь на питання коли?)

L – обставина місця (відповідь на питання де?)

Focus – мовний образ у фокусі уваги

Типи асоціативних зв'язків:

P_x – асоціативний зв'язок синтагматичного або невідомого походження

P_e – внутрішньообразна асоціація

P_y – асоціація типу окреме–загальне (гіпо–гіперонімія)

P_y – асоціація типу частина–ціле (меронімія)

P_c – синонімічна асоціація

P_o – омонімічна асоціація (окремий випадок паронімії)

P_a – антонімічна асоціація

P_p – асоціація за типом «у риму»

Операції формальної теорії:

\oplus – операція об'єднання образних конструкцій «PLUS OK».

\backslash – зв'язок типу «головний–підлеглий» в асоціативній парі МО.

\times – зв'язок типу «підмет–присудок» в асоціативній парі МО.

Інші формальні операції:

\cup – об'єднання множин.

\cap – перетин множин.

\neg – доповнення множини.

\vee – логічна диз'юнкція.

\wedge – логічна кон'юнкція.

ПЕРЕДМОВА

Безпрецедентний розвиток новітніх інформаційних технологій та розбудова всесвітньої мережі Інтернет привели до появи загального інформаційного простору планетарного масштабу, що має ознаки абсолютно нового соціально-технічного утворення. Вражаючі можливості прозорого обігу інформації не тільки розмивають кордони між державами й скорочують відстань між людьми, але й відкривають шляхи до переходу людства у суспільство знань. Окреслюються нові перспективні підходи до підтримки вільного спілкування, швидкого доступу до інформації та неперервного навчання людини упродовж всього її життя. Щойно усталене поняття електронного контенту набуває помітного соціального й економічного значення.

Вибух технологічних досягнень на перетині тисячоліть став можливим завдяки основоположним фундаментальним працям Н. Вінера, К. Шеннона, Ф. де Соссюра, Б. де Куртене, А. Тьюринга, Д. Маккарті, А. М. Колмогорова, В. М. Глушкова, Д. А. Поспелова, Т. А. Гаврилової, Т. Бернерс-Лі тощо. Проте можливості сучасних технологічних засобів оброблення контенту не відповідають вимогам міжнародних стандартів до його семантичних властивостей. Всупереч очікуванням Інтернет-спільноти, які постійно зростають, проблемним залишається розв'язання класу семантико-залежних задач обробки природномовної інформації, що потребують залучення експертних знань для оцінки отриманих розв'язків. Моделі та методи найбільш відомих підходів – лінгвістичного, статистичного та логічного – відчутно програють інтелектуальним можливостям людини-експерта, що виразно відображається в семантико-залежних задачах пошуку, перекладу, анотування тощо. Використання надвеликих обчислювальних ресурсів для вилучення знань з неструктурованих масивів інформації не може забезпечити самовдосконалення загальної бази знань.

Значний вклад у розвиток моделей та методів автоматизованого управління інформаційними системами, аналізу природної мови та когнітивних процесів людини внесли вітчизняні дослідники О. В. Палагін, В. А. Широков, М. І. Шлезінгер, А. В. Анісімов, Ю. П. Шабанов-Кушнарєнко, С. Л. Кривий, М. Ф. Бондаренко, Ю. Р. Валькман, Н. В. Шаронова, В. П. Широчин та ін. Проте відчутна парадоксальність проблеми розв'язання семантико-залежних задач вимагає засто-

сування нових підходів і методів. Значна частина моделей лінгвістичної та когнітивної семантики будується на основі знань кваліфікованих експертів, але існуючі формальні засоби не досягають того рівня розуміння тексту, що демонструє звичайна дитина, яка ще не навчилася цей текст читати.

Недостатньо враховується в сучасних підходах до семантичного аналізу текстової інформації природний шлях отримання людиною більшості знань про навколишній світ за рахунок феноменів образного мислення. Відомі інтроспективні спроби логічного узагальнення інформаційних ознак психічних процесів людини, як правило, не відзначаються належним науковим обґрунтуванням. Теоретичного та експериментального дослідження потребує підтверджена результатами наук когнітивного напрямку гіпотеза про витoki прагматичних і семантичних аспектів сенсу з понять образу, асоціації, потреб і емоцій [31]. Проте відсутність чіткої та зрозумілої картини процесів, що відбуваються у головному мозку людини, змушує дослідників користуватися моделями непрямой аналогії, які забезпечують лише окремі результати для окремих задач, які прийнято відносити до інтелектуальної діяльності.

Структурно-функціональний підхід до образного аналізу та синтезу природно-мовних конструкцій, що пропонується, також є спробою отримати нову модель непрямой аналогії для розв'язання широкого кола семантико-залежних задач обробки природно-мовного контенту. Саме така мета дослідження змусила автора зайвий раз прискіпливо проаналізувати відповідні результати наук когнітивного блоку – філософії, фізіології вищої нервової діяльності, психології, нейропсихології та лінгвістики. Насамперед для того, щоб отримані формальні конструкції не суперечили загальноновизнаним на міждисциплінарному рівні положенням щодо образних підвалин розуміння сенсу природно-мовної інформації людиною.

Монографія, матеріал якої базується на роботах [1–75], складається з п'яти розділів. У першому з них окрім вже згаданого аналізу мультидисциплінарних основ моделювання когнітивної сфери людини для обґрунтування структурно-функціонального підходу проведено аналіз відомих методів моделювання процесів обробки природно-мовної інформації. Визначені проблема, мета і задачі дослідження, а також його провідна ідея – отримати розв'язки актуальних семантико-

залежних задач у вигляді комплексу функцій інформаційної системи обробки природно-мовного контенту.

У другому розділі розглянуто формалізацію образного аналізу текстової інформації на основі понять інфологічної системи та онтогенетичного принципу, змістовного поєднання образних та природно-мовних концептів. За допомогою формальної теорії визначено поняття образної конструкції, отримано її кількісні оцінки на основі нечіткої міри та одиниці образного сенсу з урахуванням понять ентропії та кількості інформації. Обґрунтовано методологію синтезу структурно-функціональних моделей інфологічної системи шляхом кібернетичної інтерпретації процесів інтелектуальної діяльності людини.

У третьому та четвертому розділах у розвиток запропонованого підходу розроблено методи моделювання процесів образної обробки природно-мовного контенту та синтезовано базові функції відповідної інформаційної системи, що забезпечують розв'язання семантико-залежних задач на основі образного пошуку та самовдосконалення бази знань. З цією метою застосовано математичний апарат теорій множин, графів, алгебраїчних систем і алгоритмів.

У п'ятому розділі розглянуто практичну реалізацію отриманої інформаційної технології та розв'язання актуальних семантико-залежних задач, визначених вимогами до інфологічної системи. За допомогою розробленого програмного забезпечення проведено експериментальні дослідження, що демонструють переваги отриманих та впроваджених розв'язків у порівнянні з існуючими аналогами.

Автор буде дуже вдячний за відгуки на цю книгу, які можна надсилати за адресою: кафедра АІВТ, ВНТУ, Хмельницьке шосе, 95, м. Вінниця, Україна, 21021 або на E-mail: obisikalo@gmail.com.

РОЗДІЛ 1

АНАЛІЗ МЕТОДІВ МОДЕЛЮВАННЯ ПРОЦЕСІВ ОБРОБКИ ПРИРОДНО-МОВНОЇ ІНФОРМАЦІЇ ТА ОБҐРУНТУВАННЯ СТРУКТУРНО-ФУНКЦІОНАЛЬНОГО ПІДХОДУ

На основі аналізу відомих методів моделювання процесів обробки природно-мовної інформації та загально визнаних результатів наук когнітивного напрямку визначаються проблема, мета і задачі дослідження. Обґрунтовуються характерні особливості структурно-функціонального підходу, що пропонується, та провідна ідея дослідження – отримати розв’язки актуальних семантико-залежних задач у вигляді комплексу функцій інформаційної системи обробки природно-мовного контенту (СОПМК).

1.1 Аналіз існуючих методів моделювання процесів обробки природно-мовної інформації та побудови баз знань

Актуальним завданням більшості задач штучного інтелекту та інтелектуальних інформаційних технологій є визначення семантичних характеристик процесів, що моделюються. Проте, не дивлячись на значні зусилля дослідників, семантичний аналіз як природно-мовних конструкцій (ПМК), так і інших продуктів інтелектуальної діяльності людини лишається найбільш проблемним і до цього часу. Корінь цієї надзвичайно складної проблеми насамперед пов’язаний з важкою формалізованістю предметної області внаслідок відсутності науково обґрунтованого спільного погляду на семантику в науках, які відносять до когнітивного напрямку досліджень [31].

Дослідження в галузі оброблення ПМК проводяться вітчизняними науковцями Українського мовно-інформаційного фонду НАН України [76], Інституту кібернетики та Інституту проблем математичних машин і систем, Міжнародного науково-навчального центру інформаційних технологій і систем, Національного університету ім. Т. Шевченка, Національного технічного університету України «Київський політехнічний інститут», Національного університету «Львівська політехніка», Харківського національного університету радіоелектроніки, Донецького інституту проблем штучного інтелекту тощо.

Серед провідних закордонних наукових закладів, що приділяють велику увагу проблемі розпізнавання природної мови, можна виділити й такі відомі, як Інститут проблем інформатики РАН, РосНДІ штучного інтелекту, Санкт-Петербурзький державний політехнічний університет, Ульяновський державний технічний університет, Massachusetts Institute of Technologies, Carnegie Mellon University, Stanford University, Princeton University, компанії IBM, Xerox, Sun, Microsoft, Google та багато інших.

Проблема розуміння сенсу природної мови тісно пов'язана з когнітивним напрямом штучного інтелекту та моделюванням пізнавальної діяльності людини. Побудова експертних систем, у тому числі інтелектуальних електронних підручників, базується на використанні уніфікованих структур бази знань у вигляді семантичних мереж, фреймів, онтологій, матриць семантичних ознак, реляційних моделей тощо. Використання сучасних баз знань все ще обмежено вузькими предметними областями, оскільки ускладнюється питаннями «вилучення знань» з експерта і відсутністю в системі загальних представлень «здорового глузду» [77]. В той же час мовленнєва діяльність людини природним чином забезпечує навчання та отримання нових знань на основі вербальної інформації у вигляді тексту [78].

Сучасний електронний контент, семантика якого лежить у площині проблеми дослідження, на даний час, безумовно, став мультимедійним. До складу контенту входять не тільки текст і окремі ПМК, але й, згідно зі стандартами W3C, графічні, аудіо-, відео- та анімаційні об'єкти різних типів (форматів). ПМК потрібно вважати провідною складовою контенту хоча б тому, що в метадані всіх об'єктів записуються слова, словосполучення або речення. Отже, інтегральним поняттям для аналізу семантики залишається образ – мовний, графічний, музичний, анімаційний тощо.

Починаючи з моменту свого зародження у штучному інтелекті історично склалися два підходи до моделювання інтелектуальних процесів людини. Так, на основі досягнень фізіології і генетики виник так званий висхідний або м'який напрям (нейронні мережі, методи розпізнавання образів і генетичні алгоритми) [79]. Дослідники низхідного або жорсткого напрямку в штучному інтелекті запропонували математичну інтерпретацію усвідомлених процесів абстрактного мислення (алгебра логіки та нечітка логіка, продукційні моделі, семантичні ме-

режі, фрейми, онтології, об'єктно-орієнтоване програмування, формальні граматики у лінгвістиці, евристичні моделі тощо) [80]. Дещо осторонь, як на наш погляд, розташовані в другому напрямі інфологічне моделювання та нормування відношень у базах даних, а також *Data mining* – пошук семантичної інформації у великих масивах даних. Окрім цього, порівняно молодим і ще до кінця несформованим напрямом досліджень можна вважати моделювання образного мислення людини [81].

Розглянемо існуючі підходи до моделювання семантики в окреслених вище напрямках та формальних методах. Якщо генетичні алгоритми частіше використовуються для розв'язання оптимізаційних задач, то розпізнавання образів за допомогою штучних нейронних систем найбільш тісно пов'язане з первинним семантичним аналізом навколишньої інформації. З математичної точки зору коло задач розпізнавання образів перетинається з задачами кластеризації, статистичної ідентифікації, побудови аксонометрій, факторним аналізом тощо [82]. У нейронних мережах традиційно моделюються принципи асоціативного сприйняття інформації людиною, проте, як правило, поняття «образ» застосовується у вузькому обмеженому сенсі, наприклад, візуальних об'єктів. Такий підхід не дозволяє поєднати основи низхідного та висхідного напрямів штучного інтелекту з точки зору когнітивної семантики.

Предметна область більшості розробок в галузі інженерії знань, як вже було зазначено раніше, має безпосереднє відношення до парадигматичної будови мови. До цього напрямку досліджень належать логічні та евристичні моделі представлення знань [81]. У групу логічних моделей включають числення предикатів, псевдофізичні та багатомодальні логіки, реляційні моделі, а евристичними вважаються мережеві, фреймові і сценарні моделі. До останньої групи відносять також такі популярні останнім часом напрями, як онтології та об'єктно-орієнтоване програмування (ООП) [80,83].

Терміном «семантична мережа», що має безпосереднє відношення до предмету дослідження, позначається множина представлень, побудованих, як правило, на графах. Такі представлення відрізняються, головним чином, іменами вузлів, зв'язків та висновками, які можна робити в таких структурах [84]. З самого початку семантичні мережі моделювали поняття, в основному, абстрактного характеру та відношен-

ня між ними. Цей формалізм є візуально виразним і дозволяє представити будь-який вид знань, проте збільшення типів відношень викликає необхідність програмування зростаючої у геометричній прогресії кількості фактів та правил. Тому в сучасних розробках мереж намагаються досягти стандартизації відношень шляхом вибору мінімального набору семантичних примітивів.

Сценарії [85] та фрейми [86] являють собою оригінальні евристичні моделі, що мають явні витoki з результатів когнітивної семантики. Обравши за основу психологічні дослідження процесів мислення та пам'яті людини, означені моделі дають непогані результати розуміння семантики у вузьких предметних областях, проте поняття сенсу в них має чітко виражений функціональний характер. Варто відмітити, що для всіх трьох розглянутих напрямів евристичного моделювання на сьогоднішній день немає прикладів успішного застосування у випадках розуміння довільного тексту природної мови.

Одним з найбільш відомих результатів мінімізації семантичної мережі можна вважати парадигму об'єктно-орієнтованого програмування (ООП), що передбачає використання таких концептів, як об'єкт, метод та якість. Дуже важливе практичне значення з точки зору автоматизації та візуалізації програмування має використання трьох головних принципів ООП: наслідування, інкапсуляції та поліморфізму [87]. Ці принципи можна вважати такими, що відображають особливості парадигматичного мислення людини. Але важливу також роль грає абстрактне поняття, що узагальнює інші концепти, а якість властива як для об'єктів, так і для методів. Об'єднавчим концептом для поняття, об'єкта, методу та якості (об'єкта і методу) можна вважати образ.

Виявлення семантики у великих масивах інформації від самого початку супроводжує розвиток популярних в програмних технологіях реляційних моделей даних. З цією метою в практику створення та експлуатації СУБД разом з реляційною алгеброю та реляційним численням було впроваджено теорію нормалізації відношень, що застосовується на етапі інфологічного моделювання предметної області бази даних [88]. Проте в цьому випадку знаходять формальне представлення лише знання щодо структури сутностей предметної області та відношень між ними [89]. Зрештою всі інші семантичні властивості, у

т. ч. статистичні характеристики наявних зв'язків потребують додаткового моделювання засобами реляційної алгебри [90, 91].

Проблема розуміння природних мов як одна з ключових проблем штучного інтелекту завжди була рушійною силою досліджень в області представлення знань [92, 93, 94]. На відміну від формалізації синтаксису та морфології природно-мовних конструкцій, моделювання семантики у комп'ютерній лінгвістиці має значно скромніші досягнення. Ця ситуація є наслідком того, що основними моделями цього напрямку досліджень є ті ж самі семантичні мережі у найбільш складному своєму варіанті [95], оскільки природна мова вільна від будь-яких обмежень, що можуть накладатися на штучні системи на зразок [96]. Окрім цього, існуючі системи орієнтуються на статистичний аналіз ключових слів тексту, а не асоціативних зв'язків між ними.

Розв'язання проблеми явно потребує нових підходів, що моделюють природний процес поступового накопичення знань про навколишній світ [97, 98]. На відміну від традиційних лінгвістичних методів сучасні інформаційні технології, які підтримуються такими велетнями комп'ютерної індустрії як Google, започаткували розвиток альтернативного статистичного підходу до аналізу ПМК [99]. Отримані результати розв'язку задач комп'ютерної лінгвістики, наприклад, у перекладі наближаються до існуючих досягнень, проте, ще далекі до рівня людини-експерта [77].

Принципово суб'єктивний характер будь-якого мовного висловлювання спирається на семантичні засоби лексики, що відображається у його змісті (значенні). З іншого боку, висловлювання завжди спрямоване на досягнення певної внутрішньої мети (мотиву, потреби) суб'єкта, тобто має прагматичну характеристику у вигляді сенсу цього акту мовлення. І хоча семантика та прагматика мовного висловлювання тісно пов'язані між собою, історична традиція математичної лінгвістики віддає безумовну перевагу першій складовій [100, 101]. Чи не вперше поняття змісту з'явилося в роботах Московської семантичної школи, де закладено витoki моделі «Зміст–текст». Під змістом висловлювання розумілося послідовне накопичування лінгвістичних відношень для ланцюга морфологія–синтаксис–семантика в умовах інтегрального опису словарної та граматичної компонент мови [102, 103]. Проте формальне визначення змісту і на сьогодні значно поступається природному практично в усіх семантико-залежних задачах

комп'ютерної лінгвістики. Це яскраво ілюструється парадоксом розуміння «смысла речи» неосвіченими людьми, які зовсім не тямлять у лінгвістиці, але чудово орієнтуються в бажаннях співрозмовника.

Розглянемо термінологію лінгвістичної семантики згідно з [104]. Основними поняттями, що характеризують предмет цієї дисципліни є зміст (*содержание*), значення (*значение*) та смисл/сенси (*смысл*). Оскільки предмет дослідження охоплює як семантику, так і прагматику природно-мовних виразів, то будемо притримуватися в роботі терміну сенси для перекладу російського *смысл*. Окрім цього, покажемо відмінності понять смисл та сенси.

Для предметної області «СЛОВО» [109] поняття смислу пов'язано з екстенсіоналом слова – множиною вказівників на ті сутності, які позначаються цим словом та асоціаціями і конотаціями – асоційованими у свідомості того, хто застосовує слово представленнями як фактичного, так оцінювального характеру. В асоціативному розумінні смисл отримує всю гаму визначень, що характеризують відчуття та емоції суб'єкта, викликані усвідомленням асоціативно пов'язаної з словом інформації [104].

На відміну від смислу поняття значення закріплює за певною одиницею мови відносно стабільний у часі стійкий зміст, інваріантний для всіх носіїв мови. Отже, значення X -а – це інформація, пов'язана з X -ом конвенційно, згідно з загальноприйнятими правилами використання X -а як засобу передачі інформації. В той же час смисл X -а для Y -а – це інформація, пов'язана з X -ом у свідомості Y -а в період часу T , коли Y застосовує або сприймає X як засіб передачі інформації [104]. Виходячи з цього та задач дослідження будемо вважати, що сенси X -а для Y -а – це інформація, пов'язана з X -ом у когнітивній сфері Y -а (не тільки у свідомості, але й на рівні рефлексів та підсвідомості) в період часу T , коли Y застосовує або сприймає X як засіб передачі інформації.

Прийняті визначення через екстралінгвістичні знання розширюють кордони предмета дослідження з «чистої» лінгвістики до близьких наук когнітивного напрямку – філософії, психології, фізіології, педагогіки та прагматики. Такий підхід добре ілюструється цитатою з [104]: «...численні напрямки сучасної семантики можна звести до двох концепцій, що протистоять одна одній, існування яких об'єктивно обумовлене двоїстістю подвійністю) предмета семантики. Ці дві концепції семантики можна умовно назвати вузькою і широ-

кою. Вузька концепція семантики робить своїм предметом значення одиниць мови і побудованих із них мовних виразів. В широкій концепції семантики її предметом, крім того, є і сенс мовних виразів в конкретних умовах їхнього вживання.»

Проблема дослідження також тісно пов'язана з теорією інформації та спорідненими з нею теоріями ймовірностей і нечітких множин. Цінність інформації за Харкевичем пропорційна збільшенню імовірності досягнення системою мети [105], проте цей показник відображає, насамперед, прагматичну сторону повідомлення і лише опосередковано враховує семантичну. Функція належності в нечіткій логіці акумулює суб'єктивне бачення експерта відносно належності певного елемента до деякого класу елементів [106, 107], але будується на описовому принципі і не супроводжується конструктивним алгоритмом свого визначення поза участю експертів. Окрім цього, відомі нечіткі міри [108] характеризують саме ступінь належності, можливості, достатності, необхідності тощо, але при цьому лише опосередковано відображають той зростаючий обсяг знань, який накопичується у людини і, зрештою, дає підстави зробити той чи інший логічний висновок. З іншого боку, існуючі технології побудови словників на основі лексикографічної системи вводять поняття псевдотопології, за рахунок якого встановлюється ступінь близькості слів та/або лексем, а також обґрунтовують необхідність побудови специфічної формальної метамови для опису лексико-семантичних відношень природно-мовних конструкцій [109].

Формальною структурою для оцінки змісту речення на природній мові може служити штучна нейронна мережа, при цьому нейрон застосовується як математична основа кванта змісту, що масштабується для символу, частини слова, слова, словосполучення, речення, абзацу, всього тексту [110, 111]. Такий підхід передбачає, що зміст тексту або іншого мовного поняття закладено у предикат, що приймає значення «істина» чи «хибність», проте окрім верифікації повідомлень залишаються відкритими питання синтезу, наприклад, генерація адекватного відгуку системи на зовнішні впливи.

В роботі [112] запропоновано використати для вивчення природної мови таку комп'ютерну модель «дитини», в яку не закладено лінгвістичні знання. Опосередковано ця ідея використовується і в статистичному підході до аналізу ПМК [99], хоча варто зазначити, що схожі думки висловлювалися А. Тьюрингом ще на початку становлення

штучного інтелекту як наукового напрямку [226]. Оpubліковані підходи до теоретико-модельної формалізації процесів мислення та рефлексії, де враховуються поняття свідомості, представлення, образу, онтології, але лишається поза розглядом онтогенез психічних функцій людини [113]. В літературі висувуються ідеї моделювати комп'ютерну особистість шляхом використання спеціальних математичних методів для генерації нових знань та реалізації цілеспрямованої поведінки [114, 115]. Проте багато авторів відзначають, що існуючі підходи до моделювання операцій образного мислення мають поки що скоріше концептуальний, ніж практичний характер [116].

Не можна не відмітити, що навіть за наявного теоретично недостатнього рівня формалізації процесів семантичного аналізу вже існують приклади масштабних та амбітних технологічних проєктів. Актуальність цього напрямку досліджень підтверджують не тільки академічні розробки [117], але й фінансово та ресурсно вражаючі зусилля лідерів комп'ютерної індустрії IBM, ABBYY, GOOGLE [118, 119, 120].

Отже, загальний аналіз існуючих підходів до моделювання процесів обробки природно-мовної інформації та побудови баз знань дозволяє зробити такі висновки:

1. Не дивлячись на надзвичайну актуальність семантичного аналізу, в технологіях обробки тексту відсутні концептуальні підходи та відповідні теорії, за головну мету яких покладено визначення сенсу як прагматичної першооснови змісту.
2. Відсутнє поняття одиниці сенсу, не розроблено критерії для кількісної оцінки сенсу в ПМК та інших продуктах когнітивної діяльності людини.
3. Моделювання образного мислення на відміну від логічних та евристичних моделей має поки що, переважно, концептуальний характер і не доведено до практичної реалізації у комп'ютерній лінгвістиці.
4. Ідею Тьюринга про онтогенетичний характер моделювання продуктів інтелектуальної діяльності практично не розвинуто у низхідному (жорсткому) напрямку штучного інтелекту.
5. Відомі формальні методи не забезпечують прийняттого рівня розв'язку широкого кола актуальних семантико-залежних задач, а тому потребують розробки нові комплексні підходи та методи, що враховують недоліки існуючих.

1.2 Мультидисциплінарні основи моделювання когнітивної сфери людини

Дослідження когнітивної сфери людини мають глибокі історичні корені та пов'язані з видатними діячами науки. Ще давньогрецький вчений Платон, зацікавлений феноменологічним характером процесів мислення, розглядав асоціацію як основу людської пам'яті і поведінки. Його співвітчизника Аристотеля вважають автором найпершої класифікації асоціацій за схожістю (червоне – пурпурне або кішка – тигр), за часовою послідовністю (день – ніч) та за контрастом (велике – маленьке або холодне – гаряче), яка стала основою для наступних численних класифікацій і типологій [121].

Протягом віків окремими філософськими системами у поняття асоціації вкладався різний зміст. Це поняття по-різному трактували та вивчали відомі дослідники. Так, Р. Декарт використовував асоціацію для розуміння процесів оволодіння власними пристрастями, а Б. Спіноза пояснював асоціаціями певні особливості «руху думок». Т. Гоббс, у свою чергу, створив першу систему механістичної психології, де елементи свідомості (відчуття та представлення) взаємодіють на основі механістичних за своєю суттю зв'язків за суміжністю відчуттів у просторі та часі [122].

Власне сам термін «асоціація» було введено в науковий світ в XVII сторіччі Дж. Локком з метою пояснення причин виникнення забобонів та «помилкових ідей». Якщо Дж. Берклі пояснював за допомогою асоціацій сприйняття простору, то у Д. Юма асоціація стає наріжним каменем всієї пізнавальної сфери психіки. Юм також розрізняв три типи переходів від однієї ідеї до іншої – подібність, суміжність у просторі (шия – голова) та часі, причинно-наслідковий зв'язок.

XVIII та перша половина XIX століть вважаються періодом класичного асоціанізму, який супроводжувався низкою гучних імен Д. Гартлі, Д. Прістлі, Джеймса Мілля і Т. Брауна. З другої половини XIX сторіччя предмет дослідження цікавив Джона Стюарта Мілля, А. Бена, Г. Спенсера, Г. Еббінгауза, представників англійської школи В. Вундта, Т. Цігена, Г. Мюллера [123, 124]. Протягом перших десятиліть XX сторіччя гостра криза асоціанізму привела до його остаточного зникнення як цілісного напрямку психології та асиміляції його ідей в різних галузях психологічної теорії та практики. Важкий час

кризи сприяв появі нових гучних імен і напрямів в психології – Зигмунда Фрейда, Карла Г. Юнга, Альфреда Адлера (психоаналіз), німецької школи гештальтпсихології, американської школи біхевіоризму, Ж. Піаже та Л. С. Виготського, російської школи фізіології І. П. Павлова та багатьох інших [122, 125–128].

Сучасний період розвитку вільного асоціативного експерименту пов'язаний з виникненням психолінгвістики та зміщенням фокуса досліджень на проблеми мовленнєвої діяльності людини і формування її мовної здатності. Визначення асоціативного значення слова ввів у сучасну наукову парадигму Дж. Діз [129], а Ч. Осгуд застосовував метод шкал для вимірювання смислових полів і показав, що афективне значення слова представляє собою координати в багатовимірному просторі [130]. За образним виразом Х. Хермана, «значення не є асоціація, але знання асоціації» [131]. Радянська школа психолінгвістики була створена зусиллями вітчизняних дослідників А. А. Залевської, І. Г. Овчиннікової, Н. О. Золотової, Ю. М. Караулова [132–136].

До наук когнітивного спрямування належить філософія. Практично всі філософи погоджуються з тим, що психічне – якісно своєрідна форма буття [137–142]. Життєздатність суб'єкта забезпечується властивостями живої системи використовувати носіїв інформації, що входять до складу суб'єкта, про стани зовнішнього і внутрішнього середовища при регуляції поведінки. Тоді психіку можна вважати саме такою властивістю суб'єкта. Аналогічної точки зору дотримувався російський математик О. Ляпунов, який стверджував, що життя – це стійкий стан речовини, який використовує для вироблення реакцій самозбереження інформацію, що кодується складом елементів цієї речовини [140].

Відомо, що людина відрізняється від тварин наявністю мови як системи кодів, що позначають предмети і такі відносини, за допомогою яких предмети вводяться у певні системи і категорії. Ця система кодів веде до формування абстрактного логічного мислення і формування «категоріальної» свідомості [137]. Здатність людини переходити за межі наочного, безпосереднього досвіду до відвернутого, раціонального досвіду є фундаментальною особливістю її свідомості. Інтроективно або самоспостереженням суб'єкт принципово може визначити такі психологічні феномени, як: відчуття, сприйняття образу,

представлення образу, поняття образу, думку, висновок, мислення [143–148].

Важливу роль у соціогуманітарних дослідженнях відіграє структурно-функціональний підхід або метод структурно-функціонального аналізу складних соціальних явищ та систем [149]. Релевантний для соціогуманітарного знання метод задає принципи системно-організованої структури цілісності, в якій кожний елемент має визначене функціональне значення або функцію у межах цілісності. Поняття функції є центральним для структурно-функціонального аналізу та розглядається у двох аспектах: 1) як «призначення» («роль») «одного» з елементів деякої цілісності по відношенню до «другого» або до цілісності (системи) в цілому; 2) як така залежність у межах даної цілісності, при якій зміни «одного» виявляються похідними (функцією) від змін «другого». Проте до цього часу структурно-функціональний підхід не використовувався для побудови систем обробки ПМК [150].

Об'єктивні результати експериментальних досліджень когнітивної сфери отримано фізіологією вищої нервової діяльності – становлення цієї науки пов'язують з іменами видатних російських вчених І. М. Сеченова («Рефлекси головного мозку», 1863 р.) та І. П. Павлова. Науковою школою Нобелівського лауреата (1903 р.) Павлова доведено факт об'єктивного існування нервової моделі стимулу як фізіологічної основи психічної діяльності людини [151, 152].

У нервовій моделі відбувається неперервне прогнозування зовнішніх роздратувань. Якщо прогноз збігається з реальним роздратуванням, реакція відсутня; якщо вони не збігаються, то виникає сигнал неузгодження, який вже викликає орієнтувальну реакцію. Таким чином, нервова модель стимулу не з'являється миттєво. Характерно, що колатералі як периферична нервова система при формуванні нервової моделі стимулу виконують функцію зворотного зв'язку, тобто дозволяють в результаті орієнтувальних реакцій прискорити процес її розпізнавання [153].

Експериментально доведено, що у людини зберігається підвищена збудливість протягом 15–30 секунд після закінчення попереднього стимулу. Проте слід подразника, що впливав на організм у певний час, зберігається на тривалий термін у вигляді нервової моделі, яка в загальному випадку забезпечує вибіркоче гальмування орієнтувальних реакцій та вибіркоче посилення умовних реакцій. По суті свого форму-

вання нервова модель стимулу фіксує деяку послідовність сигналів у часі [154] та створює єдине ціле або гештальт, що складається з окремих ознак та параметрів, але є більшою сутністю ніж сума його складових.

Досить характерним є той факт, що І. П. Павлов дуже виразно висловлював своє переконання щодо розуміння асоціації як універсального фундаментального психологічного явища. «...Тимчасовий нервовий зв'язок, – писав Павлов, – є найбільш універсальним фізіологічним явищем у тваринному світі та в нас самих. А разом з тим воно ж і психічне – те, що психологи називають асоціацією, чи буде це утворенням з'єднань з усіляких дій, вражень або з літер, слів і думок» [152]. За Павловим, в основі асоціацій за одночасністю лежать умовні нервові зв'язки, генералізація (узагальнення) цих самих зв'язків відповідає асоціаціям за схожістю, а основні розумові операції зводяться до аналізу і синтезу умовних зв'язків.

Сила збудження, швидкість розгортання реакцій і суб'єктивні оцінки пов'язані з лабільністю та силою нервової системи, а також з її індивідуальними особливостями [155, 156]. З погляду організації (архітектури) вищої нервової системи розрізняють принцип ієрархічної підпорядкованості різних структур мозку, який приводить до зменшення ступенів свободи будь-якої системи і здійснює управління. Експериментально підтверджено можна також вважати наявність принципу багаторівневої взаємодії вертикально організованих (підкірково–кіркових) і горизонтально організованих (кірково–кіркових) шляхів проведення збудження. Внаслідок цього матеріальним субстратом вищих психічних функцій стають не окремі кіркові ділянки або центри, а функціональна система кіркових зон, що спільно працюють [157, 158].

З метою визначення ролі нервової моделі стимулу в мовній діяльності розглянемо роль слова, як основоположного поняття цієї діяльності в процесі онтогенезу. Можна вважати, що онтогенетичне значення слова послідовно змінюється відповідно до таких ролей [159]:

1. Інтегратор 1-го порядку (сигнал).
2. Інтегратор 2-го порядку (сигнал сигналів) – якщо одному слову відповідає не менше 15-ти умовних зв'язків.
3. Складне парадигматичне поняття, наприклад, «іграшка», «квіти», «тварини».

Припустимо, що слово, з фізіологічної точки зору, є одним з елементів обмеженої множини природних ознак нервової моделі стимулу. Досить переконливим доказом такого припущення виглядають етапи генези регулюючої функції мовлення, що наводяться в [160]:

1. Природний орієнтувальний рефлекс, що визначається силою або новизною подразника.

2. Назва предмета матір'ю і вказівний жест на нього. Проте спочатку нестійка спеціальна орієнтувальна реакція змінюється під впливом сильніших подразників – на неї до 3,5 років сильно діє інертність. Отже, спонукаюча функція мовлення виникає раніше, ніж її гальмівна функція. Функція спонукання дитини вже безумовно є інтерпсихологічною.

3. Егоцентричне мовлення, яке описує ситуацію і планує можливий вихід з неї згідно з Жаном Піаже. Наступним етапом для дитини є шепітне мовлення, а далі – внутрішнє мовлення. Всі ці види мовлення можна віднести до його інтелектуальної і регулюючої функції.

4. Внутрішнє мовлення з часом набуває ознак предикативного, особливо у випадках, коли тема повідомлення відома, а слід з'ясувати рему. Як приклад тут можна навести короткий план лекції. На рівні фізіології інтелектуальна і регулююча функції мови забезпечуються лобовими долями головного мозку з 2-м та 3-м шарами асоціативних нейронів [161].

Несуперечність образної основи вербальної діяльності людини підтверджує патопсихологічний аналіз, що розкриває якісну відмінність у мисленні дорослого з ураженнями мозку і здорової дитини [162]. Саме для дитячого мозку характерна висока пластичність, яка дозволяє оптимізувати процес адаптації індивідуума до умов навколишнього середовища. З іншого боку, в результаті хвороб мозку перш за все уражаються молоді утворення кори головного мозку, тобто такі, що розвинулися найпізніше у філогенезі. Завдяки могутньому орієнтувальному рефлексу, спілкуванню з тими, хто її оточує, дитина швидко засвоює різноманітні знання про предмети, накопичує і синтезує їх.

Відсутністю загальноновизнаних психологічних теорій пояснюється слабкий прогрес у моделюванні особистості програми або комп'ютерного «Я» (Его) [31]. Типові з представлених в Інтернеті комп'ютерні моделі особистості [163, 164], як правило, спрямовані на психологічні дослідження мотиваційної структури у вигляді ієрархіч-

ної організації елементів, що відповідають за обробку окремих стимулів того або іншого роду. Відмітимо, що Алан Тюрінг вважав перспективним шлях до еволюційної програмної імітації розуму дитини [165], який аж ніяк не може залишити осторонь комп'ютерне «Я». З іншого боку, практично всі відомі сучасні психологічні напрями та школи мають власні теорії особистості, а в окремі з них, наприклад, теорію поля Курта Левіна або п'ятифакторну модель особистості, закладено математичну нотацію [166].

Основною підставою для створення психоаналітичної теорії особистості, що розділяє структури психічної сфери на рівні несвідомого, підсвідомого та свідомого були клінічні спостереження [167]. Проте більш достовірно валідність теоретичних положень в психології можна довести за допомогою або статистичних методів, або експериментальним шляхом. Кожен з трьох окреслених методів має свої переваги та недоліки, проте більшість відомих теоретичних концепцій намагалися перевірити експериментально, що не суперечить ідеї моделювання комп'ютерного Его [168]. Найбільш відомими теоріями особистості в психології, окрім класичного психоаналізу, можна вважати:

- Его-психологію та пов'язані з нею напрями (Ерік Еріксон, Еріх Фромм та Карен Хорні);
- диспозиційний напрямок (Гордон Олпорт, Реймонд Кеттел та Ганс Айзенк) [169];
- біхевіористичний напрямок (Б. Ф. Скіннер);
- соціально-когнітивний напрямок (Альберт Бандура та Джуліан Роттер);
- когнітивний напрямок (Джордж Келлі);
- гуманістичний напрямок (Абрахам Маслоу) [170];
- феноменологічний напрямок (Карл Роджерс) [171];
- марксистські та радянські теорії особистості (Ж. Політцер, А. Н. Леонтьєв, С. Л. Рубінштейн, Л. С. Виготський, Б. Г. Ананьєв, Д. Н. Узнадзе) [172, 173].

Здійснено спроби розглядати велику кількість теорій особистості, що існують сьогодні, з позицій єдиної класифікації найбільш значущих початкових передумов [174–176]. Якщо залишити поза розглядом спірні питання розвитку та наповнення Я-концепції, то при різноманітні поглядів на психологічні чи фізіологічні конструкти, що склада-

ють ядро та периферію особистості, практично всі автори теорій звертають увагу на розкриття понять потреб, мотивів та емоцій людини у взаємодії з вже відомими з психоаналізу поняттями інстинктів [177].

Визначальна роль потреб виразно проявляється у сфері емоційних реакцій вищих тварин і людини, оскільки емоції по суті відображають величину і вірогідність задоволення цієї потреби в даний момент, а особливості потреби додають емоції якісно специфічного забарвлення [178]. Що ж до ділення емоцій на біологічні нижчі та соціально детерміновані вищі (так звані інтелектуальні, етичні або естетичні відчуття), то за подібною класифікацією ховається далеко не завжди виразно усвідомлювана класифікація потреб, надзвичайно різноманітних за їх складністю, походженням, індивідуальною, видовою і соціальною цінністю [179].

Отже, аналіз мультидисциплінарних результатів дослідження когнітивної сфери людини засвідчує ретроспективну актуальність та важливість цієї тематики. Отримано достовірні результати щодо окремих аспектів мислення людини, які, однак, поки не сформуvalи єдину теорію інтелекту. Проте доцільно об'єднати відомі результати як об'єктивні передумови нового структурно-функціонального підходу до побудови СОПМК. Пропонується будувати підхід на основі таких положень:

1. В онтогенезі з синтагматичних зв'язків найбільш «енергетичні» фокусуються у своєрідний центр або ядро, навколо якого розподіляється решта зв'язків. У нормі образні (синтагматичні) зв'язки витісняються смисловими (парадигматичними).

2. Визначення свідомості походить від здатності суб'єкта виділяти себе з навколишнього світу, з можливостями самостереження і рефлексії. Психіка пов'язана з властивостями живої системи використовувати інформаційні носії про стани зовнішнього і внутрішнього середовища при регуляції її поведінки.

3. У нервовій моделі стимулу закладено природний механізм неперервного прогнозування зовнішніх роздратувань за допомогою орієнтувальних реакцій. Ефект ідентифікації, що виникає після одноразового пред'явлення еталонної нервової моделі стимулу, володіє ризами стійкої образної пам'яті. На думку І. П. Павлова, тимчасовий нервовий зв'язок відповідає психологічному поняттю асоціації.

4. Використання понять інстинкту, психічної енергії, рівнів не-свідомого, підсвідомого та свідомого необхідне для розділення структури психічної сфери, емоцій, потреб та мотивів діяльності. По суті емоції відображають величину і вірогідність задоволення цієї потреби у даний момент, а особливості потреби додають емоції якісно специфічного забарвлення.

1.3 Нейропсихологічні основи структурно-функціонального підходу

Нейропсихологія – це розділ психології, в якому вивчаються мозкові механізми вищих психічних функцій на матеріалі спостережень за локальними ураженнями головного мозку. За своїм предметом дослідження нейропсихологія дуже близька до патопсихології [162]. Основи вітчизняної нейропсихології були закладені А. Р. Лурія, який розробив теорію системної динамічної локалізації психічних процесів. Як провідний метод дослідження у нейропсихології використовується синдромний аналіз порушень вищих психічних функцій, в якому дається якісна кваліфікація можливих порушень та розгорнутий діагноз ураження мозку [180].

Вищі психічні функції – введене Л. С. Виготським теоретичне поняття, яке позначає складні психічні процеси, соціальні за своїм формуванням, опосередковані і, за рахунок цього, довільні. На думку Виготського, психічні явища можуть бути «натуральними», детермінованими переважно генетичним чинником, і «культурними», надбудованими над першими. Власне вищими психічними функціями прийнято вважати «культурні» психічні явища, які цілком формуються під впливом соціальних дій [181]. Основною ознакою вищих психічних функцій є їх опосередкованість певними «психологічними знаряддями» або знаками, що виникли в результаті тривалого суспільно-історичного розвитку людства і до яких насамперед належить мова.

Спочатку вища психічна функція реалізується як форма взаємодії між людьми, між дорослим і дитиною, як інтерпсихологічний процес, і лише потім – як внутрішній, інтрапсихологічний. При цьому зовнішні засоби, що опосередковують цю взаємодію, переходять у внутрішні, тобто відбувається їх інтеріорізація. Якщо на перших етапах свого формування вища психічна функція є розгорнутою формою наочної діяльності, що спирається на відносно прості сенсорні і моторні про-

цеси, то надалі складові функції ущільнюються, стаючи автоматизованими розумовими діями [182]. Онтогенез вищих психічних функцій демонструє єдину образну основу всіх типів мислення, в тому числі вербального та логічного [183].

Локалізація вищих психічних функцій – важливий для нейропсихології теоретичний концепт, що виражає собою більш-менш однозначну відповідність певних ділянок мозку психічним процесам. Проблема зв'язку психіки і мозку ставилася ще в працях Гіппократа та Галена. У ХХ столітті дослідження цієї проблеми продовжили праці представників вузького локалізаціонізму та антилокалізаціонізму. Адекватне рішення було запропоноване в теорії системної динамічної локалізації вищих психічних функцій, розробленої Л. С. Виготським і А. Р. Лурія. Згідно з цією теорією, психічні функції людини необхідно розглядати як складні системні утворення, що формуються прижиттєво, є довільними і опосередкованими, перш за все, мовою. Фізіологічною основою психічних функцій виступають функціональні системи, що включають достатньо динамічні та взаємозамінні аферентні та еферентні ланки. Кожна ланка такої системи пов'язана з певною мозковою структурою, при цьому у різних функціональних систем існують загальні ланки, які можуть брати участь в реалізації багатьох психічних функцій. При нейропсихологічних ураженнях цих ланок виникають характерні поєднання порушень психічних функцій, які об'єднуються в певні нейропсихологічні синдроми [184].

Основні результати нейропсихології переконливо доводять, що інтелектуальна діяльність – це організоване розв'язання задач, яке спирається на логічну програму з 6-ти взаємопов'язаних операцій. Не становить винятку і мовлення, що так само, як і будь-яка інша інтелектуальна діяльність, керується центральною нервовою системою та характеризується [185]:

- метою або питанням, безпосередня відповідь на яке неможлива;
- орієнтуванням в умовах задачі, що супроводжують мету: аналіз отриманої інформації, виділення істотних даних та їх зіставлення одне з одним;
- виникненням загальної схеми (стратегії) рішення;
- виділенням тих операцій (тактики), які з найбільшою вірогідністю приведуть до досягнення поставленої мети;

- виконанням окремих дій, які завжди носять вибіркового характеру і закінчуються певною відповіддю;

- перевіркою отриманого результату або звіренням отриманих результатів з початковими умовами задачі.

Розглядаючи питання про управління розумовими процесами з фізіологічної точки зору, використовують таку класифікацію [186]:

- відстрочене відтворення і безпосереднє відтворення;
- довільний рівень управління психічними функціями і автоматизований рівень управління психічними процесами.

Фізіологічною основою безпосереднього відтворення і автоматизованого рівня управління психічними процесами виступає безумовний рефлекс. Умовний рефлекс – ключове поняття, згідно з Павловим, на стику наук про вищу нервову діяльність людини – можна розглядати на фізіологічному рівні, як тимчасовий нервовий зв'язок, а на психологічному рівні – як асоціацію. За ієрархією рівнів розгляду проблеми можна виділити такі: фізіологічний, психологічний, соціальний. Термінологічними синонімами двох основних понять на цих рівнях вважаються:

- (першосигнальне, несвідоме, базальне) – для безумовних рефлексів;

- (другосигнальне, свідоме, надбудоване) – для умовних рефлексів побудованих на їх основі вищих психічних функцій [187].

Отже, організація мислення як діяльності, у тому числі і мовленнєвої (на основі вербальних ознак), також характеризується властивостями нервової моделі стимулу і безпосередньо пов'язаним з нею поняттям умовного рефлексу.

З точки зору системного підходу, окремі ділянки головного мозку відіграють визначальну роль при регуляції розумових процесів. О. С. Адріановим була розроблена концепція структурно-системної організації мозку, що складається з проєкційних і оперативних (асоціативних), інтеграційно-пускових і лімбіко-ретикулярних систем [187]. Якщо проєкційні системи забезпечують аналіз і переробку відповідної за модальністю інформації, то асоціативні системи – аналіз і синтез різномодальних збуджень. Головні функції цих двох груп систем можуть безпосередньо визначатися на основі базового поняття нервової моделі стимулу. Інтеграційно-пускові системи, з одного боку, ініцію-

ють синтез збуджень різної модальності у відповідь на біологічно значущі сигнали і мотиваційні впливи, а з іншого боку, остаточно трансформують всі аферентні сигнали та впливи в якісно нову форму діяльності (вихід збудження на периферію). Лімбіко-ретикулярні системи, у свою чергу, забезпечують вплив енергетичних, мотиваційних і емоційно-вегетативних нервових процесів на організацію інтелектуальної діяльності в цілому.

Інші автори пропонують не роз'єднувати проекційні і оперативні системи та розглядати мозок людини у вигляді структури з 3-х основних блоків [189]. Це дозволяє визначити основні макрофункції головного мозку відповідно до таблиці 1.1.

Таблиця 1.1 – Макрофункції основних блоків мозку

№ з/п	Блоки мозку	Основні функції
1	Стовбурові відділи мозку, апарати медіальної кори або лімбічної області	Підтримка робочого тону кори, забезпечення тривалих селективних форм діяльності
2	Вторинні відділи лівої скроневої та тім'яно-потиличної області	Отримання, переробка і зберігання екстероцептивної інформації
3	Лобові відділи мозку	Програмування, регуляція та контроль діяльності мислення

Перший блок за своїми функціями можна вважати енергетичним, про що свідчать дослідження Н. А. Аладжалової та Є. Н. Соколова. У ньому об'єднані система регуляції ритмічної активності, що діє повільно [190] та швидкодійна активаційна система, що забезпечує фізіологію різного роду орієнтувальних реакцій [191]. Блок має безпосередній вплив на такі процеси, як увага, пам'ять, емоційні стани, мотивація тощо [192]. До складу другого блока входять основні системи аналізаторів (зорова, слухова, шкірно-кінестетична). Кора ділиться на первинні, вторинні і третинні (неядерні) поля. Найважливішу роль у роботі блока виконує зона *ТРО* – перекриття скроневої (*temporalis*), тім'яної (*parietalis*) і потиличної (*occipitalis*) кори [151].

Третій блок забезпечує програмування, регуляцію і контроль інтелектуальної діяльності. Слід зазначити, що 24 % поверхні великих півкуль займає саме кора лобових часток мозку. У складі блока виділяють моторні, премоторні та префронтальні відділи.

Важливу інформацію про специфічну роль різних відділів і областей головного мозку в складному процесі їх взаємодії з метою забезпечення вищих психічних функцій дозволяє отримати аналіз випадків травматичних уражень [131]. Основний клінічний матеріал та статистичний масив даних нейропсихології у цьому напрямі було зібрано в шпиталях СРСР під час Великої вітчизняної війни 1941–1945 рр. Отримані результати такого аналізу можна вважати історично унікальними та достовірними.

Потилично-тім'яні відділи мозку є кірковим кінцем зорово-просторового аналізатора [153]. При ураженні розпадається можливість об'єднувати елементи інформації, що послідовно надходить, в симультанно (просторово) осяжні схеми, орієнтуватися в просторі, об'єднувати сукупність окремих вражень в єдине ціле, розрізняти праву і ліву сторони, розумітися в географічній карті, засвоювати значення складних логіко-граматичних структур [130].

При ураженнях лобових часток мозку порушується [160]:

- предикативна будова мови (можливо, внутрішньої мови);
- процес попереднього аналізу умови задачі і випадання або нестійкість орієнтувальної основи дій;
- можливість скласти план розв'язання задачі;
- здійснення систематичних операцій, ієрархічно підлеглих програмі розв'язання задачі;
- усвідомлення того, як протікають поточні операції, порушується здатність самостійно коректувати допущені помилки.

Розрізняють такі варіанти лобового синдрому [156]:

1. Ураження базальних або медіо-базальних відділів лобової області приводять до значних змін емоційної сфери і поведінки хворого.
2. Ураження задньо-лобових відділів (насамперед всієї провідної півкулі) викликають деавтоматизацію процесів мислення і підвищення інертності раніше виниклих стереотипів.
3. Масивні ураження передніх відділів лобової області приводять до порушення програмування і контролю інтелектуальних процесів.

Більшу частину досліджень у нейропсихології було проведено на вербальному матеріалі, а меншу – на символічно-графічному, що підкреслює пріоритет мовлення серед усіх видів інтелектуальної діяльності.

ті [193] при спільності єдиного асоціативно-образного механізму мислення. Тому моделювання саме лінгвістичних процесів може бути вирішальним для досягнення суттєвого прогресу в створенні інтелектуальних технологій [194]. Зведено інформацію на основі всієї сукупності нейропсихологічних даних щодо мозкової організації для забезпечення лінгвістичних процесів [160], представлено в таблиці 1.2.

Таблиця 1.2 – Мозкова організація для забезпечення лінгвістичних процесів

Лінгвістичні процеси і функції		Відділи і області мозку
Мотиваційна основа		Стовбурові відділи мозку
Розуміння підтексту		Лобові частки, що займають близько 1/3 півкуль головного мозку
Програмування і контроль діяльності регуляції		
Синтагматична будова мови	Предикативна внутрішня мова	Задньолобні або лобово-скроневі відділи лівої півкулі
	Синтагматичні конструкції	Нижні відділи передмоторної області лівої півкулі (руйнування «кінестетичних мелодій», «телеграфний стиль»); найгрубіший випадок (зона Брока: задня третина нижньої лобової звивини лівої півкулі) – моторна афазія, неможливість вимовити окремі слова
Артикуляція звуків		Постцентральна область мовної зони
Парадигматична будова мови	Фонемна система мови, розуміння мовного вислову	Відділи скроневої області лівої півкулі (відчуження значення окремих слів, проте загальний смисл речення розуміється) – смисло-розрізнявальне значення фонемних ознак мови (зона Верніке)
	Правильний відбір лексичних і семантичних зв'язків, розуміння комунікації відносин	Третинні, тім'яно-скронево-потиличні відділи лівої півкулі

Дані таблиці свідчать про те, що енергетичний блок забезпечує лише мотиваційну основу лінгвістичних процесів, проєкційні та оперативні (асоціативні) системи підтримують парадигматичну будову мови, а інтеграційно-пускові системи відповідають за найбільшу кількість провідних функцій мовлення як інтелектуальної діяльності.

Кібернетичне узагальнення сукупності нейропсихологічних [156] та патопсихологічних [162] даних дозволяє визначити 9 відносно незалежних функціональних блоків на основі досліджених відділів головного мозку. В таблиці 1.3 показано відповідність функціональних блоків відділам головного мозку, що не суперечить відомим результатам досліджень вищої нервової системи.

Отже, нейропсихологія відрізняється, насамперед, науково обґрунтованим методом – синдромним аналізом мозкових механізмів вищих психічних функцій на матеріалі спостережень за локальними ураженнями головного мозку. Отримані результати такого аналізу дозволяють запропонувати такі особливості структурно-функціонального підходу до побудови СОПМК:

1. Онтогенез вищих психічних функцій демонструє єдину образну основу всіх типів мислення – як вербального, так і логічного. Кожна вища психічна функція жорстко не прив'язана до якогось одного мозкового центру, а є результатом системної діяльності мозку, в якій різні мозкові структури (аферентні та еферентні ланки) здійснюють власний специфічний внесок у побудову цієї функції.

Таблиця 1.3 – Відповідність функціональних блоків та відділів головного мозку

№ з/п	Відділи головного мозку	Функціональні блоки
1	Стовбурові відділи мозку	Ефективна і афективна складові нервової системи, блок емоцій
2	Моторні відділи мозку	Виконавчі механізми
3	Скронева область	Надоперативна пам'ять (буфер для створення складних образів)
4	Передмоторна область	Кінестетичні мелодії (буфер для створення складних рухів або образу-дії)
5	Зона ТРО (перекриття скроневої, тім'яної і потиличної кори)	Оперативна пам'ять (ансамбль з 7±2 образів)
6	Потилична область	Довготривала пам'ять
7	Тім'яна область	Асоціативна пам'ять
8	Постцентральні (лобово-скронева або задньолобні відділи)	Блок моделювання або формування і вибору стратегій (управління фокусом уваги і вибір образу-мети)
9	Лобові долі	Блок програмування, регулювання і контролю (багаторівневий стек)

2. Інтелектуальна діяльність – це організоване розв’язання задач, що спирається на логічну програму з 6-ти взаємопов’язаних операцій від постановки питання, безпосередня відповідь на яке неможлива, аж до перевірки отриманого результату (звіренням результатів з початковими умовами задачі).

3. Головний мозок людини за своїми макрофункціями складається з трьох основних блоків: енергетичного (лімбіко-ретикулярні системи), проєкційні та оперативні (асоціативні) системи, блок програмування, регуляції та контролю інтелектуальної діяльності.

4. Енергетичний блок забезпечує лише мотиваційну основу лінгвістичних процесів, проєкційні та оперативні (асоціативні) системи підтримують парадигматичну будову мови, а інтеграційно-пускові системи відповідають за найбільшу кількість визначальних функцій мовлення як інтелектуальної діяльності.

5. На основі нейропсихологічної інтерпретації відокремлюються 9 відносно незалежних функціональних блоків – лімбіко-ретикулярні системи (блок емоцій), виконавчі механізми, надоперативна пам’ять, блок кінестетичних мелодій, оперативна пам’ять, довготривала пам’ять, асоціативна пам’ять, блок формування і вибору стратегій, блок програмування, регулювання і контролю).

1.4 Аналіз лінгвістичних основ моделювання мовленнєвої діяльності

1.4.1 Загальнолінгвістичне представлення процесів мовлення

Основні результати дослідження механізму розуміння людиною природно-мовних конструкцій отримано лінгвістикою – наукою про мову. Розглянемо базові лінгвістичні представлення за цією проблематикою з огляду на припущення структурно-функціонального підходу до побудови СОПМК (п. 1.3).

Істотні кроки лінгвістичної науки, спрямовані на вивчення процесів мовленнєвої діяльності були зроблені в кінці 20-х – початку 30-х років ХХ сторіччя низкою видатних лінгвістів різних країн (Бодуен де Куртене, Ф. де Соссюр, Л. Блумфілд тощо), але отримали своє повне оформлення лише в 50-ті – 70-ті роки у роботах американської трансформаційної лінгвістики (Н. Хомський і його школа), так званій

породжувальній семантиці (Ч. Філлмор, Дж. Макколі, Дж. Лакофф), роботах Т. Бівера, Г. Кларка та ін., дослідженнях провідних радянських лінгвістів (І. А. Мельчук, А. К. Жолковский, Ю. Д. Апресян, А. А. Леонтьєв тощо) [195]. Проведені дослідження висвітлили декілька підходів до представлення мовленнєвої діяльності людини.

Найбільш загальною є характеристика мови як множини слів, кожне з яких має своє словникове значення. Будь-яке відоме слово позначає річ, якість, дію або відношення, що має назву наочної віднесеності або референтної функції слова. Можливі форми таких функцій для частин мови:

- іменник – предмет;
- дієслово – дія;
- прикметник – властивість;
- прийменник – зв'язок;
- союз – відношення.

З іншого боку, слово – це особлива форма віддзеркалення дійсності, завдяки йому світ людини подвоюється, оскільки вона може керувати цим другим світом внаслідок вольової дії. У цьому виявляється регульовальна функція слова. Крім того, за допомогою слова людина має можливість засвоєння досвіду поколінь – тут слово служить джерелом інформації [196].

Кожне слово має своє «семантичне поле», яке розкриває його значення та відповідає лінгвістичному терміну полісемія. З цієї точки зору у слова розрізняють безліч асоціативних (пов'язаних) та безліч денотативних (найближчих) значень. Вся сукупність значень є конотативним значенням, з яким пов'язують відомий психофізіологічний феномен «*Tip of tongue phenomenon*» – вибір слова «на кінчику язика». Як було вже відмічено, асоціативні значення формуються у слова раніше внаслідок реагування на синтагматичні мовні конструкції, а денотативні значення (синонімія) з'являються пізніше в результаті накопичення асоціативних значень – у цьому проявляється парадигматична будова мови.

Парадигматичне мислення дорослої людини виражається також в тому, що слово може отримати «категоріальне» або «понятійне» значення. Аналізуючи предмети, речі та їх властивості, абстрагуючи і узагальнюючи їх ознаки, слово вводить складові навколишньої дійсності в систему складних зв'язків і відношень. Внаслідок цього слово

є знаряддям мислення і засобом спілкування. Наприклад, слово «чорнильниця» має корінь чорн- (чорнити, фарбувати), суфікс – иль (засіб, знаряддя), 2-й суфікс – ниц (вмістище), флексія – я, ею та ін. Такими морфологічними засобами задається відношення, яке цей предмет має до навколишньої ситуації. Отже, слово є могутнім знаряддям аналізу цього світу і дозволяє проникнути в сферу раціонального [197].

Валентності слова є ще одним важливим чинником, який визначає вибір потрібного (пов'язаного з ними) іншого слова. Лексичні зв'язки засновані на тому, що багато слів суб'єктивно відчуються як неповні та вимагають доповнення іншими словами (до 3-х – 4-х, максимум до 5-ти в російській мові). Апресян виділяє такі категорії лексичних функцій:

- «incip» – давати початок;
- «fin» – закінчувати;
- «func» – функціонувати;
- «caus» – казувати, бути причиною.

Таким чином, значенням слова можна вважати однакову для всіх людей стійку систему узагальнень, що стоїть за словом та зберігає незмінне «ядро» – певний набір зв'язків. Тоді сенс – це індивідуальне (афектне) значення слова, наслідок контекстного випадку, коли із загальної системи зв'язків виділяються ті, які мають відношення до даного моменту і даної ситуації [198].

На початкових етапах розвитку мови слово було вплетене в практику (носило симпрактичний характер), тобто набувало свого значення з конкретної ситуації (жесту, інтонації). Далі слово все більше перетворювалося на самостійну систему кодів, в результаті мова стала містити в собі всі необхідні засоби для позначення предмета і виразу думки (синсемантична система). Дослідження проблеми в онтогенезі показали, що первинне «агукання» дитини відображає лише його афектний стан. Перші справжні слова дитини звернені до предмета і позначають предмет, проте тісно вплетені в ситуацію (мають симпрактичний характер). У 1 рік 6 місяців – 1 рік 8 місяців у дитини відбувається засвоєння граматики рідної мови і, внаслідок цього, спостерігається величезне збільшення її словника (з'являється синсемантична мова).

Характерно, що предметна віднесеність слова формується у дитини поступово (до 2-х років) через ситуацію, жест, міміку, інтонацію,

зберігаючи тісний зв'язок з практичною дією та істотними ознаками предмета. Значення слова розвивається в онтогенезі: афектний сенс, конкретний досвід з ряду практичних наочних образів і ситуацій, система ієрархічно зв'язаних і взаємно підпорядкованих категорій. Усвідомлюються слова дитиною поступово: іменники (спочатку конкретні предмети), дієслова, прикметники, допоміжні слова. Значно складніше засвоюється значення відносних слів і словосполучень («сестра», «в», «на», «під», «брат батька» тощо). Ієрархія понять кожної людини за Л. С. Виготським має індивідуальну «довготу» (від елементарного до складнішого) і «широту» (об'єкти, які зустрічалися або не зустрічалися в досвіді). З лінгвістичних позицій можна стверджувати, що свідомість розвивається таким чином: афектний характер, наочно-дієвий характер, відвернений вербально-логічний характер [199].

Якщо слово є одиницею мови, то речення вважається одиницею живого мовлення. Фраза служить виразом цілої думки певного судження, яку розглядають як єдину серійно організовану систему (К. Лешлі) або складну динамічну структуру (гештальтпсихологія). Симпрактичне мовлення дитини і його дійсні зв'язки: «бажання–ім'я» і/або «ім'я–дія» і/або «ім'я–дія–об'єкт» переходить в синтагми або зчеплення декількох зв'язаних слів в єдину мовну систему. Парадигматична організація цього процесу об'єднує фонетичні і лексичні елементи з морфологічними і семантичними формами слова.

Грецькі терміни синтагми та парадигми зустрічаються в слов'янському мовознавстві ще у відомій праці М. Смотрицького (1618 р.) «Грамматіки Славенския правилное Свнтагма» [200]. Сучасне лінгвістичне поняття синтагми походить від Бодуена де Куртене та Л. В. Щерби та означає, насамперед, лінійну одиницю, що виникає в результаті природного членування потоку мовлення. Філологи ще сперечаються щодо інтонаційних або семантичних принципів визначення кордонів синтагми, яка складається із слів (іноді одного слова), але в той же час вона не збігається зі словосполученням і не дорівнює йому в синтаксичному аспекті [201].

Синтагматичний принцип заснований на плавних переходах від одного слова до іншого. З'ясовано, що у дитини раніше з'являються предикативні відповіді, ніж асоціативні, це можна формально представити як $S \rightarrow P \rightarrow O$ (суб'єкт \rightarrow предикат \rightarrow об'єкт). Такий стан речей пояснюється тим, що синтагми з точки зору мети мовлення мають

предикативну структуру, яка, власне, і характеризує думку чи судження. Асоціативні зв'язки між окремими словами вже становлять результат накопичення у досвіді дитини багатьох синтагм, що повністю відповідає механізму умовного рефлексу.

Різноманітність синтагм може бути класифікована за К. Свіделіусом як:

- комунікація подій, найбільш проста форма – з'єднання простих речень за допомогою союзу «і» (паратак西斯);
- комунікація (логічних) відношень (гіпотак西斯).

Онтогенетичний розвиток мовлення, скоріш за все, також відповідає такій класифікації. Сприйняття навколишнього світу у вигляді послідовності подій справедливо вважається головним джерелом інформації не тільки для людини, але й для вищих тварин. Якщо предикативний характер найперших речень дитини характеризує важливі для неї події, то послідовність таких подій найлегше поєднувати союзом «і». Історичний аналіз писемності теж демонструє вживання паратак西斯у на ранніх стадіях розвитку мови, наприклад, у перших давньоруських літописах. Гіпотак西斯 є наслідком пізнішого за часом значного накопичення асоціативних зв'язків, які служать основою для появи стійких парадигматичних відношень між лексичними одиницями.

Формальний аналіз граматики синтаксичних структур привів Н. Хомського до створення трансформаційної лінгвістики. У ній виділяються поверхневі синтаксичні структури будь-якої мови (ступені свободи або перефразування одного і того ж речення; омонімічність при розумінні речення) і глибинні синтаксичні структури. Теорія трансформаційної породжувальної граматики [202], заснована на таких базових положеннях:

1. Вражаюча легкість оволодіння дитиною 1,5–2,5 років практично нескінченною мінливістю реальних варіантів висловів приводить до необхідності розрізняти принаймні два різні рівні організації мовного вислову: поверхневі і глибинні граматичні структури мови. Глибинні структури характеризуються порівняно нечисленними правилами побудови і є проміжною ланкою між думкою і розгорнутим висловом.

2. Разом з основними правилами побудови глибинних синтаксичних структур мови Хомський описав і формальні правила, за якими одні структури переходять в інші.

Критика Хомського в лінгвістичній літературі заснована на проблемах семантичного аналізу складних випадків побудови фрази та врахування важливих її характеристик. Маються на увазі:

- «залежні» або контекстні фрази;
- лексичні одиниці (вдівець, неодружений), що виводять фразу з ряду позаконтекстних;
- слова, що позначають логічні функції;
- тема (\perp) і рема (R) фрази;
- «топiк» і «фокус» фрази;
- просодичні засоби;
- «пре-супозиція» – прихована вибіркова установка.

Варто відмітити, що апарат та багато теоретичних постулатів генеративізму за чотири десятиліття змінилися майже до невпізнання. Теорія, яка в 1960-ті роки називалася найчастіше трансформаційною і формулювалася як система правил, в даний час ні поняття трансформації, ні навіть поняття правила не використовує. З новітньої версії генеративізму також були виключені поняття глибинної і поверхневої структур, що колись займали центральне місце в теорії. Незмінними у Хомського залишилися лише постулати про автономність граматики, природженість мовної здатності людини та про єдність базових принципів побудови різних мов, що розрізняються лише установкою деяких часткових параметрів [203]. Хоча Хомський одним з перших оголосив лінгвістику частиною когнітивної психології, проте тенденція співвіднести мовну здатність з іншими когнітивними здібностями людини з'явилася в інших теоріях.

У кінці 1960-х років Дж. Макколі, Дж. Росом, П. Посталом та ін. була запропонована «породжувальна семантика» – модель семантичного опису, яка протистоїть так званому інтерпретуючому підходу. Основна ідея нового підходу зводилася до заперечення уявлення про автономний синтаксис, відповідальний за породження виключно формальної глибинної структури, яка лише потім наповнюється деяким семантичним змістом (інтерпретується). Натомість була запропонована концепція, згідно з якою породження виразів природної мови починається з деякої семантичної структури. Термін «природна логіка» в теорії отримав популярність в розширювальному розумінні – для поз-

начення логіки повсякденного міркування, відмінної від аристотелівської силогістики і формальної логіки [204].

З середини 1970-х років Дж. Лакофф та Г. Томпсон вперше ввели поняття «когнітивної граматики», затвердженої надалі для позначення такого типу лінгвістичного дослідження, який не тільки визнає, але й заохочує використання відомостей про побудову і функціонування людської свідомості для пояснення мовних явищ [205]. Автори теорії визначали виняткову роль та використовували як інструмент дослідження поняття метафори, що зайвий раз підкреслює значне зближення «чистої» лінгвістики з семіотикою та іншими когнітивними науками наприкінці ХХ століття.

Метафора є не стільки фігурою поетичної мови, скільки найважливішим механізмом освоєння світу людським мисленням, що забезпечує перенесення знань про краще відомі з наочного досвіду понятійні сегменти на сфери, відомі гірше. Осмислення менш наочних сфер, зокрема вже задане системою неусвідомлюваних конвенціональних («мертвих») метафор, несе на собі риси первинних уявлень. Так, у виразах типу *висока мораль* або *висока філософія* простежується представлення про «верх» і «низ» у людській фізіології та ін. Уявлення про виняткову роль метафори у формуванні понятійної системи людини і структури природної мови знайшло відображення в неоднозначно сприйнятій концепції «втіленого розуму», в рамках якої вивчається залежність людських розумових здібностей і уявлень про світ від особливостей побудови людського тіла і людського мозку [206].

За своє суттю розуміння метафоричних висловлювань базується на асоціативному мисленні людини з обов'язковим застосуванням складових парадигматичної будови мови [207]. У повній мірі користуватися метафоричним апаратом людина може вже у досить зрілому віці за умови відповідного досвіду та освіти. Практика досліджень показує, що найбільш складними не тільки для моделювання [208], але і для розуміння самими носіями мови є випадки використання ще цілої низки лексичних конструкцій, таких як [209]:

a) флексивні поєднання (шматок хліба, брат батька);

b) службові слова (прийменники і союзи) : «весна перед літом»; ще важче розуміються оборотні відношення (конструкції) типу «круг під хрестом»;

c) порядок слів (дійсний і пасивний залог, незбіжність з реальним порядком подій);

d) порівняльні конструкції (Соня світліша Олі, але темніша Каті);

e) складні синтаксичні структури (складнопідрядні речення);

f) смислові інверсії (не звик не підкорятися правилам).

Окреслені проблеми розуміння пояснюються вторинним парадигматичним характером цих конструкцій, що не дозволяє носію мови застосувати синтагматичні асоціативні зв'язки, а натомість, вимагає обов'язкового використання логічного мислення. Варто також відмітити, що домінуючий на сьогодні метод моделювання лінгвістичного процесору на семантичному рівні за допомогою формалізації складових парадигматичної будови мови не призвів до бажаного результату [198, 208]. У той же час природний шлях онтогенезу від предикативної структури висловлювання та синтагматичного мислення розв'язує для потенційних носіїв мови проблему розуміння обмеженого, проте універсального за складом, мовного простору протягом 3–5 років від рівня новонародженого.

Отже, лінгвістичні дослідження мовленнєвої діяльності людини протягом останніх 40 років істотно наблизилися до задач інформаційних технологій. Проте до цього часу лінгвістичними теоріями не розв'язано у повній мірі проблему семантичного аналізу та синтезу ПМК. Визнаними положеннями лінгвістики, що мають бути враховані у новому підході, вважатимемо такі:

1. Кожне слово мови має референтну, регулюючу та інформаційну функції. Окрім того, кожне слово має своє «семантичне поле», яке визначається множинами асоціативних (синтагматичних за походженням) та денотативних (парадигматичних) значень.

2. Значенням слова можна вважати однакову для всіх людей стійку систему узагальнень, що стоїть за словом та зберігає незмінне «ядро» – певний набір зв'язків з іншими словами. Тоді сенс – це індивідуальне значення слова, наслідок відокремлення із загальної системи зв'язків того, що має відношення до даного моменту і даної ситуації.

3. Розвиток значення слова в онтогенезі проходить через такі етапи: афектний сенс, конкретний досвід з низки практичних наочних образів і ситуацій, система ієрархічно зв'язаних і взаємопідпорядкованих категорій.

4. Одиницею живого мовлення вважається речення, в онтогенезі якого розрізняють предикативну структуру, що характеризує окремі події, синтагматичну будову на основі асоціативних зв'язків та, насамкінець, парадигматичну організацію для об'єднання фонетичних та лексичних елементів з морфологічними і семантичними формами слова.

5. Складності проблем семантичного аналізу цілої низки випадків побудови фрази пояснюються вторинним парадигматичним характером відповідних лексичних конструкцій, що не дозволяє носію мови скористатися елементарними асоціативними зв'язками, а натомість, вимагає обов'язкового застосування логічного мислення.

6. Домінуючий метод моделювання лінгвістичного процесору на семантичному рівні за допомогою формалізації складових парадигматичної будови мови не привів до бажаного результату, хоча в той же час природний шлях онтогенезу від предикативної структури висловлювання та синтагматичного мислення розв'язує проблему розуміння природно-мовних конструкцій за 3–5 років.

1.4.2 Формування мовних висловлювань

Дослідниками в області лінгвістики, психології та суміжних наук питання розуміння сенсу (аналізу) натурально-мовних конструкцій небезпідставно поєднувалося з формуванням (синтезом) мовних висловлювань. Вперше наголосили на тісному зв'язку між синтезом та аналізом лінгвістичних процесів представники Вюрцбургської школи вивчення мислення (О. Кюльпе, Н. Ах, О. Зельц), розглядаючи процес формування мовного вислову як «шлях від думки до мови», а процес розуміння як «шлях від мови до думки», причому початкова «думка» розумілася як чистий духовний акт [166].

Розвинув цей погляд, але вже з матеріалістичних позицій Л. С. Виготський – думка є складним узагальненим віддзеркаленням дійсності, що прямує відомими мотивами, причому думка не виходить зі слова, а знаходить вираження в слові. Думка обов'язково має афективні джерела, які грають визначальну роль в перетворенні суспільно сформованих «значень» на індивідуальні, тобто такі, що відповідають цим афективним потребам, «сенси». На відміну від представників ідеалістичного напрямку Виготський неодноразово підкреслював, що думка – не остання інстанція. Сама думка народжується не з іншої

думки, а з мотивуючої сфери нашої свідомості, яка охоплює наші потреби і потреби, наші інтереси і спонуки, наші афекти і емоції. У джерела думки у будь-якому випадку знаходиться афективна і волюва тенденція. Тільки вона може дати відповідь на останнє «чому» в аналізі мислення [182].

Таким чином, первинним в процесі виникнення вислову є його мотив або, іншими словами, потреба цієї психічної дії. Будь-яка інтелектуальна психічна діяльність людини, в тому числі мовлення, отримує свою характеристику через мотивацію. Тому одна з тез діяльнісної концепції радянської психології стверджувала, що формування розумових дій включає на першому етапі виявлення мотиву до вирішення задачі [210].

Як вже було відмічено, проблема мотивації поведінки тварин і людини вивчається практично в усіх напрямках сучасної психології, найбільш помітними з яких є [166]:

- біхевіоризм;
- фрейдизм;
- гештальтизм;
- персоналізм.

При всій різниці у підходах більшість дослідників виділяє зовсім невелику кількість можливих мотивів мовної діяльності. Так, згідно з теорією оперантного навчання одного з апологетів необіхевіоризму Б. Скінера до мотивів мовного вислову відносяться лише [130]:

- вимога, бажання попросити що-небудь – *mand (demand)*;
- бажання передати що-небудь, звернення інформаційного характеру – *tact (contact)*;
- бажання ясніше сформулювати власну думку, зрозуміти (уточнити) щось або ввести це в систему понять – *sept (concept)*/(за Лурією).

Кожний з цих мотивів утілюється в задумі вислову, який також можна вважати первинним семантичним записом. У задумі виникає визначення теми і реми вислову, що психологічно означає формування загального суб'єктивного сенсу вислову. Під темою розуміється відоме, отримане попередньо, а рема – це те нове, що додається до теми та структурує висловлювання, а також передає комунікативну мету висловлювання. Виходячи саме з таких міркувань Л. С. Виготський

наголошував на тому, що «думка не втілюється в мові, а здійснюється в мові, формується за допомогою слова або мови» [181].

Основною психологічною проблемою взаємозв'язку мислення і мови є проблема переходу від суб'єктивного, ще словесно не оформленого і зрозумілого лише самому суб'єктові смислу, до словесно оформленої і зрозумілої будь-якому слухачеві системі значень, яка формується в мовному вислові. Характерна риса цього процесу – єдність або зв'язаність вислову, який перетворює його на замкнуту смислову конструкцію. Виготський застосовував термін «смысл» для позначення афективних відношень суб'єкта до змісту, що позначався в слові. Такі, інколи унікальні відношення виникають виключно в процесі індивідуального досвіду, в них виражаються особливо важливі для суб'єкта сторони позначуваних явищ [182].

Наступним етапом формування вислову є внутрішня мова, яка перетворює симультанну, семантичну схему в сукцесивний мовний вислів, що розгортається на основі послідовної організації. Внутрішня мова має згорнутий, аморфний характер і відрізняється, як вже було сказано раніше, предикативним складом. Для внутрішньої мови характерне збереження всіх валентностей або потенційних зв'язків кожного предиката [211].

За своєю роллю у процесі синтезу висловлювання внутрішня мова є проміжною ланкою між думкою та мовою. Функціональною ознакою внутрішньої мови є предикативний зміст, а морфологічними – її згорненість, скороченість і граматична аморфність. У онтогенезі внутрішня мова людини проходить послідовне перетворення від «егоцентричної мови» дитини (за Ж. Піаже – розгорнута зовнішня мова, що влітається в дію і допомагає вирішити задачу) до «шепітної мови», а потім поступово згортається [212]. Таким чином, внутрішня мова є саме тим механізмом, який перетворює внутрішні суб'єктивні смисли на систему зовнішніх розгорнутих значень.

Остаточним етапом мовотворення є вже, власне, формування розгорнутого мовного вислову. Таким можна вважати цілий ланцюг взаємно зв'язаних речень, який характеризується єдиною замкнутою структурою (*coherence*). Незалежних від контексту фраз (*context-free*) у мовному вислові не існує, а тому кожна фраза має не тільки референтну природу (*referential meaning*) внаслідок посилання на певну подію, але і соціально-контекстне значення (*social-context meaning*). Те-

ма і рема в розгорнутому мовному вислові мають бути істотно розширені і розділені на цілий ланцюг кроків певної програми. Але вихідна установка, створена завданням передати відповідну інформацію, повинна протягом тривалого часу протистояти побічним впливам (асоціаціям), які відволікають. Звідси мовний вислів за А. А. Леонтьєвим є формою психічної діяльності [213].

Таким чином, для вислову як складної форми психічної мовної діяльності характерними є [199]:

- мотив, мета, завдання (мета в певних умовах);
- окремі етапи або дії для реалізації завдання;
- операції, що виконують ці дії;
- достатньо широкий обсяг оперативної пам'яті;
- складна система стратегій, яка забезпечує можливість:
 - виділяти істотний сенс вислову;
 - гальмувати побічні асоціації;
 - вибирати відповідні завданню мовні формулювання;
- постійний контроль за реалізацією вислову;
- свідомий вибір потрібних мовних компонентів з багатьох альтернатив.

Достатньо інформативним з точки зору моделювання виглядає аналіз формування розгорнутого мовного вислову в онтогенезі. Лише перед шкільним віком як мотив, так і програма висловлювання набувають міцного характеру і вислів починає перетворюватися на замкнуту систему складної, обмеженої певним завданням оповіді. Даний факт можна пояснити лише тим, що внутрішня мова дитини формується значно пізніше за його зовнішню мову [212]. Предикативний склад внутрішньої мови досягає сталого вигляду тільки після оволодіння індивідом морфологічними, лексичними і синтаксичними формами, які є композиційними компонентами розгорнутої мови [213].

Альтернативний вибір мовних компонентів також суттєво залежить від онтогенезу мовлення. Вказівка на те місце, яке займає лексична одиниця в синтаксичній структурі, аніж не вичерпує всього багатства зв'язків, що супроводжують відповідну словоформу. Як вже було показано, при аналізі лексичних одиниць розрізняють ієрархічно побудовані парадигматичні системи та потенційні синтагматичні відносини з іншими лексичними одиницями. Парадигматичне співвідно-

шення окремих лексичних значень утворює поняття і є актом симуль-танного синтезу окремих одиниць інформації. Синтагматичне об'єднання окремих слів у цілісні вислови виступає як серійна органі-зація мовних процесів, «валентність» або «місткість» слів. Тому син-тагма як комунікація подій, наприклад «хмиз горить», «собака гав-кає», виникає набагато раніше і будуються набагато легше, ніж пара-дигматичні комунікації відношень [160].

Все багатство можливих семантичних зв'язків можна звести до 40–60 основних лексичних функцій. Згідно з діяльнісним підходом, застосування цих функцій для синтезу мовного вислову відбувається при визначенні операцій, що забезпечують реалізацію первинного за-думу (мети діяльності). Наприклад, відома стратифікаційна модель мови І. А. Мельчука «смысл ↔ текст», де вперше введено поняття ле-ксичної функції, припускає наявність таких основних рівнів форму-вання висловлювання [214]:

- мотив і загальний задум, який із самого початку відомий суб'єктові в найзагальніших рисах;
- семантичні уявлення (симультанні схеми, семантичні графи / модель смислу початкової думки / використання «внутрішньої мови»);
- глибинно-синтаксичні структури (ієрархічно організоване дере-во залежностей, у вузлах якого знаходяться символи глибинних син-таксичних одиниць);
- поверхнево-синтаксичні структури (морфологічна, фонологічна і фонетична розгортка).

Уявна роздрібність, громіздкість і складність синтаксичного запи-су на основі лексичних функцій є необхідною умовою для переходу до наступного рівня. Якщо глибинно-синтаксичних відношень для ро-сійської мови приблизно 10, то вже предикативних – лише 4–6. Від-ношення на рівні внутрішньої мови відповідають питанням хто? що?, кого? чого?, кому? чому? тощо, одне атрибутивне, одне сурядне і одне приєднувальне відношення [156].

Значну увагу лінгвісти приділяють темі (те, що вже відоме тому, хто говорить) і ремі (те, що ще має бути повідомлене) висловлювання. Саме виділення реми припускає побудову обмеження до того неви-значеного числа асоціацій, які спливають при застосуванні певного слова. Реалізація реми у синтагмі приводить до вибору тих систем

зв'язків, які в контексті вислову представляються найбільш істотними та найповніше відповідають намірам, з яких суб'єкт виходить в своєму вислові. Специфіка різних видів мовлення проявляється у тому, що вираз внутрішніх «смислів» може здійснюватися через інтонації, розподіл пауз, акценти (просодика), а іноді стає зрозумілим лише з «підтексту» вислову (теорія режисури) [199].

У лінгвістиці використовується така класифікація висловів [211]:

- афективні вигуки («Ой», «Галю, ти!» тощо);
- діалогічні вислови:
 - відповідь цілком повторює питання або відтворює частину питання;
 - відповідаючи слід сформулювати щось нове;
- монологічне мовлення (присутні мотив та задум).

Кожний з цих типів висловлювання має власні особливості – так, наприклад, для усної діалогічної мови характерним є:

- можливість відсутності власного мотиву і задуму вислову при відповіді на питання;
- знання загальної теми і змісту бесіди, ситуації, в якій відбувається бесіда;
- можливість використання позамовних засобів (міміка, жести, засоби інтонації, паузи).

Внаслідок цього в усній діалогічній мові допускається значна граматична неповнота (еліпси або елізії), що створює найбільші проблеми при моделюванні діалогу [215].

Усне монологічне мовлення відрізняється від усного діалогічного і складається з декількох смислових груп (*chunks*), зв'язаних між собою в цілісну «замкнену» структуру. Граматична неповнота усного монологу можлива, якщо співбесідник досить добре знає загальну ситуацію і розділяє мотиви того, хто говорить. Розглянемо як найбільш показові приклади усного монологічного мовлення:

- Лектор: представлення матеріалу в найбільш послідовному і логічно стрункому вигляді, збереження чіткого логічного переходу між істотними частинами висловлюваного матеріалу, позамовні засоби виділення окремих частин (синтагм, словосполучень та слів) в тексті.

- Актор: опанування образом – зрозуміти мотиви, переживання, тип мислення героя, вжитися в роль на основі надзадачі, а потім вже відбиття створеного образу в монологічному мовленні.

Усне монологічне мовлення може коригуватися ззовні, переходячи в приховану форму діалогічного, а також доповнюватися цілою низкою просодичних (інтонації, паузи) та позамовних (міміка, жести) маркерів. По відношенню до практичної дії усне монологічне мовлення згідно з театральною традицією може бути «драмою» або «епосом» [216]. Найбільш складне з усіх типів письмове монологічне мовлення має спиратися лише на достатньо повне використання розгорнутих граматичних засобів мови. Письмова мова із самого початку є свідомим довільним актом, в якому засоби виразу виступають як основний предмет діяльності. Лише письмова мова у повній мірі забезпечує свідомий контроль за всіма процесами мислення, а тому може вважатися основним знаряддям удосконалення розумової діяльності [156].

Отже, особливості кожного з типів висловлювання повинні мати відповідне відображення при моделюванні мовної діяльності. Парадоксальним виглядає той факт, що традиція комп'ютерної лінгвістики основну увагу приділяє найскладнішим варіантам мовлення. Так, на першому місці у розробників лінгвістичних систем знаходиться переклад – складна форма аналітичної діяльності, найважливішим завданням якої є усвідомлення логічного ладу первинної думки, логічної структури початкового задуму [217]. Необхідно визнати, що онтогенетичний розвиток мовлення у людини відбувається з «точністю до навпаки» – від афективних вигуків до діалогу та усного і писемного монологу. Найбільш переконливою щодо цього виглядає думка Л. С. Виготського про те, що «думка народжується не з іншої думки, а з мотивуючої сфери нашої свідомості, яка охоплює наші потяги і потреби, наші інтереси і спонуки, наші афекти і емоції». Виходячи з цієї концепції, можна стверджувати:

1. Первинним в процесі виникнення вислову є його мотив або, іншими словами, потреба цієї психічної дії. Кількість типів мотивації висловлювання обмежена вимогою або бажанням чого-небудь, зверненням інформаційного характеру та необхідністю уточнити щось (увести це в систему понять).

2. Первинний семантичний запис вислову зосереджено у його задумі, де вперше виявляється тема і рема вислову, що психологічно

означає формування загального суб'єктивного сенсу цього висловлювання.

3. Перехід від суб'єктивного та афективного смислу до словесно оформленої і зрозумілої будь-якому слухачеві системи значень, що формується у мовному вислові, забезпечує внутрішня мова. Головна функція внутрішньої мови полягає в перетворенні симультанної, семантичної схеми в суцесивний мовний вислів, що розгортається на основі послідовної організації.

4. Формування розгорнутого мовного вислову здійснюється через ланцюжок взаємно пов'язаних речень, який характеризується єдиною замкнутою структурою. Тема і рема в розгорнутому мовному вислові мають бути істотно розширені, але вихідна установка, створена завданням передати відповідну інформацію, повинна протягом тривалого часу протистояти побічним, відволікаючим впливам (асоціаціям).

5. Аналіз формування розгорнутого мовного вислову в онтогенезі демонструє шлях: зовнішня мова—«шепітна мова»—внутрішня мова—стійкі мотиви і програма висловлювання. Синтагматична комунікація подій виникає набагато раніше і будується набагато легше, аніж парадигматичні комунікації відношень.

1.4.3 Розуміння мовних висловлювань

Лінгвістична традиція у діалектичному протиставленні синтезу та аналізу завжди віддавала перевагу першому, враховуючи, що повна ясність щодо формування мовного висловлювання дозволить визначити механізм розуміння цього самого висловлювання [199]. Спочатку у комп'ютерній лінгвістиці взагалі панував погляд певної групи авторів про те, що розуміння повідомлення формально можна досягти на основі використання словника як системи понять у поєднанні з системою граматичних правил. Проте створені за таким принципом автоматичні перекладачі своїми незадовільними практичними результатами дуже швидко продемонстрували помилковість простого вирішення проблеми розуміння. Цим самим було отримано експериментальне підтвердження думки А. Р. Лурія: «Немає ніяких підстав сумніватися в тому, що обидва перераховані моменти абсолютно необхідні для розуміння мовного повідомлення. Проте навряд чи ці дві умови є достатніми для розшифрування його сенсу, яку б форму – усного або письмового повідомлення – воно не носило» [130].

На сьогоднішній день здобула визнання зовсім інша точка зору групи відомих лінгвістів та психологів – процес розуміння починається з пошуків загальної думки вислову, що складає зміст цієї форми психічної діяльності. У роботах таких авторів, як Р. Роммейтвейт, Ч. Філлмор, Дж. Мак-Колі, Дж. Лакофф, Дж. Уертч тощо розглядалася така стратегія декодування: пошук загальної думки → лексикофонемний рівень (окремі слова) → синтаксичний рівень (окремі фрази). Проведені психологічні експерименти показали, що реальний процес розуміння розгорнутого мовного повідомлення не збігається з тим порядком, за яким надходить інформація і згідно з яким той, хто слухає спочатку сприймає сенс окремих слів, а потім цілих фраз. Отже діяльність, спрямована на розуміння повідомлення (наприклад, викладеного в певному тексті) має складний характер, її забезпечують різні процеси, лише частина яких пов'язана із сприйняттям значення слів та декодуванням синтаксичних правил їх поєднання [218].

Інша група процесів розуміння пов'язана з тим, що вже на перших етапах сприйняття повідомлення у суб'єкта виникають гіпотези або припущення (*pre-suppositions*) про сенс повідомлення. Звідси центральним процесом механізму декодування є пошук смислу, що приводить до вибору з низки виниклих альтернатив. Той, хто слухає (або читає) ніколи не ставить своїм завданням зрозуміти окремі слова або ізольовані фрази – обидва ці процеси грають роль підлеглих, допоміжних операцій і лише в деяких випадках перетворюються на спеціально усвідомлені дії. Підлеглість розуміння окремих слів або фраз особливо виразно проявляється при сприйнятті інформації, яка надходить на чужій, погано знайомій мові.

За Уертчем механізм декодування побудовано на таких поняттях, як значення всього повідомлення або внутрішній сенс, що створює загальну зв'язність (*external coherence*) і «підтекст» або «глибина» повідомлення (*internal coherence*). Акт розуміння характеризується спробами розшифрувати значення всього повідомлення (визначити його внутрішній сенс) і надати повідомленню глибину або «підтекст». Ці спроби завжди спрямовані на пошук у контексті вислову (синсемантичному або ситуаційному), без якого ні розуміння цілого тексту, ані правильна оцінка його складових неможливі. Тому стверджується, що не існує ніяких вільних від контексту елементів вислову, а на перший план висувається процес пошуку і формування відповідних гіпотез

або припущень, які визначають конкретні значення слів або фраз – складових мовного вислову [219]. Отже, декодування можна вважати активним за природою і складним за структурою психічним процесом.

Виготський та Міллер також вважали, що ключовою ланкою процесу декодування є перехід від поверхневих граматичних структур мови до глибинних або базисних семантичних, оскільки внутрішній сенс вислову може розходитися з його зовнішнім значенням. Зв'язний текст обов'язково включає загальну думку і підтекст. Завдання повного розуміння сенсу вислову або його «підтексту» полягає якраз в тому, щоб не обмежуватися розкриттям лише зовнішнього значення повідомлення, але абстрагуватися від нього і від поверхневого тексту перейти до глибинного підтексту, від значення – до смислу, а потім і до мотиву, що закладено в основу цього повідомлення.

Таким чином, в основі стратегії декодування лежить така логічна схема: слово → фраза → текст → зовнішнє значення → підтекст (основний внутрішній сенс) → мотиви [125]. Як можна побачити, отримана схема є зворотною від ланцюга формування мовних висловлювань. Проте на відміну від процесу синтезу, часова послідовність роботи механізму декодування не збігається з логічною схемою, оскільки представляє собою складну взаємодію процесів аналізу, що діють паралельно.

Послідовно розглянемо основні складові елементи розуміння тексту у відповідності до логіки процесу. Розуміння окремого слова лише ззовні здається простим, але насправді семантичний вибір адекватного значення слова є складним психологічним процесом, вивчення якого може розкрити основні механізми, що визначають розуміння або декодування отриманого повідомлення. Значення слова завжди багатозначне (різні форми омонімії), а для його розуміння необхідно завжди робити вибір з урахуванням як того контексту, в якому дається слово, так і тієї ситуації, в якій робиться повідомлення. Умовами вибору адекватного значення вважаються [220]:

- частотність цього слова в мові, у практиці людини;
- мовний контекст (коса в косаря або дівоча коса);
- ситуаційний контекст (перший укіс або візит до перукарні).

Є всі підстави вважати, що семантичне уявлення є багатовимірною симультанною структурою, яку можна зобразити у вигляді семантичного графа. Ігнорування факту багатозначності слова і вирішаль-

ної ролі контексту при виборі потрібного значення з низки альтернатив може істотно ускладнити процес розуміння мовного повідомлення. Аналогічна ситуація має місце і в тих випадках, коли навчання іноземної мови починається не з засвоєння контекстної мови, а з вивчення словникового значення окремих слів. Такий шлях опанування мовою неминуче приводить до низки труднощів, які значною мірою усуваються, якщо змінити метод навчання та виходити з мовного контексту, а лише, в другу чергу, звертатися до словникового значення ізольованих слів [130].

У звичайних умовах організований, вибірковий характер обрання одного варіанта з багатьох вірогідних можливостей з метою виконання поставленого завдання (декодування) забезпечується автоматично, тобто так, що потрібне слово, зміцнене минулим досвідом, спливає у свідомості з більшою за побічні слова ймовірністю. Навколо кожної лексичної одиниці дійсно створюється багатовимірною мережа зв'язків, причому в нормі переважають смислові зв'язки (семантичне поле), які гальмують звукову схожість, але в патологічних станах витіснені примітивні (звукові) зв'язки розгальмовуються і зрівнюються за своїм значенням із смисловими зв'язками [221]. Отже, особливості синтагматичної та парадигматичної будови мови відіграють вирішальну роль при контекстному розумінні значення слова.

Розуміння фрази за всіма ознаками спирається на внутрішню мову, а тому легко здійснюється у разі прямої комунікації подій, наприклад:

- суб'єкт → предикат («листя опадає»);
- суб'єкт → предикат → об'єкт («листя опадає на землю»);
- ознака (суб'єкта) → суб'єкт → предикат → ознака (об'єкта) → об'єкт («золоте листя опадає на прохолодну землю»).

Певні ускладнення виникають при сприйнятті значення фраз у разі, коли описується зворотний порядок подій. У цих випадках для розуміння значення фрази необхідна відома трансформація, пов'язана з переходом від поверхневої синтаксичної структури до глибинної (логічної) синтаксичної структури. Наприклад, маленькі діти безпомилково розуміють фрази на зразок «Петрик дав іграшку», але набагато важче відбувається декодування фрази при необхідності трансформації – «Петрик попросив іграшку». У другому випадку декодування

вимагає виділення уявного логічного підмета, а тому розуміння цієї фрази під силу вже старшим дітям [130].

Ще важчим є процес розуміння комунікації відношень і фраз з складною системою управління, наприклад при використанні взаємно підлеглих граматичних структур. Так, для розуміння комунікації відношень необхідно трансформувати:

- флексивні конструкції (родовий атрибутивний відмінок – «брат батька» або «батько брата»);
- незмінні флексії, значення яких залежить від порядку слів, що застосовуються («вбрання зачепило весло» або «весло зачепило вбрання»);
- складні оборотні конструкції з прийменником («квадрат під кругом» або «круг під квадратом»);
- конструкції, що співвідносять за допомогою службового слова «який»;
- дистантні конструкції з множинною ієрархією відношень;
- порівняльні конструкції («Петрик сильніший за Дмитрика»);
- семантична інверсія («я не звик не підкорятися правилам»).

Для декодування таких конструкцій необхідна внутрішня допоміжна трансформація речення, перетворення симультанної схеми, що виражає відношення, в розгорнутий ланцюг послідовних положень [156].

Розуміння сенсу тексту характеризується за Виготським впливом (вливанням) сенсів окремих фраз. Фрази не можуть бути ізольованими ланками колишнього ланцюга, оскільки кожна подальша фраза «вливає» або включає значення попередньої – цей феномен визначає механізм розуміння основного змісту тексту [182]. Для визначення цього механізму Лурія використовував поняття «смислових ядер», які виділяються суб'єктом при читанні окремих фраз та пошукової орієнтувальної діяльності людини, що декодує мовне повідомлення. Пошукова орієнтувальна діяльність, що ставить своїм завданням виявити ті ланки вислову, які повинні бути співвіднесені один з одним при сприйнятті окремої фрази, набуває ще складнішого характеру при орієнтуванні в цілому тексті, коли доводиться зближувати іноді дуже далеко віддалені один від одного елементи складного вислову [185].

Хоча в психології ще не до кінця досліджено процеси визначення «сміслових ядер» в окремих фразах, припускається, що на рівні підсвідомості важливу роль відіграють тема і рема тексту, а також умовний рефлекс на значущі слова. При декодуванні відбувається неперервний процес аналізу при зверненні «сміслових ядер» і, надалі, синтез загального сенсу.

На процес розуміння тексту також має вплив імовірнісний характер вибору можливого закінчення вже декодованого матеріалу. Існують тексти, в яких створюється загальна установка на один певний смисл, і тоді багато частин тексту не мають вирішального значення для розуміння загального сенсу. У цих випадках вступає в дію здогадка, створена установкою або припущенням, що виникло при читанні попередньої частини тексту, а тому активний аналіз виявляється непотрібним. Проте в деяких інших текстах вірогідність появи тільки одного певного сенсу дуже низька. І щоб розібратися в цьому тексті, потрібен тривалий активний аналіз, звернення окремих елементів тексту один з одним, створення гіпотез про загальний сенс, складний вибір з низки альтернатив, які виникають при смислового аналізу тексту [125].

Таким чином, чим менша ймовірність появи викликаних контекстом асоціацій, тим більше потрібен активний аналіз тексту, тим більші ускладнення викликає його розуміння. Процес декодування мовного вислову використовує механізм «впливу (вливання) сенсів» для єдності сприйняття смислу (з використанням оперативної пам'яті) і механізм загальної установки або програми для збереження основного напрямку сприйнятого смислу.

Підтекст або внутрішній сенс присутній далеко не в усіх мовних висловах – певні вербальні конструкції не передбачають підтексту, але інколи навіть прості висловлювання можуть мати поруч з зовнішнім також і внутрішній сенс (наприклад, «Вже десять!» може означати «Пора йти ...»). Майже в кожній репліці, в кожному вислові актора в п'єсі завжди, поряд із зовнішнім значенням тексту, існує і внутрішній підтекст або приховане значення. Тому основне завдання актора полягає в тому, щоб донести до глядача і текст, і його підтекст, інакше кажучи, надати глядачу можливість глибшого «прочитання» тексту.

Підтекст обов'язково присутній та краще всього виявляється в таких вербальних конструкціях [130]:

- вирази з переносним сенсом («золоті руки»);

- конструкції порівняння («очі, як безмежні озера»);
- прислів'я;
- байки.

Для розуміння всіх цих конструкцій необхідно абстрагуватися від безпосередньої системи значень для того, щоб виявити та проаналізувати внутрішній сенс, який має прихований вираз у системі розгорнутих зовнішніх значень. Характерно, що внутрішній сенс має яскраво виражену емоційну оцінку, а глибина прочитання тексту більше залежить від емоційної витонченості людини, ніж від її формального інтелекту. Так, можна зустріти людей, які, з великою повнотою і ясністю розуміючи логічну структуру зовнішнього тексту і аналізуючи його значення, майже не сприймають того смислу, який стоїть за цими значеннями, не розуміють підтексту і мотиву, залишаючись тільки в межах аналізу зовнішніх логічних значень. Отже, досягти глибокого розуміння внутрішнього сенсу тексту можна виключно на основі образного мислення.

На відміну від усіх інших живих істот засоби мови дозволяють людині забезпечити процес продуктивного логічного виведення, який відбувається на вербально-логічному рівні. Наявність мови та її складних логіко-граматичних структур дає можливість людині робити припущення та висновки на основі логічних міркувань, не звертаючись кожного разу до свого безпосереднього чуттєвого досвіду та обмежуючись лише тими засобами, які має в своєму розпорядженні сама мова [222]. Ця властивість мови забезпечує появу складних форм дискурсивного (індуктивного і дедуктивного) мислення, які є основними формами продуктивної інтелектуальної діяльності людини. В результаті стає можливим феномен силогізмів, тобто отримання знань раціональним шляхом у результаті виникнення відчуття логічної неповноти.

У найпростішій формі силогізму перша (велика) теза містить у собі відому думку, що має загальний характер. Друга (мала) теза вказує на те, що об'єкт відноситься саме до тієї категорії, яка була сформульована у великій тезі. В результаті поєднання великої і малої тез на підставі загального правила формулюється висновок – ті закономірності, які справедливі по відношенню до певної категорії об'єктів, повинні розповсюджуватися на всі об'єкти, які відносяться до цієї категорії. Сило-

гізм як апарат логічного мислення, створений в процесі суспільної історії, дозволяє зробити висновок тільки з цієї системи логічних зв'язків, не спираючись ні на які додаткові матеріали наочного досвіду.

У психології мислення у свій час склалося уявлення про те, що «логічні переживання», або «логічні відчуття», є якимись природженими особливостями духовного життя. Так, деякими психологами Вюрцбургської школи подібні «логічні відчуття» трактувалися як первинна властивість психічного життя, що не спирається ні на наочні образи, ні на мову, однаково зустрічається у всіх людей і носить універсальний характер. Аналогічні переконання розділяли і представники гештальтпсихології, які розцінювали відчуття «логічної неповноти» як потребу доповнити до цілого незавершену логічну побудову [223].

Сучасні психологи вважають, що незалежне від особистого практичного досвіду силогічне мислення виникло лише на тих ступенях розвитку культури, коли діяльність, здійснювана за допомогою мови, відокремилася в особливий самостійний вид теоретичної діяльності. Такий погляд підтверджують дослідження онтогенезу – дитина 3-х–4-х років робить основою для висновку не логічне узагальнення, а свій власний практичний досвід, тісно пов'язаний з безпосередніми враженнями [224]. У цьому віці ще не сформовані базові операції індукції та дедукції, а виділення логічних відношень неможливе. За Піаже поява окремих логічних операцій спостерігається у віці від 7-ми років, проте, вони можливі лише за наявності наочного досвіду, а не поза ним; стадія формальних операцій настає лише з 11-ти–14-ти років.

Результати досліджень Лурії показали, що операції логічного висновку є продуктом історичного розвитку, оскільки на ранніх стадіях еволюції суспільства з переважаючими наочно-дієвими формами практики формально-логічні операції висновку обмежуються лише кордонами такої практики. Радикальна перебудова суспільно-економічного устрою, ліквідація неписьменності, включення в культуру приводять не тільки до розширення кола понять і оволодіння складнішими формами мови, але і до формування апаратів логічного мислення, що дозволяє виходити за межі безпосереднього досвіду [130].

Отже, комп'ютерна лінгвістика за допомогою програмного експерименту вже довела хибність погляду на те, що розуміння мовних висловлювань можна досягти лише на основі використання словника як системи понять у поєднанні з системою граматичних правил. Пробле-

ма розуміння полягає у складному характері взаємодії психічних процесів, що одночасно відбуваються у напрямках пошуку загальної думки та декодування на лексико-фонемному і синтаксичному рівнях. Психологічні дослідження виявили такі особливості процесів розуміння мовного вислову:

1. Вже на перших етапах сприйняття повідомлення у суб'єкта виникають гіпотези або припущення про сенс повідомлення. Центральним процесом механізму декодування є пошук загального смислу, що приводить до вибору з низки альтернатив, що виникаючих. Процеси розуміння окремих слів або ізольованих фраз грають роль підлеглих, допоміжних операцій і лише в окремих випадках перетворюються на спеціально усвідомлювані дії.

2. Декодування можна вважати активним за природою та складним за структурою психічним процесом, який характеризується спробами розшифрувати значення всього повідомлення та визначити його глибину або «підтекст». Такі спроби завжди спрямовані на пошук у синсемантичному або ситуаційному контексті вислову, без якого ні розуміння цілого тексту, ані правильна оцінка його складових неможливі.

3. Ключовою ланкою декодування є перехід від поверхневих граматичних структур мови до глибинних семантичних. В основі стратегії декодування лежить логічна послідовність слово → фраза → текст → зовнішнє значення → підтекст (внутрішній сенс) → мотиви, проте часова послідовність роботи механізму декодування не збігається з логічною.

4. Розуміння фрази спирається на внутрішню мову, а тому легко здійснюється у разі прямої комунікації подій, важче при зворотному порядку подій і ще важче – для комунікації відношень і фраз зі складною системою управління. Для декодування таких конструкцій необхідна внутрішня допоміжна трансформація речення, перетворення симультанної схеми, що виражає відношення, в розгорнутий ланцюг послідовних положень.

5. Механізм розуміння тексту використовує поняття «сміслових ядер» та «впливу (вливання) сенсів» на основі пошукової орієнтувальної діяльності і припущень людини, що декодує мовне повідомлення. Імовірнісний характер вибору можливого закінчення вже декодованого матеріалу впливає на складність неперервного процесу аналізу при зверненні «сміслових ядер» і, надалі, синтезу загального сенсу за допомогою оперативної пам'яті.

6. Внутрішній сенс має яскраво виражену емоційну оцінку, а глибина прочитання тексту більше залежить від емоційної витонченості людини, ніж від його формального інтелекту. Досягти глибокого розуміння внутрішнього сенсу тексту можна виключно на основі образного мислення.

1.5 Вибір напрямку, мети та постановка завдань дослідження

Будемо виходити з того, що існуючі методи моделювання природної мови, в які вкладено всі відомі правила мовотворення, програють у розумінні висловлювань дитині, яка не тільки не знає цих правил і відповідної термінології, але навіть не вміє читати. Отже, розуміння природно-мовних конструкцій є результатом пізнавальної, у тому числі мовленнєвої діяльності людини, а тому більше залежить від психологічного механізму її асоціативного образного мислення, ніж від формальних ознак та властивостей текстової інформації. Потребують розроблення моделі, методи та відповідні технологічні засоби, що обґрунтовують образний аналіз та синтез, інтерпретуючи образний сенс як чисельні результати формальних процедур оброблення ПМК. Адекватність такого підходу можна підтвердити шляхом порівняння нових та існуючих розв'язків цілої низки актуальних задач, в яких семантичний компонент є домінантним, тобто оцінку розв'язку можна отримати лише за допомогою експерта. Такого типу задачі у подальшому будемо вважати семантико-залежними.

Проведений аналіз існуючих методів семантичного аналізу та мультидисциплінарних основ моделювання дає можливість визначити основну ідею дослідження та особливості підходу до розв'язання науково-прикладної проблеми у вигляді мети та задач дослідження. Останні сформулюємо шляхом обґрунтування функціональних вимог до інформаційної технології та відповідної інформаційної системи, що забезпечує обробку природно-мовного контенту на основі вилучення з текстів та накопичення загальних знань образного сенсу.

Мета дослідження полягає у підвищенні якості розв'язання комплексу семантико-залежних задач обробки інформації шляхом створення методології образного аналізу та синтезу структурно-функціональних моделей систем на основі природно-мовної інформації. Досягнення мети має забезпечити побудова інформаційної технології образного аналізу тексту та синтезу структурно-функціональних

моделей СОПМК, що орієнтована на загальні знання образного сенсу, має здатність до їх самовдосконалення та складом власних функцій забезпечує розв'язання комплексу семантико-залежних задач.

Оскільки мовленнєва та пізнавальна діяльність людини знаходяться у відношенні окреме–загальне [224], дамо необхідні визначення. Сенсом у широкому значенні можна вважати визначену за мету пізнавальної діяльності людини деяку властивість вербальної інформації. Тоді образний сенс – це чисельний параметр, в основу якого покладено міру співвідношення нових знань (вмінь, навичок) отримувача цієї інформації до формальної одиниці сенсу. В цьому випадку результат розуміння сенсу-властивості полягає у визначенні чисельної оцінки (вимірюванні) образного сенсу. Запропонований підхід дозволяє побудувати інформаційну технологію оброблення електронного контенту на основі формальної метамови у вигляді семантичної мережі як моделі сенсу-властивості, простору образного сенсу на основі відповідної одиниці сенсу, штучної системи, яку в подальшому будемо називати інфологічною і яка грає роль абстрактної конструктивної моделі СОПМК.

Самовдосконаленням бази знань будемо вважати врахування нової текстової інформації через формальні характеристики топології та сили первинних асоціативних зв'язків між образами семантичної мережі, що, зрештою, приводить до виявлення в ній нових парадигматичних зв'язків, які не задані явно у вихідних текстах. З огляду на формалізацію психодидактичного середовища [225] – відносно нового педагогічного концепту, який до певної міри забезпечує виконання функцій викладача в навчальному процесі – структурно-функціональний підхід дозволяє виключити з контуру управління навчальним процесом або значно обмежити роль експерта, що створює та супроводжує загальну базу знань предметної області. Функції викладача в системі дистанційного навчання повністю або частково виконує СОПМК з технологічними засобами аналізу та синтезу елементів ЕК, що позначені на рисунку 1.1 текстурною заливкою.

Запропонований підхід базується на моделюванні розглянутих раніше ключових понять когнітивного простору особистості – образу, асоціації та події з урахуванням онтогенетичного принципу побудови системи, на яку покладено роль експерта з предметної області та викладача. За умови відповідності елементів ЕК стандарту SCORM функ-

ції підсистеми адміністрування дистанційним навчанням (ПАДН) може забезпечити будь-яке з популярних та поширених в навчальних закладах середовищ з розвинутим рядом програмних можливостей (Moodle, Lotus Learning, Прометей, e-Learning Server тощо). Ефективність системи можна оцінити підвищенням успішності студентів, що вимірюється відомими статистичними методами. Проте необхідно бути впевненим у тому, що покращення успішності отримано завдяки вищій якості семантичного аналізу ЕК на основі поняття образного сенсу.

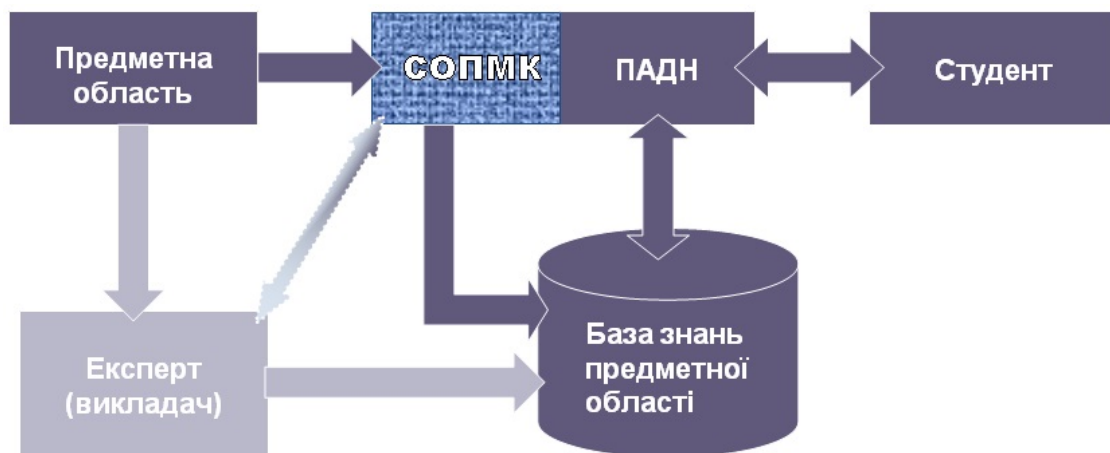


Рисунок 1.1 – Структура системи дистанційного навчання на основі СОПМК

Вибір критеріального апарата для оцінки досягнення мети дослідження побудуємо на основі аналізу широкого кола семантико-залежних задач, актуальних для комп'ютерної лінгвістики, електронного навчання та інших предметних областей. З цією метою визначимо функції та можливості штучної системи, які дозволяють стверджувати про певну якість розуміння цією системою контенту за такими функціональними групами:

- 1) підтримка діалогу;
- 2) лінгвістичні можливості;
- 3) дидактичні можливості;
- 4) додаткові можливості (пошук/переклад).

З часу появи у 1950 році відомої статті Алана Тьюринга [226] діалог на природній мові вважається універсальним засобом перевірки штучної системи на інтелектуальність. Жодна з сучасних програм, навіть після 60-річної історії досліджень не витримує перевірку тестом

Тьюринга, що свідчить про глибину проблеми розуміння сенсу ПМК. Проте в умовах навчального процесу та в багатьох інших прикладних задачах немає потреби вимагати від системи здатності вести універсальний діалог. Достатньо, щоб діалог, при потребі, підтримувався системою у межах ЕК, наприклад, з метою забезпечення пізнавального процесу тих, хто навчається. Тому пропонується вважати корисними функціями ЕК варіанти тесту Тьюринга з певними обмеженнями згідно з таблицею 1.4, в якій відтінками сірого відмічено позиції, що відрізняються вже існуючими досягненнями.

Рівень сучасних продуктів комп'ютерної лінгвістики привчив користувачів до високої культури обробки текстової інформації, наприклад, у вигляді перевірки орфографії, синтаксису тощо. Тому розуміння сенсу ПМК має базуватися на вже відомих напрацюваннях, для цього в таблиці 1.5 наведено відповідні можливості. Практично недосяжною для сучасних лінгвістичних процесорів лишається інваріантна функція 2.3.

Таблиця 1.4 – Окремі варіанти тесту Тьюринга з певними особливостями та обмеженнями

№ з/п	Назва критерію/тесту	Особливості/обмеження
1	Тест Тьюринга (ТТ)	Універсальний діалог (без обмежень)
1.1	ТТ у певній предметній області з обмеженим словником	Вузька предметна область, відповідь обмежено тільки тими словами, що увійшли до словника
1.2	ТТ: сценарний варіант	Застосування певної множини сценарних заготовок для відповідей
1.3	ТТ: «дельфійський оракул»	Відповіді у вигляді множини слів, асоціативно пов'язаних з питанням
1.4	ТТ: «магістр Йода»	Відповіді цитатами з літературних джерел, що входять до складу ЕК
1.5	ТТ: «Basic English»	Недотримання синтаксичних правил при відповіді

Таблиця 1.5 – Характеристики лінгвістичних можливостей обробки ЕК

№ з/п	Назва критерію/тесту	Особливості/обмеження
2.1	Володіння морфологічними правилами	Розрізняються правильно побудовані слова мови
2.2	Володіння синтаксичними правилами	Розрізняються правильно побудовані речення мови
2.3	Визначення однакового сенсу в перефразованих виразах	Забезпечується вживання перефразованих речень

До розуміння сенсу ЕК безумовно мають відношення специфічні для пізнавального процесу особливості цього контенту. Збільшення дидактичних можливостей навчальної системи згідно з таблицею 1.6 та поліпшення їх якості опосередковано свідчить про якість розуміння закладеного в систему ЕК. Сучасні системи підтримки електронного навчання на тому чи іншому рівні забезпечують всі окреслені можливості за винятком 3.3.

Про якість розуміння сенсу ЕК також свідчить низка функцій штучної системи, наведених в таблиці 1.7 у вигляді додаткових можливостей пошуку, перекладу та застосування мультимедійних об'єктів.

Таблиця 1.6 – Характеристики дидактичних можливостей систем електронного навчання на основі природно-мовного контенту

№ з/п	Назва критерію/тесту	Особливості/обмеження
3.1	Презентація знань студенту	Забезпечується представлення студенту ЕК з потрібним обсягом знань
3.2	Оцінювання знань студента	Забезпечується оцінка кінцевих та проміжних результатів пізнавальної діяльності студента
3.3	Самовдосконалення бази знань ЕК	Забезпечується удосконалення бази знань на основі оброблення нових навчальних текстів
3.4	Побудова алгоритмів адаптивного навчання	Забезпечується реакція на дії студента у вигляді зміни його траєкторії навчання
3.5	Генерація повідомлень щодо стану та потреб системи	Застосовується сценарне оброблення значимих подій навчальної системи

Таблиця 1.7 – Додаткові можливості інфологічної системи

№ з/п	Назва критерію/тесту	Особливості/обмеження
4.1	Розуміння сенсу мультимедійних об'єктів	В ЕК розрізняються мультимедійні об'єкти, визначається їхня роль
4.2	Пошук за маскою слів	Результати пошуку повністю відповідають словам з маски
4.3	Пошук з урахуванням правил морфології та синтаксису	Результати пошуку розширюють сукупність слів з маски морфологічно і синтаксично правильними словами та словосполученнями
4.4	Семантичний пошук	Результати пошуку релевантні сукупності слів з маски за змістом
4.5	Багатомовність	Забезпечується застосування природно-мовних конструкцій на різних мовах

Потрібно відмітити, що значний внесок у функцію 4.1 забезпечують такі стандарти контенту, як IMS, SCORM, XML тощо. Всі інші можливості на достатньому рівні забезпечуються сучасними пошуковими системами за винятком 4.4 та, нерідко, 4.5. Загальний граф розглянутих функціональних вимог до інфологічної системи представлено на рисунку 1.2. Наведені у графі функції будемо вважати такими, що задають вимоги при синтезі інфологічної системи. В електронному навчанні на цю систему частково покладено функції викладача – експерта з предметної області знань, на рисунку 1.1 для неї застосовано позначення СОПМК у зв'язку з провідним значенням ПМК у електронному контенті.

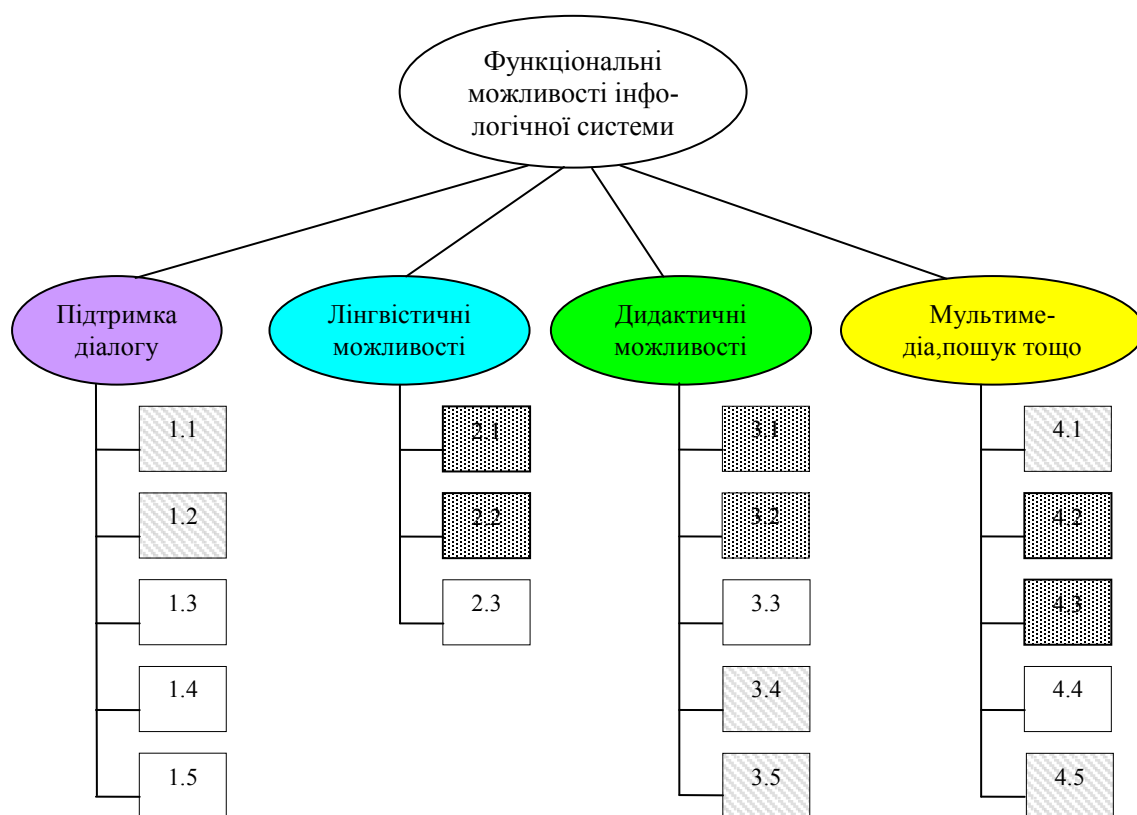


Рисунок 1.2 – Функціональні вимоги до інфологічної системи

Отже, виходячи з функціональних вимог до інфологічної системи, основними задачами дослідження будемо вважати:

1. Побудову теоретичних основ образного аналізу текстової інформації через:

1.1. Визначення концептуальних основ представлення образного сенсу природно-мовних конструкцій.

1.2. Формалізація комутативної напівгрупи образних конструкцій на основі прикладної теорії першого порядку.

1.3. Розроблення онтогенетичного методу побудови нечіткого відношення образного сенсу з обґрунтуванням одиниці сенсу.

2. Синтез структурно-функціональних моделей системи обробки природно-мовного контенту (СОПМК), а саме:

2.1. Створення методології синтезу структурно-функціональних моделей інфологічної системи.

2.2. Побудова методів моделювання процесів образної обробки природно-мовного контенту з обґрунтуванням архітектури бази знань СОПМК.

2.3. Розроблення функціональної моделі СОПМК на основі класифікації типів образного пошуку.

3. Побудову інформаційної технології образної обробки природно-мовного контенту на основі СОПМК, зокрема:

3.1. Розроблення моделей і алгоритмів реалізації базових типів образного пошуку.

3.2. Формалізація методу моделювання механізму оперативної пам'яті СОПМК.

3.3. Створення методу моделювання складових парадигматичного устрою мови.

4. Розв'язання семантико-залежних задач на основі інформаційної технології.

5. Практичну реалізацію й аналіз результатів впровадження отриманих методологічних і технологічних засобів.

РОЗДІЛ 2

ОСНОВИ ТЕОРІЇ ОБРАЗНОГО АНАЛІЗУ ТЕКСТОВОЇ ІНФОРМАЦІЇ ТА ФОРМАЛІЗАЦІЯ ПОНЯТТЯ ІНФОЛОГІЧНОЇ СИСТЕМИ

Підхід до образного аналізу текстової інформації будується шляхом введення понять інфологічної системи та онтогенетичного принципу, змістовного поєднання образних та природно-мовних концептів. Поняття образної конструкції визначається за допомогою формальної теорії і отримує кількісні оцінки на основі нечіткої міри та одиниці образного сенсу з урахуванням понять ентропії та кількості інформації. Обґрунтовується методологія синтезу структурно-функціональних моделей інфологічної системи шляхом кібернетичної інтерпретації процесів інтелектуальної діяльності людини.

2.1 Концептуальні поняття інфологічної системи та онтогенетичного принципу її побудови

Протягом останніх років у світі та в Україні спостерігається помітне зростання інтересу наукової спільноти до формалізації значимих аспектів образного мислення людини. Свідченням цього є реалізація Державної науково-технічної програми «Образний комп'ютер» на 2000–2010 роки, що затверджена постановою кабміну № 1652 від 08.11.2000 р. Проте, не дивлячись на помітні успіхи в розпізнаванні образів та створенні технічних засобів штучного бачення [227], існуючі підходи до моделювання операцій образного мислення мають на сьогодні скоріше концептуальний, ніж практичний характер [31].

Кібернетикою, як наукою про загальні принципи управління, закладено основоположні концептуальні поняття чорного ящика, інформаційних потоків, зворотного зв'язку та самоорганізації, що стали основою системного аналізу багатьох галузей діяльності людства та природних явищ з метою побудови відповідних інформаційних технологій. Абстрактна природа кібернетичних систем фокусує увагу дослідження на задачах обробки зовнішньої інформації, реагуванні на неї та можливостях зміни внутрішньої структури системи з метою підвищення ефективності виконання перших двох задач [228, 229].

Потрібно відмітити, що результати застосування кібернетичного підходу до систем довільної природи все ж таки відрізняються для

неживих та живих, особливо розумних елементів системи з огляду на ефективність керуючих дій [230]. Так, очевидні здобутки технічної кібернетики набагато вагоміші за реальні результати економічної, медичної, і, тим більше, біологічної кібернетики. Такий стан речей може бути наслідком того, що, наприклад, теорія автоматів і теорія оптимального керування, які входять до теоретичного ядра кібернетики, жодним чином не враховують відмінності між неживими та розумними складовими системи. Отже, не ставлячи під сумнів основоположні принципи кібернетики, виникає потреба врахувати суттєві особливості процесів життєзабезпечення та свідомості для моделювання задач біологічної кібернетики і штучного інтелекту.

Певним рухом в означеному напрямку можна вважати помітне концептуальне розширення понять так званого інфологічного підходу. Сам термін було запозичено з теорії баз даних, де під інфологічним рівнем розуміється інформаційно-логічна модель предметної області, з якої виключено надлишковість даних, а, натомість, відображено інформаційні особливості об'єкта управління. Інфологічне представлення даних не враховує особливості та специфіку конкретної СУБД, а орієнтовано, насамперед, на людину, яка проектує або використовує базу даних. Інфологічне проектування покликане виключити аномалії та протиріччя зовнішнього опису даних, такі як дублювання, надлишковість та непогодженість елементів даних [89]. Отже, інфологічна модель є засобом структуризації предметної області і деталізації концептуальних понять семантики даних.

З'являються публікації, автори яких використовують інфологічний підхід значно ширше, а саме для моделювання психологічних аспектів особистості [163, 231]. Тобто можна спостерігати явище застосування методології баз даних до нових предметних областей, формалізація яких вважається проблемною, оскільки вимагає потужних міждисциплінарних досліджень. Не представлене до цього часу в літературі формальне кібернетичне відображення філософських поглядів, що беруть початок у софістиці та суб'єктивному ідеалізмі, а на даний час найбільш яскраво представлені у феноменології (Е. Гуссерль, М. Хайдеггер). Вперше застосування інфологічного підходу до кібернетичного аналізу образного мислення запропоновано у [34, 37].

Постановка задачі: виходячи з того, що предметом кібернетичного дослідження є інтерпретація психічної діяльності людини, роз-

робити інфологічний підхід до моделювання образної обробки ПМК. Для досягнення мети потрібно:

- обґрунтувати і визначити зміст поняття інфологічної системи (ІС);
- розглянути формальні схеми взаємодії незалежних ІС в процесах пізнавальної діяльності;
- дати характеристику онтогенетичного принципу побудови ІС як моделі когнітивного простору особистості та визначити критерії ефективності ІС.

Поняття інфологічної системи. Визначимо клас ІС, як окремий випадок більш загального класу кібернетичних систем, що орієнтований на моделювання образної основи психічної діяльності людини. Будемо вважати, що основні відмінності між кібернетичною системою та ІС полягають в такому:

1. Кібернетичне дослідження спрямоване на управління системою, а інфологічне – на логіку взаємодії інформаційних потоків системи.

2. Обов'язкова присутність в кібернетичній системі об'єкта управління надає можливість:

а) дослідження об'єкта управління за допомогою моделі «чорного ящика»;

б) введення поняття зворотного зв'язку;

в) формалізації зовнішніх критеріїв за кінцевими результатами об'єкта управління.

3. В ІС принципово немає зовнішнього об'єкта управління, натомість:

а) «чорним ящиком» служить весь навколишній світ;

б) критеріями потрібно вважати виключно внутрішні параметри, в першу чергу, пов'язані з поняттями мотивів (потреб), емоцій, рефлексії та асоціативної мережі образів (АМО) як головного когнітивного ресурсу бази знань людини [31];

в) діалог та інші форми мовленнєвої діяльності можуть розглядатися як результат взаємодії двох чи декількох схожих інфологічних систем.

Розглянемо систему S , яку в подальшому будемо називати інфологічною, з точки зору суттєвих процесів її функціонування [43]. Нехай S здатна розпізнавати образи з нескінченної множини $I = \{i_1, i_2, \dots, i_n, \dots\}$ аналогічно тому, як людина розпізнає гештальт. ІС

також сприймає асоціативні зв'язки між парами образів як елементи множини $\omega \in \Omega$, де $\Omega \subseteq I \times I$ – довільна множина (простір) упорядкованих (асоціативних) пар. Для визначення образної конструкції (ОК) застосуємо поняття \mathbf{F} – сигма-алгебри (σ -алгебри) підмножин з Ω . Далі будемо вважати ОК будь-яку підмножину $\gamma \subseteq \Omega$, що має властивість $\gamma \in \mathbf{F}$. Якщо, згідно з властивостями σ -алгебри [232], множини $A, B \in \mathbf{F}$, то об'єднання, перетин і різниця A та B у теоретико-множинному сенсі також належать \mathbf{F} .

Припустимо, що система S обмінюється інформацією із зовнішнім світом як чорним ящиком виключно у вигляді ОК, з яких розрізняють послідовність вхідних подій $X = \{x_1, x_2, \dots\}$ та множину образних реакцій системи $Y = \{y_1, y_2, \dots\}$, причому $x_i \in \mathbf{F}$, $y_i \in \mathbf{F}$. На рисунку 2.1 зображено схему інфологічної моделі психічної діяльності людини, що включає у свій склад зовнішній «чорний ящик» та внутрішню інфологічну систему, на вхід якої неперервно подається множина образів подій у вигляді потоку X .

На виході ІС з'являються образи Y , які є реакцією цієї системи на зовнішню ситуацію X згідно з підходом до моделювання образного мислення людини [53]. Отже, введене поняття інфологічної системи повністю представлено у термінах моделі образного мислення та є формальним відображенням філософських поглядів феноменології в межах кібернетичного підходу.

Взаємодія інфологічних систем. З точки зору моделювання пізнавальної діяльності людини, інфологічні системи можуть взаємодіяти одна з одною та групами. Схематично процес взаємодії (діалог) двох інфологічних систем зображено на рисунку 2.2, де вхідна частина кожної такої системи має темний колір, як сховище інформації з «чорного ящика», а вихідна – світлий (символізує «висвітлені» завдяки заломленню через внутрішні мотиви системи фрагменти АМО).

Такий підхід дозволяє відокремити зовнішні канали передачі інформації у вигляді потоків образів та внутрішні функції перетворення інформації на основі АМО в інфологічній системі. Принциповим моментом формалізації «діалогу» є повна незалежність обробки образної інформації кожною інфологічною системою, що також відповідає феноменологічним поглядам.

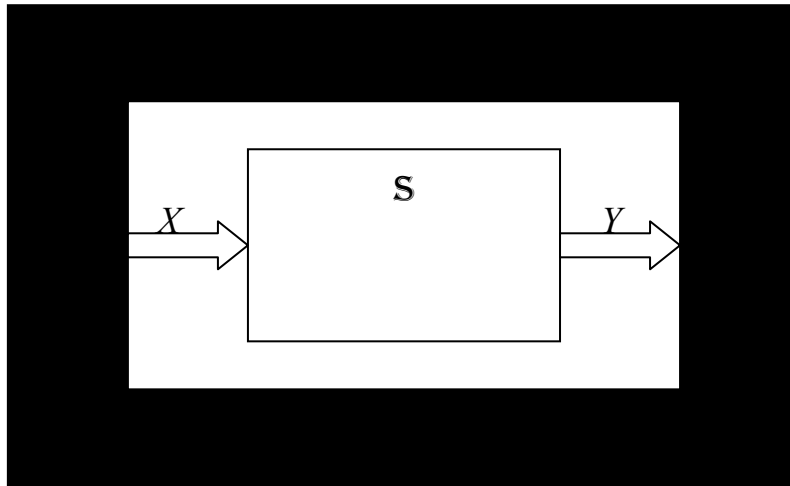


Рисунок 2.1 – Схема інфологічної моделі

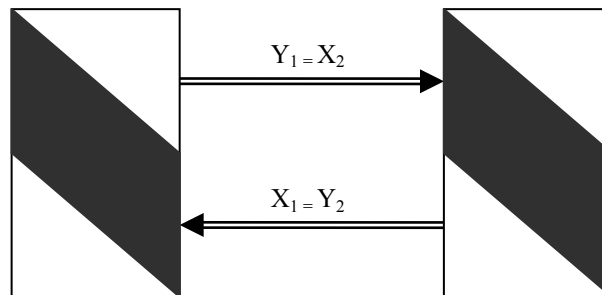


Рисунок 2.2 – «Діалог» двох інфологічних систем

Зрозуміло, що діалог може вважатися базовою, але далеко не єдиною формою пізнавальної та мовленнєвої діяльності людини у соціумі. На рисунку 2.3 схематично зображено процес взаємодії (спілкування) 3-х інфологічних систем за принципом «кожний з кожним».

У процесі «спілкування» виникає необхідність розгалуження інформаційних потоків з ОК, що може бути досягнуто формальними алгебраїчними засобами – оператором «PLUS» та еквівалентним перетворенням. На рисунку 2.3 використано схематичні зображення для вузлів \oplus – об'єднання інформаційних потоків з різних джерел (вузол «PLUS ОК») та \bullet – дублювання інформаційного потоку між різними інфологічними системами (вузол еквівалентного перетворення). За допомогою запропонованих універсальних схематичних елементів можна представити і більш складні процеси взаємодії інфологічних систем.

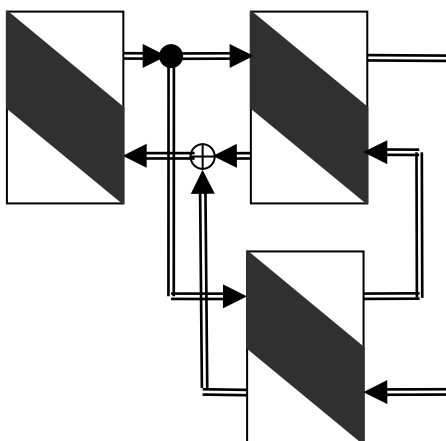


Рисунок 2.3 – «Спілкування» 3-х інфологічних систем

Так, на рисунку 2.4 зображено приклад взаємодії, що моделює такий процес навчання трьох учнів, в якому вони не спілкуються між собою, а вчитель та екзаменатор є незалежними особами. Відомо, що такий спосіб навчання вважається одним з найбільш ефективних для широкого кола задач пізнавальної діяльності – від оволодіння бойовими мистецтвами до наукових досліджень.

Нарешті, на рисунку 2.5 схематично зображено процес взаємодії (спілкування) n інфологічних систем за принципом «кожний з кожним», де інформаційні потоки з'єднуються тим же оператором \oplus та еквівалентним перетворенням \bullet . Такий тип спілкування також можна вважати традиційним для пізнавальної діяльності, оскільки історичними прикладами його ефективності вважають школу Сократа та круглий стіл короля Артура.

Критерії ефективності інфологічних систем. Суттєвою відмінністю запропонованого інфологічного підходу до моделювання образної основи психічної діяльності людини від класичного кібернетичного підходу є використання рефлексії як внутрішнього зворотного зв'язку. Формалізувати процес «висвітлення» отриманих з «чорного ящика» зовнішніх образних комбінацій до певних внутрішніх образних реакцій інфологічної системи [37] дозволяють такі принципи рефлексії:

- адекватність результатів образного пізнання реальному зовнішньому світу («чорному ящику»);
- використання внутрішніх мотивів та потреб як об'єктивних передумов фізичного існування інфологічної системи;

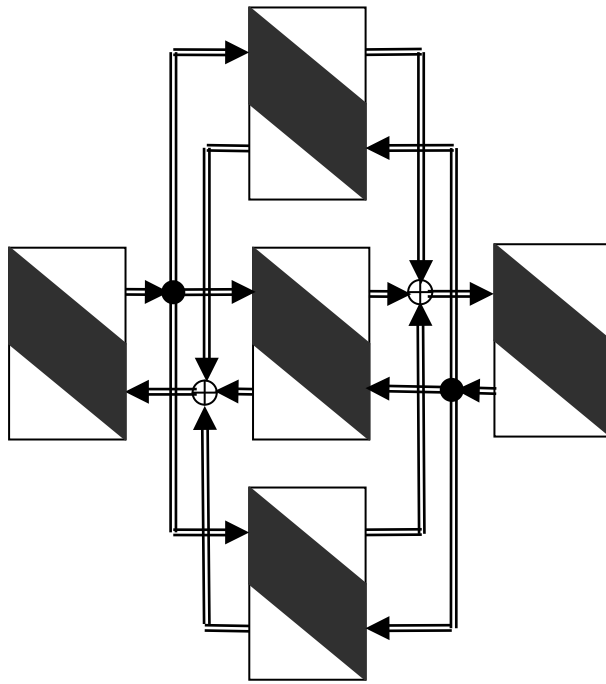


Рисунок 2.4 – «Навчання» з незалежним екзаменатором

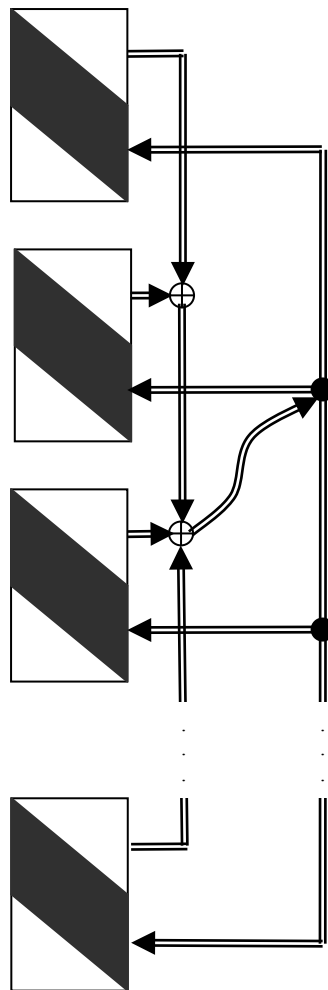


Рисунок 2.5 – «Сплкування» n інфологічних систем

- застосування механізму емоцій як сигнального апарата щодо стану мотиваційної сфери та вектора емоцій як критерію досягнення мети;
- застосування орієнтувального рефлексу як основного інструменту пізнавальної діяльності людини.

Якщо природним інструментом рефлексії виступає механізм емоцій, який забезпечує моніторинг досягнення мотиваційної мети, то в умовах штучного інтелекту досягти повної імітації значної частини комплексу мотивів та емоцій неможливо [31]. Проте закладемо як базовий *онтогенетичний принцип* побудови інфологічної системи – когнітивний ресурс Ω системи S , що визначає сенс її функціонування, отримується виключно шляхом послідовного накопичення параметрів чергових ω з зовнішнього «чорного ящика» та подальшого самовдосконалення множини Ω . Формально онтогенетичний принцип відображається в тому, що базу знань системи S будемо як $B = \bigcup_{i=1}^{m'} x_i$, де m' – загальна кількість сприйнятих системою на даний час вхідних образних конструкцій.

Одним з перспективних рішень для застосування інфологічного підходу в задачах дослідження є спрощення моделі мотиваційно-емоційної сфери людини за допомогою формального обмеження критеріального апарата. Пропонується взяти за основу модель так званої «чистої» інтелектуальної діяльності, що найбільш виразно проявляється у процесах пізнавальної (когнітивної) сфери людини. Тому критерієм ефективності взаємодії інфологічної системи з зовнішнім «чорним ящиком» пропонується обрати подвійний критерій для потоку образів Y «розв'язання» активної ситуації. З метою досягнення домінантного мотиву пізнавальної діяльності необхідно:

- а) застосувати мінімальні внутрішні енергетичні зусилля (наслідок принципу мінімальних витрат енергії на адекватне функціонування живої системи);
- б) забезпечити максимальну надійність бажаного інформаційного результату.

Як приклад окреслимо можливість формалізації критерію ефективності взаємодії інфологічної системи з зовнішнім «чорним ящиком» в умовах мовленнєвої діяльності людини. Запропонований подвійний критерій представимо як вибір такої з можливих відповідей на певне

питання (запит) до інфологічної системи, в якій враховано поточний стан системи та:

- досягнуто максимальної лаконічності висловлювання (потужність підмножини вихідних образів мінімальна);
- для забезпечення мети вихідного висловлення як реакції на вхідне використано когнітивний ресурс АМО з найбільшою кількістю сенсу.

Таким чином, для «чистої» інтелектуальної діяльності вектор емоцій будемо вважати пропорційним критерію сенсу образних, у тому числі природно-мовних конструкцій. Цей підхід моделює ситуацію, коли людина отримує задоволення від добре виконаної справи, причому кількість позитивних емоцій залежить від якості досягнення кінцевого результату роботи.

Отже, розглянутий структурно-функціональний підхід дозволяє удосконалити кібернетичну модель пізнавальних процесів психічної діяльності людини за рахунок моделювання мотиваційно-емоційної сфери та використання критеріїв сенсу ІС через поняття ОК. Отриманий на основі застосування запропонованого онтогенетичного принципу критеріальний апарат забезпечує концептуальну можливість формальної оцінки ефективності інфологічної системи.

2.2 Концепція визначення образного сенсу природно-мовних конструкцій

Для завершення побудови концептуальної моделі образного сенсу ПМК розглянемо зміст фундаментальних понять, закладених в основу теорії, що пропонується. Вперше цей підхід було висвітлено в роботах [37, 49]. Взавши до уваги всі проаналізовані раніше джерела та врахувавши вербальну основу об'єктів ЕК, пропонується прийняти у вигляді складових концепції такі визначення.

З філософської точки зору, образ є результатом відображальної (пізнавальної) діяльності суб'єкта, власне відображення у свідомості предметів, явищ об'єктивної дійсності [142]. Отже, образ – це стійкий відбиток у мозку людини (нервова модель) фрагмента навколишньої дійсності (зовнішнього стимулу). Формально будемо представляти образи як елементи нескінченної множини $I = \{i_1, i_2, \dots, i_n, \dots\}$, що розрізняються людиною як природним праобразом ІС у вигляді гештальтів (п.1.2), а під образною конструкцією (ОК) будемо розуміти довільну підмножину $\gamma \subset \mathbf{F}$, де \mathbf{F} – σ -алгебра підмножин з $\Omega \subseteq I \times I$.

Враховуючи характер предмету дослідження, основну увагу будемо приділяти таким образам (живих та неживих об'єктів, явищ, подій тощо), які мають свій мовний еквівалент у вигляді вербальних ознак. З цією метою введемо поняття мовного образу (МО), що відрізняється головною символічною ознакою у вигляді кореня слова. Іншими словами МО – це сукупність однокореневих слів, які характеризують окремий образ, виходячи з морфемної класифікації [109]. Таке поняття більш загальне за синсет [233] з відомого WordNet, словарну статтю чи лексему [234], у формі яких закладаються поняття в онтології. Формально підмножину $I' \subset I$ доступних для сприйняття людиною МО представимо за допомогою четвірки основних змістовних концептів:

$$I' = \langle N; O; M; Q' \rangle, \quad (2.1)$$

де N – поняття; O – об'єкт; M – метод; Q' – якість.

У свою чергу, деталізація базових концептів приводить до появи таких концептів:

$$N = \langle ON; QN; MN \rangle, \quad (2.2)$$

де ON – поняття об'єкта; QN – поняття якості; MN – поняття методу;

$$Q' = \langle OQ; MQ \rangle, \quad (2.3)$$

де OQ – якість об'єкта; MQ – якість методу;

$$M = \langle Ev; C \rangle, \quad (2.4)$$

де Ev – подія (процес); C – стан;

$$MQ = \langle H; T; L \rangle, \quad (2.5)$$

де H – власне обставина (відповідь на питання як?); T – обставина часу (відповідь на питання коли?); L – обставина місця (відповідь на питання де?).

Взаємозв'язок розглянутих елементарних концептів МО (2.1)–(2.5) у вигляді дерева графа представлено на рисунку 2.6.

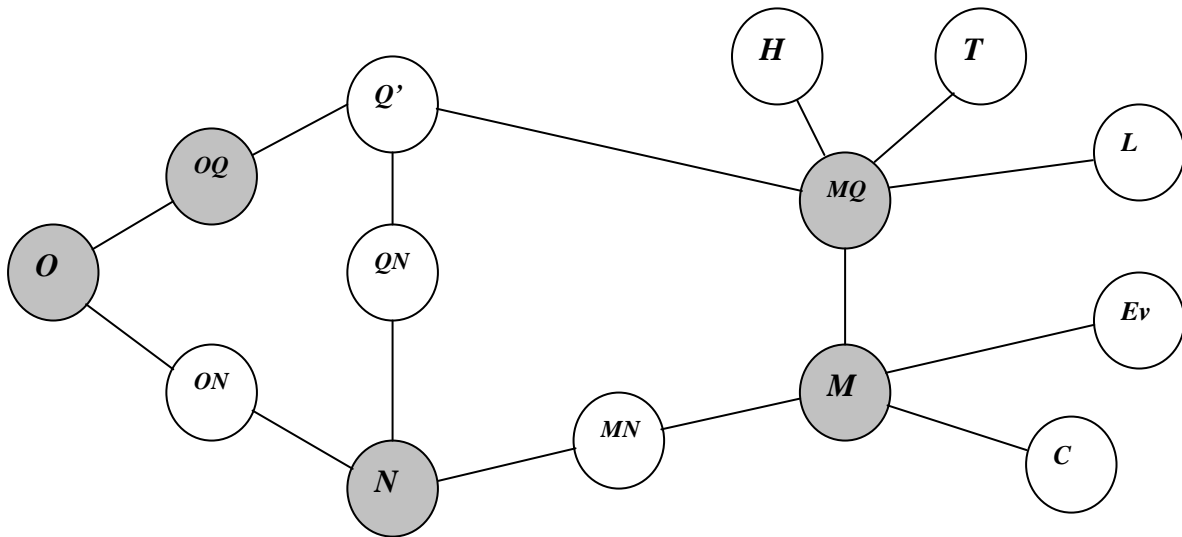


Рисунок 2.6 – Граф дерева взаємозв'язків елементарних концептів МО

Поява саме таких елементарних концептів (складових МО) пояснюється врахуванням значимих, тобто таких, що мають постійно визначений сенс, частин мови [235]. Сірим кольором на графі рисунку 2.6 позначено 5 змістовних концептів МО, які мають стійкі відображення у таких частинах української мови, як прикметник, іменник, дієслово та обставина. Поняття образу (N) додано до вибраної п'ятірки концептуальних складових з метою демонстрації єдності як наочно-предметних (синтагматичних), так і абстрактних (парадигматичних) ознак мовних образів.

У таблиці 2.1 представлено приклад російськомовних змістовних концептів для МО гіпотетичної інтроспективної моделі «мисливець та звір».

Слід зауважити, що використання у таблиці 2.1 п'яти вказаних концептів МО з 14-ти потенційних має певний універсальний характер для слов'янських мов (приклад такого опису на українській мові наведено в таблиці 2.2). Проте, в загальному випадку, вибір підмножини змістовних концептів мовних образів залежить від мови та мети дослідження.

Таблиця 2.1 – Приклад російськомовних змістовних концептів

<i>ОQ</i> (прикметник)	<i>О</i> (іменник)	<i>N</i> [поняття МО]	<i>M</i> (дієслово)	<i>H</i> (обставина)
смелый	смельчак	смелость	смелеть	смело
охотничий	охотник	охота	охотиться	охотно
стреляный	стрелок	стрельба	стрелять	
дневной	день			денно
белый	беляк	белизна	белеть	бело
звериный	зверь	зверство	звереть	зверски
пугливый		пугливость	пугать	пугливо
разъяренный	разъярен- ность	ярость	разъяряться	разъяренно
		кушанье	кушать	
ночной		ночь		ночно
темный		темнота	темнить	темно
страшный		страх	страшить	страшно
Какой?	Кто? Что?	Что?	Что делать?	Как?

Асоціація – це реалізований у мозку людини в результаті її унікального життєвого досвіду парний зв'язок між двома образами [194]. Від кількості та обставин (якості) повторення в житті однакових поєднань образів безпосередньо залежить спрямованість та сила асоціативного зв'язку [154]. Спільне використання понять образу I та асоціації як елемента множини $\omega \in \Omega$ створює когнітивне за своєю основою поняття простору асоціативних пар, яке будемо представляти за допомогою відношення $\Omega \subseteq I \times I$.

Як вже було зазначено у п.1.4, за характером виникнення асоціативні зв'язки можуть бути синтагматичними та парадигматичними. Синтагматичний (первинний) або мінімальний асоціативний зв'язок між двома образами виникає внаслідок сприйняття людиною події, де ці образи беруть участь. У мовному еквіваленті такі асоціації відтворюють комунікацію подій. Парадигматичний (вторинний) асоціативний зв'язок виникає на основі накопичення критичної кількості первинних зв'язків та характеризує певні відношення двох образів. Тому у мові парадигматичні асоціації, як правило, відтворюють комунікацію відношень.

Таблиця 2.2 – Приклад МО з україномовними змістовними концептами

<i>ОQ</i>	<i>О</i>	<i>N</i>	<i>М</i>	<i>Н</i>
сміливий	сміливець	сміливість	сміливішати	сміливо
мисливський	мисливець	мисливство		
стріляний	стрілець	стрільба	стріляти	
денний	день		днювати	денно
білий	біляк	білизна	біліти	біло
звіриний	звір	звірство	звірити	звіряче
лякливий	ляканий	лякливість	лякати	ляжливо
розлючений	розлюченість	лють	розлютитися	розлючено
харчовий	харч	харчування	харчувати	
нічний	ніч			
темний	темінь	темнота	темнити	темно
страшний	страховище	страх	страшити	страшно
Який?	Хто? Що?	Що?	Що робити?	Як?

Позначимо мінімальний асоціативний зв'язок синтагматичного або невідомого походження як P_x , а внутрішньообразну асоціацію як P_e . Введемо також позначення для найбільш відомих парадигматичних асоціацій:

- P_q – асоціація типу окреме–загальне (гіпо–гіперонімія);
- P_y – асоціація типу частина–ціле (меронімія);
- P_c – синонімічна асоціація;
- P_o – омонімічна асоціація як частковий випадок паронімії;
- P_a – антонімічна асоціація;
- P_p – асоціація типу «у риму».

Тоді визначимо множину асоціативних операцій P_i над концептами мовних образів як

$$P_i = \{P_x, P_e, P_q, P_y, P_c, P_o, P_a, P_p\}. \quad (2.6)$$

Формально подія є елементом множини ОК $E\nu \in \gamma$ з певними обмеженнями на кількість образів (7 ± 2) та зв'язність, що будуть розглянуті пізніше. У мовленнєвій діяльності подія має своїм аналогом син-

тагму [31], яка, зазвичай у вигляді простого оповідного речення, описує певну подію. Взаємозв'язок концептів подія–синтагма–речення формалізуємо на основі введення понять конструкції мовних образів (КМО) та природно-мовної конструкції (ПМК). Під КМО будемо розуміти таку підмножину ОК $\gamma' \subset \gamma$, до складу якої входять лише мовні образи. Тоді синтагмою будемо вважати елемент множини $Sy \in \gamma'$. Відповідно ПМК – це така підмножина КМО, за елементи якої взято слова з урахуванням морфологічних та синтаксичних правил певної природної мови.

Російське слово «смысл», з одного боку, позначає внутрішній логічний зміст, значення чого-небудь, що досягається розумом («мыслью») [104]. Природно-мовний ЕК, оброблення якого входить до об'єкта дослідження, фактично і є змістом з точки зору семантики. Оскільки ПМК складається не з окремих слів, а, як правило, з речень та природно-мовних конструкцій, то в понятті сенсу запропоновано розрізняти:

- образний сенс – параметр або чисельна оцінка таких елементів ПМК, які розрізняються ІС;
- сенс-властивість як знання/вміння/навички, що ставляться за мету пізнавальної діяльності отримувача природно-мовної інформації.

У відповідності до онтогенетичного принципу (п. 2.1) первинною основою сенсу є прагматична компонента, а семантична складова вторинна за своєю природою [53], тому будемо шукати образний сенс як деяку функцію від аргументу (ОК–КМО–ПМК), а значення цієї функції інтерпретувати як розуміння сенсу.

Застосування асоціативної мережі образів як моделі когнітивної системи у вигляді ІС дозволяє запропонувати єдиний підхід до розуміння сенсу на основі сенссполучень (термін введено за аналогією до словосполучення). Під сенссполученням будемо розуміти виявлений в межах події зв'язок між двома або трьома образами. Кожна синтагма в розгорнутому вигляді може (але не обов'язково) включати до свого складу такі змістовні концепти образів, що відповідають певній події:

1. Якість об'єкта чи суб'єкта дії – OQ або SQ .
2. Об'єкт дії – O .
3. Метод (власне дія) – M .
4. Місце дії – L .

5. Час події – T .
6. Якість методу – H .
7. Суб'єкт дії – S .

Враховуючи таку нумерацію, введемо 7-м базових синтагматичних типів асоціацій (всього 6 односпрямованих та 1 двоспрямована) як складових P_x . Тоді МО можуть пов'язуватися у дереві синтагми через свої змістовні концепти так, як зображено на рисунку 2.7.

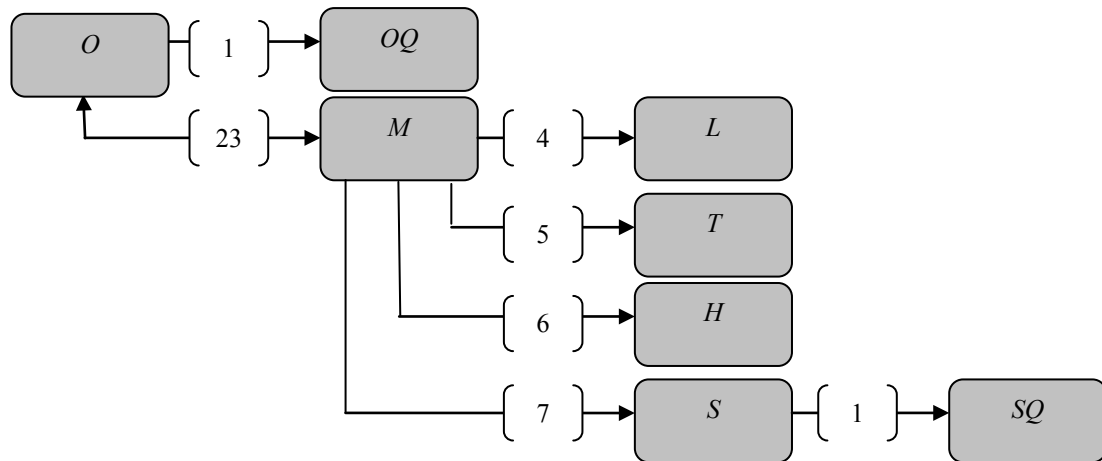


Рисунок 2.7 – Дерево можливих асоціацій синтагми

Оскільки сам термін [по-дія] (по російськи – [со-бытие]) передбачає центральну роль саме дії, попередній граф можна зобразити в розгорнутому вигляді навколо головної дії (рисунок 2.8).

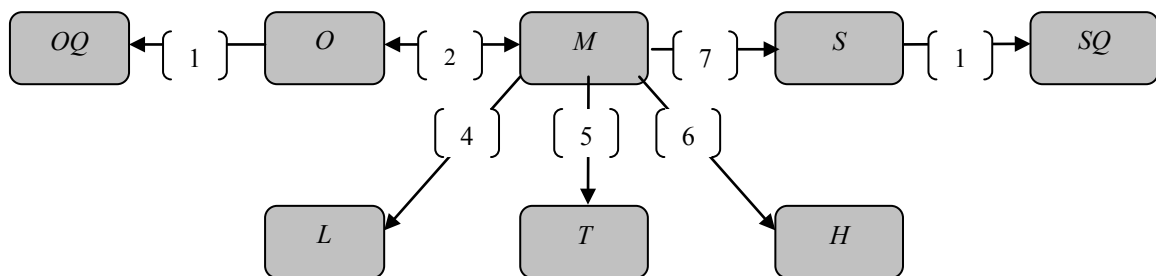


Рисунок 2.8 – Розгорнутий граф синтагматичного поєднання асоціацій навколо головної дії

Згідно з концепцією, що пропонується, кожна з цих КМО може представляти собою такий же граф або частину його у різного типу складнопідрядних реченнях, у механізмі впливу (вливанні) сенсу тощо. Зв'язки спрямованих синтагматичних асоціацій, у свою чергу, можна розкласти для ПМК у вигляді підмножин питальних займенни-

ків з метою побудови зручного лексичного процесора. Наприклад, для російської мови:

1. Какой? Какая? Какое?
2. Что делает? Что делают? Что делают? Что делаешь? тощо.
3. Кто? Что?
4. Где? Откуда? Куда?
5. Когда? В какое время? Сколько времени? Как долго?
6. Как? Каким образом?
7. Кого? Что? Чего? Кому? Чему? Кем? Чем? О ком? О чем?

На рисунку 2.9 показано результати формального поєднання образних концептів зі складовими КМО та ПМК. Запропонований підхід дозволив виявити взаємозв'язок між значимими концептуальними поняттями ІС та конструктами лексичного процесора через нові визначення МО, сенссполучення, синтагми і КМО. Для формалізації позначених кольором лексичних конструктів ПМК–речення–словосполучення–слово мають застосовуватися морфологічні, синтаксичні та семантичні відношення з урахуванням пріоритету образної основи концепції.

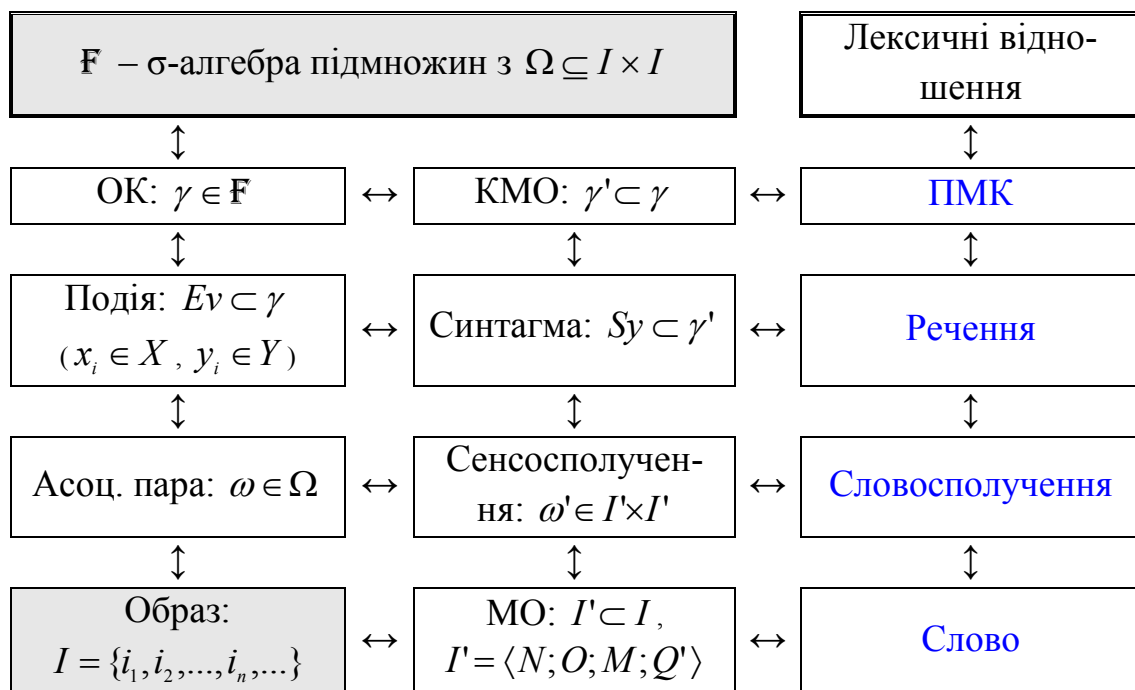


Рисунок 2.9 – Взаємозв'язок формальних понять образного аналізу ПМК

Отже, в запропоновану концепцію визначення формального взаємозв'язку між складовими конструкцій з образів, мовних образів та ПМК закладено пріоритет понять образного мислення для аналізу природно-мовної інформації. Первинну прагматичну основу сенсу-властивості отримано за допомогою 14-ти змістовних концептів образу, в тому числі понять події та стану.

2.3 Формалізація комутативної напівгрупи образних конструкцій на основі прикладної теорії першого порядку

З метою визначення оператора \oplus для ОК $\gamma \in \mathbb{F}$ у межах концептуального поняття ІС (п. 2.1) задамо формальну теорію Th як прикладну теорію першого порядку. З цією метою візьмемо за основу відомі результати теорії формальних систем [236, 248, 259] та врахуємо вигоди запропонованої концепції образного сенсу ПМК (п. 2.2).

1. Введемо скінченний алфавіт, до складу якого входять символи, що будуть використані як позначення:

- а) $Al = \{A, B, \dots, Z, x_1, x_2, \dots, x_n, t_1, t_2, t_3\}$ – змінних;
- б) $Con = \{\emptyset, 1, \dots, n\}$ – констант;
- в) $\{\backslash, \oplus\}$ – символів бінарних операцій, визначення яких дамо нижче;
- г) $\{=\}$ – бінарного предикатного символу «дорівнює» у значенні теорії множин;
- д) $\{\neg, \rightarrow, \forall\}$ – логічних зв'язок та кванторів, де \neg – заперечення (ні), \rightarrow – логічне слідування (якщо ..., то ...), \forall – квантор загальності;
- е) дужок «(», «)» та коми «,».

У відповідності до концепції розуміння сенсу елементів ЕК будемо вважати, що символи з в) позначають: \backslash – зв'язок між двома образами в асоціативній парі $\omega \in \Omega$, значення якого буде наведено нижче; \oplus – операція об'єднання образних конструкцій «PLUS ОК».

2. Визначимо процедури побудови термів (рядків символів) та формул (допустимих виразів) формальної теорії Th . Терми отримуємо за допомогою процедури конкатенації символів алфавіту:

- а. $\langle Терм \rangle := x_i j \mid x_i \in Al, j \in Con$;
- $\langle Терм \rangle := \langle Терм \rangle \langle Терм \rangle$.

Будемо позначати літерами $t_1, t_2, t_3 \in Al$ таким чином побудовані терми у асоціативній нормальній формі (АНФ)

$$\langle ANF\omega \rangle ::= x_i \setminus x_j \mid x_i, x_j \in Al;$$

$$\langle ANFтерм \rangle ::= \langle ANF\omega \rangle;$$

$$\langle ANFтерм \rangle ::= \langle ANFтерм \rangle \oplus \langle ANFтерм \rangle,$$

де $\langle ANF\omega \rangle$ будемо називати елементарним термом у АНФ.

б. Для спрощення сприйняття літерами $A, B, \dots, Z \in Al$ будемо

окремо позначати побудовані таким чином формули

$$\langle Формула \rangle ::= \langle ANFтерм \rangle;$$

$$\langle Формула \rangle ::= (\langle Формула \rangle);$$

$$\langle Формула \rangle ::= \neg \langle Формула \rangle;$$

$$\langle Формула \rangle ::= \langle Формула \rangle \rightarrow \langle Формула \rangle;$$

$$\langle Формула \rangle ::= (\forall x) \langle Формула \rangle.$$

Для зручності користування до складу алфавіту теорії Th введемо ще 3 логічні зв'язки, квантор та функціональний символ

$$A \& B ::= \neg(A \rightarrow \neg B);$$

$$A \vee B ::= \neg A \rightarrow B;$$

$$A \Leftrightarrow B ::= (A \rightarrow B) \& (B \rightarrow A);$$

$$(\exists x)(A) ::= \neg(\forall x)(\neg A);$$

$$x_i \times x_j ::= (x_i \setminus x_j) \oplus (x_j \setminus x_i),$$

де $\&$ – логічне «і»; \vee – логічне «або»; \Leftrightarrow – тоді і тільки тоді; \exists – квантор існування; \times – прикладний функціональний символ, визначення якого буде дано нижче через символ \setminus . У подальшому формулу A , в якій змінна $x_i \in Al$ або терм $t_j \in Al$ є зв'язаними по відношенню до одного з кванторів, будемо позначати через $A(x_i)$ або $A(t_j)$.

3. Виділимо множину формул, які вважаються схемами аксіом.

Логічні аксіоми (3.1–3.3 – числення висловлювань, 3.4 і 3.5 – числення предикатів першого порядку [248]):

$$3.1. A \rightarrow (B \rightarrow A).$$

$$3.2. (A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C)).$$

$$3.3. (\neg B \rightarrow \neg A) \rightarrow ((\neg B \rightarrow A) \rightarrow B).$$

$$3.4. \forall x_i A(x_i) \rightarrow A(t_i) \text{ [де } A(x_i) \text{ є формула з } Th \text{ і } t_i \text{ є терм з } Th, \text{ вільний для } x_i \text{ в } A(x_i)\text{]}.$$

3.5. $\forall x_i(A \rightarrow B) \rightarrow (A \rightarrow \forall x_i B)$ [за умови, якщо формула A не містить вільних входжень x_i].

Власні аксіоми (3.6–3.11 – аксіоми комутативної напівгрупи [259], 3.12–3.14 – прикладні аксіоми (продукції) теорії):

3.6. $\forall t_1 \forall t_2 \forall t_3 (t_1 \oplus (t_2 \oplus t_3) = (t_1 \oplus t_2) \oplus t_3)$ (асоціативність).

3.7. $\forall t_1 (t_1 = t_1)$ (рефлексивність).

3.8. $\forall t_1 \forall t_2 (t_1 = t_2 \rightarrow t_2 = t_1)$ (симетричність).

3.9. $\forall t_1 \forall t_2 \forall t_3 (t_1 = t_2 \rightarrow (t_2 = t_3 \rightarrow t_1 = t_3))$ (транзитивність).

3.10. $\forall t_1 \forall t_2 \forall t_3 (t_2 = t_3 \rightarrow (t_1 \oplus t_2 = t_1 \oplus t_3) \& (t_2 \oplus t_1 = t_3 \oplus t_1))$ (підстановка).

3.11. $\forall t_1 \forall t_2 (t_1 \oplus t_2 = t_2 \oplus t_1)$ (комутативність).

3.12. $\forall x_i, x_j, x_k (x_i j x_k \rightarrow x_j \setminus x_i \oplus x_k)$ (перетворення рядка на терми в АНФ).

3.13. $\forall x_i, x_j (x_i j \rightarrow x_j \setminus x_i)$ (кінцеве перетворення рядка на терм в АНФ).

3.14. $\forall x_i, x_j (x_i \setminus x_j \oplus x_i \setminus x_j \rightarrow x_i \setminus x_j)$ (скорочення терма в АНФ).

4. Визначимо скінченну множину правил виведення, які дозволяють отримати з деякої скінченної множини формул іншу множину формул

$A, A \rightarrow B \vdash B$ «*Modus ponens*»,

$A \vdash (\forall t)A$ «*правило узагальнення*»,

де запис $\Gamma \vdash A$ означає, що A є наслідком множини формул Γ .

Окрім теорем формальної теорії предикатів першого порядку, в теорії *Th* справедливими є такі власні теореми:

Теорема 1. $\langle \text{Терм} \rangle \rightarrow \langle \text{АНФтерм} \rangle$.

Доведення індукцією за довжиною виведення $B_1, B_2, \dots, B_k = B$:

а) $\langle \text{Терм} \rangle$ – гіпотеза;

б) $x_1 j$ – база індукції: згідно з 1-м визначенням терма (2.а.);

в) $x_j \setminus x_1$ – 3.13 до б);

г) $\langle \text{АНФтерм} \rangle$ – згідно з 1-м визначенням терма у АНФ;

д) $x_1 j x_2 i$ – або згідно з 2-м визначенням терма;

е) $x_j \setminus x_1 \oplus x_2 i$ – 3.12 до д);

- ж) $x_j \setminus x_1 \oplus x_i \setminus x_2 - 3.13$ до е);
 з) $\langle \text{АНФтерм} \rangle -$ згідно з 2-м визначенням терма у АНФ;
 и) $\underbrace{x_1 j x_2 i \dots x_k l}_{k-1} -$ індукційний перехід: згідно з 2-м визначен-

ням терма;

- к) $\langle \text{АНФтерм} \rangle \oplus x_k l - 3.12$ до и) $k - 1$ раз;
 л) $\langle \text{АНФтерм} \rangle \oplus x_l \setminus x_k - 3.13$ до к);
 м) $\langle \text{АНФтерм} \rangle -$ згідно з 2-м визначенням терма у АНФ.

Теорема 2. $\langle \text{АНФтерм} \rangle \rightarrow \langle \text{АНФ}q \rangle \oplus \langle \text{АНФ}?\rangle \oplus \langle \text{АНФ}a \rangle,$

де $\langle \text{АНФ}\omega \rangle = x_i \setminus x_j \mid x_i, x_j \in Al$ для зручності користування позначимо як $\langle \text{АНФ}?\rangle;$

$\langle \text{АНФ}a \rangle -$ всі елементарні терми з $\langle \text{АНФтерм} \rangle,$ в яких символ x_j є першим, потім рекурсивно підставляється наступний символ за принципом пошуку у глибину, але, якщо в рекурсії знаходиться $\langle \text{АНФ}?\rangle = x_j \setminus x_i,$ то ця гілка пошуку переривається (символ x_i та всі наступні за ним не враховуються);

$\langle \text{АНФ}q \rangle -$ всі інші окрім $\langle \text{АНФ}?\rangle \oplus \langle \text{АНФ}a \rangle$ елементарні терми, що складають $\langle \text{АНФтерм} \rangle.$

Доведення за всіма можливими варіантами побудови терма в АНФ:

- а) $\langle \text{АНФтерм} \rangle -$ гіпотеза;
 б) $x_i \setminus x_j \mid x_i, x_j \in Al -$ елементарний варіант: згідно з 1-м визначенням терма в АНФ;
 в) $\langle \text{АНФ}?\rangle -$ за визначенням у теоремі 2;
 г) $\emptyset \oplus \langle \text{АНФ}?\rangle \oplus \emptyset -$ згідно з аксіомою 3.15;
 д) $\langle \text{АНФ}q \rangle \oplus \langle \text{АНФ}?\rangle \oplus \langle \text{АНФ}a \rangle -$ за умови $\langle \text{АНФ}q \rangle = \emptyset, \langle \text{АНФ}a \rangle = \emptyset;$
 е) $x_j \setminus x_1 \oplus \langle \text{АНФ}?\rangle -$ перше можливе ускладнення варіанта б) згідно з 2-м визначенням терма в АНФ;
 ж) $\langle \text{АНФ}?\rangle \oplus x_j \setminus x_1 -$ згідно з аксіомою 3.11;
 з) $\langle \text{АНФ}?\rangle \oplus \langle \text{АНФ}a \rangle -$ за умови $\langle \text{АНФ}a \rangle = x_j \setminus x_1;$
 и) $\emptyset \oplus \langle \text{АНФ}?\rangle \oplus \langle \text{АНФ}a \rangle -$ згідно з аксіомою 3.15;
 к) $\langle \text{АНФ}q \rangle \oplus \langle \text{АНФ}?\rangle \oplus \langle \text{АНФ}a \rangle -$ за умови $\langle \text{АНФ}q \rangle = \emptyset;$

- л) $x_1 \setminus x_i \oplus \langle AN\Phi? \rangle$ – друге можливе ускладнення варіанту б) згідно з 2-м визначенням терма у АНФ;
- м) $\langle AN\Phi q \rangle \oplus \langle AN\Phi? \rangle$ – за умови $\langle AN\Phi q \rangle = x_1 \setminus x_i$;
- н) $\langle AN\Phi q \rangle \oplus \langle AN\Phi? \rangle \oplus \emptyset$ – згідно з аксіомою 3.15;
- о) $\langle AN\Phi q \rangle \oplus \langle AN\Phi? \rangle \oplus \langle AN\Phi a \rangle$ – за умови $\langle AN\Phi a \rangle = \emptyset$;
- п) $\langle AN\Phi a \rangle \oplus x_j \setminus x_2 \oplus x_1 \setminus x_3$ – знімаємо умову $\langle AN\Phi a \rangle = x_j \setminus x_1$ для з) згідно з 2-м визначенням терма у АНФ;
- р) $\langle AN\Phi a \rangle$ – за визначенням $\langle AN\Phi a \rangle$ у теоремі 2;
- с) $\langle AN\Phi q \rangle \oplus x_2 \setminus x_i \oplus x_3 \setminus x_1$ – знімаємо умову $\langle AN\Phi q \rangle = x_1 \setminus x_i$ для м) згідно з 2-м визначенням терма у АНФ;
- т) $\langle AN\Phi q \rangle$ – за визначенням $\langle AN\Phi q \rangle$ у теоремі 2, а саме тоді, коли терм в АНФ $x_2 \setminus x_i \oplus x_3 \setminus x_1 \oplus \langle AN\Phi? \rangle \oplus \langle AN\Phi a \rangle$ не допускає скорочення згідно з аксіомою 3.14.

Розглянемо модель формальної теорії *Th* як комутативну напівгрупу образних конструкцій. В межах моделі будемо вважати, що функціональні символи позначають такі зв'язки між двома мовними образами [56]: \setminus – зв'язок «головний–підлеглий», \times – зв'язок типу «підмет–присудок». Під термом будемо розуміти образну конструкцію простого речення, а під формулою теорії – образний аналог логічного природно-мовного виразу. Літерами x_1, x_2, \dots, x_n будемо позначати окремі образи з множини $I = \{x_1, x_2, \dots, x_n\}$, літерами t_1, t_2, t_3 – терми в АНФ, A, B, \dots, X – формули, Y – невідомий підмет (об'єкт дії), Z – невідомий присудок (метод). Елементарний терм в АНФ $\langle AN\Phi\omega \rangle \mid \langle AN\Phi? \rangle$ будемо називати асоціативною парою образів, де \mid – позначення оператора АБО в нотації Бекуса-Наура. Терми або образні конструкції будуються з природно-мовних речень на основі такого правила 1: речення з k слів записується у вигляді рядка з $2 \cdot k$ символів, де кожному i -му слову речення ставиться у відповідність мовний образ $x_i \in Al$, а після нього записується $j \in Con$ як вказівник на інший образ x_j цього речення, що є головним до підлеглому образу x_i . Якщо у реченні зустрічаються однорідні члени, то можливі випадки $(x_1 \& x_2)j \rightarrow x_j \setminus x_1 \oplus x_j \setminus x_2$ або

$$(x_1 \& x_2)j \oplus \langle \text{АНФтерм} \rangle \oplus x_j \setminus x_1 \rightarrow x_j \setminus x_1 \oplus x_j \setminus x_2 \oplus \\ \langle \text{АНФтерм} \rangle \oplus x_1 \setminus x_j \oplus x_2 \setminus x_j.$$

Обмеження розглянутої моделі:

- природно-мовні речення обов'язково мають і підмет і присудок, інакше їх штучно вводять за допомогою символів Y та/або Z ;
- правило 1 застосовується тільки до значимих слів речення, що відповідають мовним образам, а розділові знаки, прийменники та службові слова речення не враховуються.

В межах моделі доведені теореми формальної теорії Th отримують таку інтерпретацію:

Теорема 1. Будь-який терм, що відповідає природно-мовному реченню (синтагмі) та створений на основі правила 1, можна представити у вигляді терма в АНФ: $\langle \text{Терм} \rangle \rightarrow \langle \text{АНФтерм} \rangle$.

Теорема 2. Якщо з речення, що представлене у вигляді терма в АНФ $\langle \text{АНФтерм} \rangle$, вибрати одну асоціативну пару як питальний займенник, то всі безпосередньо залежні від цієї пари елементарні терми в АНФ складуть відповідь, а всі інші елементарні терми з $\langle \text{АНФтерм} \rangle$ – питальне речення: $\langle \text{АНФтерм} \rangle \rightarrow \langle \text{АНФ}q \rangle \oplus \langle \text{АНФ} ? \rangle \oplus \langle \text{АНФ}a \rangle$.

Для зручності застосування моделі формальної теорії Th в ПМК введемо правило 2: $\langle \text{АНФтерм} \rangle ::= \langle \text{АНФ} ? \rangle \langle tQ \rangle ? \langle tA \rangle$, де

$$\langle tQ \rangle ::= (x_i \mid \langle \text{АНФ}q \rangle = \emptyset) \mid \\ (x_i x_l \dots x_m x_k \mid \langle \text{АНФ}q \rangle = x_i \setminus x_l \oplus \dots \oplus x_m \setminus x_k); \\ \langle tA \rangle ::= (x_j \mid \langle \text{АНФ}a \rangle = \emptyset) \mid \\ (x_j x_l \dots x_m x_k \mid \langle \text{АНФ}a \rangle = x_j \setminus x_l \oplus \dots \oplus x_m \setminus x_k);$$

? – додатковий знак, що позначає закінчення питальної частини $\langle \text{АНФтерм} \rangle$.

Отримані для $\langle tQ \rangle$ та $\langle tA \rangle$ рядки символів $x_i x_l \dots x_m x_k$ переписуються шляхом вилучення зліва направо тих символів, що раніше повторювалися. Формально для другого символу

$$x_1 x_2 \rightarrow ([x_2 = x_1] x_1, x_1 x_2), \text{ для } k\text{-го символу}$$

$$x_1 x_2 \dots x_k \rightarrow ([x_k = x_1 \mid x_k = x_2 \mid \dots \mid x_k = x_{k-1}] x_1 x_2 \dots x_{k-1}, x_1 x_2 \dots x_k).$$

З метою демонстрації можливостей комутативної напівгрупи ОК як моделі формальної теорії Th розглянемо приклади речень на українській мові.

Приклад 1. Рече (та) стогне Дніпр широкий $(x_{11}x_{12}x_{13}x_{14})$.

Згідно з правилом 1, будуємо терм

$$(x_{11} \& x_{12})13x_{13}11x_{14}13 \rightarrow x_{11}13x_{12}13x_{13}11x_{13}12x_{14}13;$$

продукція 3.12 до підрядка $x_{11}13x_{12}$ веде до

$$x_{13} \setminus x_{11} \oplus x_{12}13x_{13}11x_{13}12x_{14}13;$$

продукція 3.12 до підрядка $x_{12}13x_{13}$ веде до

$$x_{13} \setminus x_{11} \oplus x_{13} \setminus x_{12} \oplus x_{13}11x_{13}12x_{14}13;$$

продукція 3.12 до підрядка $x_{13}11x_{12}$ веде до

$$x_{13} \setminus x_{11} \oplus x_{13} \setminus x_{12} \oplus x_{11} \setminus x_{13} \oplus x_{13}12x_{14}13;$$

продукція 3.12 до підрядка $x_{13}12x_{14}$ веде до

$$x_{13} \setminus x_{11} \oplus x_{13} \setminus x_{12} \oplus x_{11} \setminus x_{13} \oplus x_{12} \setminus x_{13} \oplus x_{14}13;$$

продукція 3.13 до підрядка $x_{14}13$ веде до

$$x_{13} \setminus x_{11} \oplus x_{13} \setminus x_{12} \oplus x_{11} \setminus x_{13} \oplus x_{12} \setminus x_{13} \oplus x_{13} \setminus x_{14} - \text{маємо терм в АНФ.}$$

Отже, початкова природно-мовна конструкція в АНФ має такий вигляд:

рече \ Дніпр \oplus

стогне \ Дніпр \oplus

Дніпр \ рече \oplus

Дніпр \ стогне \oplus

Дніпр \ широкий .

Позначимо $\langle AN\Phi? \rangle := x_{13} \setminus x_{14}$ словом $\langle \text{який?} \rangle$. За теоремою 2 $\langle AN\Phi a \rangle \rightarrow \emptyset$, $\langle AN\Phi q \rangle \rightarrow x_{13} \setminus x_{11} \oplus x_{11} \setminus x_{13} \oplus x_{13} \setminus x_{12} \oplus x_{12} \setminus x_{13}$.

Тоді, згідно з правилом 2, $\langle tA \rangle \rightarrow x_{14}$, а $\langle tQ \rangle \rightarrow x_{13}x_{11}x_{12}$. Отже, маємо такий результат:

який? Дніпр рече стогне ? широкий.

Неважко довести еквівалентність графової моделі та представленої в цій роботі теорії Th , що дозволяє використати відомі алгоритми пошуку на графах для розв'язання прикладних задач знаходження оптимального шляху, обходу дерева графа та пошуку при обробленні ПМК. Ілюструє приклад застосування моделі теорії Th дерево графа речення, представлене на рисунку 2.10.

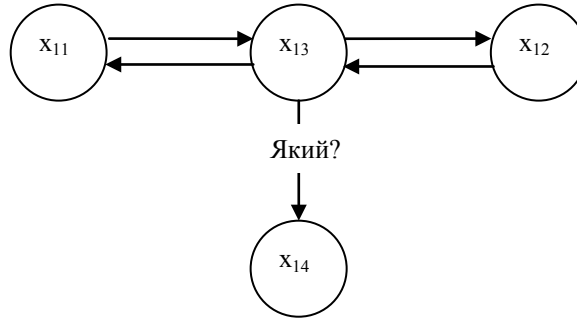


Рисунок 2.10 – Дерево графа терма $x_{11}13x_{13}11x_{12}13x_{13}12x_{14}13$ з виділенням асоціативної пари $x_{13} \setminus x_{14}$

Приклад 2. Мені (аж) страшно (як) згадаю оту хатину край села ($x_1x_2x_3x_4x_5x_6x_7$).

Згідно з правилом 1, будуємо терм

$$x_12x_21x_32x_45x_53x_65x_76;$$

6 разів продукція 3.12 до підрядка $x_12x_21x_32x_45x_53x_65x_7$

$$x_2 \setminus x_1 \oplus x_1 \setminus x_2 \oplus x_2 \setminus x_3 \oplus x_5 \setminus x_4 \oplus x_3 \setminus x_5 \oplus x_5 \setminus x_6 \oplus x_76;$$

продукція 3.13 до підрядка x_76

$$x_2 \setminus x_1 \oplus x_1 \setminus x_2 \oplus x_2 \setminus x_3 \oplus x_5 \setminus x_4 \oplus x_3 \setminus x_5 \oplus x_5 \setminus x_6 \oplus x_6 \setminus x_7 - \text{маємо}$$

терм в АНФ.

Отже, початкова природно-мовна конструкція в АНФ має такий вигляд:

$$\begin{aligned} &\text{страшно} \setminus \text{мені} \oplus \\ &\text{мені} \setminus \text{страшно} \oplus \\ &\text{страшно} \setminus \text{згадаю} \oplus \\ &\text{хатину} \setminus \text{оту} \oplus \\ &\text{згадаю} \setminus \text{хатину} \oplus \\ &\text{хатину} \setminus \text{край} \oplus \\ &\text{край} \setminus \text{села}. \end{aligned}$$

На рисунку 2.11 представлено дерево графа речення з виділенням двох пар.

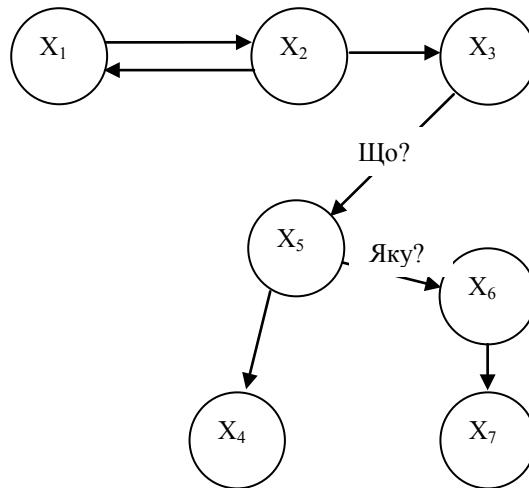


Рисунок 2.11 – Дерево графа терма $x_1 2x_2 1x_3 2x_4 5x_5 3x_6 5x_7 6$ з виділенням асоціативних пар $x_3 \setminus x_5$ та $x_5 \setminus x_6$

Позначимо $\langle AN\Phi? \rangle ::= x_3 \setminus x_5$ словом $\langle \underline{\text{що?}} \rangle$. За теоремою 2
 $\langle AN\Phi a \rangle \rightarrow x_5 \setminus x_4 \oplus x_5 \setminus x_6 \oplus x_6 \setminus x_7$,
 $\langle AN\Phi q \rangle \rightarrow x_2 \setminus x_1 \oplus x_1 \setminus x_2 \oplus x_2 \setminus x_3$.

Тоді, згідно з правилом 2, $\langle tA \rangle \rightarrow x_5 x_4 x_6 x_7$, а $\langle tQ \rangle \rightarrow x_2 x_1 x_3$.
 Отже, маємо такий результат:

що? страшно мені згадаю ? хатину оту край села.

Тепер позначимо $\langle AN\Phi? \rangle ::= x_5 \setminus x_6$ словом $\langle \underline{\text{яку?}} \rangle$. За теоремою 2
 $\langle AN\Phi a \rangle \rightarrow x_6 \setminus x_7$,
 $\langle AN\Phi q \rangle \rightarrow x_2 \setminus x_1 \oplus x_1 \setminus x_2 \oplus x_2 \setminus x_3 \oplus x_3 \setminus x_5 \oplus x_5 \setminus x_4$.

Тоді, згідно з правилом 2, $\langle tA \rangle \rightarrow x_6 x_7$, а $\langle tQ \rangle \rightarrow x_2 x_1 x_3 x_5 x_4$.
 Зрештою, маємо:

яку? страшно мені згадаю оту хатину? край села.

Отже, наведені приклади демонструють інтуїтивну зрозумілість результатів застосування комутативної напівгрупи ОК як моделі формальної теорії *Th* до ПМК у вигляді речень на українській мові. Приклади застосування теорії *Th* для російськомовних речень наведено у Додатку А. На відміну від відомих формальних теорій подальший розвиток теорії *Th* здійснено за рахунок бінарного оператора спрямованого асоціативного зв'язку та поняття АНФ згідно з підходом до об'єктивного аналізу текстової інформації. Побудована на основі моделі при-

кладної теорії першого порядку Th комутативна напівгрупа образних конструкцій забезпечує представлення ОК природно-мовної синтагми у вигляді 3-х складових питальної конструкції мовних образів.

2.4 Метод побудови нечіткого відношення образного сенсу

Реалізуємо розглянуту раніше концепцію представлення сенсу ПМК за допомогою введення функціонального простору з мірою. Вперше підхід до відокремлення у понятті сенс пізнавальної діяльності сенсу-властивості, який визначає нові знання, вміння та навички від образного сенсу, що вимірюється шляхом порівняння з одиницею сенсу на основі сили асоціативного зв'язку, було опубліковано в роботах [31, 34, 37, 41, 43]. На відміну від запропонованого підходу в роботах з моделювання актуальних аспектів образного мислення, розглянутих у [237] та, особливо, в найбільш близькому дослідженні [238], не враховано онтогенез пізнавальної, у тому числі мовленнєвої діяльності людини.

Феноменологічний характер підходу дозволяє припустити, що в основу простору має бути покладена нечітка міра. Розглянемо можливості формалізації природного методу побудови функцій належності для тієї нечіткої сукупності (множини) основних образних понять, які складають когнітивний простір особистості. Запропонована у [41] функція образного сенсу як характеристика сили елементарного асоціативного зв'язку між парою образів за формальними ознаками близька до функції належності нечіткого відношення, а формалізований у [43] онтогенетичний спосіб побудови таких функцій моделює природний шлях накопичення знань людиною. Нерозв'язаними лишаються задачі визначення особливостей ОК як бінарних нечітких відношень, для яких функцією належності є функція образного сенсу. Отже, задачами дослідження є побудова бінарних нечітких відношень на основі онтогенетичного підходу до представлення образного сенсу, визначення нечіткого відношення сенсу та його основних властивостей, інтерпретація базових операцій з нечіткими відношеннями сенсу з точки зору задач пізнавальної діяльності людини, формалізація простору з нечіткою мірою.

У відповідності до [239] бінарне нечітке відношення, що задане на одній базисній множині (універсумі) образів I , визначимо як нечітке відношення

$$Q = \{ \langle i_l, i_j \rangle, \mu_Q(\langle i_l, i_j \rangle) \}, \quad (2.7)$$

де $\mu_Q(\langle i_l, i_j \rangle)$ – функція належності бінарного нечіткого відношення, що задається як відображення $\mu_Q : I \times I \rightarrow [0, 1]$. У виразі (2.7) через $\langle i_l, i_j \rangle$ позначено елементарний терм в АНФ $\langle \text{АНФ}\omega \rangle$ як кортеж з двох елементів, причому $i_l \in I, i_j \in I$. Якщо носій Q_s нечіткого відношення Q є скінченним, то потужність цього нечіткого відношення чисельно дорівнює кількості кортежів його носія і позначається як $\text{card}(Q_s)$.

На практиці найчастіше використовуються такі способи формального визначення бінарних нечітких відношень (2.7) на універсумі I , як [240]:

1. Список з безпосереднім перерахуванням всіх кортежів нечіткого відношення та відповідних цим кортежам значень функцій належності – у випадку скінченного та невеликого числа таких кортежів.
2. Аналітично у формі деякого математичного виразу, що забезпечує можливість обчислення значення функцій належності для кожного з кортежів. У цьому випадку нечітке відношення записують у вигляді (2.7), де

$$\mu_Q(\langle i_l, i_j \rangle) = f(i_l, i_j) - \quad (2.8)$$

певна задана функція двох змінних з універсуму I , що задовольняє стандартні вимоги до функції належності. Функцію (2.8) називають також [108] функцією, що породжує нечітке відношення (2.7).

3. Графічно у формі певної поверхні або сукупності окремих точок у тривимірному просторі, в якому дві координати (незалежні змінні) відповідають значенням елементів i_l та i_j з універсуму I , а третя координата – функції належності зі значенням з інтервалу $[0, 1]$.

4. У формі квадратної матриці A_Q нечіткого відношення Q , рядкам якої відповідають перші елементи кортежів $i_l \in I$, а стовпцям – другі елементи кортежів $i_j \in I$. Елементами (l, j) матриці A_Q є відповідні значення функції належності $\mu_Q(\langle i_l, i_j \rangle)$ цього відношення.

5. У формі орієнтованого нечіткого графа $G = (V, E, \mu_G)$, де $V = \{v_1, v_1, \dots, v_n\}$ – множина вершин нечіткого графа, а $E = \{e_1, e_1, \dots, e_m\}$ – множина дужок нечіткого графа, кожній з яких приписано значення функції належності. Натуральні числа n та m визначають кількість вершин та дужок нечіткого графа ($card(Q) = m$), а кожному елементу універсуму $i_l \in I$ відповідає окрема вершина $v_l \in V$ ($l = \overline{1, n}$) орієнтованого нечіткого графа. Довільному k -му кортежу нечіткого відношення $\langle i_l, i_j \rangle \in Q$ відповідає дужка графа $e_k = \langle v_l, v_j \rangle$ з початком у вершині v_l , закінченням у вершині v_j та значенням функції належності $\mu_Q(\langle i_l, i_j \rangle)$.

Для задач дослідження оберемо способи 2, 4 та 5 визначення нечітких відношень.

Бінарне нечітке відношення (2.7) застосуємо як онтогенетичну характеристику множини Ω , тоді функцію належності (2.8) можна вважати природною чисельною мірою сенсу. Згідно з [41], значення $\mu_Q(\langle i_l, i_j \rangle) = 1$ називатимемо одиницею сенсу розміром один *Saw* (Синтагматичної асоціації вага, рос. – Синтагматической ассоциации вес) або *Saw* (Syntagmatic association weight). Окрім цього, англійське слово *Saw* (друга форма неправильного дієслова *to See*) перекладається як «побачене» і, тим самим, вказує на образний шлях появи сенсу в когнітивній системі. При такому підході значення елемента (l, j) матриці A_Q або k -ї дужки $e_k = \langle v_l, v_j \rangle$ графа G залежить від статистики появи зв'язку для кортежу $\langle i_l, i_j \rangle$ за час спостереження L вхідних ОК. Отже, в загальному вигляді функцію належності (базовий рівень) нечіткого відношення сенсу для пар образів задамо як

$$\mu_Q(\langle i_l, i_j \rangle) = f(k_{lj}, t_L), \quad (2.9)$$

де k_{lj} – кількість зафіксованих інфологічною системою зв'язків між l -м та j -м образами на момент часу t_L . Значення k_{lj} неважко отримати шляхом підрахунку кількості виконаних операцій скорочення терма в АНФ для кортежу $\langle i_l, i_j \rangle$ у відповідності до аксіоми 3.14 формальної теорії (п. 2.3). Додатково відношення сенсу має вра-

ховувати такі важливі властивості ІС, як емоційний стан, потреби (мотиви) та рефлекси або інші корисні для функціонування системи результати зовнішнього навчання.

У відповідності до концептуального підходу (п. 2.1) деталізуємо функцію належності, що породжує бінарне нечітке відношення сенсу (2.7) на таких чотирьох послідовних рівнях, побудованих на базовому (2.9):

1. Рівень імовірнісного прогнозування – з метою нормування функції належності у проміжку $[0, 1]$ передбачено розрахунок статистичної оцінки λ (математичного сподівання): якщо $k_{\Sigma} = \sum_{l=1}^n \sum_{j=1}^n k_{lj}$, а m – кількість ненульових елементів матриці A_Q , то $\lambda = k_{\Sigma} / m$ – в цьому випадку застосуємо відому сигмоїдальну функцію

$$\mu_Q(< i_l, i_j >) = f_1(k_{lj}, \lambda) = 1 / (1 + e^{-k_{lj} + \lambda}). \quad (2.10)$$

Внаслідок нормування з'являється характерна властивість функції належності, отриманої за онтогенетичним методом – середнє значення

$$\overline{\mu_Q} = \frac{1}{m} \sum_{j=1}^m \mu_{Qj} = 0,5 .$$

2. Рівень врахування емоційного стану – введено можливість врахування бінарної моделі емоцій ІС за рахунок показника $\mu = \{\dots, -2, -1, 1, 2, \dots\}$, тоді

$$\mu_Q(< i_l, i_j >) = f_2(k_{lj}, \lambda, \mu) = 1 / (1 + e^{\frac{k_{lj} - \lambda}{|\mu|}}). \quad (2.11)$$

При $\mu = -1 \vee 1$ емоції не впливають на сенс функціонування ІС, а функція належності (2.11) вироджується у функцію (2.10). Збільшення показника μ симетрично згладжує сигмоїдальну функцію f_2 , що продемонстровано на рисунку 2.12.

3. Рівень врахування мотиваційної компоненти на основі образів-центрів потреб – запропоновано моделлю мотиву інфологічної системи на момент часу t_l вважати досягнення образу-центру потреби j' , а

також розрахувати дисперсію та середньоквадратичне відхилення результатів спостережень k_{lj} як $D = \frac{1}{m} \sum_{l=1}^n \sum_{j=1}^n (k_{lj} - \lambda)^2 \mid k_{lj} > 0$ і $\sigma = \sqrt{D}$.

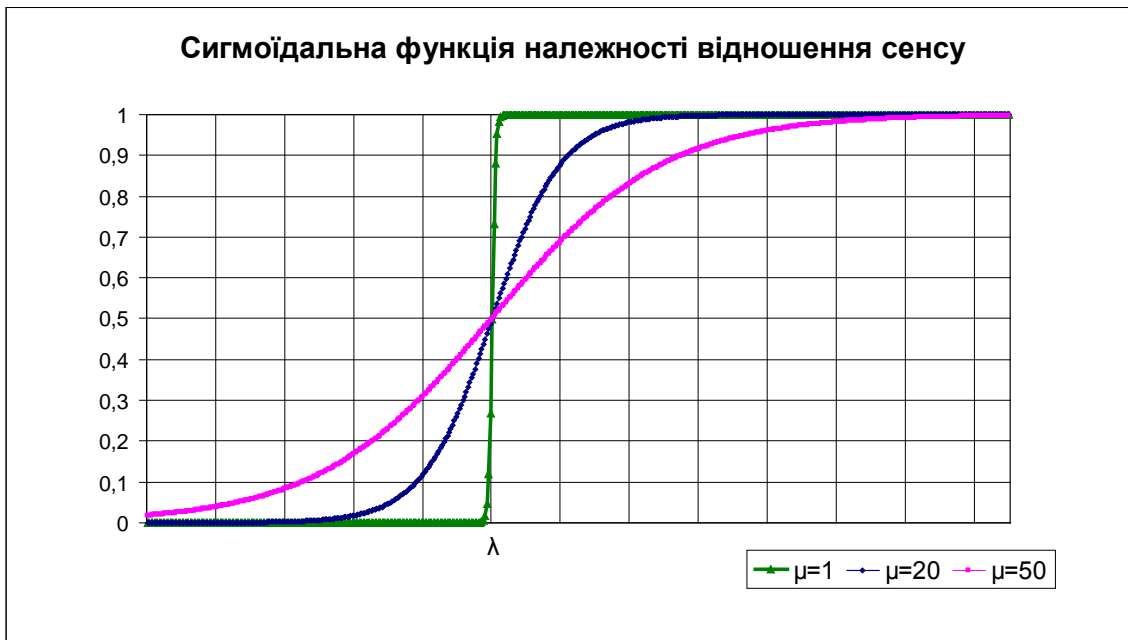


Рисунок 2.12 – Вплив показника μ на функцію належності (2.11)

Тоді, в залежності від ступеня наближення r пари образів $\langle i_l, i_j \rangle$ до j' , функцію (2.11) можна зміщувати вліво за віссю абсцис шляхом зменшення математичного сподівання для цієї пари $\lambda_{lj} = \lambda - r \cdot \sigma$, де $r = \{0, 1, 2, 3\}$, зрештою маємо

$$\mu_Q(\langle i_l, i_j \rangle) = f_3(k_{lj}, \lambda_{lj}, \sigma, \mu, i') = 1 / (1 + e^{\frac{k_{lj} - \lambda_{lj}}{|\mu|}}). \quad (2.12)$$

Розглянутий підхід передбачає застосування окремого алгоритму для визначення ступеня наближеності r пари $\langle i_l, i_j \rangle$ до образу-потреби j' , наприклад:

- а) якщо $j = j'$, то $r = 3$, інакше
- б) якщо $\exists e_k \in E \mid e_k = \langle v_j, v_{j'} \rangle$, то $r = 2$, інакше
- в) якщо $\exists e_k \in E \mid e_k = \langle v_l, v_{j'} \rangle$, то $r = 1$, інакше $r = 0$.

4. Рівень врахування рефлексів та результатів зовнішнього навчання – запропоновано фіксувати в ІС статистику результатів образ-

ної реакції y'_i на вхідну ОК x'_i при мотиваційній меті i' як зміну емоційного параметра $\Delta\mu$. Тоді, за наявності у попередньому досвіді схожих ситуацій $\langle i', x'_i \rangle$ для певної підмножини пар $e_{ij} \in E'$, де $E' \subseteq y'_i$ застосовується збільшення або зменшення математичного сподівання $\lambda_{ij} = \lambda \pm r \cdot \sigma$. Вибір $r = \{0, 1, 2, 3\}$ залежить виключно від знака та значення $\Delta\mu$ та потребує окремого алгоритму, а функція належності матиме вигляд

$$\mu_Q(\langle i_l, i_j \rangle) = f_4(k_{ij}, \lambda_{ij}, \sigma, \mu, i', x', \Delta\mu) = 1 / (1 + e^{\frac{k_{ij} - \lambda_{ij}}{|\mu|}}). \quad (2.13)$$

Зауважимо, що, на відміну від (2.10) та (2.11), у вищих рівнях функції належності відношення сенсу (2.12) та (2.13) внаслідок локальних зсувів математичного сподівання зникає властивість $\overline{\mu_Q} = 0,5$, що, на думку автора, свідчить про належну формальну інтерпретацію відомих фактів з психології та фізіології щодо протиріч між загальноприйнятим (середньостатистичним) сенсом і діями під впливом сильних мотивів або набутих рефлексів.

Властивості відношення сенсу. Відомо [241], що кількість кортежів нечіткого відношення Q може розглядатися як змінна від мінімального (ядро Q_p) до максимального (носій Q_s), а проміжні значення цієї змінної узагальнює нечітке відношення α -рівня. Будемо вважати відношенням сенсу α -рівня звичайним відношенням $Q_\alpha = \{\langle i_l, i_j \rangle \mid \mu_Q(\langle i_l, i_j \rangle) \geq \alpha\}$, де $\langle i_l, i_j \rangle \in I \times I$, $\alpha \in [0, 1]$. Зрозуміло, що $card(Q_p) \leq card(Q_\alpha) \leq card(Q_s) = m$.

Особливістю онтогенетичного методу моделювання сенсу-параметра є те, що рівність двох нечітких відношень сенсу може бути досягнута лише в одному випадку – коли дві інфологічні системи отримували абсолютно однакові множини вхідних подій $X = \{x_1, x_2, \dots\}$ і то лише на рівні (2.10) визначення функції належності. Поняття нечіткого домінування принципово відсутнє для цього методу.

Бінарне нечітке відношення сенсу Q згідно з [239] не є рефлексивним, проте є антирефлексивним, оскільки $\mu_Q(\langle i_j, i_j \rangle) = 0$, $i_j \in I$ –

всі елементи головної діагоналі матриці A_Q дорівнюють 0. Така властивість відношення Q пов'язана з тим, що визначити образ можна лише завдяки множині його зв'язків з іншими образами, але не з самим собою – вирази на зразок «масло масляне» не додають сенсу.

Також для Q можна вважати відсутнім властивості симетричності, асиметричності та антисиметричності [108], що виразно проявляється в матриці A_Q . З іншого боку, розглянуті вище особливості онтогенетичного методу не дають підстав стверджувати про транзитивність або котранзитивність нечіткого відношення сенсу Q . Отже, не варто говорити про подібність (толерантність), відмінність та еквівалентність, якщо в основу покладено відношення образного сенсу – множини образів та асоціативних зв'язків між ними не мають цих властивостей.

Отже, внаслідок дослідження отримано формальні характеристики онтогенетичного методу створення бінарного нечіткого відношення образного сенсу Q інфологічної системи S_Q шляхом моделювання понять мотиваційної мети та емоційного стану. Запропоновано принципи послідовної багаторівневої побудови функції належності $\mu_Q(<i_1, i_j >)$, що породжує нечітке відношення Q , визначено основні властивості Q , в тому числі характерну властивість онтогенетичного методу $\overline{\mu_Q} = 0,5$.

2.5 Дослідження простору образного сенсу з нечіткою мірою

Розглянемо можливості аксіоматизації простору з мірою на основі запропонованого нечіткого відношення образного сенсу Q . Відомо, що мірою називають функцію множини $m: P(X) \rightarrow R^+$, яка відповідає таким трьом аксіомам [232]:

1. $\forall A \subseteq X \Rightarrow m(A) \geq 0, m(\emptyset) = 0$.

2. $A \subseteq B \Rightarrow m(A) \leq m(B)$.

3. Якщо $A, B \in P(X)$ та $A \cap B = \emptyset$, то $m(A \cup B) = m(A) + m(B)$, де $P(X)$ – множина всіх підмножин X (σ -алгебра), $R^+ = [0, \infty]$ – множина додатних дійсних чисел.

При $R^+ = [0, 1]$ ці аксіоми визначають ймовірнісну міру, в основу якої покладено базове поняття події з можливостями експерименталь-

ної перевірки. В теорії нечітких мір замість події застосовується поняття суб'єктивного ступеня впевненості людини в цій події, що також зручно обмежити відрізком $[0, 1]$. За рахунок узагальнення поняття міри виникає можливість уникнути обмежувальної вимоги адитивності, що забезпечують відомі λ -нечіткі міри Сугено для визначення ступеня необхідності, довіри, правдоподібності, можливості тощо [240].

Відповідно до концепції образного сенсу складових ПМК та моделювання пізнавальної діяльності простір образного сенсу неможливо обмежити мірою на відрізку $[0, 1]$. Не дивлячись на те, що вчитель-експерт ставить оцінку студенту за будь-якою системою оцінювання знань саме на основі особистого суб'єктивного ступеня впевненості, проте об'єктом порівняння для нього при визначенні оцінки служать власні знання. Отже, результати пізнавальної діяльності мають властивість необмеженого або обмеженого фізичною тілесністю зростання, а відповідний до них сенс-параметр будемо вимірювати на $R^+ = [0, \infty]$. Також зауважимо, що властивості операції \oplus об'єднання ОК, введеної у формальній теорії (п. 2.3), не суперечать аксіомі адитивності.

На основі базового поняття знань/вмін/навичок як результатів моделювання пізнавальної діяльності та властивостей бінарного нечіткого відношення сенсу з урахуванням позначень концепції (п. 2.1) задамо такі аксіоми *нечіткої міри сенсу* $Se : \mathbf{F} \rightarrow [0, \infty]$:

1. $\forall \gamma \subseteq \Omega \Rightarrow Se(\gamma) \geq 0, Se(\emptyset) = 0, Se(\Omega) \leq \infty$.
2. $\forall \omega \in \Omega \Rightarrow Se(\omega) \leq 1, Se(\omega) = \mu_Q(\omega)$.
3. Якщо $\gamma_1, \gamma_2 \in \mathbf{F}$ та $\gamma_1 \subseteq \gamma_2$, то $Se(\gamma_1) \leq Se(\gamma_2)$ (монотонність).
4. Якщо $\gamma_i \in \mathbf{F}$, де $\{\gamma_i, i = 1, 2, \dots\}$ є монотонною послідовністю $\gamma_1 \supseteq \gamma_2 \dots \supseteq \gamma_i \supseteq \dots$, то $\lim_{i \rightarrow \infty} Se(\gamma_i) = Se(\lim_{i \rightarrow \infty} \gamma_i)$ (неперервність).
5. Якщо $\{\gamma_i\}_{i=1}^{\infty} \in \mathbf{F}$ – зліченне сімейство множин з \mathbf{F} , що попарно не перетинаються, тобто $\gamma_i \cap \gamma_j = \emptyset, i \neq j$, то $Se(\bigcup_{i=1}^{\infty} \gamma_i) = \sum_{i=1}^{\infty} Se(\gamma_i)$ (σ -адитивність).

Під *простором сенсу з нечіткою мірою* будемо розуміти трійку (Ω, \mathbf{F}, Se) , де Se – нечітка міра сенсу або сенс-параметр, що визнача-

ється на основі аксіом 1–5. Запропонована міра Se вважається σ -скінченною, оскільки існує зліченне сімейство множин, що вимірюється $\{\gamma_i\}_{i=1}^{\infty} \in \mathbf{F}$, таке що $Se(\gamma_i) < \infty, i \in N^+$ та $\Omega = \bigcup_{i=1}^{\infty} \gamma_i$, де N^+ – множи-

на додатних натуральних чисел. Використання σ -скінченної міри Se приводить до того, що весь простір сенсу може бути представленим у вигляді зліченного об'єднання множин, що вимірюється на основі скінченної міри. На основі аксіоми 5. задамо означення 1: якщо ОК x та y перетинаються, то $Se(x \cup y) = Se(x) + Se(y) - Se(x \cap y)$.

Покажемо, що можливість подання сенсу у числовому вигляді також дозволяє розглядати простір упорядкованих пар образів Ω як топологічний і квазіметричний. Простір асоціативних пар Ω є топологічним, оскільки множина \mathbf{F} її підмножин є топологією на Ω , для якої виконуються умови [242]:

- 1) $\Omega \in \mathbf{F}, \emptyset \in \mathbf{F}$ – множина Ω і пуста множина належать \mathbf{F} ;
- 2) об'єднання та перетин довільного сімейства множин, що належать \mathbf{F} , належить \mathbf{F} (згідно з властивостями σ -алгебри).

Простір асоціативних пар Ω є квазіметричним простором [242], оскільки будь-яким двом елементам $\omega_i, \omega_j \in \Omega$ ставиться у відповідність невід'ємне число q таке, що $q(\omega_i, \omega_i) = 0$, а для будь-якої трійки $\omega_1, \omega_2, \omega_3 \in \Omega$ виконується нерівність трикутника $q(\omega_1, \omega_3) \leq q(\omega_1, \omega_2) + q(\omega_2, \omega_3)$ за умови, що квазівідстань $q(\omega_i, \omega_j) = |Se(\omega_i) - Se(\omega_j)|$. Простір Ω не можна вважати метричним простором, оскільки не виконується умова про те, що $q(\omega_i, \omega_j) = 0$ тоді і тільки тоді, коли $\omega_i = \omega_j$. Також неважко показати, що множина ОК з \mathbf{F} також є квазіпростором з квазівідстанню q .

Операції з відношеннями сенсу. Розглянемо пізнавальну інтерпретацію основних операцій над бінарними нечіткими відношеннями сенсу $Q = \{ \langle i_l, i_j \rangle, \mu_Q(\langle i_l, i_j \rangle) \}$ та $R = \{ \langle i_l, i_j \rangle, \mu_R(\langle i_l, i_j \rangle) \}$, що задані на декартовому добутку $I \times I$ того ж універсуму образів, тобто $\langle i_l, i_j \rangle \in I \times I$. Будемо вважати, що відношення Q відповідає інфологічній системі S_Q , а відношення R відповідає іншій ІС S_R .

Результатом *перетину* нечітких відношень сенсу Q та R назвемо таке нечітке відношення сенсу $P = \{ \langle i_l, i_j \rangle, \mu_P(\langle i_l, i_j \rangle) \} = Q \cap R$, яке задане на тому ж самому декартовому добутку $I \times I$ та функція належності якого визначається за формулою

$$\mu_P(\langle i_l, i_j \rangle) = \min \{ \mu_Q(\langle i_l, i_j \rangle), \mu_R(\langle i_l, i_j \rangle) \}. \quad (2.14)$$

Операція перетину «І» визначає нове нечітке відношення сенсу P , що має спільні знання двох незалежних інфологічних систем S_Q та S_R .

Результатом *об'єднання* нечітких відношень сенсу Q та R назвемо таке нечітке відношення сенсу $P = \{ \langle i_l, i_j \rangle, \mu_P(\langle i_l, i_j \rangle) \} = Q \cup R$, яке задане на тому ж самому декартовому добутку $I \times I$ та функція належності якого визначається за формулою

$$\mu_P(\langle i_l, i_j \rangle) = \max \{ \mu_Q(\langle i_l, i_j \rangle), \mu_R(\langle i_l, i_j \rangle) \}. \quad (2.15)$$

Операція об'єднання («АБО») визначає нове нечітке відношення сенсу P , що має знання ІС S_Q або ІС S_R .

Результатом *різниці* нечітких відношень сенсу Q та R назвемо таке нечітке відношення сенсу $P = \{ \langle i_l, i_j \rangle, \mu_P(\langle i_l, i_j \rangle) \} = Q \setminus R$, яке задане на тому ж самому декартовому добутку $I \times I$ та функція належності якого визначається за формулою

$$\mu_P(\langle i_l, i_j \rangle) = \max \{ \mu_Q(\langle i_l, i_j \rangle) - \mu_R(\langle i_l, i_j \rangle), 0 \}. \quad (2.16)$$

Операція різниці («МІНУС») визначає нове нечітке відношення сенсу P , що має такі знання ІС S_Q , які відсутні у ІС S_R .

Результатом *композиції* нечітких відношень сенсу Q та R назвемо таке нечітке відношення сенсу $P = \{ \langle i_l, i_j \rangle, \mu_P(\langle i_l, i_j \rangle) \} = Q \otimes R$, яке задане на тому ж самому декартовому добутку $I \times I$ та функція належності якого визначається за формулою

$$\mu_P(\langle i_l, i_j \rangle) = \max_{i_k \in I} \{ \min \{ \mu_Q(\langle i_l, i_k \rangle), \mu_R(\langle i_k, i_j \rangle) \} \}. \quad (2.17)$$

Операція композиції («MAX-MIN») визначає нове нечітке відношення сенсу P , яке є відгуком (реакцією) інфологічної системи S_R на знання ІС S_Q . Якщо узагальнити функцію належності $\mu_Q(\langle i_l, i_j \rangle)$ на міру сенсу Se певної окремої події чи синтагми як ОК, то відношення сенсу α -рівня над результатом операції «MAX-MIN» можна вважати відповіддю ІС S_R , що безумовно потребує подальших досліджень.

На відміну від операцій (2.14)–(2.17) операції симетричної різниці та доповнення нечітких відношень сенсу в задачах пізнавальної діяльності, на думку автора, корисної прикладної інтерпретації не мають. Потрібно зауважити, що внаслідок застосування операцій (2.14)–(2.17) в отриманих скінчених відношеннях сенсу також зникає властивість $\overline{\mu_Q} = 0,5$, але це пояснюється зникненням суб'єктивного характеру накопичення сенсу з втратою єдиного носія. Отже, з метою підтримки адекватності реакцій інфологічної системи, потрібно зберігати її базове відношення сенсу з функцією належності (2.9).

Чисельна оцінка образного сенсу через поняття ентропії. Поняття інформації, як і поняття знань не мають однозначного трактування, що підтверджується існуванням значної кількості різних визначень. Цінність застосування цих понять у сучасних інформаційних технологіях базується на формальних обмеженнях і, головне, кількісних оцінках існуючих баз даних та баз знань. Очевидно, що найбільш загальний характер має класична міра інформації К. Шеннона, в основу якої покладено поняття ентропії [243]. Проте оцінка знань у вигляді наукового тексту або бази знань в одиницях інформації виглядає неінформативно чи, навіть, незрозуміло. Тому, в залежності від типу бази знань, використовують такі показники, як кількість аксіом, правил за типом ЯКЩО-ТО, вузлів семантичної мережі, фреймів тощо [81].

Шлях від загального поняття інформації до більш складного поняття знань має проходити через накладення певних формальних обмежень. Якщо в межах запропонованої теорії було формалізовано поняття образного сенсу деякої АМО на основі відповідної одиниці Sav , то варто провести їх кількісне порівняння з ентропійною оцінкою цієї ж самої АМО та одиницею інформації Bit . Така постановка задачі пе-

редбачає, що, власне, АМО і є тією системою обмежень, що породжує образний сенс з понять ентропії та інформації.

Задамо деяку АМО на певний момент часу такими параметрами: k_{lg} – кількість виявлених системою зв'язків між l -м та g -м образами, m – кількість ненульових елементів матриці A_Q . Також маємо статистичну оцінку математичного сподівання кількості повторень одного зв'язку як $\lambda = k_{\Sigma} / m$, де $k_{\Sigma} = \sum_{l=1}^n \sum_{g=1}^n k_{lg}$. Внаслідок нормування образного сенсу кожної (l, g) -пари згідно з (2.10) сигмоїдальною функцією $\mu_Q(< i_l, i_g >) = 1/(1 + e^{-k_{lg} + \lambda})$ отримуємо оцінку середнього значення образного сенсу однієї пари цієї АМО

$$\overline{\mu_Q} = \frac{1}{m} \sum_{j=1}^m \mu_{Qj} = 0,5 \quad [Cav]. \quad (2.18)$$

Згідно з запропонованим підходом одиниця образного сенсу розміром один *Cav* характеризує максимальну вагу (l, g) -пари АМО як $\mu_Q(< i_l, i_g >) = 1$. У той же час факт появи на вході системи S кожної j -ї пари мовних образів з імовірністю $p(x_j)$ дозволяє оцінити ентропію цієї системи. Для отримання верхньої межі ентропії будемо вважати, що ОК складається з незалежних пар образів, хоча в реальних природно-мовних конструкціях це не зовсім так. Відомо, що у цьому випадку загальна ентропія або кількість інформації [243] системи S дорівнює

$$H = - \sum_{j=1}^m n_j \cdot \log p(x_j), \quad (2.19)$$

де значення n_j відповідає k_{lg} як кількості зв'язків між l -м та g -м образами.

Також можна визначити середню ентропію, що припадає на одну пару. З цієї метою розділимо (2.19) на k_{Σ} :

$$H_1 = - \sum_{j=1}^m \frac{n_j}{k_\Sigma} \cdot \log p(x_j).$$

Врахуємо, що для великих значень n_j та k_Σ імовірність j -ї пари мовних образів $p(x_j) = \lim_{k_\Sigma \rightarrow \infty} \frac{n_j}{k_\Sigma}$. Тоді середня ентропія однієї пари дорівнює

$$H_1 = - \sum_{j=1}^m p(x_j) \cdot \log p(x_j). \quad (2.20)$$

Оцінка (2.20) вже завищена умовою незалежності образних пар, але максимального значення середня ентропія пари досягає за додатковою умовою [243]: якщо поява кожної з m можливих пар мовних образів на вході S рівноімовірна, то

$$\overline{H_1} = \log_2 m \quad [Bim]. \quad (2.21)$$

Зрозуміло, що побудова матриці A_Q на основі реального текстового матеріалу не приведе до максимального значення ентропії (2.21). Але, з суто формальної точки зору кількісні оцінки $\overline{\mu_Q}$ та $\overline{H_1}$ є різними інтерпретаціями тієї ж самої чисельної характеристики АМО – середньої ваги однієї пари. Отже, з урахуванням (2.18), можна отримати верхню оцінку співвідношення одиниць образного сенсу та інформації як логарифмічну згортку

$$1 [Cav] = 2 \log_2 m \quad [Bim]. \quad (2.22)$$

Тепер визначимо нижню межу згортки інформації в образний сенс (2.20)–(2.22), взявши до уваги те, що кожне речення та відповідні до нього КМО та ОК є підграфами–деревами загального орієнтованого графу АМО. Тому поява тільки першої (l, g) -пари образів для S може бути незалежною, а всі наступні вже мають бути пов'язаними з l -м чи g -м образами або (рекурсивно) з новоприєднаними до підграфу образами. Для оцінки середньої ентропії однієї пари образів у таких умовах скористаємося виразом

$H = -\sum_{i=1}^n p(x_i) \sum_{j=1}^n p(x_j / x_i) \cdot \log p(x_j / x_i)$, де $p(x_j / x_i)$ – умовна імовірність появи в системі S пари x_j , якщо попередньою парою була x_i [243].

Поставлена задача зводиться до рекурсивного визначення $p(x_j / x_i)$ з елементів матриці A_Q . Позначимо (l, g) -пару як x_i , а у парі x_j разом з h -м має бути або l -й або g -й образ. Тоді маємо

$$4 \text{ варіанти } p(x_j / x_i) = \lim_{k_{\Sigma} \rightarrow \infty} \left(\frac{k_{hg} \mid k_{lh} \mid k_{gh} \mid k_{hl}}{k_{\Sigma hg} + k_{\Sigma lh} + k_{\Sigma gh} + k_{\Sigma hl}} \right).$$

Отже, вперше обґрунтовано аксіоматику нечіткої міри сенсу, на основі якої побудовано топологічний і квазіметричний простір упорядкованих пар образів Ω . Надано пізнавальну інтерпретацію таким операціям з нечіткими відношеннями образного сенсу, як перетин, об'єднання, різниця та композиція. Результати інтроспективного аналізу отриманих математичних співвідношень свідчать про належну формальну інтерпретацію відомих фактів з психології та фізіології людини стосовно мотивів діяльності, набутих рефлексів та відсутності низки властивостей нечітких відношень.

Запропоновану чисельну міру образного сенсу 1_{Cav} можна вважати логарифмічною згорткою ентропії образних пар КМО, що обумовлена системою обмежень у вигляді семантичної мережі АМО.

2.6 Підхід до формалізації механізму функціонування інфологічної системи

У розділі 1 розглянуто існуючі підходи до визначення свідомості, мислення та інтелекту в науках когнітивного напрямку, що мають бути враховані в процесі синтезу інфологічної системи на основі поняття образного сенсу. З цією метою пропонується визначити механізм функціонування ІС шляхом моделювання природного аналогу – образного мислення людини. В основу концептуальної моделі образного аналізу ПМК було покладено [12] три головні ідеї:

1. Базовими концептами образного мислення є образ, асоціація та подія.
2. При побудові феноменологічних моделей мають бути враховані філогенез та онтогенез психічної діяльності.

3. Концептуальна модель на основі припущень щодо інтерпретації нейропсихологічних даних повинна імітувати функціональну активність 9-ти відділів головного мозку згідно з табл. 1.3.

Викладемо ці та інші ідеї, отримані в результаті огляду значимих когнітивних понять, у вигляді припущень щодо концептуальної моделі ІС. Метою формалізації висунутих припущень є послідовне нарощування в концепції можливостей відомих психічних феноменів з точки зору філогенезу. Вперше суттєві припущення та особливості такого підходу було сформульовано в [12].

Припущення 1. Образне мислення є первинним, а логічне або формальне мислення до певної міри можна вважати окремим випадком образного. Цей окремий випадок з унікальними характеристиками став доступним людині, на відміну від інших живих істот з розвинутою нервовою системою, виключно завдяки придбаній та генетично закріпленій здатності знакового (символьного) представлення образів.

Припущення 2. Завдання синтезу концептуальної моделі образного аналізу ПМК можна сформулювати таким чином: методом послідовного ускладнення побудувати траєкторію проходження функціонально досліджених відділів головного мозку такими психофізіологічними феноменами, як простий рефлекс, складний рефлекс, умовний рефлекс, інтелектуальна діяльність (з розділенням несвідомого, підсвідомого та свідомого). Врахування наявних наукових даних полягає у використанні концептуальною моделлю функціональних особливостей відомих (достатньо досліджених) відділів головного мозку [244].

Припущення 3. Розрізнятимемо і відповідним чином позначатимемо в концептуальній моделі такі функціональні блоки на основі відділів головного мозку [156]:

1. Стовбурові відділи мозку (ефективна і афективна складові нервової системи, блок емоцій).
2. Моторні відділи мозку (виконавчі механізми).
3. Скронева область (надоперативна пам'ять або буфер для створення складних образів).
4. Передмоторна область (блок створення складних рухів або кінестетичних мелодій).
5. Зона ТРО (оперативна пам'ять).
6. Потилична область (довготривала пам'ять).
7. Тім'яна область (асоціативна пам'ять).

8. Постцентральні області (блок моделювання або формування і вибору стратегій).

9. Лобові долі (блок програмування, регулювання і контролю).

Фізіологічні контурні межі 9-ти функціональних блоків [245] мають графічне відображення на карті цитоархітектонічних полів головного мозку за К. Бродманом [0]. Пропонується побудувати концептуальну модель образного аналізу ПМК, що дозволяє інтерпретувати такі види діяльності нервової системи людини:

1. Безумовні рефлексі (нервові моделі стимулу як відчуття), у тому числі:

1.1. Припущення 4: прості рефлексі – найкоротший нервовий ланцюг, в рамках якого збудження із стовбурових відділів замикається безпосередньо в лімбіко-ретикулярній системі через моторні області.

1.2. Припущення 5: складні рефлексі – збудження через стовбурові відділи потрапляє в надоперативну пам'ять і, потім, в передмоторній області формується складний рух, що реалізовується моторними відділами.

2. Припущення 6: умовні рефлексі – нервовий ланцюг включає стовбурові відділи–надоперативну пам'ять–асоціативну пам'ять–передмоторні області–моторні відділи. Для умовних рефлексів нервові моделі стимулів вже знаходяться на рівні представлень образів, що дозволяє фіксувати в асоціативній пам'яті складні образи. Умовний рефлекс більшістю дослідників, починаючи з І. П. Павлова [153], безпідставно вважається базовим феноменом «розумної» поведінки. Також умовний рефлекс може бути наслідком (скороченням) інтелектуальної діяльності, в цьому випадку відбувається випадання окремих ланок у довшому (див. припущення 7) нервовому ланцюгу, а саме оперативної пам'яті і блока моделювання (постцентральних відділів).

3. Припущення 7: інтелектуальна діяльність – ґрунтується на образах (гештальтах), що виникають в процесі передачі нервової моделі стимулу в оперативну пам'ять з надоперативної або, згодом, з інших блоків головного мозку, що беруть участь в локальних циклах. В оперативній пам'яті в кожний момент часу складається активна ситуація з 7 ± 2 образів (ансамбль образів), причому стовбурові відділи мозку генерують для всіх образів вектор емоцій, а один з них (найбільший за вектором) потрапляє до фокусу уваги. На відміну від стійкого умовного рефлексу у блоці моделювання відбувається вибір одного з можли-

вих рішень активної ситуації. Ще однією обов'язковою складовою базової інтелектуальної діяльності є проходження нервового імпульсу через лобові долі, чим забезпечується програмування, контроль і регуляція діяльності. Інтелектуальна діяльність дозволяє створити найдовший нервовий ланцюг з усіма можливими локальними циклами. Така траєкторія проходження нервових імпульсів, в чому і полягає її ключова ознака, фіксується у вигляді подій в довготривалій пам'яті.

Припущення 8: базову інтелектуальну діяльність характеризує наявність таких петель в основній траєкторії:

П1 – 3–7–3 розпізнавання образів і отримання їх уявлень;

П2 – 5–6–5 обмін подіями з довготривалою пам'яттю;

П3 – 5–7–5 отримання асоційованих образів;

П4 – 5–1–5 отримання вектора емоцій для образів активної ситуації;

П5 – 4–9–4 контроль окремих складних дій.

Є підстави вважати, що для охоплення несвідомого (за Фрейдом) достатньо базової інтелектуальної діяльності. Важливе практичне значення має визначення в рамках моделі таких складних концептів інтелекту, як свідоме та підсвідоме. З цією метою розглянемо певні психологічні феномени, які традиційно відносять до свідомої діяльності – планування (на рівні тактики та стратегії), контроль, регулювання або управління. Проте свідомою природою цих феноменів, як правило, пояснюється звичними можливостями інтроспективного аналізу відповідних психічних процесів на відміну від процесів базової інтелектуальної діяльності. Тому введемо поняття розгорнутої інтелектуальної діяльності.

Припущення 9: Розгорнуту інтелектуальну діяльність характеризує наявність таких додаткових петель до базової траєкторії:

П6 – 5–8–5 вибір альтернатив (стратегій);

П7 – 9–8–9 деталізація стратегій шляхом формування плану дій;

П8 – 5–9–5 контроль і регулювання плану дій;

П9 – 5–5 локальне замикання циклу в оперативній пам'яті для регулювання (управління) складними діями.

Варто зауважити, що виключно розгорнута інтелектуальна діяльність дозволяє людині побудувати і використовувати в подальшому:

- думку (стратегію, інтегральну програму) – представляє собою розв'язання часткової інтелектуальної задачі, пов'язаної з окремими інтелектуальними діями та забезпечується лобовими долями, а саме петлею П6;

- висновок (обґрунтування, логічний ланцюжок, розв'язання загальної задачі інтелектуальної діяльності) – забезпечується передфронтальними і лобовими долями (петлями П6, П7 і П8) та використовує розгорнуту траєкторію нервового імпульсу на підсвідомому та/або свідомому рівнях.

Актуальне питання про те, як формально відокремити свідоме від підсвідомого в інтелектуальній діяльності людини залишається відкритим і до цього часу. В рамках розглянутої концепції пропонується така формальна інтерпретація – припущення 10: якщо образам активної ситуації відповідають символічні ознаки, то розгорнута інтелектуальна діяльність потенційно може «спливати» з базального рівня (підсвідомого) в надбудовне (свідоме). Мислення будемо вважати свідомою інтелектуальною діяльністю людини. Це означає, що підсвідоме також належить до інтелектуальної діяльності, причому охоплює її базову та розгорнуту траєкторії. На рівні підсвідомого теж може бути обґрунтована стратегія вирішення певної проблеми, проте усвідомити метод вирішення людина зможе лише при знаковому виразі образів, що складають механізм розв'язання проблеми.

Згідно з останніми результатами фізіологічних досліджень [245], кількість операцій мислення складає невелику частину від кількості операцій підсвідомого та несвідомого. Ще більшими та невидимими шарами підґрунтя інтелектуальної діяльності виступають розглянуті вище три типи рефлексів. Тому небезпідставним вважається таке припущення – фізичний обсяг нервової маси головного мозку, необхідний для забезпечення вищезазначених нервових процесів, зворотно пропорційний їх кількісним показникам. Такий стан речей, який пояснюється ускладненням вищих типів психічної діяльності у порівнянні з нижчими в геометричній прогресії, проілюстровано на рисунку 2.13, де перевернуту піраміду обсягу нервової маси зображено у відтінках сірого.

Отже, сформульовано припущення, які дозволяють послідовно побудувати концептуальну модель образного аналізу ПМК на основі інтерпретацій істотних особливостей таких відомих феноменів, як безумовний рефлекс (простий та складний), умовний рефлекс, інтелектуальна діяльність (з виділенням несвідомого, підсвідомого та свідомого). В результаті надано характеристику основним функціональним блокам та інформаційним потокам образної природи, що вимагають обов'язкового врахування в інфологічній системі з метою формалізації функцій СОПМК та розв'язання семантико-залежних задач.

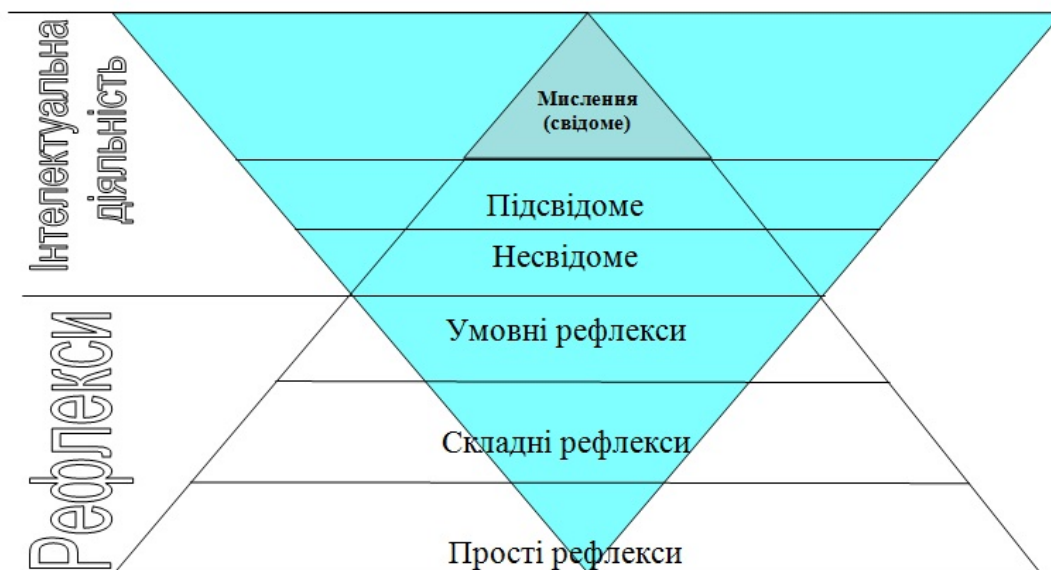


Рисунок 2.13 – Види психічної діяльності у співвідношенні до обсягу нервової маси

2.7 Структурно-функціональна модель образної обробки природно-мовного контенту

На основі запропонованих припущень щодо висунутої концепції побудуємо структурно-функціональну модель образної обробки ПМК. Формальна інтерпретація має відповідати розглянутим раніше фізіологічним даним головного мозку, згідно з якими в роботі останнього розрізняють три узагальнені блоки – енергетичний, проекційно-асоціативні системи та інтеграційно-пускові системи (див. таблицю 1.1). Згідно з фізіологічною основою узагальнених блоків мозку, задамо відповідним моделям кольорову палітру в таблиці 2.3.

Таблиця 2.3 – Палітра для моделей фізіологічних структур головного мозку

№ з/п	Узагальнені блоки мозку	Фізіологічна основа	Колір
1	Енергетичний (лімбіко-ретикулярні системи)	Стовбурові відділи мозку, апарати медіальної кори або лімбічної області	
2	Основні аналізаторні (проекційні) та асоціативні системи	Вторинні відділи лівої скроневої і тім'яно-потиличної області	
3	Інтеграційно-пускові системи	Кора лобових долей мозку (24 % поверхонь великих півкуль)	

Спираючись на фізіологічні та нейропсихологічні засади запропонованого концептуального підходу, деталізуємо моделі трьох узагальнених блоків головного мозку. Враховуючи нумерацію функціональних відділів головного мозку, згідно з припущенням 3, будемо вважати, що узагальнені складові мозку вміщують у собі такі структурні блоки:

1. Ефективна (1а) та афективна (1в) складові інфологічної (нервової) системи спільно з блоком емоцій (1б).
2. Блок пам'яті (3, 5, 6, 7).
3. Блок управління (4, 8, 9) та моторні відділи (2).

Загальна структурна схема ІС представлена на рисунку 2.14. На структурній схемі моторні відділи виділені з 3-го узагальненого блока в окремий блок для кращої демонстрації потоків образів, що відповідають безумовним і умовним рефлексам згідно з припущеннями 4–6.

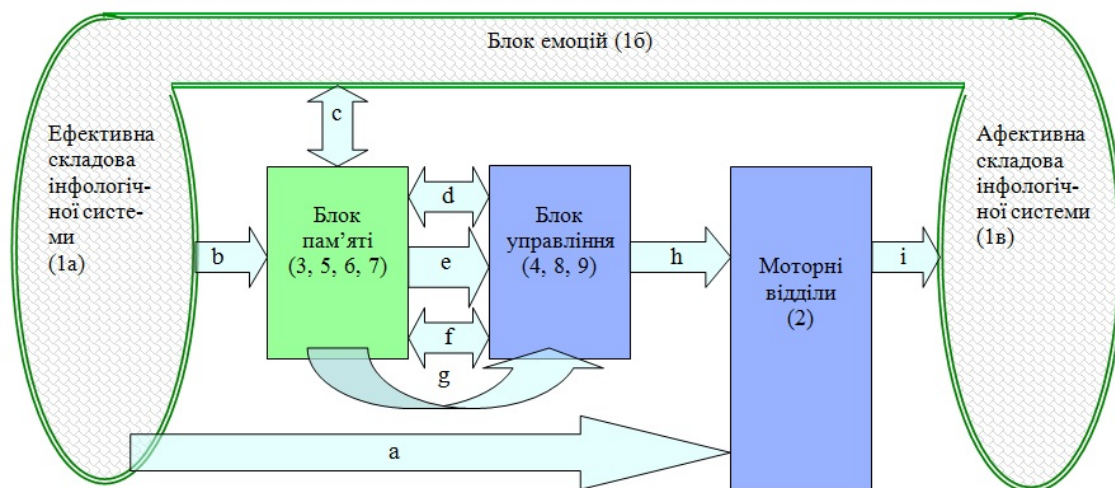


Рисунок 2.14 – Загальна структурна схема ІС

Для потоків нервових та психічних процесів, що поєднують блоки моделі, в структурній схемі використано такі позначення [31]:

- a)* пряма рефлексорна дуга;
- b)* відчуття;
- c)* петля П4 – отримання вектора емоцій для образів активної ситуації;
- d)* петля П6 – отримання образу розв'язування активної ситуації в результаті впливу (вливання) сенсів;
- e)* образ у фокусі уваги або інтегрований образ розв'язування (дії);

- f)* петля П8 – отримання образу контролю стратегії для образів активної ситуації;
- g)* сприйняття і уявлення, що беруть участь в умовних рефлексах;
- h)* складні команди (кінестетичні мелодії);
- i)* команди управління з моторних відділів.

Загальна структурна схема ІС моделює узагальнений механізм образного аналізу ПМК. Блок ефективною складовою (1а) безперервно постачає до блока пам'яті множину відчуттів *b* з периферійної нервової системи, в т.ч. суб'єктивне відчуття часу. Паралельно цьому між блоком 1а та моторними відділами (2) підтримується пряма рефлекторна дуга *a*. Блок пам'яті розрізняє в потоці відчуттів окремі образи, які у вигляді сприйняття та уявлень служать основними атомами (конструктами) моделі психічної (у т.ч. інтелектуальної) діяльності людини. Частина з цих сприйняття та уявлень образів *g* бере участь в складних та умовних рефлексах через блок управління.

Актуальна в даний момент сукупність з 7 ± 2 образів (ансамбль образів або активна ситуація) одночасно потрапляє до блока емоцій (1б) та у складові блока управління. Блок емоцій, в свою чергу, оцінює активну ситуацію у вигляді вектора емоцій *c*, який в той же такт часу передається назад в пам'ять *i*, далі, в блок управління з метою спрямування фокусу уваги тільки на один з образів ансамблю.

Головним завданням блока управління, який також можна вважати блоком уяви, є вибір і подальша активізація образу розв'язування (дії) *e*. Для цього між блоками пам'яті та управління відбувається циркуляція послідовності образів *d*, що реалізує механізм впливу (вливання) сенсів множини активних ситуацій. Образ розв'язування приводить в дію набір складних команд *h* для моторних відділів. Останні перетворюють множини команд *h* та/або *a* в послідовність елементарних команд управління *i* для афективної складової (1в) нервової системи. Оскільки регулювання є складним та багатокроковим процесом, а окремі його дії, як правило, також мають розгорнуту в часі ієрархічну програму, блок управління обмінюється образом контролю *f* з блоком пам'яті.

Можна також вважати, що між афективною і ефективною складовими нервової системи існує постійний зворотний зв'язок, оскільки результат будь-якої дії в той же такт часу відображається за допомогою відчуттів (у т.ч. і м'язових) та повертається назад у блок пам'яті.

Остання, у свою чергу, також має нетривіальну будову і складається, як мінімум, з таких функціональних блоків: надоперативної (3), оперативної (5), асоціативної (7) та довготермінової (6) пам'яті. Структурна схема блока пам'яті ІС представлена на рисунку 2.15.

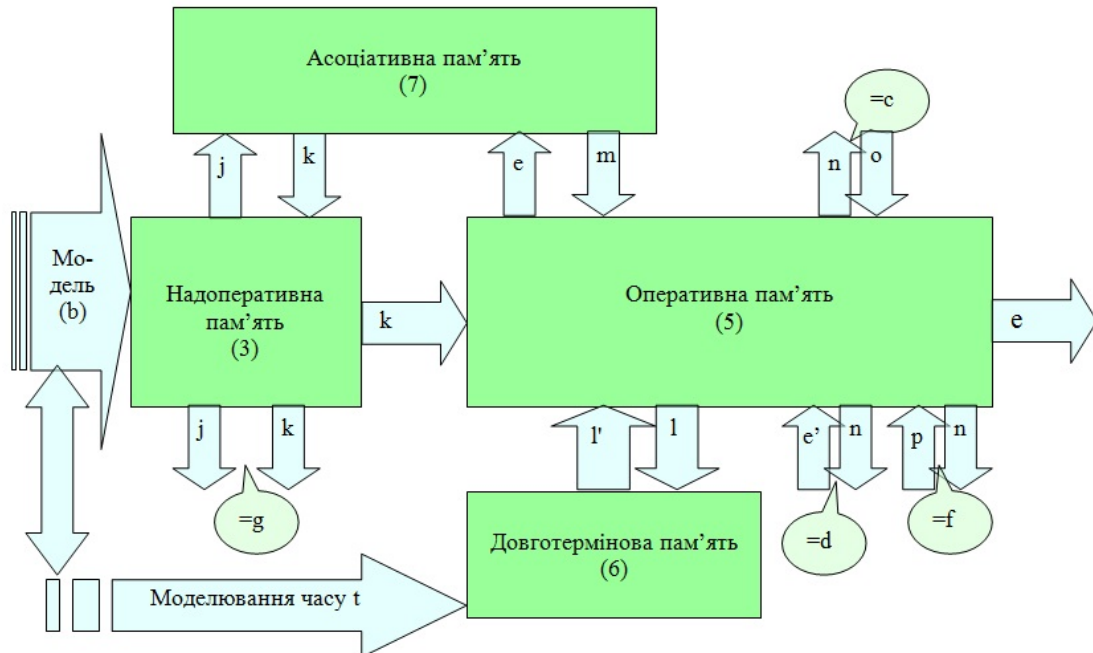


Рисунок 2.15 – Структурна схема блока пам'яті ІС

На схемі рисунка 2.15 прийняті такі позначення для внутрішніх потоків блока пам'яті:

- j*) сприйняття образів;
- k*) представлення образів;
- l*) записані та відтворені (*l'*) події у вигляді ансамблю образів;
- m*) асоційовані образи;
- n*) образи активної ситуації;
- o*) вектор емоцій для образів активної ситуації;
- p*) образ незбігу контрольної та активної ситуацій.

Механізм роботи пам'яті складається з таких процесів. У надоперативну пам'ять (3) ззовні безперервно надходить потік відчуттів *b*, які представляють собою, в загальному випадку, сукупність певних характеристик (властивостей, ознак і параметрів). Як буфер накопичення надоперативна пам'ять формує з цього потоку сприйняття образів *j*, які вже в дискретні проміжки часу передаються далі в асоціативну пам'ять (7), де фактично і відбувається розпізнавання образів. Розпізнані образи у вигляді уявлень *k* вертаються у надоперативну

пам'ять і передаються далі в оперативну пам'ять (5), яка збирає і в кожний такт часу утримує в собі 7 ± 2 образи. Ця сукупність образів n з фокусом уваги на одному з них ϵ , з одного боку, активною ситуацією, а, з іншого боку, подією – внаслідок ієрархічної структуризації зафіксованого в певний момент часу ансамблю образів.

Згадана структуризація образів активної ситуації з оперативної пам'яті здійснюється в довготерміновій пам'яті (6), яку можна вважати безперервним ланцюжком самозапису подій l на осі часу t , що суб'єктивно відчувається. У відповідь на знайому (близьку за конфігурацією) активну ситуацію довготермінова пам'ять може знайти у себе і повернути в оперативну пам'ять подію l' , записану раніше. Проте ключовим аспектом механізму поновлення активної ситуації в оперативній пам'яті є фізіологічна можливість отримання з лімбіко-ретиккулярних систем вектора емоцій o для складових ансамблю образів. Головним завданням блока емоцій є оцінка активної ситуації і визначення мотивації для подальшої реакції мозку. Тому в кожний такт часу фокус уваги в оперативній пам'яті спрямовується на такий образ e , який має найбільше значення вектора емоцій. Цей образ e у той же такт часу передається до блока управління як образ розв'язування (дії).

Такий же взаємний обмін даними, як з довготерміновою пам'яттю, відбувається між оперативною пам'яттю та асоціативною. Оперативна пам'ять передає образ з фокусом уваги e у вже активізовану підмережу (павутину) образів асоціативної пам'яті, а та повертає навзаєм нові (асоційовані з e) образи m . Оновлений в результаті такої реакції асоціативної пам'яті на зміну фокусу уваги поточний варіант активної ситуації знову передається з оперативної пам'яті в блок емоцій з метою чергового отримання вектора емоцій і т. д. Ключову роль відіграє оперативна пам'ять також в процесі планування і регулювання складних дій, оскільки безперервно обмінює з блоком управління свою активну ситуацію n на бажаний образ розв'язування (дії) e' та на образ незбігу контрольної і активної ситуацій p .

Якщо робота пам'яті відповідає, в цілому, за відтворення адекватного набору образів у відповідь на зовнішню ситуацію, то реагування ІС на стороннє подразнення забезпечується блоком управління, структура якого представлена на рисунку 2.16.

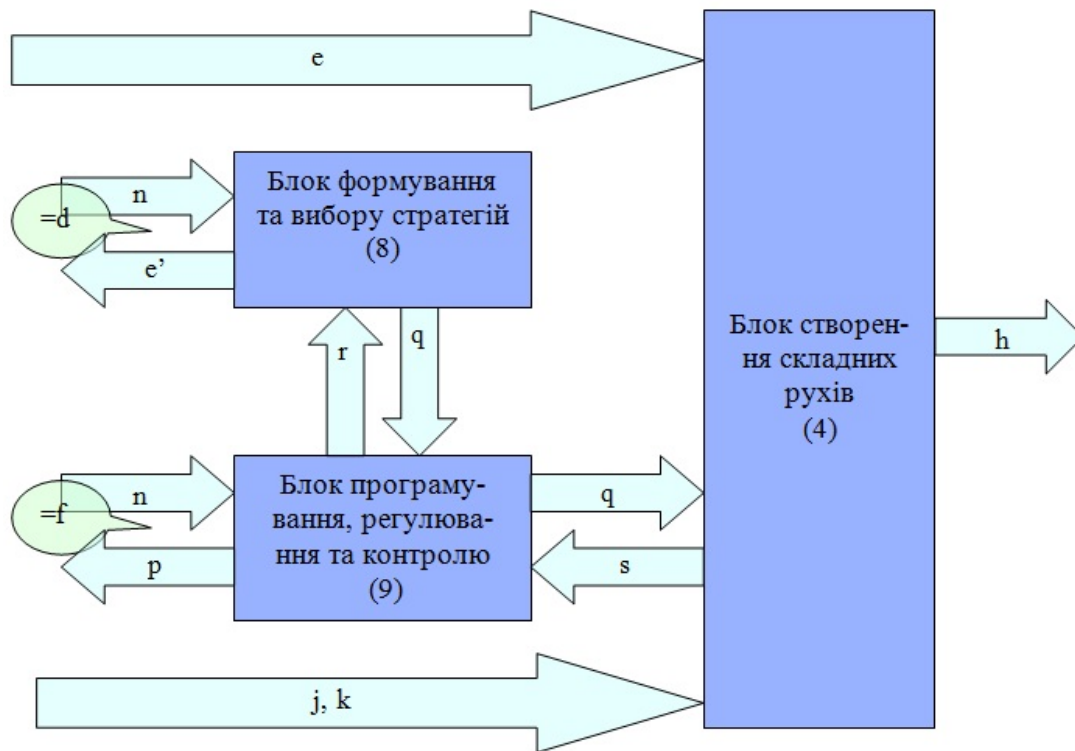


Рисунок 2.16 – Структурна схема блока управління ІС

На схемі рисунка 2.16 прийняті такі позначення для внутрішніх потоків блока управління:

- q*) образ мети (стратегія);
- r*) план (програма) дій у вигляді послідовності образів;
- s*) набір стійких рефлекторних автоматизмів для виконання інтелектуальних дій.

Моделі фізіологічних структур блока управління відповідають за ухвалення (вибір) адекватних рішень у відповідь на активну ситуацію (постцентральні області), планування і контроль регуляційної діяльності (лобові доли), а також декодують образ розв'язування в набір складних команд (кінестетичні мелодії) за допомогою передмоторних відділів. Останній блок (4) також може самостійно реагувати на сприйняття та уявлення з надоперативної пам'яті, закриваючи цим самим ланцюжки складних безумовних і умовних рефлексів.

Ключову роль в процесі управління інтелектуальною діяльністю відіграє блок моделювання або формування і вибору стратегій (8). Оцінка активної ситуації *n* цим блоком приводить до генерації образу розв'язування *e'* внаслідок впливу (вливання) смислу попередніх активних ситуацій на поточну. Окрім цього, при появі образу

розв'язування автоматично визначається шлях (стратегія) отримання потрібного розв'язку, а тому блок моделювання одночасно генерує образ найближчої мети q , що веде до реалізації стратегії.

Блок програмування, регулювання і контролю (9) у відповідь на образ мети (стратегію) q передає назад до блока моделювання програму дій r , побудовану як послідовність елементарних і вже засвоєних до автоматизму дій s блока створення складних рухів (4). В цьому полягає планівна функція блока програмування, регулювання і контролю. Контролююча функція цього блока забезпечується порівнянням образів активної ситуації n з даними стеку, де знаходяться образи плану дій. Образи p , що стали результатом незбігу контрольної та активної ситуацій, відразу спрямовуються до оперативної пам'яті з метою регулювання відхилень від плану.

Образна природа поняття траєкторії нервового збудження потребує подальшої деталізації у вигляді функціональної моделі та вимагає формального визначення понять для всіх складових інформаційних потоків. Згідно з [157], будемо вважати, що нервові збудження можуть передавати такі типи сигналів:

- I) кількісний параметр;
- II) якісний параметр (властивість, ознака) як множина кількісних параметрів;
- III) образ (гештальт) як множина якісних параметрів;
- IV) подія як множина образів з обмеженням на зв'язність (п. 2.3).

Припускається, що нумерація типів сигналів відповідає логічному принципу від простого до складного та, скоріш за все, має філогенетичну хронологію. Інформаційні потоки кількісної природи безперервно надходять через всі органи чуття людини і так само безперервно розпізнаються в автоматичному режимі. Базовий тип нервового збудження задається виключно кількісними параметрами – силою та тривалістю сигналу [154], а також специфічними для кожного органу чуття характеристиками, наприклад, світлових або звукових хвиль тощо. Результатом первинного розпізнавання можна вважати ідентифікацію певної множини або комбінації кількісних параметрів у вигляді якісного параметра, що має відношення до одного з раніше розпізнаних образів. Звідси якісні параметри (інакше – відчуття) представляють собою ознаки або властивості, яким може ставитися у відповідність вербальний (символьний) еквівалент, наприклад, «зелений»,

«великий», «голосний», «приємний», «повільний», «небезпечний», «солодкий» тощо.

Згідно з нейропсихологічними основами моделювання [185] первинне розпізнавання відбувається на несвідомому рівні, проте в результаті інтроспективних процесів може згодом «спливати» у свідомість саме завдяки вербальним ознакам відповідних образів. Наступним етапом розпізнавання, що спочатку також реалізується на несвідомому рівні, є феномен сприйняття людиною цілісного образу або гештальту [189]. Будь-який образ об'єднує в єдине ціле досить значну, але обмежену кількість якісних ознак та властивостей, що пов'язують його з іншими образами. Представлення образів у пам'яті людини визначаються певною множиною найбільш значимих (у першу чергу – статистично) зв'язків, що можна продемонструвати відомими загадками, наприклад: українськими – «Круглий, як куля, зелений, як трава, червоний, як кров, солодкий, як мед» «Червона, солодка, пахуча, росте низько, до землі близько» або російськими – «Сам – алий, сахарный, кафтан – зеленый, бархатный» або «Два кольца, два конца, посередине гвоздик».

Третім природним типом розпізнавання для людини можна вважати одночасне сприйняття ансамблю з 7 ± 2 взаємодіючих образів у вигляді події. Під час події кожний образ може грати одну з ролей, найголовнішими з яких є об'єкт дії, дія (метод) та суб'єкт дії. Інші образи активної ситуації за необхідності характеризують ознаки (властивості) об'єкта та/або суб'єкта, та/або обставини дії. Такий тип розпізнавання належить до природних, тому що, як і попередні два, на початку здійснюється поза свідомістю. У певних випадках, пов'язаних з небезпечними для людини обставинами, первинна здатність людини до рефлексивного сприйняття та реагування на події є найбільш ефективним або навіть єдиним шляхом до самозбереження. Мовленнєва діяльність людини, хоча і належить до вищої нервової діяльності, так само побудована на комунікації подій та, відповідно, синтагматичному устрої мови.

Виходячи з запропонованої класифікації, проведемо аналіз використаних у структурній моделі образної обробки ПМК інформаційних потоків. Загальна кількість можливих об'єктів аналізу дорівнює 17-ти, оскільки у структурній схемі використано позначення для 3-х петель, що мають змішану природу сигналу:

c) петля П4 – отримання вектора емоцій для образів активної ситуації;

d) петля П6 – отримання образу розв’язування активної ситуації в результаті впливу (вливання) сенсів;

f) петля П8 – отримання образу контролю стратегії для образів активної ситуації.

Сигнал з кількісними параметрами передається у 4-х таких потоках:

- a) пряма рефлекторна дуга;
- i) команди управління з моторних відділів;
- o) вектор емоцій для образів активної ситуації;
- t) суб’єктивне відчуття часу.

Сигнал з якісними параметрами передається у 3-х таких потоках:

- b) відчуття;
- h) складні команди (кінестетичні мелодії);
- s) набір стійких рефлекторних автоматизмів для виконання інтелектуальних дій.

Сигнал з цілісними образами передається у 5-х таких потоках:

- e) образ у фокусі уваги або інтегрований образ розв’язування (дії);
- g) сприйняття й уявлення, що беруть участь в умовних рефлексах;
- j) сприйняття образів;
- k) представлення образів;
- m) асоційовані образи.

Сигнал з подіями передається у 5-х таких потоках:

- l) записані та відтворені (l') події у вигляді ансамблю образів;
- n) образи активної ситуації;
- p) складовий образ незбігу контрольної та активної ситуацій;
- q) складовий образ мети (стратегія);
- r) план (програма) дій у вигляді послідовності образів.

Узагальнимо результати проведеного аналізу в таблиці 2.4.

Таблиця 2.4 – Відповідність інформаційних потоків моделі типам сигналу нервового збудження

№ з/п	Тип сигналу	Позначення інформаційних потоків	Кількість потоків
I	Кількісний параметр	a), i), o), t)	4
II	Якісний параметр	b), h), s)	3
III	Образ	e), g), j), k), m)	5
IV	Подія	l), n), p), q), r)	5

Побудуємо граф структурно-функціональної моделі образної обробки ПМК шляхом послідовного визначення основних психічних феноменів згідно з *припущенням 2*. З цією метою формально опишемо для моделей кожного психічного феномену відповідну траєкторію з урахуванням позначень інформаційних потоків розглянутих вище структурних схем.

1. Згідно з *припущенням 4*, траєкторією простого рефлексу вважатимемо петлю 1–2–1. Граф відповідної траєкторії з темнішим позначенням у відтинках сірого найбільш давньої з філогенетичної точки зору моделі блока-вузла 1 зображено на рисунку 2.17 (інші блоки, починаючи від 2-го, представляють собою моделі областей кори головного мозку).

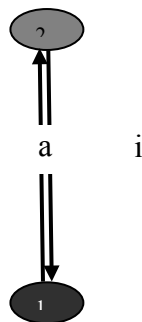


Рисунок 2.17 – Граф функціональної моделі простого рефлексу

2. Згідно з *припущенням 5*, траєкторію складного рефлексу формально позначимо як 1–3–4–2–1. Граф відповідної траєкторії зображено на рисунку 2.18, причому для позначення порівняно нових у філогенезі моделей блоків-вузлів 3 і 4 використано світліші від простого рефлексу відтінки сірого.

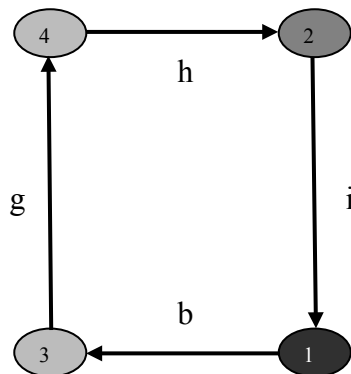


Рисунок 2.18 – Граф функціональної моделі складного рефлексу

3. Згідно з припущенням 6, траєкторію умовного рефлексу формально позначимо як 1–3–7–3–4–2–1 з урахуванням петлі П1(3–7–3). Граф відповідної траєкторії зображено на рисунку 2.19, причому для позначення філогенетично нової у порівнянні з попередніми рефлексами моделі блока-вузла 7 використано ще світліший відтінок сірого.

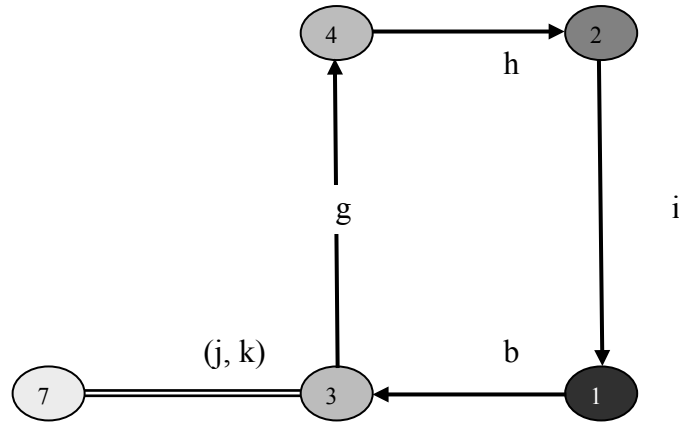


Рисунок 2.19 – Граф функціональної моделі умовного рефлексу

4. Згідно з припущеннями 7 та 8, траєкторію базової інтелектуальної діяльності формально позначимо як 1–3–7–3–5–6–5–7–5–1–5–4–9–4–2–1 з урахуванням петель П1(3–7–3), П2(5–6–5), П3(5–7–5), П4(5–1–5) та П5(4–9–4). Граф відповідної траєкторії зображено на рисунку 2.20, причому філогенетично нові у порівнянні з рефлексами моделі блоків-вузлів 5, 6 та 9 позначені білим кольором. У позначеннях цієї та попередньої траєкторій ключові блоки-вузли врахованих замкнених петель виділено підкресленням – 7, 6, 1, 9.

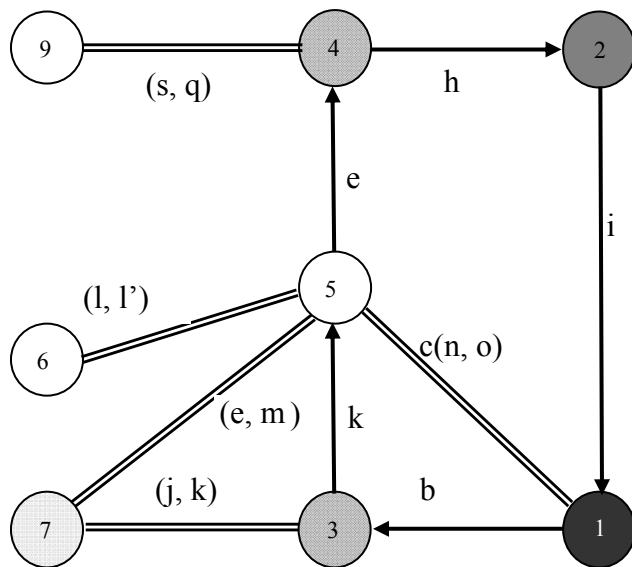
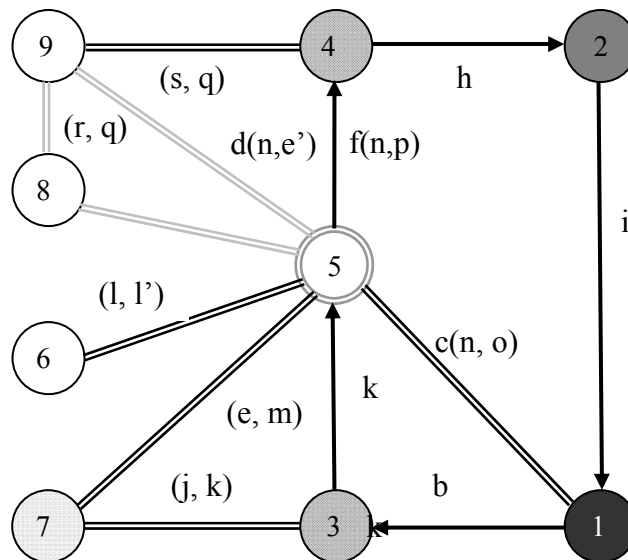


Рисунок 2.20 – Граф функціональної моделі базової інтелектуальної діяльності

5. Згідно з припущенням 9, траєкторія розгорнутої інтелектуальної діяльності відрізняється від базової траєкторії наявністю петель П6(5–8–5), П7(9–8–9), П8(5–9–5) та П9(5–5). Граф траєкторії розгорнутої інтелектуальної діяльності, в якому додаткові 4 петлі освітлені, а модель блока-вузла 8 позначена білим кольором, представлено на рисунку 2.21. $d(n, e')$



Рисунк 2.21 – Граф функціональної моделі розгорнутої інтелектуальної діяльності

Загальний граф структурно-функціональної моделі має об'єднувати графи відомих психічних феноменів (рисунки 2.17–2.21) з урахуванням моделі суб'єктивного відчуття часу t . У позначеннях вузлів та потоків графа збережені градації кольору з метою демонстрації філогенетичного віку відповідних моделям нервових структур головного мозку – від темнішого (найстарші структури та ділянки кори) до світлішого (наймолодші ділянки кори).

Загальну структурно-функціональну модель образної обробки природно-мовного контенту у вигляді графа представлено на рисунку 2.22.

Отже, запропонована структурна схема образної обробки природно-мовного контенту враховує відомі фізіологічні і нейропсихологічні дані та відповідає 10-м формальним припущенням інфологічної концепції. Структурна схема представляє процеси обробки у вигляді інтелектуальної діяльності на основі взаємодії 9-ти моделей функціональних блоків головного мозку та 19-ти інформаційних потоків, що мають образну природу.

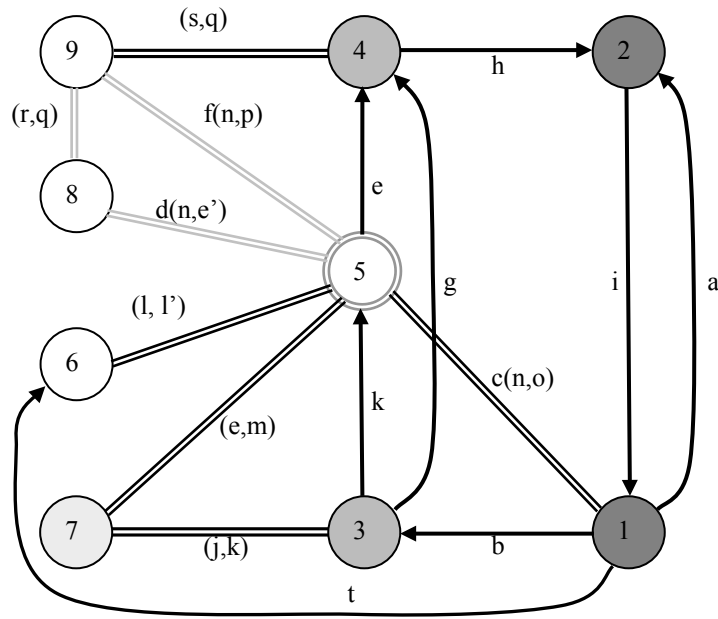


Рисунок 2.22 – Загальний граф структурно-функціональної моделі образної обробки ПМК

На основі інтерпретації процесів образного мислення та формального визначення інформаційних потоків створено граф загальної структурно-функціональної моделі образної обробки природно-мовного контенту. Запропонована модель враховує інфологічні припущення щодо побудови концепції образного аналізу текстової інформації та формальні визначення складових траєкторії нервового збудження, що передають сигнали 4-х типів між 9-ма функціональними блоками моделі. Закладена в позначеннях вузлів та потоків графічної моделі філогенетична хронологія відповідних нервових структур головного мозку забезпечує несуперечливість методології синтезу структурно-функціональних моделей ІС результатам інтерпретації даних фізіології.

РОЗДІЛ 3

МЕТОДИ МОДЕЛЮВАННЯ ПРОЦЕСІВ ОБРАЗНОЇ ОБРОБКИ ПРИРОДНО-МОВНОГО КОНТЕНТУ

Формалізація основних концептів теорії образного аналізу текстової інформації пов'язується з теоретико-множинною інтерпретацією простору сенсу з нечіткою мірою у вигляді орієнтованих та неорієнтованих графів. Вводиться новий клас графів на основі застосування булеану і поняття «піраміди сенсу», що приводить до визначення булевої алгебри сенсу як двоосновної алгебраїчної системи. Обґрунтовується архітектура бази знань СОПМК, визначаються її особливості у порівнянні з відомими аналогами.

3.1 Формалізація асоціативної мережі образів за допомогою графів

Важливим теоретико-експериментальним поняттям психолінгвістики вважається асоціативно-вербальна мережа (АВМ) носія мови [135]. Сучасні Інтернет-технології забезпечують зручний інструментарій для розробки та підтримки АВМ з метою проведення масштабних психолінгвістичних досліджень на основі вільного асоціативного експерименту [247]. Отримані в цьому напрямку напрацювання варто використати як об'єктивне підґрунтя для створення асоціативної мережі образів (АМО) – важливого поняття концепції образного аналізу текстової інформації [26], яке узагальнює відомі семантичні моделі тезаурусів [233] на основі поняття мовного образу [57].

В загальному випадку АМО, так само, як і АВМ, враховує спрямованість зв'язків між словами мови та відповідними МО, проте для низки задач предмета дослідження вирішальне значення має сила асоціативного зв'язку, а не його напрямок [31]. Отже, для інфологічного моделювання значимих понять образної обробки ПМК (пп. 2.6, 2.7) будемо застосовувати орієнтовані та неорієнтовані зважені графи, оскільки на фіксований момент часу множини I образів та ω асоціативних зв'язків (п. 2.1) можна вважати скінченними.

Розглянемо задачу формального опису АМО з урахуванням особливостей моделі образної обробки ПМК. Вперше підхід до графового представлення АМО з урахуванням змісту поняття МО запропоновано у [39]. Будемо виходити з того, що початкова траєкторія нервового збудження може передавати такі типи сигналів (п. 2.6):

- кількісний параметр – базовий тип інформаційного сигналу;
- якісний параметр як скінченна множина кількісних параметрів;
- образ (гештальт) як скінченна множина якісних параметрів.

Постановка задачі. Будемо вважати вузлами АМО певні вербальні ознаки n образів, які можуть поєднуватися між собою зв'язками \setminus , що введені у формальній теорії. В процесі функціонування ІС фіксує кількість k_{ij} виявлених зв'язків між i -м та j -м образами [34], причому у кожний момент часу відоме поточне значення k_{ij} , $i, j \in [1, n]$. З метою врахування аксіоматичних основ простору сенсу з нечіткою мірою (п. 2.4) припустимо, що вага асоціативного зв'язку, згідно з (2.9)–(2.13)

$$c_{ij} = \mu_Q(\langle i_i, i_j \rangle) = F_l(k_{ij}) = Se_{ij}, \quad l \in \{1, 2, 3, 4\}. \quad (3.1)$$

де l – формальний рівень функції сенсу Se_{ij} для пари образів.

Кількість ненульових зв'язків від кожної вершини графа до інших вершин набагато менша за $n-1$, оскільки, згідно з практикою вільного асоціативного експерименту [121], при великих значеннях n людина може підібрати не більше 7 ± 2 значимих слів-асоціатів на слово-стимул. Необхідно обґрунтувати властивості орієнтованого зваженого графа для формалізації асоціативної мережі образів з урахуванням відомих фізіологічних, нейропсихологічних і психолінгвістичних даних.

Визначення графа АМО. Побудуємо граф $G = \{V, E\}$ з метою формалізації концептуальної моделі процесів образної обробки ПМК у відповідності до пп. 2.5–2.7. З постановки задачі випливає, що граф G є орієнтованим графом. Множину вершин графа визначимо у вигляді

$$V = \{X, K, Y\},$$

де X – вхідні кількісні параметри, $X = \{X_i\}$, $i = \overline{1, r}$; K – проміжні якісні параметри, $K = \{K_j\}$, $j = \overline{1, p}$; Y – образи з власними вербальними позначеннями, $Y = \{Y_i\}$, $i = \overline{1, n}$; r – загальна кількість каналів нервового збудження органів чуття; p – загальна кількість якісних параметрів, які розрізняє людина.

На рисунку 3.1 представлено приклад графа для формалізації фізіології процесів образного мислення людини.

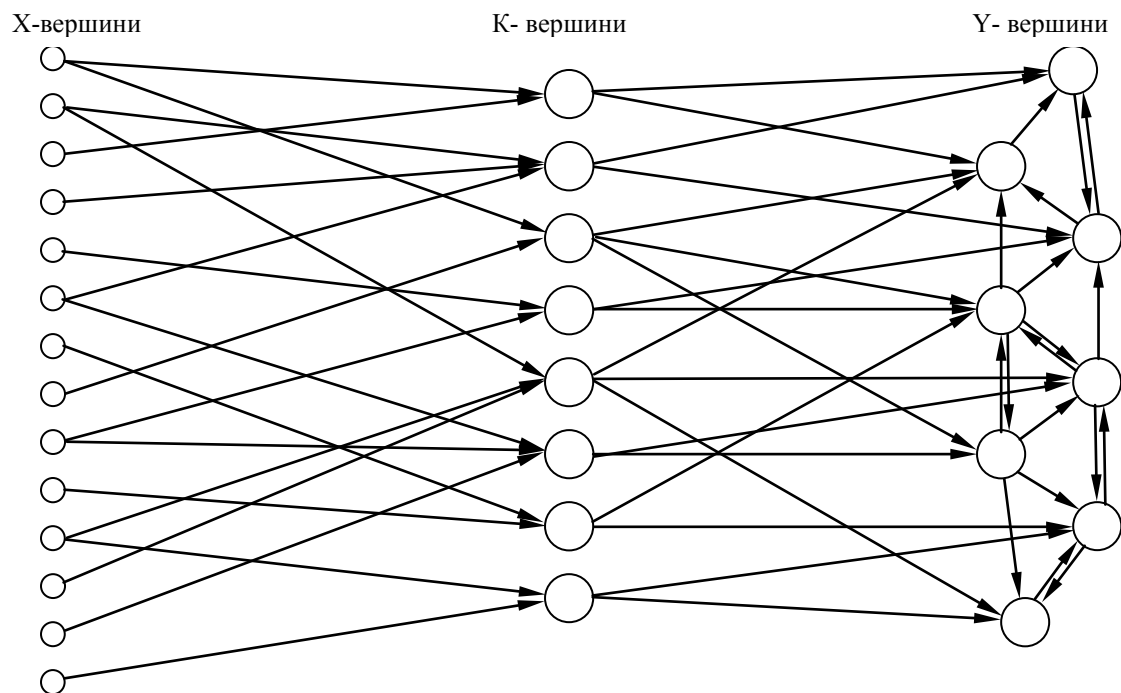


Рисунок 3.1 – Граф формального представлення процесів образного мислення людини

Оскільки певні якісні параметри відчуттів мають вербальний еквівалент, то очевидно, що

$$X \cap K = \emptyset, \quad (3.2)$$

$$X \cap Y = \emptyset, \quad (3.3)$$

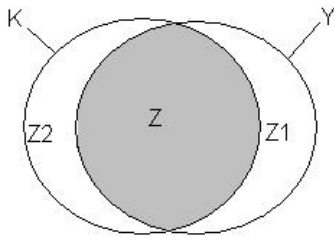
$$K \cap Y \neq \emptyset. \quad (3.4)$$

Введемо такі позначення:

$Z = K \cap Y$ – множина відомих мовних образів та відповідних їм якостей;

$Z_1 = Y \setminus Z$ – невідомі слова без відчуттів – множина вербальних позначень образів, neprисутніх у чуттєвому досвіді суб'єкта, отже, для яких ще несформовані якісні параметри;

$Z_2 = K \setminus Z$ – відчуття без слів – множина якісних параметрів, що розрізняються на рівні відчуттів, проте для яких у суб'єкта ще немає вербальних позначень.



З урахуванням цих позначень співвідношення (3.4) представлено на рисунку 3.2 у вигляді діаграми Ейлера–Венна. Відмітимо, що Z , Z_1 і Z_2 – непусті множини. Загальною множиною вершин орієнтованого графа G , що об’єднує всі різні вершини множин Z , Z_1 і Z_2 будемо вважати

Рисунок 3.2 – Діаграма перетину вершин V графа G

$$W = Z_1 \cup Z_2 \cup Z.$$

Зрозуміло, що $X \cap W = \emptyset$ як наслідок з (3.2) та (3.3).

Тепер визначимо множину ребер $E = \{(a, b)\}$ графа G . Нехай $(a, b) \in E$, тоді можливі такі варіанти:

1. $a, b \in X$ – не існує відповідно до постановки задачі;
2. $a \in X$, $b \in W$ – існує, проте не розглядається при моделюванні АМО;
3. $a, b \in W$ – складається з таких варіантів:

a) $a \in Z_1$, $b \in Z$;	b) $a \in Z_1$, $b \in Z_2$;
c) $a \in Z_2$, $b \in Z$;	d) $a \in Z_1$, $b \in Z_1$;
e) $a \in Z_2$, $b \in Z_2$;	f) $a_i \in Z$, $b_j \in Z$.

Варіанти a) та c) існують, проте їх можна віднести до окремого класу моделей орієнтувального рефлексу, а тому в дослідженні МО та КМО такі ребра не розглядаються. Варіанти b), d) та e) також існують, але для моделювання вже сформованої АМО їх можна не враховувати. Отже, тільки варіант f) відповідає АМО та в подальшому буде розглядатися як основний тип ребер моделі. З вищезазначеного випливає, що для моделювання АМО з графа $G = \{V, E\}$ мають бути відібрані підмножини вершин $V_z \subseteq V$ та ребер $E_z \subseteq E$. Фактично нас цікавить підграф $G_z(V_z, E_z) \subset G$, де $V_z \subseteq V$ та $E_z \subseteq E$. Для нього справедливо $V_z = Z$, а $(a, b) \in E_z$ (відмітимо, що $E_z = K \cap Y$).

Розглянемо спрямованість ребер $(a, b) \in E_z$. Будемо вважати, що a – початок ребра, а b – його кінець. Для ребра (a, b) принципово можливі такі випадки:

1. $a \in K, b \in K$ – не існує, виходячи з графа процесів образного мислення людини (див. рисунок 3.1);
2. $a \in K, b \in Y$ – існує, проте не розглядається, оскільки відноситься до класу моделей розпізнавання образів;
3. $a \in Y, b \in K$ – існує, проте не розглядається, оскільки у мовній практиці не прийнято визначати сутність образу через його власну якість («масло масляне»);
4. $a \in Y, b \in Y$ – існує та вважається основним типом ребер орієнтованого підграфа G_z .

Представимо множину $Y = Y_+ \cup Y_-$, причому $a \in Y_+, b \in Y_-$ для будь-якого ребра $(a, b) \in E_z$. Припустимо, що в загальному випадку, $Y_+ \cap Y_- \neq \emptyset$. Тоді розглянемо граф $G_z(Y_z, E_z)$, $Y_z = Y_+ \cup Y_- = Y$. Приклад графа G_z , що цікавить нас з метою моделювання АМО, зображено на рисунку 3.3 у вигляді двостороннього орієнтованого графа.

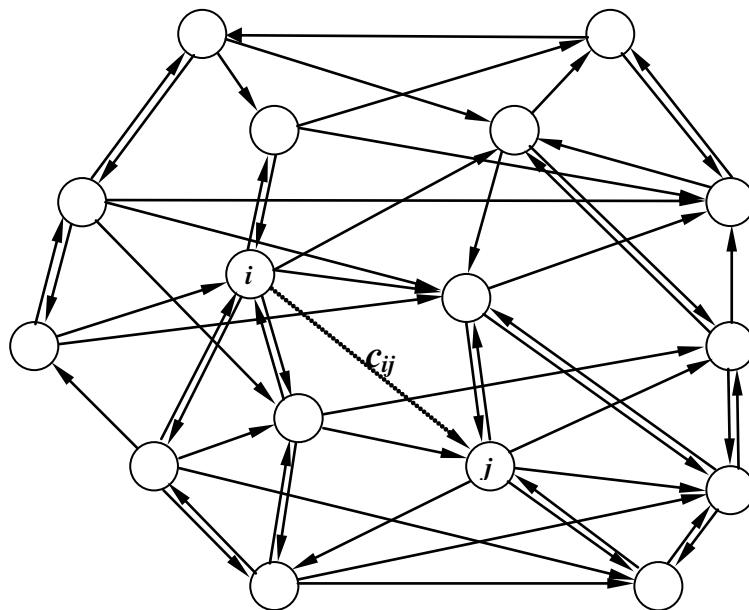


Рисунок 3.3 – Загальний вигляд орієнтованого графа G_z

Введемо поняття ваги та шляху для цього графа. Нехай $(a_i, b_j) \in E_z$, $a_i \in Y_+, b_j \in Y_-$. Тоді вага c_{ij} для ребра (a_i, b_j) визначається згідно з (3.1). Особливість задач на графі G_z у відповідності до мовної практики полягає в тому, що вершина не може одночасно бути початковою та кінцевою в одному шляху.

Як наскрізний приклад для апробації в подальшому моделей та алгоритмів розглянемо АМО, що представляє собою зв'язний підграф загального графа мовних образів елемента ЕК на тему «WEB-технології: стандарти Semantic WEB» (Додаток Б). Для наочності 27 ребер орієнтованого графа (рис. 3.4) позначені не вагами, а кількостями повторень k_{ij} , що визначені в результаті аналізу всіх КМО обраного елемента ЕК, тобто $Se_{ij} = F_1(k_{ij})$. Вершини графа в загальній кількості 12 позначені кодами мовних образів, інтерпретація яких надається в Додатку Б.

Зауважимо, що зображений на рисунку 3.4 граф АМО було автоматично згенеровано за допомогою *Graphviz* – розробленого спеціалістами лабораторії AT&T пакета утиліт з автоматичної візуалізації графів – на основі отриманих експериментальним шляхом даних.

Побудуємо для цієї АМО відповідний до неї граф G_z :

$Y_z - \{17, 34, 35, 37, 56, 71, 73, 87, 88, 93, 111, 123\}$;

$E_z - (17, 73) = 1; (34, 17) = 1; (34, 35) = 2;$

$(35, 17) = 1; (35, 34) = 2; (35, 56) = 2; (37, 37) = 2; (37, 111) = 2;$

$(56, 71) = 2; (56, 88) = 3; (56, 93) = 1; (56, 111) = 3; (56, 123) = 2;$

$(71, 56) = 3; (71, 93) = 1;$

$(73, 17) = 1; (73, 56) = 1; (73, 87) = 2;$

$(87, 123) = 1; (88, 87) = 1; (93, 71) = 1; (93, 87) = 1;$

$(111, 17) = 2; (111, 37) = 1;$

$(123, 37) = 2; (123, 56) = 2; (123, 87) = 1.$

Формальні поняття та ознаки графа асоціативної мережі образів. На відміну від довільного орієнтованого графа запропонована модель образної пам'яті у вигляді АМО дозволяє врахувати суттєві особливості його побудови з метою уникнення *NP*-повної задачі в алгоритмах пошуку на графі G_z . Для розгляду таких особливостей введемо такі формальні визначення:

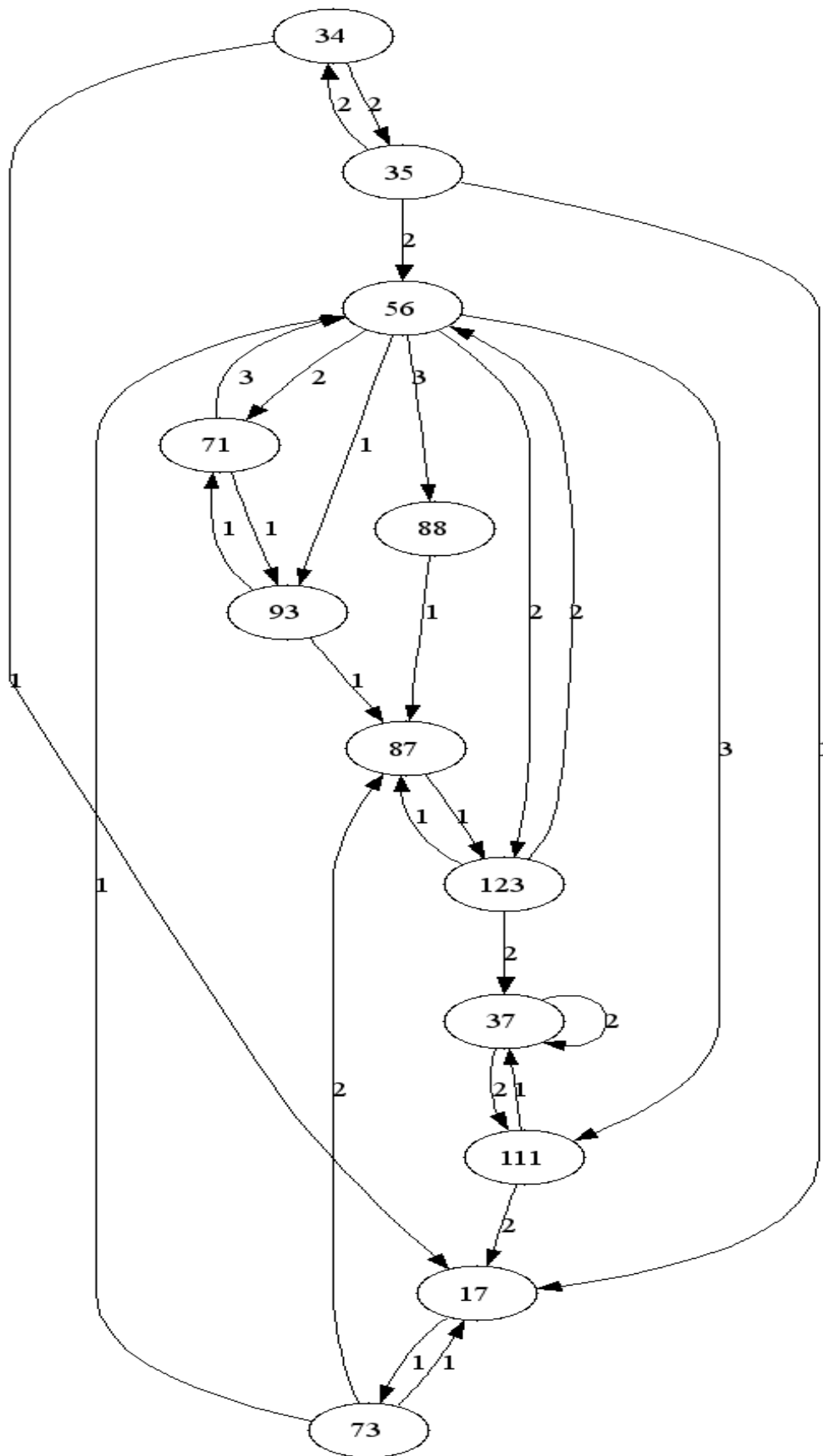


Рисунок 3.4 – Наскрізний приклад графа АМО

Мовний образ – вершина графа, що з'єднується з іншими вершинами не менше, ніж двома ребрами. Кожна вершина має унікальну множину зв'язків з іншими суміжними вершинами, тому у графі G_z не можуть існувати випадки так званих симетричних зв'язків між двома вершинами (рисунок 3.5). Це необхідно відслідковувати при додаванні нових вершин та ребер до моделі образної пам'яті, навіть без врахування спрямованості ребер.

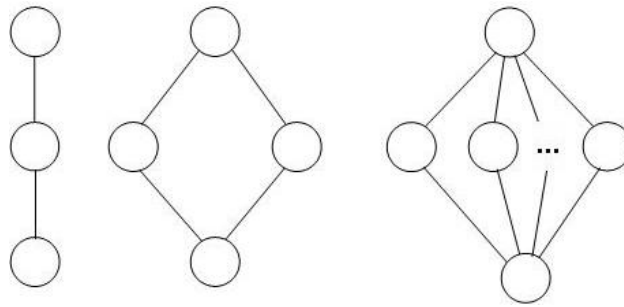


Рисунок 3.5 – Неприпустимі випадки симетричних зв'язків між двома вершинами графа G_z

Кількість ненульових (вага більша за 0) зв'язків від кожної вершини графа до інших вершин набагато менша за $n-1$. Згідно з практикою асоціативного експерименту, при великих значеннях n людина може підібрати не більше 7 ± 2 значимих слів-асоціатів на словостимул [136, 247]. Між будь-якими двома вершинами графа розмірністю n обов'язково існує скінченний шлях довжиною не більшою за $(n+1)/2$ у зв'язку з визначенням поняття МО. Найдовший скінченний шлях за відсутності симетричних зв'язків показано на рисунку 3.6.

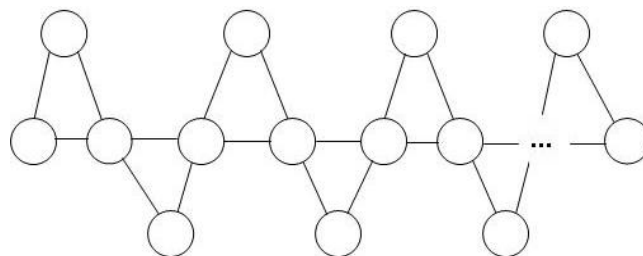


Рисунок 3.6 – Найдовший шлях між двома вершинами графа

Пара МО – ребро між двома суміжними вершинами графа з вагою, що визначається на основі (3.1). Пара МО як компонент події фіксує одиничний асоціативний зв'язок від головного образу до підлеглого [31]. При внесенні нових синтагм (подій) до моделі образної пам'яті необхідно збільшувати на 1 лічильники k_{ij} для тих пар, з яких складається синтагма (подія).

З урахуванням напрямку зв'язку вага ребра Se_{ij} в загальному випадку не дорівнює вазі ребра Se_{ji} . Проте для розв'язання певних задач образного мислення виникає необхідність врахування сумарної ваги між МО пари, в цих випадках для перетворення орієнтованого графа G_z у неорієнтований граф G'_z будемо застосовувати таке співвідношення для ваги ребра

$$c'_{ij} = \mu_Q(\langle i_i, i_j \rangle) + \mu_Q(\langle i_j, i_i \rangle) = Se_{ij} + Se_{ji}. \quad (3.5)$$

Синтагма – підграф розмірністю m загального орієнтованого графа G_z , такий що:

- може бути представлений у вигляді дерева, коренем якого є вершина-метод;
- загальна кількість спрямованих ребер (та відповідних пар), з урахуванням подвійного зв'язку між вершиною-об'єктом та вершиною-методом, $m \ll n$ (п. 2.2);
- у кожній вершині дерева кількість підлеглих вершин не може бути більшою за $m-1$ (п. 2.3).

Формально визначимо КМО як дерева $T = \{\langle V_T, E_T \rangle\}$, $V_T \subseteq V$, $E_T \subseteq E$ на основі рекурсивної процедури через асоціативну пару $(a, b) \in \Omega$. Мається на увазі:

- а) T є пара $(a, b) \in E$, тобто $V_T = \{a, b\}$, $E_T = (a, b)$;
- б) якщо T – це КМО, то $T \oplus (a, b)$ також є КМО тоді і тільки тоді, коли $a, b \in V$, $(a \vee b) \in V_T$, $(a, b) \notin E_T$, де \oplus – операція об'єднання ОК формальної теорії;
- в) ніяких інших, окрім а) та б) КМО не існує.

Отже, формалізація АМО у вигляді зваженого орієнтованого графа з двостороннім рухом дозволила врахувати суттєві особливості природної АВМ людини. Проведено формальне відокремлення класів

задач моделювання орієнтувального рефлексу та розпізнавання образів від класу загальних задач образного мислення. В таких термінах графа АМО як вершина, ребро, вага ребра та підграф отримано нову інтерпретацію базових концептів структурно-функціонального підходу до побудови СОПМК – мовного образу, асоціативного зв'язку, сили зв'язку та синтагми (події), визначено формальні обмеження цих концептів і шляху в графі.

3.2 Інтерпретація простору образного сенсу на основі булеану

Складність концепції образного сенсу ЕК та моделі образної обробки текстової інформації, що виникає під впливом необхідності врахування багатьох феноменологічних факторів, може бути певним чином зменшена за рахунок вибору максимально простих складових методології моделювання. Обмеження методу побудови нечіткого відношення сенсу (п. 2.4) та особливостей структурно-функціональної моделі (п. 2.7) все ж таки не дозволяють інтерпретувати частину важливих концептів образного мислення тільки на основі орієнтованих та неорієнтованих зважених графів (п. 3.1). У роботі [18] вперше обґрунтовано застосування відомого з теорії множин поняття булеану для формальної інтерпретації когнітивного простору асоціативних пар образів. Тим самим введено особливий клас графів, в якому враховано специфіку об'єкта дослідження.

Розглянемо можливість інтерпретації значимих концептів образного аналізу та синтезу шляхом ігнорування напрямку асоціативного зв'язку між парою образів та визначення його загальної сили згідно з (3.5). Будемо позначати через $P(I)$ множини-ступінь (булеан) множини I . За визначенням [248] ґратка $P(I)$ є скороченим позначенням для

$$P(I) = \{Y \mid Y \subseteq I\}. \quad (3.6)$$

Якщо будь-яка скінчена множина X містить в собі n елементів, то $P(I)$ очевидно має в своєму складі 2^n елементів (звідси, власне, назва – множина-ступінь). Пропонується зобразити булеан на площині у вигляді графа, який ілюструє відношення часткового порядку ґратки та складається з $n+1$ шару.

Застосуємо до $P(I)$ відношення порядку з таким визначенням: якщо $\langle I, \leq \rangle$ є скінченою частковою впорядкованою множиною, тоді

$I \leq Z$ рівнозначно тому, що існує ланцюг вигляду $I = I_1 \leq I_2 \leq \dots \leq I_n = Z$, в якому кожний I_{i+1} покриває I_i . Будь-яку скінченну впорядковану множину I можна представити у вигляді наочної схеми, де її елементи зображено у вигляді вузлів графа. При цьому вузол, що зображає Z , розташуємо вище вузла, що зображає I , лише в тому і тільки в тому випадку, коли $I \leq Z$, причому, якщо Z покриває I , то I та Z з'єднуються прямолінійним відрізком.

Такого роду схеми використовуються [249] як для того, щоб зобразити вже задані певним чином частково впорядковані множини, так і для первинного визначення частково впорядкованих множин. Відношення порядку в останньому випадку, за визначенням, є відношенням, що з'єднує елементи-вузли схеми, які поєднані між собою висхідними ламаними.

Множину-ступінь $P(I)$ множини образів I , що містить в собі n елементів, зручно представити у вигляді сукупності бінарних послідовностей (чисел), яка побудована на основі таких правил:

- кількість розрядів кожного числа відповідає n ;
- кількість всіх чисел дорівнює 2^n ;
- якщо i -й елемент I входить у цю підмножину $P(I)$, то в i -му розряді відповідного коду знаходиться 1, а інакше – 0;
- пуста множина \emptyset також входить в $P(I)$ та позначається як $000\dots 0$ (кількість розрядів дорівнює n).

Будемо вважати графічне зображення множини-ступеня $P(I)$, що задовольняє попередній опис, одиничним кубом n -вимірного простору (в n -вимірній системі координат). Закономірності зростань n -вимірного простору, що можуть бути виявлені на основі цього підходу, дозволяють розробити рекурсивний алгоритм графічної інтерпретації множини-ступеня $P(I)$. Фактично задача зводиться до розрахунку координатної проекції кожної точки n -вимірного простору на двовимірну площину у випадку $n > 2$.

На основі аналізу випадків з малими значеннями n за методом математичної індукції визначимо такі співвідношення для багаточислової графічної проекції ґратки булеану (n -вимірного багатогранника) на площину:

- кількість шарів: $n + 1$ (від 0 до n);
- кількість вузлів: 2^n ;

- кількість зв'язків від вузла i -го шару ($i = \overline{0, n}$):
 - вверх: $n - i$;
 - вниз: i ;
- кількість вузлів i -го шару (дорівнює числу сполучень з n по i) [250]:

$$C_n^i = \frac{n!}{i!(n-i)!}; \quad (3.7)$$

- загальне число зв'язків між вузлами графа:

$$\sum_{i=0}^n C_n^i \cdot (n-i) = \sum_{i=0}^n \frac{n!(n-i)}{i!(n-i)!}. \quad (3.8)$$

Запропонована графічна інтерпретація булеану проілюстрована на рисунку 3.7 для випадків тривимірної та чотиривимірної ґратки.

На рисунку 3.8 зображено фрагмент графа булеану для випадку $n = 10$, отриманого за допомогою пакета *GraphViz* в автоматичному режимі розташування вузлів на площині.

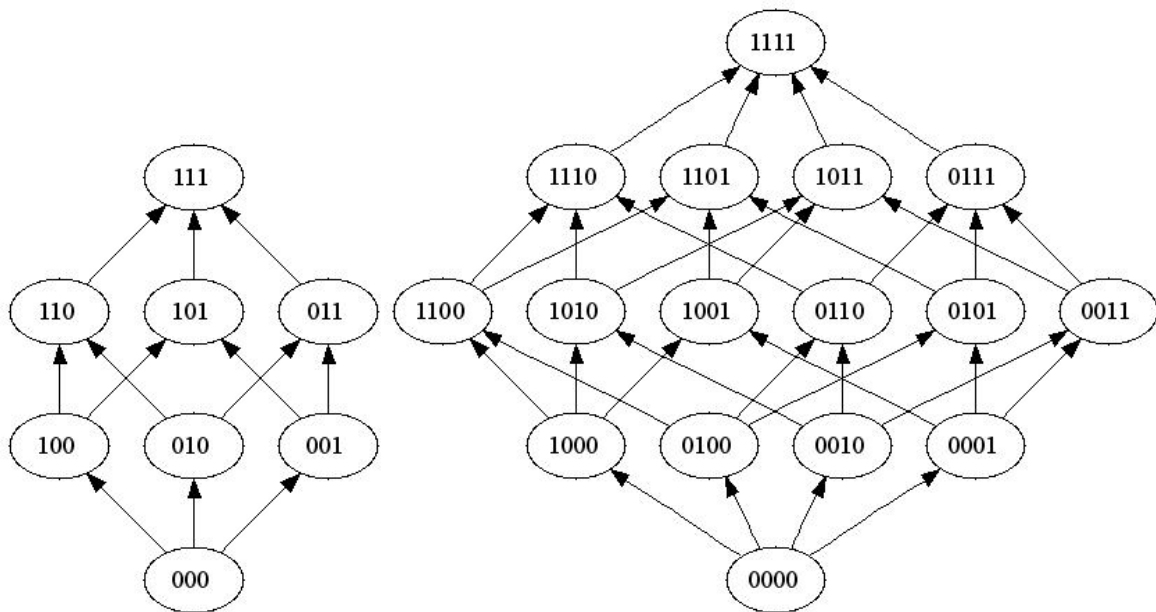


Рисунок 3.7 – Багатошарова графічна проекція ґратки булеану на площину для випадків $n = 3$ та $n = 4$

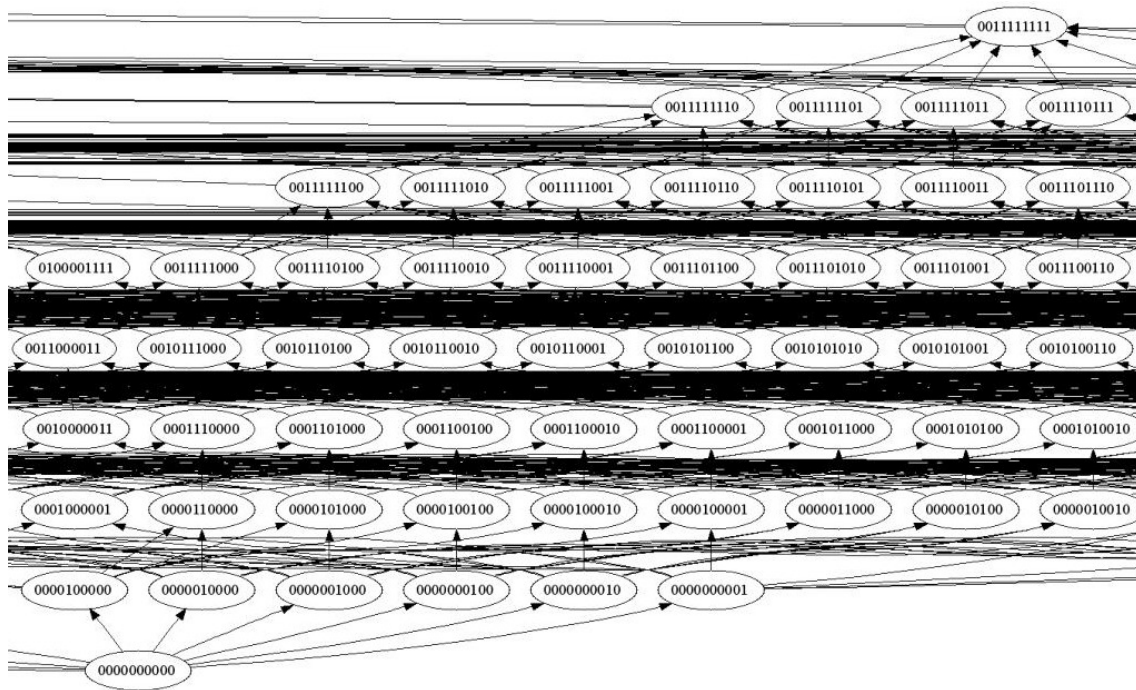


Рисунок 3.8 – Фрагмент графа булеану для випадку $n = 10$

Згідно з логікою моделі образного аналізу ПМК [14], на нульовому шарі графа знаходиться єдиний вузол (позначається кодом $000\dots000$ або символом Θ), який будемо вважати глобальним нулем. На першому шарі завжди знаходяться n вузлів з кодами $100\dots000$, $010\dots000$, $001\dots000$, ..., $000\dots010$, $000\dots001$ – можливі образи, з яких складається простір асоціативних пар. На другому шарі також завжди маємо множину всіх можливих асоціативних пар, які можна скласти з n образів – позначення кожного вузла має у двох розрядах одиниці, а в усіх інших розрядах коду знаходяться нулі. За такою самою схемою на i -му шарі кількість одиниць у коді кожного вузла дорівнює i , а за логікою моделі образного сенсу вузли третього шару та вище, до n -го, відповідають синтагмам та/або цілим текстам. Зрозуміло, що на n -му шарі знаходиться один вузол з одиницями в усіх n розрядах коду, який будемо вважати глобальною одиницею та позначати як Δ .

Розглянемо на кодах вузлів булеану унарну операцію доповнення вузла, а також такі дві бінарні операції, як об'єднання двох вузлів або знаходження їх найближчого спільного нащадка та перетин двох вузлів або знаходження їх найближчого спільного предка. Формально ці операції задамо у вигляді таблиці істинності (таблиця 3.1) для кожного з розрядів коду.

Таблиця 3.1 – Таблиці істинності для порозрядних операцій доповнення, об'єднання та перетину

Аргументи	Доповнення \neg	Об'єднання \cup		Перетин \cap	
		1	0	1	0
1	0	1	1	1	0
0	1	1	0	0	0

Необхідно зауважити, що, на відміну від бінарних операцій, доповнення або унарна операція знаходження діагонального вузла ґратки булеану не має предметної інтерпретації в моделі образного сенсу [18]. На графі ґратки координати вузлів з кодами, які є доповненням один до одного, розташовані симетрично відносно умовної центральної точки площини графа.

Формальну інтерпретацію когнітивного простору асоціативних пар на основі булеану будемо вважати дистрибутивною обмеженою ґраткою з доповненням до кожного елемента [250], оскільки для будь-яких вузлів ґратки a , b , c справедливо:

- а) за визначенням ґратки (властивості ідемпотентності, комутативності, асоціативності та поглинання)

$$a \cup a = a, \quad a \cap a = a; \quad (3.9)$$

$$a \cup b = b \cup a, \quad a \cap b = b \cap a; \quad (3.10)$$

$$a \cup (b \cap c) = (a \cup b) \cap c, \quad a \cap (b \cup c) = (a \cap b) \cup c; \quad (3.11)$$

$$(a \cap b) \cup a = a, \quad (a \cup b) \cap a = a; \quad (3.12)$$

- б) за дистрибутивністю

$$a \cup (b \cap c) = (a \cup b) \cap (a \cup c), \quad a \cap (b \cup c) = (a \cap b) \cup (a \cap c); \quad (3.13)$$

- с) згідно з властивістю обмеженості

$$a \cup \Delta = \Delta, \quad a \cap \Theta = \Theta; \quad (3.14)$$

$$a \cup \Theta = a, \quad a \cap \Delta = a; \quad (3.15)$$

d) згідно з існуванням та властивостями доповнення

$$a \cup \neg a = \Delta, \quad a \cap \neg a = \Theta; \quad (3.16)$$

$$\neg(\neg a) = a; \quad (3.17)$$

$$\neg(a \cap b) = \neg a \cup \neg b, \quad \neg(a \cup b) = \neg a \cap \neg b. \quad (3.18)$$

Вираз (3.18) фактично представляє собою закони де Моргана, які можуть бути застосованими до простору асоціативних пар у розглянутій інтерпретації.

Отже, перевага запропонованої у [18] інтерпретації когнітивного простору асоціативних пар на основі графа булеану та системи бінарних кодів полягає у можливості використання відомих властивостей (3.9)–(3.18) дистрибутивної обмеженої ґратки з доповненням до кожного елемента.

3.3 Поняття «піраміди сенсу»

Однією з важливих задач моделі образної обробки природномовного контенту [50] є накопичення зв'язків в асоціативній пам'яті на основі первинної інформації подій та/або синтагм з природномовних текстів. В останньому випадку в модель простору асоціативних пар необхідно ввести додаткові дані щодо параметрів об'єднання підмножини пар в синтагму, в тому числі час запису синтагми, тип і напрямок зв'язку в кожній її парі, вид і правила представлення кожного мовного образу в залежності від його ролі в певному асоціативному зв'язку. Крім того, модель має визначати і зберігати значення сили прямого і зворотного асоціативного зв'язку в кожній парі, що було вперше запропоновано в [54]. Окремі аспекти природних аналогів такого підходу формалізовані в роботах [108, 114, 115].

З метою ілюстрації підходу, що пропонується, як приклади КМО розглянемо українські речення «Студент дистанційно завдання отримав» з 4-х слів та «Багато мають нових ідей вчені» з 5-ти слів. Обидва обраних приклади демонструють випадки зміни звичного порядку слів речення, що, однак, зовсім не заважає читачеві зрозуміти його зміст. У таблиці 3.2 зображено вербальні ознаки мовних образів цих конструкцій, з яких створюються впорядковані пари за допомогою питальних займенників [15] від головного образу (рядки таблиці) до підлеглого (стовпчики таблиці).

Кожному МО в таблиці присвоєно свій бінарний код, який відповідає вузлу першого шару в графі булеану. Дані таблиці ілюструють підхід до кодування синтагм на основі графа булеану для випадків $n = 4$ та $n = 5$. Згідно з цим підходом, на другому шарі має бути активізовано стільки вузлів (асоціативних пар), скільки питальних займенників знаходиться у таблиці.

Таблиця 3.2 – Приклад асоціативних пар КМО

		студент	дистан- ційно	завдання	отримав	багато	мають	нових	ідей	вчені
		1000	0100	0010	0001	10000	01000	00100	00010	00001
студент	1000				Що зробив?					
дистан- ційно	0100									
завдання	0010									
отримав	0001	Хто?	Як?	Що?						
багато	10000									
мають	01000					Скільки?			Чого?	Хто?
нових	00100									
ідей	00010							Яких?		
вчені	00001						Що роблять?			

Також треба врахувати, що у кожному реченні може бути суміщення двох симетричних відносно головної діагоналі таблиці питальних займенників в один «подвійний» вузол (підмет і присудок речення). На рисунках 3.9 та 3.10, де наведено графи булеану для розглянутих прикладів, такі «подвійні вузли» на другому шарі разом з підлеглими вузлами першого шару позначено сірим кольором.

Наведені рисунки також ілюструють формальну можливість на основі інформації другого шару побудувати так звану «піраміду сенсу» окремої синтагми. Поняття «піраміди сенсу» вперше було введено в [22] на відміну від схожого поняття зростаючих пірамідальних мереж [114] при розгляді задач побудови лінгвістичного процесора на основі моделі образного сенсу ПМК.

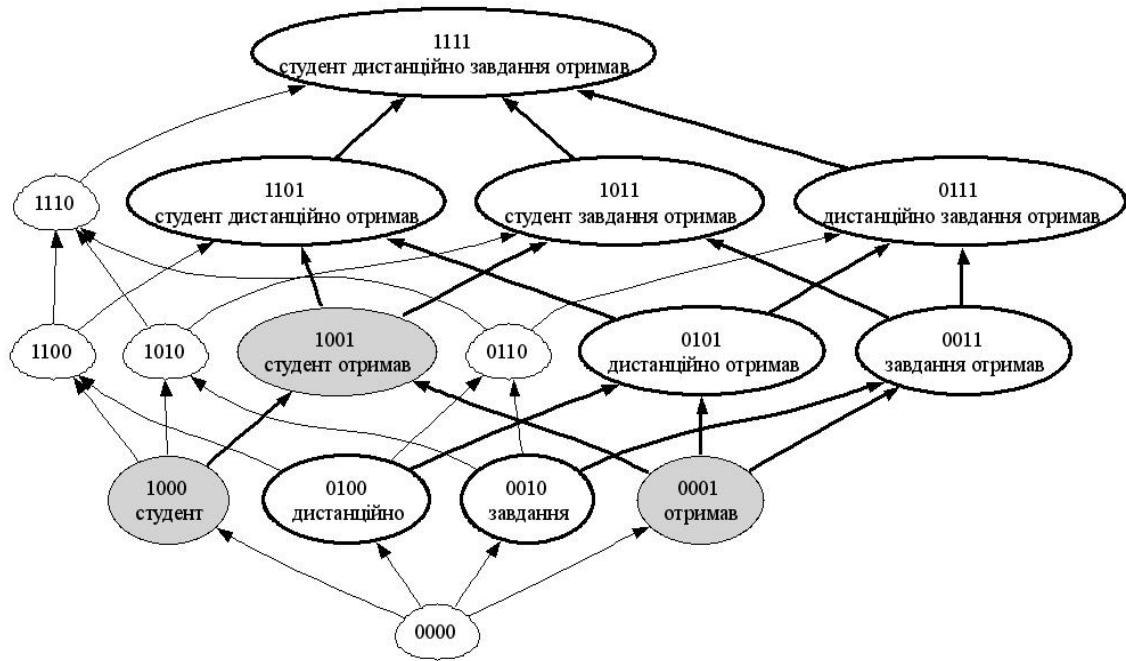


Рисунок 3.9 – Приклад графа булеану синтагми для випадку $n = 4$

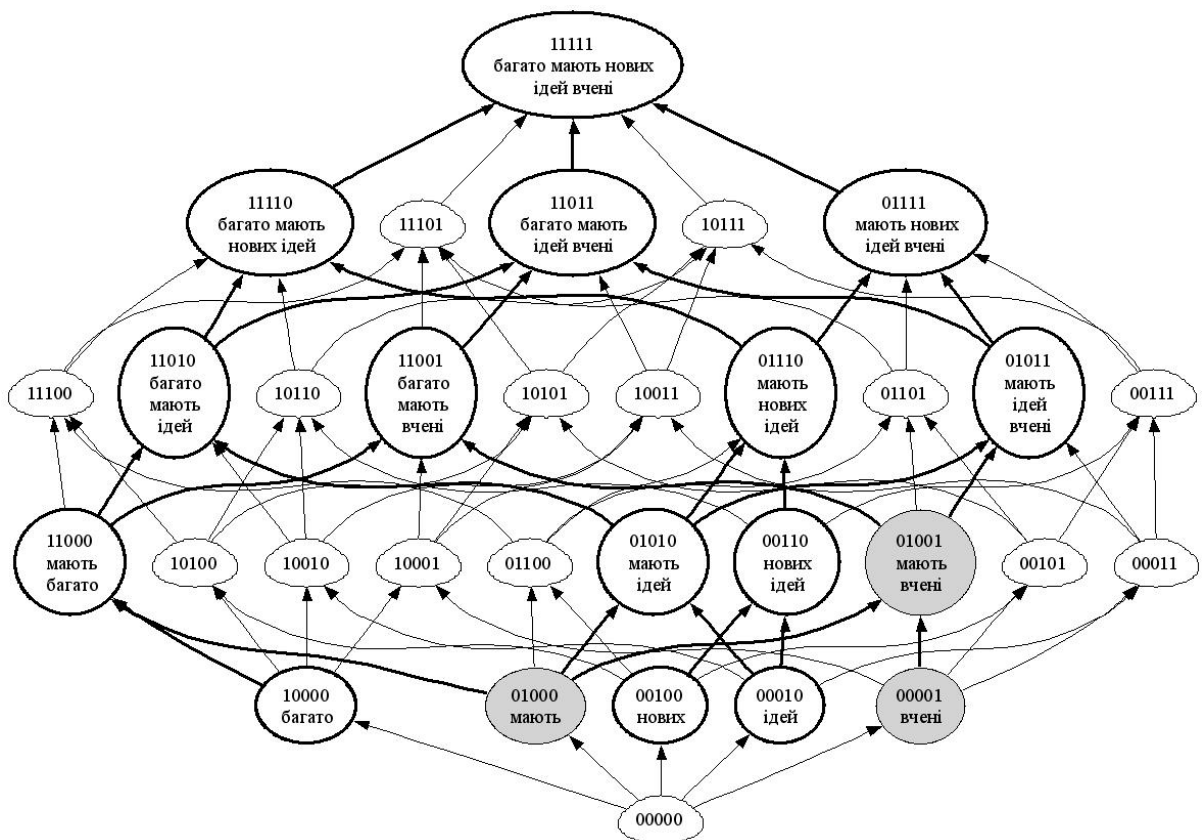


Рисунок 3.10 – Приклад графа булеану синтагми для випадку $n = 5$

Позначення в графі булеану тих вузлів та зв'язків між ними, що належать до окремої піраміди сенсу (виділені на рисунках напівжирними контурами та лініями), відбувається за таким алгоритмом:

1. Активізувати всі вузли першого шару, які відповідають МО певної синтагми *Syntagma*'.

2. На другому шарі активізувати тільки ті вузли, які відповідають асоціативним парам *Syntagma*', тобто у піраміді фіксується склад та напрямок зв'язків між мовними ознаками образів відповідної події.

3. На всіх шарах булеану, вищих другого, активізувати вузли, коди яких є результатом такої операції об'єднання \cup кодів довільної пари активованих на попередньому шарі вузлів, що тільки в одному з розрядів отриманого коду на відміну від кожного з початкових одиниць замінює нуль.

4. Для всієї «піраміди сенсу» активізувати тільки ті зв'язки, які можуть вести до активного вузла i -го шару з активних вузлів $i-1$ -го шару.

За допомогою поняття «піраміди сенсу» послідовність подій шляхом оброблення КМО синтагм фізично заноситься до довготермінової пам'яті системи у вигляді множини МО (1-й шар) та множини асоціативних зв'язків між ними (2-й шар). Проте з'являється принципово нова можливість формально відновити всі значимі вузли третього та наступних шарів, які можна інтерпретувати як такі поєднання образів (МО), що вже мали сенс у попередньому досвіді інфологічної системи.

Характерною для першого шару «піраміди сенсу» є така ознака підлеглості МО в даній синтагмі – вузол першого шару, що відповідає цьому образу є підлеглим тоді і тільки тоді, коли на другий шар від нього активізовано лише один зв'язок, причому не в «сірій», а в «білий» вузол (наприклад, вузли 10000 – «багато» та 00100 – «нових» на рисунку 3.10).

Ще одна важлива ознака «піраміди сенсу», що буде використана в подальшому, впливає з особливостей побудови синтагми – якщо «сірій» вузол вважати за два у зв'язку з його синтаксичною «подвійністю», то кількість задіяних в піраміді вузлів другого шару дорівнює кількості вузлів основи (першого шару). Цим самим забезпечується алгоритмічна можливість скоротити для запису в пам'ять кількість елементів списку булеану кожної синтагми від 2^n до $2 \times n$ практично без втрати змістовної інформації.

Таким чином, в результаті використання алгоритмічного поняття «піраміди сенсу» для кожної синтагми збільшується час її внесення в систему, проте з'являється можливість оминати *NP*-повну проблему при розв'язуванні тих задач образного аналізу та синтезу, що для свого розв'язання потребують поєднання образів, які вже мали сенс у попередньому досвіді інфологічної системи.

3.4 Булева алгебра сенсу та формалізація концептів теорії

Представимо модель процесів образної обробки ПМК за допомогою двоосновної алгебраїчної системи [31, 51]. З цією метою застосуємо інтерпретацію простору асоціативних пар у вигляді графа булеану з урахуванням поняття «піраміди сенсу» (п. 3.3). Будемо в подальшому вважати *булевою алгеброю сенсу* (БАС) двоосновну алгебраїчну систему:

$$BAS = \langle B; \Omega_b \rangle, \quad (3.19)$$

$$\text{де } B = \{Word, Number\} \text{ – основи,} \quad (3.20)$$

$$\text{а } \Omega_b = \{OP, RE, IF\} \text{ – сигнатура системи.} \quad (3.21)$$

В межах першої з основ *Word* варто розрізняти такі види символічних послідовностей (слів):

$$Word = \{Image - Word, Link - Type, Re - Word\}, \quad (3.22)$$

де слова, що позначають мовні образи *Image-Word* будемо класифікувати в залежності від їх ролі в синтагмі (п. 2.2), а саме *Object - Quality* – якість об'єкта; *Object* – об'єкт; *Notion* – поняття; *Method* – метод; *Method - Quality* – якість методу:

$$Image - Word = \left\{ \begin{array}{l} Object - Quality, Object, Notion, \\ Method, Method - Quality \end{array} \right\}. \quad (3.23)$$

Словами *Link-Type* будемо позначати ролі МО у реченні, що відповідає синтагмі за допомогою множини типів синтагматичних зв'язків,

$$Link - Type = \{r1, r2, r3, r4, r5, r6, r7\}, \quad (3.24)$$

де $r1$ – визначення, $r2$ – присудок, $r3$ – підмет, $r4$ – обставина місця, $r5$ – обставина часу, $r6$ – обставина, $r7$ – додаток. Подальша вербальна деталізація основних типів синтагматичних зв'язків може бути здійснена у вигляді підмножин відповідних питальних займенників *Pronoun*.

Словами з множини *Re-Word* будемо позначати атрибути відношень, що мають бути внесеними в блок пам'яті системи.

Як другу основу *Number* будемо розглядати числа таких видів:

$$Number = \{Bi, Logic, Integer, Time\}, \quad (3.25)$$

де Bi – n -вимірний бінарний код вузлів булеану; $Logic$ – логічні значення 0 (False) та 1 (True):

$$Logic = \{0,1\}; \quad (3.26)$$

$$Integer - \text{цілі невід'ємні числа: } Integer = \{x \mid x \in Z^+\}; \quad (3.27)$$

$$Time - \text{дійсні невід'ємні числа: } Time = \{x \mid x \in R^+\}. \quad (3.28)$$

Для n -вимірного бінарного коду будемо позначати через Bi_i код з i одиницями у будь-яких розрядах та, відповідно, $n-i$ нулями в усіх інших розрядах. Очевидно, що всі вузли з кодом Bi_i належать до i -го шару графа булеану. Формально

$$Bi = \{Bi_i, i = \overline{1, n}\}. \quad (3.29)$$

Отже, двоосновна алгебраїчна система БАС дозволяє застосувати інтерпретацію простору упорядкованих пар образів у вигляді графа булеану та врахувати як основи всі значимі концепти моделі образного мислення.

Запропонований формалізм надає можливість отримати теоретико-множинні визначення складових концептуальної моделі образного мислення, незважаючи на їх складний феноменологічний характер. Закладемо таку формальну ієрархію основних понять в алгебраїчну модель образного мислення [31]:

- Синтагма – множина бінарних кодів *Syntagma* відповідає таким вузлам булеану, що як вершини дерева графа «піраміди сенсу» є образ-

ним представленням події та формальним аналогом простого розповідного речення

$$Syntagma = \{x \mid x \in Bi_i, i = \overline{3, n}\}; \quad (3.30)$$

- Текст – множина бінарних кодів *Text* відповідає вузлам булеану, які є об'єднанням деякої послідовності *m* синтагм у таку зв'язну підмножину, де кожна з синтагм має з усіма іншими хоча б один спільний образ

$$Text = \{x \mid (x \in Bi_i, i = \overline{3, n}), (Bi - Sy_j \cap x \neq \emptyset)\}, \quad (3.31)$$

де $Bi - Sy_j \in Syntagma$ – *j*-та синтагма з низки об'єднаних у текст;

- Довготермінова пам'ять – вузол булеану Δ з бінарним кодом *Long – memory* (глобальна одиниця) об'єднує синтагми всіх текстів в загальну модель часової послідовності (ланцюжка) подій за допомогою поняття «піраміди сенсу»

$$Long - memory := x \in Bi_n \mid \bigcup_j Bi - Te_j \subseteq x, \text{ де } Bi - Te_j \in Text, \quad (3.32)$$

причому очевидно, що, за визначенням *n*-вимірного булеану, в кожний момент часу

$$Long - memory = \overbrace{1 \ 1 \ \dots \ 1}^n; \quad (3.33)$$

- Асоціативна пам'ять – множина бінарних кодів *Assoc – memory* всіх вузлів другого шару булеану

$$Assoc - memory = \{x \mid x \in Bi_2\}; \quad (3.34)$$

- Надоперативна пам'ять – множина бінарних кодів *Super – memory* складається з *n* вузлів першого шару булеану, кожний з яких може бути активізований в будь-який момент часу внаслідок моделювання процесу сприйняття відповідного образу органами відчуттів

$$Super - memory = \{x \mid x \in Bi_1\}; \quad (3.35)$$

- Фокус уваги – вказівник на один вузол булеану з бінарним кодом *Focus* з множини образів *Super – memory* в кожний з дискретних тактів часу

$$Focus := x \in Super - memory; \quad (3.36)$$

Оперативна пам'ять – множина бінарних кодів *Oper – memory* складається з таких вузлів булеану від 5-го шару до 9-го, кожний з яких представляє ансамбль активних образів на даний момент часу

$$Oper - memory = \{x \mid (x \in Bi_i, i = \overline{5,9})\}. \quad (3.37)$$

Для кращого розуміння сутності формальних виразів БАС введемо такі позначення:

$$\begin{aligned} Bi - I &\in Super - memory, \\ Bi - Tw &\in Assoc - memory, \\ Bi - OM &\in Oper - memory. \end{aligned} \quad (3.38)$$

Тоді для кожного з дискретних тактів часу можна пов'язати образ, що знаходиться у фокусі уваги з іншими образами оперативної пам'яті таким співвідношенням

$$Bi - OM \cap Focus \neq \emptyset. \quad (3.39)$$

Отже, у межах алгебраїчної системи БАС отримано формальні теоретико-множинні визначення для таких феноменологічних понять концептуальної моделі образної обробки ПМК (розділи 1 та 2), як синтагма, текст, фокус уваги, а також види пам'яті – довготермінова, асоціативна, оперативна та надоперативна.

3.5 Операції, предикати та відношення БАС

Функціональну складову концептуальної моделі образної обробки природно-мовного контенту відобразимо у сигнатурі алгебраїчної системи Ω_b , що складається з операцій *OP*, відношень *RE* та предикатів *IF* [22]. До складу базових операцій БАС будемо відносити:

$$OP = \left\{ \begin{array}{l} Add - Image, Add - Twice, Evaluate - Level, Find - Image, \\ Level - Up, Level - Down, Change - Image, Sense - Pyramid \end{array} \right\}, \quad (3.40)$$

де *Add-Image(Bi-I)* – додати новий образ – у перший шар булеану додається ще один вузол «крайнім з правого боку», код *(Bi-I)* для якого генерується шляхом додавання зліва нуля до коду попереднього вузла,

після цього до кодів всіх $n-1$ попередніх вузлів першого шару та всіх інших занесених до пам'яті інфологічної системи вузлів булеану справа додається нуль;

Add-Twice($Bi-I_1, Bi-I_2, Bi-Tw$) – додати нову пару – до другого шару булеану додається ще один вузол, код для якого $Bi-Tw$ будується шляхом

$$Bi - I_1 \cup Bi - I_2 \xrightarrow{Add-Twice} Bi - Tw; \quad (3.41)$$

Evaluate-Level(Bi, x) – визначити рівень вузла – підраховується кількість одиниць x у коді Bi вузла;

Find-Image($Bi-I_i, Bi-I_j$) – елементарний пошук асоційованого образу – знаходиться код $Bi-I_j$ образу, який має найбільше сумарне за модулем значення сили зв'язку (ваги Se_{ij}) з образом, що має код $Bi-I_i$;

Level-Up($Bi-OM, Bi-I$) – перейти на рівень уверх – до вузла оперативної пам'яті з кодом $Bi-OM$ додається образ з кодом $Bi-I$

$$Bi - OM \cup Bi - I \xrightarrow{Level-Up} Bi - OM; \quad (3.42)$$

Level-Down($Bi-OM, Bi-I$) – перейти на рівень вниз – від вузла оперативної пам'яті з кодом $Bi-OM$ віднімається образ з кодом $Bi-I$

$$Bi - OM \cap (\neg Bi - I) \xrightarrow{Level-Down} Bi - OM; \quad (3.43)$$

Change-Image($Bi-OM, Bi-I_1, Bi-I_2$) – навігація між вузлами того самого шару булеану – представляє собою композицію двох попередніх операцій

$$(Bi - OM \cap (\neg Bi - I_1)) \cup Bi - I_2 \xrightarrow{Change-Image} Bi - OM; \quad (3.44)$$

Sense-Pyramid – операції відновлення «піраміди сенсу», а саме

$$Sense - Pyramid = \left\{ \begin{array}{l} Pyramid - Up, \\ Pyramid - Images, Pyramid - Twice \end{array} \right\}, \quad (3.45)$$

де *Pyramid-Up*($Bi-Sy$) – відновити проміжні вузли піраміди – за відомими кодами вузлів другого шару по черзі визначаються коди $Bi-Sy$ наступного шару піраміди на основі послідовного використання такої операції об'єднання \cup за принципом «кожний з кожним», внаслідок якої отриманий код вузла наступного шару відрізняється від поперед-

ніх кодів тільки на одиницю в одному з розрядів; рекурсивне використання операції *Pyramid-Up* дозволяє для синтагми з i образами дістатися вершини піраміди (вузол i -го шару);

Pyramid-Images(Bi-Sy) – визначити основу піраміди – за відомим кодом вершини піраміди (код синтагми $Bi-Sy$) визначаються вузли її першого шару (образи, що входять до складу синтагми); організується цикл за довжиною коду вершини – якщо знаходиться одиниця, то до списку підлеглих вузлів першого шару додається ще один елемент з одиницею саме в цьому розряді та нулями в усіх інших – цикл можна обмежити за кількістю одиниць x у коді вершини;

Pyramid-Twice(Event-Id) – визначити множину пар піраміди – на основі додаткових відношень *Construct* та *Event* за кодом події $Event-Id$ визначаються вузли другого шару «піраміди сенсу», причому можна обмежити пошук коду асоціативних пар серед множини сполучень тих образів, що знайдено за результатом операції *Pyramid-Images(Bi-Sy)*.

Отже, за допомогою алгебраїчних операцій БАС реалізовано всі базові можливості навігації між вузлами булеану та формально задано три функції, що розкривають поняття «піраміди сенсу».

За допомогою предикатів, що є невід’ємною частиною алгебраїчної системи, визначимо значимі для концептуальної моделі образного мислення стани булеану. Задамо базовий склад предикатів БАС як [56]:

$$IF = \{If - Image, If - A - Twice, If - Event\}, \quad (3.46)$$

де $If-Image(Bi-I)$ – перевірка існування образу з кодом $Bi-I$

$$\exists Bi - I \mid Bi - I \in Bi_1 \rightarrow If - Image; \quad (3.47)$$

$If-A-Twice(Bi-Tw)$ – перевірка існування асоціативної пари з кодом $Bi-Tw$

$$\exists Bi - Tw \mid Bi - Tw \in Assoc - Memory \rightarrow If - A - Twice; \quad (3.48)$$

$If-Event(Bi-Sy)$ – перевірка існування синтагми з кодом $Bi-Sy$

$$\exists Bi - Sy \mid Bi - Sy \in Syntagma \rightarrow If - Event. \quad (3.49)$$

Відношеннями алгебраїчної системи, які дозволяють розглядати складові БАС у термінах реляційної моделі даних, будемо вважати [56]:

$$RE = \left\{ \begin{array}{l} Image, Assoc - Twice, Construct, Event, \\ Interrogative - Pronoun, Link, Text, Words, Role \end{array} \right\}, \quad (3.50)$$

де *Image* – образ – ставить у відповідність коду з множини *Bi-I* слова з підмножин, що складають множину *Image-Word*

$$\begin{aligned} Image \subset Bi - I \times Object - Quality \times Object \times \\ Notion \times Method \times Method - Quality; \end{aligned} \quad (3.51)$$

Assoc-Twice – асоціативна пара образів – пов’язує коди головного *Bi-I₁* і підлеглого *Bi-I₂* образів з кодом пари *Twice-Id* та значеннями сил прямого *Force⁺* та зворотного *Force⁻* зв’язку між ними

$$Assoc - Twice \subset Bi - I_1 \times Bi - I_2 \times Twice - Id \times Force^+ \times Force^-; \quad (3.52)$$

за допомогою відношення *Assoc-Twice* можна поставити у відповідність кожному образу відсортований за силою зв’язку підсписок асоційованих з ним образів (хеш-таблиця), що формально забезпечує ключову роль асоціативної пам’яті в образному аналізі та синтезі;

Construct – складова синтагми – пов’язує код складової *Construct-Id* з кодом питального займенника *Pronoun-Id*, кодом пари *Twice-Id*, кодами головного *Word-Id₁* та підлеглого *Word-Id₂* слів пари, кодом події *Event-Id* та кодом ролі підлеглого слова у цій парі *Role-Id*

$$\begin{aligned} Construct \subset Construct - Id \times Pronoun - Id \times Twice - Id \times \\ Word - Id_1 \times Word - Id_2 \times Event - Id \times Role - Id; \end{aligned} \quad (3.53)$$

При створенні нового кортежу у відношенні *Construct* паралельно у відповідний кортеж відношення *Assoc-Twice* обов’язково додається 1 або до значення в атрибуті *Force⁺* або до значення в атрибуті *Force⁻* залежно від того, в якому порядку (прямому або у зворотному) було реалізовано такий асоціативний зв’язок – цим самим за допомогою відношень враховується така інформація орієнтованого графа, що забезпечує отримання значення (3.5);

Event – подія – пов’язує код події *Event-Id* з кодом синтагми *Bi-Sy*, кодом відповідного тексту *Text-Id* та, власне, текстом речення *Sentence*

$$Event \subset Event - Id \times Bi - Sy \times Text - Id \times Sentence; \quad (3.54)$$

Interrogative–Pronoun – питальний займенник – пов’язує код займенника *Pronoun–Id* з кодом зв’язку *Link–Id*, до множини якого цей займенник належить та словом *Pronoun*, що, власне, є вербальною ознакою займенника

$$Interrogative - Pronoun \subset Pronoun - Id \times Link - Id \times Pronoun; \quad (3.55)$$

Link – асоціативний зв’язок – пов’язує код асоціативного зв’язку *Link–Id* зі словом *Link–Type*, що позначає цей тип зв’язку, та його вербальним описом *Specification*

$$Link \subset Link - Id \times Link - Type \times Specification; \quad (3.56)$$

Text – текст – пов’язує унікальний код *Text–Id* з бінарним кодом тексту *Bi–Te*, його назвою *Title*, автором *Author* та часом створення *Time*

$$Text - RE \subset Text - Id \times Bi - Te \times Title \times Author \times Time; \quad (3.57)$$

Words – вербальні ознаки образів – пов’язує унікальний код слова *Word–Id* з вербальним позначенням *Word* та кодом відповідного образу *Bi–I*

$$Words \subset Word - Id \times Word \times Bi - I; \quad (3.58)$$

Role – роль образу в події – пов’язує унікальний код ролі *Role–Id* зі словом *Role–Type*, що позначає роль образу у події

$$Role \subset Role - Id \times Role - Type. \quad (3.59)$$

Застосовані та описані вище назви атрибутів для відношень (3.51)–(3.59) визначимо таким чином через основи БАС:

$$Re - Word = \left\{ \begin{array}{l} Sentence, Pronoun, Specification, \\ Title, Author, Word, Role - Type \end{array} \right\}, \quad (3.60)$$

$$\begin{array}{l} Twice - Id, Force^+, Force^-, Construct - Id \in Integer, \\ Pronoun - Id, Word - Id, Event - Id \in Integer, \\ Link - Id, Text - Id, Role - Id \in Integer. \end{array} \quad (3.61)$$

Вищерозглянуті відношення моделі образної обробки ПМК у вигляді схеми бази даних зображено на рисунку 3.11, причому назви *n*-вимірних полів з кодами вузлів булеану для відношень *Image*, *Assoc–Twice*, *Words*, *Event* та *Text* починаються з префіксу *bi–*. Слід відзначити, що самостійну цінність в запропонованій реляційній моделі мають

відношення *Image i Assoc-Twice*, оскільки вони спільно утворюють асоціативну мережу образів і несуть головне змістовне навантаження. Решта всіх відношень або має допоміжне значення або фактично моделює ланцюжок подій (довготермінову пам'ять) [31].

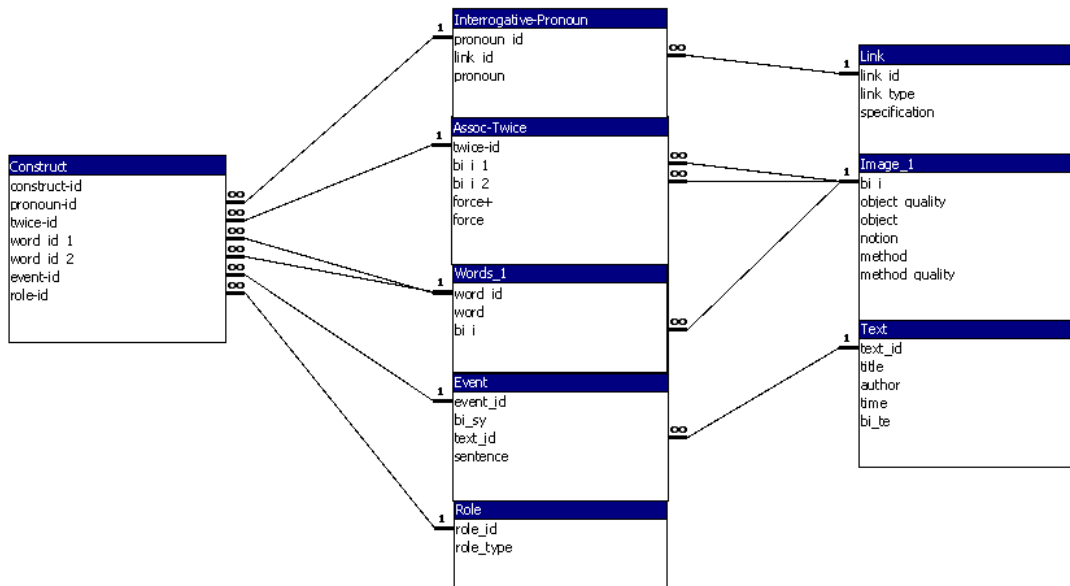


Рисунок 3.11 – Відношення моделі образної обробки ПМК у вигляді бази даних

Отже, за допомогою 3-х предикатів БАС визначено значимі для моделі образної обробки ПМК стани булеану, а введення 9-ти відношень дозволило представити складові основ алгебраїчної системи у вигляді реляційної моделі даних.

3.6 Організація бази знань інфологічної системи

Згідно з концепцією образної обробки ПМК, що пропонується, першочерговою задачею визначено моделювання онтогенетичного процесу побудови АМО на основі накопичення елементарних асоціативних зв'язків синтагматичного типу. Таке моделювання імітує природний шлях появи парадигматичних відношень між образами, коли більш складні мовні конструкції спираються на існуючий фундамент у вигляді простіших конструкцій. Розглянемо цей метод побудови парадигматичної піраміди образів як моделі природної онтології більш детально, враховуючи, що історія накреслення подібних онтологій веде свій відлік від Аристотеля [251].

Будемо виходити з того, що за основу моделі образної обробки ПМК прийнято єдиний словник вербальних ознак (тезаурус) мовних

образів, для якого визначають змістовні концепти з метою забезпечення первинного предикативного устрою мови. Внаслідок цього у складі МО розрізняють, у першу чергу, образи-об'єкти *O*, образи-методи *M*, образи-якості *OQ* та *H* (п. 2.2). Образи-поняття *N* мають допоміжне значення, оскільки вони, як правило, узагальнюють та доповнюють образи-якості та образи-методи. Запропонований підхід дозволяє аксіоматично закласти первинний (найміцніший) внутрішньообразний зв'язок P_6 між трьома ключовими концептами самого поняття «образ».

Побудова єдиного словника образів забезпечує сприйняття самого поняття «мовний образ» як неподільної сукупності об'єкт–дія–якість, що визначається на основі таких же триєдиних понять-образів. Необхідно вказати, що цим самим в модель образного мислення закладено парадигматичний парадокс. З філософської точки зору, маємо замкнене логічне коло, оскільки в концепції базове поняття (образ) визначається через себе (множину образів). Такий стан речей пояснюється феноменологічним характером образного мислення людини, проте висунута концепція не обмежується теоремою Геделя про неповноту формальних систем [252]. Аргумент для цієї тези – в запропонованій концепції застосовано метааксіоматичний рівень, на якому задаються принципи побудови моделі образного мислення та структура самих образів, аналогічно тому, як генетичний рівень визначає фізіологію людини.

За умови формального опису образу через інші образи виникає психологічна колізія видалення одного з образів. Інтроспекція та психоаналіз показують, що неможливо повністю висмикнути один образ з множини, але людині під силу як послаблювати за рахунок витіснення, так і зміцнювати певні групи зв'язків. Гуманістична літературна традиція, з свого боку, стверджує, що замкнуте коло не розірветься, проте стане біднішим – це питання добре висвітлили в своїх непересічних творах А. де Сент-Екзюпері та Е. Хемінгуей. Отже, в концепцію має бути закладено мінімальне (неподільне) коло образів. Логічно припустити, що таке коло має бути присутнє і у несвідомого немовляти і у представників вищого тваринного світу на рівні інстинктів.

Будемо вважати, що мінімальним колом є раніше окреслене Еґо-коло, первинне формування якого здійснюється на несвідомому рівні психіки. Еґо-коло з самого початку обов'язково включає два діаметрально протилежні полюси з умовними назвами «Добро» і «Зло». Решта всіх образів, формуючись з часом, лише вплітаються, знаходячи своє унікальне (внаслідок неповторності життєвого досвіду) місце в

складній павутині асоціативної мережі образів. Згідно з *припущеннями 7–10* побудови інфологічної моделі образної обробки ПМК (п. 2.6) той же самий блок асоціативної пам'яті грає ключову роль в інтелектуальній діяльності на рівнях несвідомості, підсвідомості та свідомості.

Декларативний характер парадигматичних відношень бази знань ІС забезпечимо за допомогою правил побудови класифікацій образів, множина яких складає модель когнітивного простору людини. Основними правилами, що походять з особливостей об'єкта та предмета дослідження, будемо вважати такі:

1. Вихідним матеріалом для побудови будь-якої парадигматичної класифікації образів є асоціативна мережа блока 7 структурної моделі та/або вже існуючі класифікації.

2. Метою побудови кожної парадигматичної класифікації образів є часткова задача пізнавальної діяльності.

3. За структурою парадигматичні класифікації представляють собою ієрархічні дерева. Вузлами ієрархії виступають образи-об'єкти або образи-поняття, причому кожному вузлу ставиться у відповідність певна множина образів-методів та образів-якостей.

4. Вузли ієрархії у вигляді образів-об'єктів володіють властивістю спадкоємства якостей і методів – дочірні вузли можуть наслідувати якості та методи материнських вузлів.

5. Важливою характеристикою методу є його спрямованість – розрізняють спрямованість методу на «свій» об'єкт і на «чужий» об'єкт.

6. Будемо розрізняти нетермінальні (проміжні) та термінальні (кінцеві) вузли ієрархічних дерев. Кожний нетермінальний вузол дерева характеризується множиною методів і якостей, які успадковуються дочірніми вузлами. Методи і якості дочірніх вузлів мають бути різновидами, що розкривають суть своїх попередників на материнських вузлах.

7. Термінальні об'єкти характеризуються наявністю хоча б одного різновиду хоча б одного з материнських методів або якостей, причому уособлюють цей різновид різні об'єкти з певної групи.

Актуальною задачею моделювання образної обробки ПМК є розробка методу автоматичної класифікації нових образів-об'єктів, що з'являються в асоціативній мережі. Якщо за основу такого методу береться вже наявне ієрархічне дерево парадигматичної будови мови, то маємо ознаку процедурного характеру або самовдосконалення бази знань. Природний інтелект без особливих складнощів виконує таку задачу, створюючи власні онтології як на несвідомому, так і на свідо-

тому рівнях. Приклад фрагменту подібної онтології образів-методів та образів-якостей для понять «тіло» та «істота» показано на рисунку 3.12, де методи, що спрямовані на «свій» об'єкт позначаються заливкою.

Розглянута в цьому розділі формальна інтерпретація моделі образної обробки ПМК у вигляді орієнтованого та неорієнтованого графів, у тому числі БАС, дозволяє запропонувати власну архітектуру бази знань інфо-логічної системи. Врахуємо, що розріджені графи, в яких кількість ребер $e \ll n$ економніше представляти у вигляді списків, а не матриць суміжності – наприклад, виконання відомого алгоритму Дейкстри в першому випадку обмежується часом порядку $O(e \log n)$ замість $O(n^2)$ у другому.

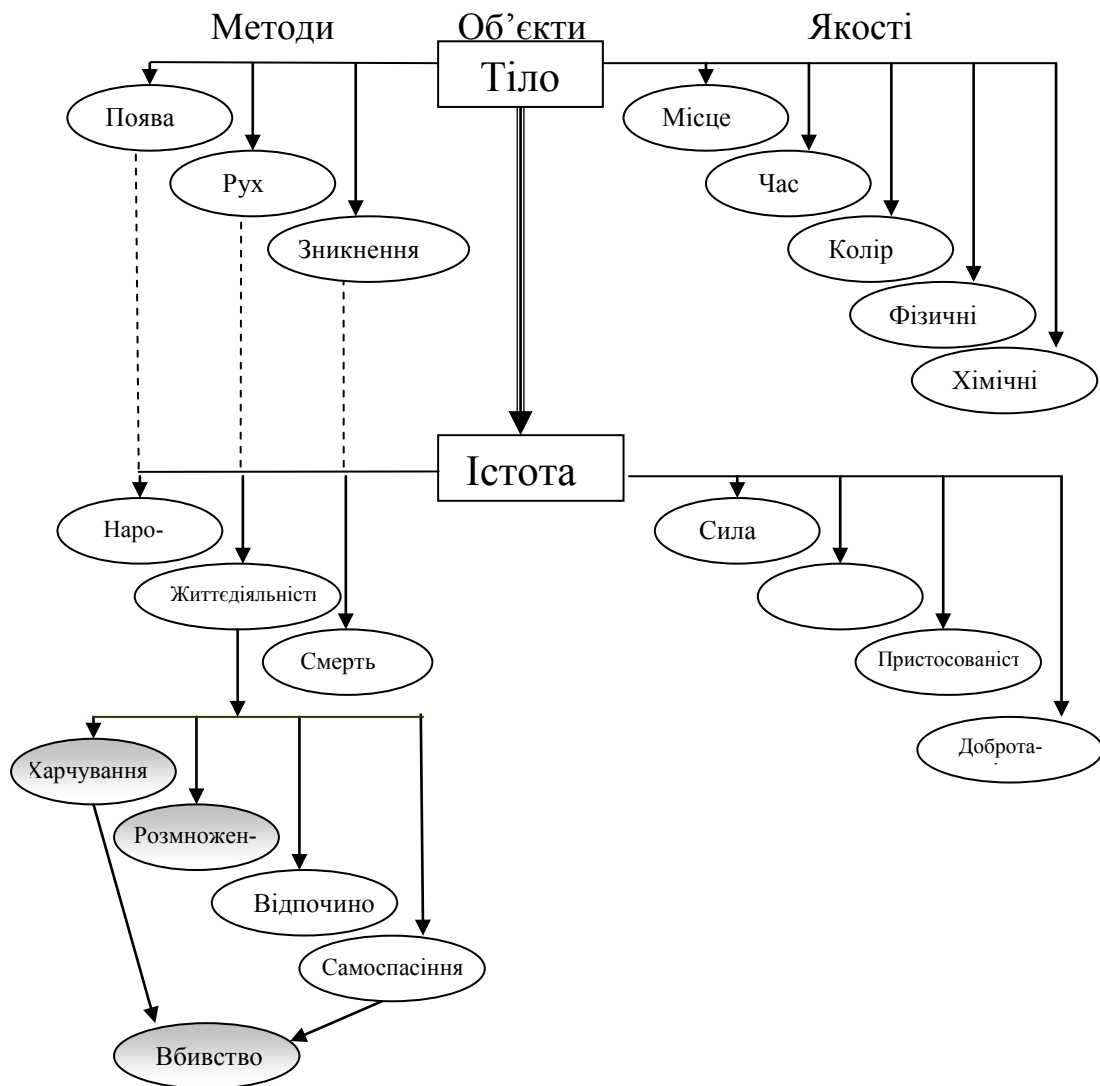


Рисунок 3.12 – Фрагмент онтології якостей та методів, отриманої інтроспективним шляхом

Виходячи з особливостей моделі, логічно застосувати такі принципи організації сховища ІС як декларативної складової бази знань ЕК за допомогою спискових структур [54]:

1. булеан-модель блока пам'яті розмірністю n образів складається з n списків вузлів на кожному шарі булеану, причому початкові списки є пустими; вхідною інформацією моделі будемо вважати послідовність (список) синтагм *Event* з ключовими елементами *Event-Id*, що відповідають непустим вузлам булеану 3-го шару та вище з бінарним кодом *Bi-Sy*;

2. булеан-модель блока пам'яті являє собою динамічну конструкцію, яка рекурсивно поповнюється (самовдосконалюється) в процесі внесення нових синтагм і збільшення кількості образів n та характеризується такими особливостями:

а) список першого шару *Image* (ключовий елемент – *Bi-I*) складається з n кодів булеану та інших ознак образів, що вже внесені в пам'ять;

б) список другого шару (власне асоціативна пам'ять *Assoc-Twice* з ключовим елементом *Twice-Id*) має найбільш складну структуру, оскільки для кожної пари образів з кодами *Bi-I₁* та *Bi-I₂* додаються значення лічильників *Force⁺* та *Force⁻*, а також відкритий підсписок *Construct* з елементів (*Construct-Id* (ключовий елемент), *Pronoun-Id*, *Twice-Id*, *Word-Id₁*, *Word-Id₂*, *Event-Id*, *Role-Id*), (... , ... , ... , ... , ... , ...), ... – довжина такого підписку визначається кількістю синтагм, в яких була задіяна пара з цим кодом;

3. якщо потрібно фіксувати належність певної послідовності синтагм окремому тексту, то для шарів від 3-го та вище можна формувати ще один список *Text* (ключовий елемент – *Text-Id*), де кожному непустому елементу з кодом *Bi-Te* може відповідати відкритий підсписок з елементів (*Event-Id*, *Bi-Sy*, *Text-Id*, *Sentence*), (... , ... , ... , ...), ... ; більш повна характеристика списку *Text-RE* буде досягнута за допомогою додаткових елементів *Title*, *Author* та *Time*.

Як допоміжні списки сховища ІС, необхідні для практичної реалізації лінгвістичної системи, пропонується застосувати:

1. *Interrogative-Pronoun* – список всіх питальних займенників (ключовий елемент – *Pronoun-Id*).

2. *Link* – список типів можливих асоціативних зв'язків у синтагмі (ключовий елемент – *Link-Id* – унікальний код типу зв'язку).

3. *Words* – список вербальних позначень (слів), що відповідають образам (ключовий елемент – *Word-Id* – унікальний код слова).

4. *Role* – список можливих ролей образу у синтагмі (ключовий елемент – *Role-Id* – унікальний код ролі).

Таким чином, запропоноване сховище ІС подібне до розширеної бази даних з текстовою інформацією, в якій:

1. Список *Text* відповідає таблиці з загальною інформацією про тексти, причому надлишковий бінарний код *Bi-Te* в кожному записі характеризує образний склад тексту.

2. Список *Event* відповідає таблиці з реченнями тексту, причому надлишковий бінарний код *Bi-Sy* в кожному записі характеризує образний склад речення.

3. Список *Image* відповідає таблиці–словнику або глосарію книги з текстами, причому надлишковий бінарний код *Bi-I* в кожному записі об'єднує навколо МО можливі його синтагматичні значення (якість об'єкта, об'єкт, поняття, метод, якість методу).

Допоміжні конструкції сховища ІС забезпечують важливі функції лексичного процесору, а саме:

1. В таблицях *Words* та *Role* вміщено слова речення (синтагми) та показано, яку семантичну роль відіграють відповідні їм мовні образи у події, що описана цим реченням.

2. В таблицях *Construct* та *Assoc-Twice* зосереджена інформація про синтагматичні зв'язки між парами слів у реченні.

3. Таблиці *Interrogative-Pronoun* та *Link* вміщують пов'язану інформацію про питальні займенники, що відповідають можливим синтаксичним зв'язкам між парами слів у реченні.

З метою ілюстрації підходу та подальшого використання в моделях та алгоритмах розділу 4 в таблиці 3.3 представлено нумерацію, кодування та інтерпретацію мовних образів графа з прикладу п. 3.1.

Таблиця 3.3 – Кодування та інтерпретація мовних образів графа з прикладу п. 3.1

№ з/п	Код $Vi-I$	Код МО	ОQ	О	N	M	H
1	000000000001	17	мережевий	мереживо	мережа	мережити	{null}
2	000000000010	34	інформаційний	{null}	інформація	інформувати	{null}
3	000000000100	35	змістовний	зміст	розміщення	розміщувати	змістовно
4	000000001000	37	визначений	{null}	означення	означати	{null}
5	000000010000	56	мовний	{null}	мова	мовити	{null}
6	000000100000	71	існуючий	{null}	існування	є	{null}
7	000001000000	73	ресурсний	{null}	ресурс	{null}	{null}
8	000010000000	87	{null}	RDF	RDF	{null}	{null}
9	000100000000	88	{null}	XML	XML	{null}	{null}
10	001000000000	93	{null}	{null}	засіб	{null}	{null}
11	010000000000	111	онтологічний	{null}	онтологія	{null}	{null}
12	100000000000	123	дозвільний	{null}	дозвіл	дозволяти	дозволено

Оскільки кінцева реалізація блока пам'яті інфологічної системи практично побудована на реляційній базі даних, варто порівняти концепцію образного сенсу ПМК з класичними поняттями dBase-машини, що відповідають реляційній моделі даних. В роботі [22] вперше запропоновано провести порівняльний аналіз двох концепцій в табличній формі, де кожному поняттю dBase-машини відповідають поняття сховища ІС на рівнях органів чуття (у вигляді ОК) та мовних засобів (як КМО та ПМК). Для зручності сприйняття інформації в таблиці 3.4, де наведено результати такого аналізу, позначено сірим кольором ті три рядки, в яких виявлено принципові відмінності двох концепцій.

Таблиця 3.4 – Порівняльний аналіз концепцій dBase-машини та сховища ІС

№ з/п	Поняття dBase-машини	Поняття сховища ІС	
		Рівень ОК	Рівні КМО та ПМК
1	Реляційна алгебра	Алгебраїчна система БАС	
2	База даних	Булеан з додатковими відношеннями	
3	Реляційний зв'язок ($1 \rightarrow 1$, $1 \rightarrow \infty$)	Прямий та зворотний асоціативний зв'язок між образами у вигляді значень полів $Force^+$, $Force^-$ відношення Assoc-Twice	

Продовження табл. 3.4

№ з/п	Поняття dBase-машини	Поняття сховища ІС	
		Рівень ОК	Рівні КМО та ПМК
4	Індексний файл	Хеш-таблиця 2-го шару булеану	
5	Відношення (таблиця)	Вузли верхніх шарів булеану з додатковими відношеннями	Текст
6	Робоча область (select)	Оперативна пам'ять (5–9 шари булеану)	
7	Активне відношення (use)	Реєстраційний ланцюжок подій	Активний текст
8	Атрибут (поле)	Образ-ознака іншого образу	Чинник структури речення
9	Домен	Можливі значення ознак, тобто множина існуючих асоціацій образу	Множина всіх можливих значень кожного з чинників структури речення
10	Кортеж (запис)	Значимий образ (об'єкт, дія, суб'єкт) події або ансамбль образів	Синтагма – речення
11	Вказівник запису	Фокус уваги на значимому образі	Активне речення тексту
12	Перший запис (top)	Перша подія реєстраційн. ланцюжка	Перше речення тексту
13	Останній запис (bottom)	Остання подія реєстраційн. ланцюжка	Останнє речення тексту
14	Перехід на наступний запис (skip)	Фокус уваги на значимому образі наступної події	Перехід до наступного речення тексту
15	Перехід на попередній запис (skip –)	Фокус уваги на значимому образі попередньої події	Перехід до попереднього речення тексту
16	Абсолютний перехід на запис (go)	Перехід до образу за сукупністю образів-ознак	
17	Індексний пошук (seek)	Перехід до образу за одним з типів асоціативного зв'язку	

Як видно з таблиці, до найвагоміших відмінностей двох концепцій можна віднести:

- Відсутність у сховищі ІС чітко визначених атрибутів, що приводить до відсутності сталої структури та розміру кортежів.
- Принципову неможливість модифікації даних в сховищі ІС, внаслідок чого помилки можна виправляти лише опосередковано,

шляхом накладання (додавання до сховища) нових подій та/або речень.

- Можливість знаходити значимий образ події за сукупністю образів-ознак, тобто абсолютний перехід на запис фактично замінено складним індексним пошуком, більш характерним для людини.

Надамо оцінку інформаційної збитковості потенційних технологічних засобів підтримки образного сенсу ПМК. З цією метою проведемо порівняння задіяних інформаційних ресурсів з відомими підходами.

1. Збитковість бази даних. Оскільки можна вважати, що потік вхідної інформації СОПМК представляє собою множину текстових файлів, то потрібний для їх зберігання обсяг інформації є пропорційним загальній кількості слів n_v в усіх цих файлах. Запропонований підхід передбачає визначення n мовних образів, довжина яких пропорційна середній довжині слова, причому, в загальному випадку $n_v \geq n$, а при $n_v \rightarrow \infty$ за рахунок ефекту дублювання слів $n_v \gg n$. Загальний обсяг інформації у Бт для відношень бази даних (п. 3.5) оцінимо з таких міркувань:

- загальна кількість записів у відношеннях *Interrogative–Pronoun*, *Link*, *Text*, *Role* дорівнює $n_{const} \ll n$, отже всі вони додають до БД обмежену деякою константою E [Бт] кількість інформації;

- кількість записів у відношенні *Image* дорівнює n , але, в зв'язку з необхідністю зберігання поля з ідентифікаційним кодом та додаткових полів, всього інформації потрібно $n \cdot (C' + \log_{256} n)$, де C' – деяка константа [Бт];

- аналогічним чином оцінимо кількість інформації у відношеннях *Words*, *Event*, *Construct* як $n \cdot (C'' + \log_{256} n)$, де C'' – константа [Бт], $C'' > C'$;

- найбільшу кількість інформації додає до БД відношення *Assoc – Twice* з кількістю записів n^2 , а саме $n^2 \cdot (A + 4 \cdot \log_{256} n)$, де A – константа [Бт], $A < C'$.

Отже, сумарний обсяг інформації для відношень БД складає $V_\Sigma = A \cdot n^2 + B \cdot n^2 \cdot \log_{256} n + C \cdot n + D \cdot n \cdot \log_{256} n + E$, де A, B, C, D, E – константи [Бт]. Зрозуміло, що з моменту досягнення співвідношення $n_v > n^2$ кількість інформації V_Σ БД наблизиться до, а потім стане

меншим від обсягу інформації, потрібної для зберігання множини вхідних текстових файлів.

2. Збитковість бази знань. На відміну від отриманої шляхом нормалізації відношень схеми БД запропоновані засоби підтримки образного сенсу ЕК використовують базу знань у вигляді модифікованих згідно з методом моделювання (пп. 3.2–3.5) відношень БАС. З інформаційної точки зору різниця між БД та базою знань полягає у додаванні до частини відношень останньої полів з n -розрядним бінарним кодом, а саме:

- відношення *Text* з кодом *Vi-Te* має розмір $n_{const} \cdot (E' + n/256) \approx n'_{const} \cdot n + E'$, де $n_{const}, n'_{const}, E'$ – деякі константи;
- відношення *Image* з кодом *Vi-I* має розмір $n \cdot (C' + n/256) \approx n^2/256 + C' \cdot n$, де C' – деяка константа [Бт];
- відношення *Event* з кодом *Vi-Sy* має розмір $n \cdot (C'' + \log_{256} n + n/256) \approx n^2/256 + C'' \cdot n + n \cdot \log_{256} n$, де C'' – константа [Бт];
- відношення *Assoc – Twice* з кодами *Vi-I₁* та *Vi-I₂* має розмір $n^2 \cdot (A' + 2 \cdot \log_{256} n + n/128) \approx F \cdot n^3 + A' \cdot n^2 + B' \cdot n^2 \cdot \log_{256} n$, де F, A', B' – деякі константи [Бт].

З урахуванням співвідношень БД загальний обсяг бази знань складає $V_{\Sigma}^{BAS} = F \cdot n^3 + A \cdot n^2 + B \cdot n^2 \cdot \log_{256} n + C \cdot n + D \cdot n \cdot \log_{256} n + E$. Отже, V_{Σ}^{BAS} збільшилося порівняно з V_{Σ} на член полінома третього степеня.

3. Збитковість інформаційного забезпечення графових моделей. Відомо, що орієнтований або неорієнтований граф розмірністю n задається матрицею суміжності розмірністю n^2 . У базі знань на основі БАС роль матриці суміжності грає відношення *Assoc – Twice*, розмірність якого, як було показано раніше, досягає $F \cdot n^3 + A' \cdot n^2 + B' \cdot n^2 \cdot \log_{256} n$, де F, A', B' – константи [Бт]. Маємо не критичне уповільнення роботи відомих алгоритмів оброблення графів, тим більше, що врахування бінарних кодів необхідне не для всіх функцій інфологічної системи.

Додатково зауважимо, що оброблення n -вимірних бінарних кодів булеану за допомогою операцій (3.40) та предикатів (3.46) потребує

значно більше пам'яті (2^n), ніж їх зберігання (n^3). Але таке обмеження рівня *NP*-повної проблеми для алгебраїчної системи *BAS* не варто вважати критичним, оскільки побітова послідовність виконання всіх складових (3.40) та (3.46) дозволяє застосувати для їх виконання будь-яку необхідну ступінь розпаралелювання відповідних обчислювальних процесів.

Отже, за допомогою структури з 9-ти взаємопов'язаних списків або відношень сховища ІС, що грає роль декларативної складової бази знань на основі БАС забезпечено реалізацію як синтагматичної першооснови, так і парадигматичних можливостей моделі образної обробки природно-мовного контенту. Природна мінливість атрибутів та кортежів даних, що врахована у СОПМК, дозволяє досягти більшого ступеня гнучкості операцій пошуку інформації. Такий підхід збільшує обсяг бази знань ІС до поліноміальної складності інформаційного забезпечення з порядком не більше третього внаслідок включення до складу відношень БАС бінарних *n*-вимірних полів з кодами вузлів булеану на рівні образів, асоціативних пар, синтагм та текстів.

РОЗДІЛ 4

СИНТЕЗ ФУНКЦІЙ ОБРАЗНОГО ПОШУКУ ТА ГЕНЕРАЦІЇ ЗНАТЬ СИСТЕМИ ОБРОБКИ ПРИРОДНО-МОВНОГО КОНТЕНТУ

Функціональна модель системи обробки природно-мовного контенту згідно з онтогенетичним принципом будується на класифікації типів образного пошуку. Розглядаються взаємопов'язані алгоритми обходу графа асоціативної мережі образів з метою моделювання інсайту і побудови ланцюга образів та формалізації орієнтувального рефлексу і механізму оперативної пам'яті. Обґрунтовується метод самодосконалення бази знань системи на основі моделювання базових складових парадигматичного устрою мови.

4.1 Функціональна модель системи обробки природно-мовного контенту на основі класифікації можливих типів образного пошуку

Як показано у розділі 1, пошук образів за допомогою асоціацій закладено в основу природних інформаційних систем, що мають ознаки інтелектуальності. Аксиоматика образного сенсу (розділ 2) побудована на головному постулаті – єдиним методом сприйняття інфологічною системою інформації з навколишнього світу є образи та асоціативні зв'язки між ними. Тому пропонується формально визначати базові функції СОПМК на основі класифікації можливих типів асоціативного образного пошуку. Варто також відмітити, що сфера використання асоціативного пошуку ПМК через КМО набагато ширша за предмет дослідження – до неї також можна віднести [253] експертні системи, бази даних та бази знань, інформаційно-пошукові системи, автоматичний переклад, програмні засоби анотування та реферування текстів тощо.

Виходитимемо з того, що кілька базових варіантів пошуку зустрічаються частіше в інтелектуальній діяльності, оскільки для їх реалізації потрібні лише первинні синтагматичні відношення між образами [254]. Всі інші варіанти образного пошуку можна вважати вторинними, похідними від базового типу внаслідок появи парадигматичних відношень (нових асоціативних зв'язків) між образами. Вперше ідея підходу до класифікації типів асоціативного пошуку образів на основі інформації з подій, що відбулися, запропонована в [19]. Сутність цієї ідеї, яка впливає з концептуальної моделі образної обробки ПМК [31], можна викласти такими тезами:

- інформація до інфологічної системи потрапляє у вигляді синтагм, якими описано послідовність подій;
- асоціативні зв'язки між образами за своєю природою діляться на синтагматичні та парадигматичні;
- елементарними складовими асоціативного пошуку вважаються мовний образ (x), сенссполучення (y) та синтагма (z);
- в класифікацію типів пошуку закладається пріоритет синтагматичного мислення;
- застосовується онтогенетичний принцип «навчання дитини» для накопичення парадигматичних зв'язків між мовними образами [30].

Розглянемо можливі варіанти образного пошуку, одиницями якого будемо вважати МО, сенссполучення та синтагму. Формально позначимо мовний образ як елемент множини $x \in X$, якщо $X \subseteq I$. Сенссполучення представимо у вигляді множини пар $Y = \{(x, a)\}$, де a – сила асоціативного зв'язку між парою МО згідно з (3.1), що встановлюється відображенням

$$A : X \times X \rightarrow a. \quad (4.1)$$

Синтагми або сукупність взаємопов'язаних у дерево КМО (п. 3.1) асоціативних пар МО представимо як елементи множини $z \in Z$. Виходячи з введених позначень, пропонується така класифікація образного пошуку, в якому базовий синтагматичний тип відмічено сірим кольором:

1. **МО \rightarrow МО** у вигляді відсортованого за зменшенням списку найбільш близьких за силою асоціативного зв'язку мовних образів:

$$AList : X \rightarrow X. \quad (4.2)$$

2. **МО \rightarrow МО**, де обидва мовних образи близькі синонімічно, наприклад, «світлий \leftrightarrow ясний», «реве \leftrightarrow стогне», «високий \leftrightarrow стрункий»:

$$SynXX : X \rightarrow X. \quad (4.3)$$

До цього парадигматичного зв'язку мовних образів близькими є такі види пошуку, як антонімія $AntXX$, окреме–загальне $PgXX$ і частина–ціле $CmXX$.

3. **МО \rightarrow пара**, в якій однією з складових знайденого сенссполучення є даний МО («Дніпро широкий», «хвилі Дніпра»):

$$PartXY : X \rightarrow Y. \quad (4.4)$$

4. MO → сенсосполучення – пошук синонімічного словосполучення до даного MO, наприклад, «обманювати ↔ водити за ніс», «ледарювати ↔ бити байдики»:

$$SynXY : X \rightarrow Y. \quad (4.5)$$

5. MO → синтагма, коли за одним MO можна згадати речення, подію або прислів'я («гривонька → ой чий то кінь стоїть»):

$$PartXZ : X \rightarrow Z. \quad (4.6)$$

6. Сенсосполучення → MO – результат впливу (влиття за Виготським) сенсу пари в MO, наприклад, (рос.) «тело жирное → пингвин», «гордо реет → буревестник»:

$$InsYX : Y \rightarrow X. \quad (4.7)$$

7. MO → X-Image → MO – пошук невідомої ланки в ланцюгу MO, що суб'єктивно може сприймається людиною як інсайт («Дніпро → (широкий) → світ»):

$$Find \cdot 1 : X \times X \rightarrow X. \quad (4.8)$$

8. Сенсосполучення → сенсосполучення, коли два словосполучення є синонімічними, наприклад, (рос.) «мы победили ↔ враг бежит» або «красна девица ↔ писанная красавица»:

$$SynYY : Y \rightarrow Y. \quad (4.9)$$

9. MO → X-Image-1 → X-Image-2 → MO – пошук двох невідомих ланок у ланцюгу MO, наприклад, «розплетена → (бо з милим я) → розлучена»:

$$Find \cdot 2 : X \times X \rightarrow X \times X. \quad (4.10)$$

10. Сенсосполучення → синтагма, де сенсосполучення є однією з складових синтагми, наприклад, (рос.) «кот ученый → и днем и ночью кот ученый все ходит по цепи кругом»:

$$PartYZ : Y \rightarrow Z. \quad (4.11)$$

11. MO → X-Image-1 → ... → X-Image-n → MO – пошук n невідомих ланок в ланцюгу MO, який найчастіше використовується для побудови відповідей або висловлювань, наприклад, «Еней → (був парубок моторний і хлопець хоч куди) → козак»:

$$Find \cdot n : X \times X \rightarrow \underbrace{X \times \dots \times X}_n. \quad (4.12)$$

12. Сенсосполучення \rightarrow синтагма – пошук синонімічної синтагми до даного словосполучення, наприклад, (рос.) «стражи порядка \leftrightarrow моя милиция меня бережет»:

$$SynYZ : Y \rightarrow Z. \quad (4.13)$$

13. Синтагма \rightarrow МО – результат впливу (влиття) сенсу синтагми в МО, наприклад, (рос.) «рука миллионопалая, сжатая в один громающий кулак \rightarrow партия»:

$$InsZX : Z \rightarrow X. \quad (4.14)$$

14. Синтагма \rightarrow сенсосполучення – результат впливу (влиття) сенсу синтагми в сенсосполучення, наприклад, (рос.) «и только небо тебя поманит \rightarrow синяя птица»:

$$InsZY : Z \rightarrow Y. \quad (4.15)$$

15. Синтагма \rightarrow синтагма, як дві синонімічні конструкції, наприклад, (рос.) «как с гуся вода \leftrightarrow Васька слушает да ест»:

$$SynZZ : Z \rightarrow Z. \quad (4.16)$$

Узагальнену класифікацію з 15-ти розглянутих варіантів образного пошуку представимо в таблиці 4.1 з урахуванням того, що три тривіальні випадки, коли результатом образного пошуку є складові частини більш об'ємної первинної ОК, а також рідко вживані випадки не розглядалися.

Таблиця 4.1 – Узагальнена класифікація образного пошуку

$\rightarrow \uparrow$	МО	Сенсосполучення	Синтагма
МО	(1) <i>AList</i> , (2) <i>SynXX</i> , <i>AntXX</i> , <i>PgXX</i> , <i>CmXX</i>	(3) <i>PartXY</i> , (4) <i>SynXY</i>	(5) <i>PartXZ</i>
Сенсосполучення	(6) <i>InsYX</i> , (7) <i>Find-1</i>	(8) <i>SynYY</i> , (9) <i>Find-2</i>	(10) <i>PartYZ</i> , (11) <i>Find-n</i> , (12) <i>SynYZ</i>
Синтагма	(13) <i>InsZX</i>	(14) <i>InsZY</i>	(15) <i>SynZZ</i>

Необхідно зауважити, що 5 позначених сірим кольором функцій таблиці 4.1, які відповідають базовому синтагматичному типу образного пошуку, відіграють визначальну роль для всіх інших функцій системи обробки природно-мовного контенту. З іншого боку, розглянуті функції образного пошуку дозволяють отримати зрозумілу інтерпретацію на 3-х типах графів (розділ 3), а тому можуть бути покладені в основу комплексу моделей та алгоритмів у предметній області дослідження. Функціональну модель СОПМК, побудовану за допомогою класифікації образного пошуку у відповідності до мети і задач дослідження, представлено у таблиці 4.2.

Таблиця 4.2 – Функціональна модель СОПМК

Складові СОПМК	Позначення	Базові алгоритми та прикладні задачі						
		Семантичний пошук			Інваріантний зміст і генерування повідомлень	ТТ: обмеження синтаксису, напрямку і словника	ТТ: дельфійський оракул	ТТ: магістр Йода
		Асоціативний пошук	Інсайтний пошук	Парадигматичний пошук				
Орієнтовані графи	$G_z(V, E)$	$AList^{out}$ $AList^{in}$	$Find \cdot 1$	$SynXX$ $AntXX$ $PgXX$ $CmXX$		$Find \cdot 2$ $Find \cdot n$		
Неорієнтовані графи	$G'_z(V, E')$	$AList'$	$Find \cdot 1'$		$InsYX$	$Find \cdot 2'$ $Find \cdot n'$		
Булева алгебра сенсу	BAS $\langle B, \Omega_b \rangle$				$InsZX$ $InsZY$ $SynZZ$	$Find \cdot n''$	$Mech$ $Or \cdot Re$	$PartXY$ $PartXZ$ $PartYZ$

Отже, у запропонованій класифікації образного пошуку враховано особливості парадигматичного мислення при безумовному пріоритеті синтагматичного, що відповідає онтогенезу розвитку мовлення. Класифікація формалізує 15 варіантів пошуку для можливих сполучень триосновних конструктів моделі образного сенсу: 5 варіантів можна віднести до базового синтагматичного типу, а 10 – до похідного парадигматичного типу. Отриману класифікацію взято за основу побудови функ-

ціональної моделі СОПМК у відповідності до задач дослідження та можливості зручної інтерпретації 3-ма типами графів, визначених методами моделювання функцій образного аналізу та синтезу (розділ 3).

4.2 Розробка алгоритмів асоціативного та інсайтного пошуку

Базовому типу образного пошуку № 1 узагальненої класифікації п.4.1 формально відповідає поняття асоціативної або хеш-таблиці, яке традиційно використовується в сучасних інформаційних технологіях для прискорення пошуку даних [82, 89]. Існує значне різноманіття відомих методів створення хеш-таблиць [255, 256, 257], але основною перевагою підходу, що пропонується, є зменшення розмірності інших алгоритмів СОПМК, залежних від базових. Формально для побудови базових алгоритмів асоціативного пошуку потрібні дані тільки одного відношення *Assoc-Twice* (3.52) сховища IC.

Розглянемо зважений орієнтований граф $G_z(V, E)$, заданий матрицею суміжності A_Q , згідно з п. 2.4. Задача дослідження полягає у побудові таких таблиць, де кожній з n вершин відповідають відсортовані списки вихідних $AList^{out}$, вхідних $AList^{in}$ та всіх неорієнтованих $AList'$ ребер. У подальшому будемо вважати вагою неорієнтованого ребра вагу $a'_{ij} = a_{ij} + a_{ji}$ будь-якого (i, j) елемента матриці A' неорієнтованого графа $G'_z(V, E')$. Позначимо кількість елементів таких списків для i -ї вершини відповідно як m_i^{out} , m_i^{in} та m_i' .

Відсортований список ребер $AList_i^{out}$, що виходять з i -ї вершини, отримаємо таким чином:

[Крок 1] Досліджується i -а вершина. У рядку i матриці A_Q визначається множина елементів $\{a_{ij} \mid a_{ij} > 0\}$.

[Крок 2] СОРтуємо множину з m_i^{out} елементів за спаданням, у результаті отримуємо список пар

$$AList_i^{out} = \left\langle (a_{ij_1}, j_1)(a_{ij_2}, j_2) \dots (a_{ij_{m_i}}, j_{m_i}) \mid \forall k < k' \ a_{ij_k} \geq a_{ij_{k'}} \right\rangle.$$

Аналогічно отримаємо відсортований список ребер $AList_j^{in}$, що входять до j -ї вершини:

[Крок 1] Досліджується j -та вершина. В стовпці j матриці A_Q визначається множина елементів $\{a_{ij} \mid a_{ij} > 0\}$.

[Крок 2] Сортуємо множину з m_j^{in} елементів за спаданням, у результаті отримуємо список

$$AList_j^{in} = \left\langle (a_{i_1j}, i_1)(a_{i_2j}, i_2) \dots (a_{i_{m_j}j}, i_{m_j}) \mid \forall k < k' \ a_{i_kj} \geq a_{i_{k'}j} \right\rangle.$$

Розмірність схожих алгоритмів $AList^{out}$ та $AList^{in}$ залежить від методу сортування, отже, наприклад, для бульбашкового методу [257] побудова обох таблиць не буде перевищувати значення $o(n \cdot \bar{m}^2)$, де \bar{m} – середня кількість вихідних або вхідних ребер однієї вершини орієнтованого графа. Для наскрізного прикладу орієнтованого графа (пп. 3.1, 3.6) результати застосування розглянутих алгоритмів представлено в таблиці 4.3 у вигляді списків пар (вага–ребра, суміжна–вершина).

Таблиця 4.3 – Результати застосування алгоритмів $AList^{out}$ та $AList^{in}$

Номер вершини	$AList^{out}$	$AList^{in}$
17	(1, 73)	(1, 34)(1, 35)(1, 73)(2, 111)
34	(1, 17)(2, 35)	(2, 35)
35	(1, 17)(2, 34)(2, 56)	(2, 34)
37	(2, 37)(2, 111)	(2, 37)(1, 111)(2, 123)
56	(2, 71)(3, 88)(1, 93)(3, 111)(2, 123)	(2, 35)(3, 71)(1, 73)(2, 123)
71	(3, 56)(1, 93)	(2, 56)(1, 93)
73	(1, 17)(1, 56)(2, 87)	(1, 17)
87	(1, 123)	(2, 73)(1, 88)(1, 93)(1, 123)
88	(1, 87)	(3, 56)
93	(1, 71)(1, 87)	(1, 56)(1, 71)
111	(2, 17)(1, 37)	(2, 37)(3, 56)
123	(2, 37)(2, 56)(1, 87)	(2, 56)(1, 87)

Алгоритм $AList'$ може будуватися на додатковому перетворенні орієнтованого графа $G_z(V, E)$ у неорієнтований $G'_z(V, E')$ згідно з (3.5). В цьому випадку розмірність прямого алгоритму створення таблиці з відсортованими списками неорієнтованих ребер збільшиться до $o(n^3 \cdot \bar{m}^2)$. Проте таку розмірність $AList'$ можна зменшити, якщо використати вже побудовані таблиці $AList^{out}$ та $AList^{in}$ і врахувати попередню відсортованість закладених в них списків.

Розглянемо можливість зменшити обсяг обчислень алгоритму $AList'$ шляхом такої модифікації бульбашкового методу:

[Крок 1] Організуємо цикл по кожній i -й вершині з n . У черговій ітерації циклу доступними є списки $AList_i^{out}$ розмірністю m_i^{out} та $AList_i^{in}$ розмірністю m_i^{in} .

[Крок 2] Задаємо $AList_i' = AList_i^{out}$, відповідно $m_i' = m_i^{out}$. Організуємо цикл кожній k -й парі $(a_{j_k i}, j_k)$ списку $AList_i^{in}$ з m_i^{in} .

[Крок 3] Перевіряємо кожну l -ту пару (a_{ij_i}, j_l) з m_i' списку $AList_i'$: якщо $j_k = j_l$, то $a_{ij_i} = a_{ij_i} + a_{j_k i}$, а новостворена «пара-бульбашка» (a_{ij_i}, j_l) пересувається уперед списку до пари з більшою від неї вагою або до початку списку; інакше вставляємо у своє місце списку $AList_i'$ нову пару $(a_{j_k i}, j_k)$ та встановлюємо $m_i' = m_i' + 1$.

Зрозуміло, що при такому підході обсяг обчислень **Кроків 2 і 3** не буде перевищувати $o((2 \cdot \bar{m})^2)$. Отже, побудова таблиці $AList'$ пропорційна $o(n \cdot \bar{m}^2)$. Результати застосування алгоритму $AList'$ до прикладу графа з пп. 3.1, 3.6 представлено в таблиці 4.4 відсортованими списками в загальній кількості 39 пар.

Таблиця 4.4 – Результати застосування алгоритму $AList'$ до графа з п. 3.1

Номер вершини	$AList'$	Вага вершини
17	(2, 111)(2, 73)(1, 34)(1, 35)	6
34	(4, 35)(1, 17)	5
35	(4, 34)(2, 56)(1, 17)	7
37	(4, 37)(3, 111)(2, 123)	9
56	(5, 71)(4, 123)(3, 88)(3, 111)(2, 35)(1, 93)(1, 73)	19
71	(5, 56)(2, 93)	7
73	(2, 17)(2, 87)(1, 56)	5
87	(2, 123)(2, 73)(1, 88)(1, 93)	6
88	(3, 56)(1, 87)	4
93	(2, 71)(1, 87)(1, 56)	4
111	(3, 37)(3, 56)(2, 17)	8
123	(4, 56)(2, 87)(2, 37)	8

Психологічне поняття інсайту має яскравий феноменологічний характер і характеризує момент раптового осяяння в процесі розв'язання складної задачі або проблеми, який і приводить до знаходження правильного розв'язку [148]. Відомі методи моделювання інсайту неналежно формалізовані та, звичайно, мають відношення до психологічних методик гештальтпсихології, нейролінгвістичного програмування, когнітивної психології тощо [189, 207].

Запропонований концептуальний підхід до моделювання образного сенсу (розділ 2) дозволяє отримати нову формальну інтерпретацію цієї задачі. З метою моделювання інсайту на неорієнтованому графі G'_z в [30] вперше запропоновано алгоритм $Find \cdot 1'$, що відповідає базовому типу образного пошуку № 7 узагальненої класифікації (див. таблицю 4.1).

Якщо визначати силу зв'язку згідно з (3.1) або (3.5), то фактично критерієм інсайту стає максимум сенсу. Тим самим, вибір можливих варіантів образів для інсайту буде тим багатшим, чим більший простір асоціативних пар закладено в систему.

Постановка задачі. Маємо орієнтований $G_z(V, E)$ та неорієнтований згідно з (3.5) $G'_z(V, E')$ графи, в яких i -й елемент множини образів I відповідає вершині графа з двійковим кодом $Bi-I_i$ у відповідності до п. 3.4. У загальному випадку множина вершин графа $V_z = \{i \mid 1 \leq i \leq n\}$. Будемо вважати, що між вершинами i та j немає прямого зв'язку, тобто $(i, j) \notin E_z$ та $(j, i) \notin E_z$. Необхідно знайти таку вершину k , для якої:

$$a) G_z(V, E): (a_{ik} + a_{kj}) \xrightarrow{Find \cdot 1} Max \text{ у відповідності до (3.1);}$$

$$b) G'_z(V, E'): (a'_{ik} + a'_{kj}) \xrightarrow{Find \cdot 1'} Max \text{ у відповідності до (3.5).}$$

Ідея алгоритмів полягає у застосуванні попередньо відомих таблиць $AList^{out}$ та $AList^{in}$ для задачі а) і $AList'$ для задачі б) з метою обмеження множини потенційно кращих з'єднуючих вершин для пари (i, j) за критерієм максимуму сумарної сили асоціативного зв'язку. Якщо прямий перебір варіантів вибору дає найкращий варіант пропорційно $o(n-2)$, то використання списків забезпечує розмірність $o(m_i^{out} \times m_j^{in})$ або подібну $o(m'_i \times m'_j)$. Для розрідженого графа, яким є

АМО, при $\bar{m} \ll n$ такий підхід вже може дати позитивний результат. Але, зважаючи на попередню відсортованість списків, пропонується додатково зменшити розмірність алгоритмів типу

$$\bigcap_{AList'_i, k = AList'_j, k} \text{Max}(AList'_i, a_{ik} + AList'_j, a_{kj})$$

внаслідок застосування методу обмежувачів.

Для розв'язання задачі а) введемо штучні верхні та нижні обмежувачі списків $AList_i^{out}$ та $AList_j^{in}$, позначених на рисунку 4.1 як *LeftList* та *RightList*. Обмежувачі лівого та правого списків позначимо відповідно мітками *LeftTop*, *LeftBottom*, *RightTop*, *RightBottom*.

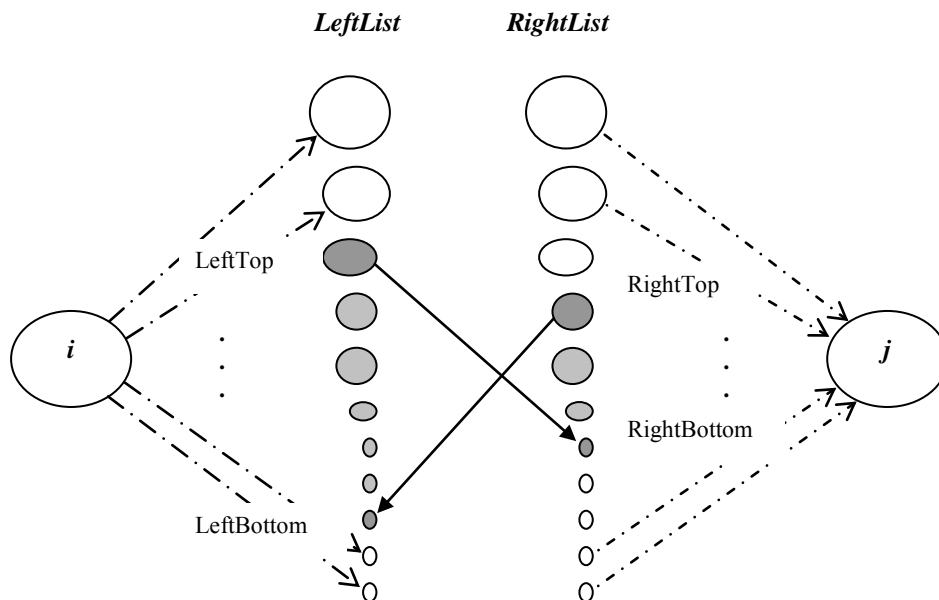


Рисунок 4.1 – Верхні та нижні обмежувачі при пошуку спільних елементів двох списків

Позначимо пари (вага–ребра, суміжна–вершина) у списках *LeftList* та *RightList* відповідно як $(Fl(i), BiL(i))$ та $(Fr(j), BiR(j))$, а пари скінченого списку *ListI* як $(Fo(k), Bi-I(k))$, де i, j, k – порядкові номери (індекси) елементів в списках. Задачу розв'яжемо за допомогою такого узагальненого алгоритму:

[Крок 1] Фіксуємо перший елемент списку *LeftList* та перебираємо список *RightList* до тих пір, доки в ньому не буде знайдено такий самий елемент. У цьому випадку номер елемента списку *LeftList* прис-

воюється змінній $LeftTop$, а номер елемента списку $RightList$ – змінній $RightBottom$, далі формуємо код і вагу першого елемента списку $ListI$ та переходимо до кроку 2.

Інакше – переходимо на наступний елемент списку $LeftList$ та знову повертаємося на початок кроку 1. Якщо список $LeftList$ вичерпано, то видається повідомлення про те, що спільних елементів у двох списках не знайдено (після найтрудомісткішого варіанта – повного перебору розмірністю $O(m_i^{out} \times m_j^{in})$).

[Крок 2] Фіксуємо перший елемент списку $RightList$ та перебираємо список $LeftList$ з номера $LeftTop+1$ до тих пір, доки в ньому не буде знайдено такий же елемент. В цьому випадку номер елемента списку $RightList$ присвоюється змінній $RightTop$, а номер елемента списку $LeftList$ – змінній $LeftBottom$, далі формуємо код і вагу другого елемента списку $ListI$ та переходимо до кроку 3.

Інакше – переходимо на наступний елемент списку $RightList$ та знову повертаємося на початок кроку 2. Якщо список $RightList$ дійшов до елемента $RightBottom$, то алгоритм закінчується і видається єдиний вже знайдений варіант спільного елемента $B_i - I(LeftTop)$.

[Крок 3] Проводимо повний перебір елементів списків $LeftList$ та $RightList$ відповідно між парами елементів $LeftTop+1$ та $LeftBottom-1$, $RightTop+1$ та $RightBottom-1$, при знаходженні спільних елементів формуються k -ті складові списку $ListI$ (починаючи з $k=3$).

Обґрунтуємо метод обмежувачів таким чином – потенційних елементів, більших за верхні обмежувачі не існує в силу вже проведеної перевірки, а елементи, менші за нижні обмежувачі перевіряти не потрібно, оскільки їх сумарна вага не може бути більшою за вже знайдений варіант. При знаходженні елементів з усіма чотирма мітками розмірність пошуку обмежується зверху значенням

$$\begin{aligned} &LeftTop \times m_j^{in} + (m_i^{out} - LeftTop) \times RightTop + \\ &(LeftBottom - LeftTop - 1) \times (RightBottom - RightTop - 1), \quad (4.17) \\ &\leq m_i^{out} \times m_j^{in} \end{aligned}$$

причому два прийнятних варіанти пошуку перед третім перебором вже існують. Формально представимо запропонований алгоритм аналітичними виразами та граф-схемами з нотацією [257] по кроках:

$$\begin{aligned}
\text{Step-1} ::= & \{[\overline{LeftList : i = 1, n}] \ (\{[\overline{RightList : j = 1, m}] \\
& ([BiL(i) = BiR(j)] \ (i \rightarrow LeftTop * BiL(i) \rightarrow Bi - I(1) * , (4.18) \\
& Fl(i) + Fr(j) \rightarrow Fo(1) * j \rightarrow RightBottom), Abort)\}\}
\end{aligned}$$

де *Step-1* – перший крок алгоритму; *Abort* – повідомлення про те, що рішення не існує;

$$\begin{aligned}
\text{Step-2} ::= & \{[\overline{RightList : j = 1, RightBottom}] \ (\{[\overline{LeftList : i = \\
& \overline{LeftTop + 1, n}] \ ([BiL(i) = BiR(j)] \ (j \rightarrow RightTop * BiR(j) \rightarrow , (4.19) \\
& Bi - I(2) * Fl(i) + Fr(j) \rightarrow Fo(2) * i \rightarrow LeftBottom), Only)\}\})
\end{aligned}$$

де *Step-2* – другий крок алгоритму; *Only* – повідомлення про єдине рішення *LeftTop*;

$$\begin{aligned}
\text{Step-3} ::= & 2 \rightarrow k * \{[\overline{LeftList : i = \overline{LeftTop + 1, LeftBottom - 1}] \\
& (\{[\overline{RightList : j = \overline{RightTop + 1, RightBottom - 1}] \ ([BiL(i) = BiR(j)] \ , (4.20) \\
& (k + 1 \rightarrow k * BiL(i) \rightarrow Bi - I(k) * Fl(i) + Fr(j) \rightarrow Fo(k), E)\}\})
\end{aligned}$$

де *Step-3* – третій крок алгоритму; *E* – пустий оператор.

В граф-схемах першого, другого та третього (рисунок 4.2, 4.3, 4.4) кроків використані такі позначення для операторів структурного програмування:

- Do – цикл за параметром;
- If (+ –) – альтернатива;
- * – композиція.

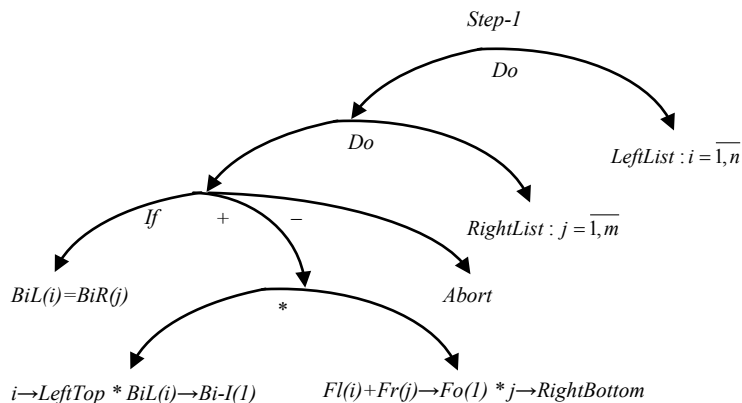


Рисунок 4.2 – Граф-схема кроку 1 алгоритму за методом обмежувачів

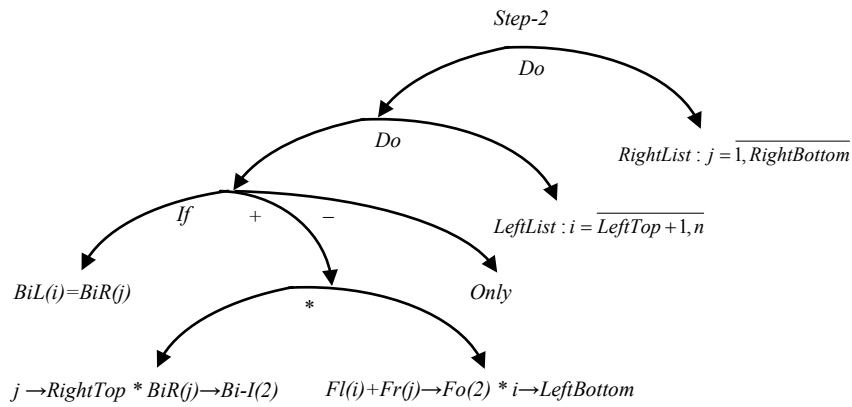


Рисунок 4.3 – Граф-схема кроку 2 алгоритму за методом обмежувачів

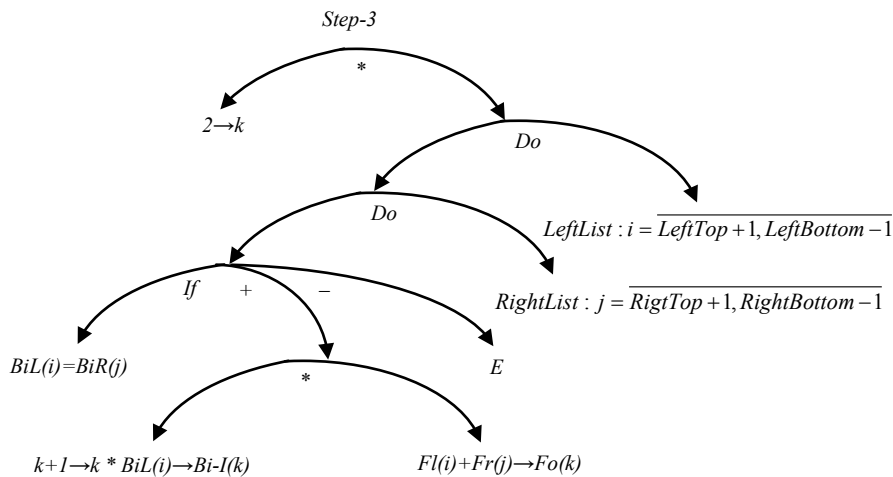


Рисунок 4.4 – Граф-схема кроку 3 алгоритму за методом обмежувачів

Кінцеве формування хеш-таблиці інсайтних образів досягається сортуванням списку $ListI$ за вагою $Fo(k)$, наприклад, за допомогою бульбашкового методу (алгоритм *Pyzyrek* з [257]):

$$\begin{aligned}
 Hash - Insight ::= & Step - 1 * [k = 0] \quad Stop, \quad (Step - 2 * \\
 & [k = 1] \quad Stop, \quad Step - 3 * Pyzyrek) \quad . \quad (4.21)
 \end{aligned}$$

Зауважимо, що розв'язок $Find \cdot 1'$ задачі b) відрізняється від попереднього алгоритму $Find \cdot 1$ тільки даними списків $AList'_i$ та $AList'_j$, що застосовуються як $LeftList$ та $RightList$. Розглянемо реалізацію запропонованих алгоритмів на графах наскрізного прикладу з пп. 3.1, 3.6 з урахуванням даних таблиць 4.3 та 4.4. З цією метою задамо пару

мовних образів (i, j) , що не мають прямого зв'язку між собою як (56, 87) або, за інтерпретацією експериментального прикладу, «мова» – «RDF».

У результаті застосування $Find \cdot 1$ було отримано такий відсортований оператором *Hash – Insight* список інсайтних образів $ListI$ з парами $(Fo(k), Bi - I(k))$: (4, 88), (3, 123), (2, 93). Тобто, найкращим інсайтним образом для обраної пари є «XML» з можливих «дозвіл» та «засіб». Застосування алгоритму $Find \cdot 1'$ для тих же самих даних призвело до побудови дещо іншого списку інсайтних образів: (6, 123), (4, 88), (3, 73), (2, 93). Таким чином, в умовах неорієнтованого графа найкращим інсайтним образом виявився «дозвіл» та у списку $ListI$ з'явився додатковий МО «ресурс». Отримані результати, наведені у Додатку В, свідчать про високу ступінь адекватності запропонованого підходу до моделювання інсайту в умовах АМО навіть для відносно невеликого обсягу експериментальних даних наскрізного прикладу.

Отже, базовий асоціативний пошук реалізовано за допомогою таблиць з відсортованими списками вихідних $AList^{out}$, вхідних $AList^{in}$ та всіх неорієнтованих $AList'$ ребер. На основі цих алгоритмів інсайтний пошук інтерпретовано для орієнтованих і неорієнтованих графів як визначення найбільш вагомої проміжної ланки для безпосередньо не пов'язаних між собою вершин. Запропонований підхід має високу ступінь адекватності та прискорює побудову інсайтних таблиць $Find \cdot 1$ та $Find \cdot 1'$ для розріджених графів у порівнянні з відомими методами, що було продемонстровано на даних наскрізного тестового прикладу.

4.3 Алгоритм визначення ланцюга образів у зваженому графі

Розвинемо запропонований у п. 4.1 підхід, що інтерпретує задачі образної обробки ПМК у вигляді задач на графах. Згідно з ним (див. таблицю 4.2) формалізація низки задач образного мислення, у т. ч. генерація лінгвістичним процесором відповідей на питання, потребує побудови ланцюга з трьох і більше вершин АМО для випадків орієнтованого G_z та неорієнтованого G_z' графів [15, 17, 32, 53]. Зауважимо, що неорієнтований граф G_z' АМО практично можна отримати шляхом «ущільнення» існуючих АВМ, коли декілька споріднених вузлів

об'єднують в один, причому відповідні асоціативні зв'язки переходять «у спадок» новому вузлу. Будемо виходити з того, що в межах моделі образного мислення побудовано алгоритми $AList'$ та $Find \cdot 1'$ (п. 4.2) для неорієнтованих графів.

За цих умов розглянемо випадок, коли необхідно знайти n невідомих ланок в шляху без врахування напрямку зв'язків у зваженому графі G'_z АМО. Задача стає актуальною тоді, коли алгоритм $Find \cdot 1'_{ij}$ не знаходить спільних елементів в списках $AList'_i$ та $AList'_j$. Відомі алгоритми теорії графів, у тому числі алгоритм Дейкстри забезпечують знаходження найкоротшого шляху між двома довільними вершинами орієнтованого або неорієнтованого графа з обчислювальними витратами $o(n^2)$ [259]. Складність розглянутої задачі полягає в тому, що для АМО актуальним є пошук мінімального за кількістю ланок шляху, але з найбільшою сумарною вагою. За таких умов алгоритм Дейкстри не працює, оскільки у G'_z не виконується нерівність трикутника (п. 2.3), навіть якщо брати зворотне значення ваги ребер для пошуку найкоротшого шляху.

Постановка задачі: на основі вже відомих алгоритмів $AList'$ та $Find \cdot 1'$ знайти в графі G'_z мінімальний за довжиною n' шлях $Find \cdot n'(Bi - I_1, Bi - I_2)$ між вершинами $Bi - I_1$ та $Bi - I_2$ з максимальною сумарною вагою ребер $Se^\Sigma = \sum_{i=1}^{n'-1} c_{i,i+1}$, якщо вага кожного ребра визначається згідно з (3.5).

Ідею розв'язання поставленої задачі проілюструємо зображенням фрагменту АМО на рисунку 4.5. Цей граф демонструє від'ємний результат застосування оператора $Find \cdot 1'(Bi - I_1, Bi - I_2)$, оскільки вершини $Bi - I_1$ та $Bi - I_2$ неможливо поєднати двома ланками.

Якщо застосувати відомі методи пошуку критичного шляху у мережевих моделях, то у загальному випадку розв'язання задачі вимагає прямого перебору розмірністю $o((n-1)!)n$ вершин графа [250]. Проте можна врахувати особливості розрідженого графа АМО (п. 3.1), що впливають з моделі образного аналізу та уникнути розв'язання NP -повної задачі:

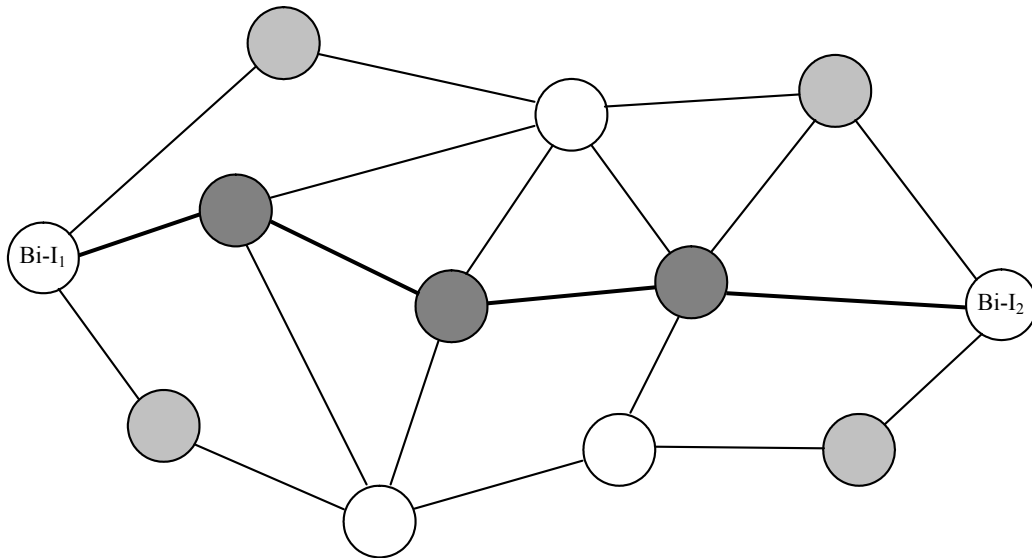


Рисунок 4.5 – Фрагмент асоціативної мережі образів з результатом дії оператора $Find \cdot n'(Bi - I_1, Bi - I_2)$

- кількість значимих зв'язків від кожного вузла графа (образу) до інших вузлів (образів) не перевищує 7 ± 2 , але не менша двох;
- не існує двох різних вузлів (образів), які б мали однаковий набір суміжних вузлів (асоційованих образів);
- між будь-якими двома вузлами графа розмірністю n обов'язково існує скінченний шлях довжиною не більшою $(n+1)/2$.

За таких умов розмірність побудови маршруту зменшується до $o(\bar{m}^L)$, де \bar{m} – середня кількість суміжних з вершиною ребер неорієнтованого графа, а $L \leq (n+1)/2$ – довжина маршруту у ребрах. Пропонується такий підхід до розв'язання поставленої задачі, що представляє собою модифікацію ідеї «жадібного» алгоритму [260] з урахуванням попередньої відсортованості списків суміжних ребер для кожної вершини:

- будується «жадібний» опорний маршрут на основі рекурсивного сортування за вагою зв'язку c'_{ij} отриманого на кожному i -му кроці списку вершин та вибору найкращої з них за винятком вже задіяних у попередніх ітераціях;
- для зберігання ланок маршруту використовується стек $Route_i$, причому внесення нової ланки маршруту (вершини) до стеку супрово-

джується перевіркою на зациклення, оскільки у стеку не повинно бути двох однакових вершин;

- отриманий опорний маршрут оптимізується з метою зменшення його довжини L (у ребрах), оскільки другою складовою цільової функції виступає лаконічність мовної конструкції (висловлювання).

Алгоритм розв'язання поставленої задачі побудуємо за допомогою таких формальних операцій та предикатів:

1. Операція $AList'(Bi - I_i)$ визначення відсортованого за силою асоціативного зв'язку S_i списку образів $List_i$ для образу-стимулу $Bi - I_i$:

$$(Bi - I_i, S_i) \xrightarrow{AList'} List_i. \quad (4.22)$$

2. Предикат $Stop$, що показує завершення побудови асоціативного маршруту:

$$Bi - I_i = Bi - I_2 \rightarrow Stop. \quad (4.23)$$

3. Операція $Find \cdot 1'(Bi - I_1, Bi - I_2)$ знаходження проміжної ланки $Bi - I_x$ (образу-інсайту) для образів $Bi - I_1$ та $Bi - I_2$:

$$(Bi - I_1, Bi - I_2) \xrightarrow{Find \cdot 1'} Bi - I_x. \quad (4.24)$$

4. Предикат $Insight$, що показує знаходження образу-інсайту в результаті застосування операції $Find \cdot 1(Bi - I_1, Bi - I_2)$:

$$\exists Bi - I_x \rightarrow Insight. \quad (4.25)$$

5. Предикат $Cycle$, що виявляє зациклення у роботі алгоритму побудови ланцюга образів за рахунок повторення у стеку $Route_i$ однакового образу:

$$Bi - I_i \in Route_j, j = \overline{1, i-1} \rightarrow Cycle. \quad (4.26)$$

Для розв'язання поставленої задачі з урахуванням формальних операцій та предикатів (4.22)–(4.26) використаємо двоосновну алгебраїчну систему (3.14), врахувавши, що вага ребер неорієнтованого графа АМО $c'_{ij} \in Time$. До визначеного у п. 3.5 складу операцій та предикатів системи (3.14) додамо розглянуті вище операції та предикати

$$OP = OP \cup \{AList', Find \cdot 1'\}; \quad (4.27)$$

$$IF = IF \cup \{Stop, Insight, Cycle\}. \quad (4.28)$$

Покажемо, що задача побудови ланцюга образів у межах алгебраїчної системи *BAS* фактично зводиться до рекурсивного алгоритму на основі операцій *OP* та предикатів *IF*. У граф-схемах для алгебраїчних конструктів, що будуть розглядатися з цією метою, використані такі ж самі позначення для операторів структурного програмування [257], як і у п. 4.2. Достатньою для розв'язання поставленої задачі є послідовна побудова таких операторів з сигнатури алгебраїчної системи з урахуванням уведених раніше позначень.

1. Оператор R_e , що збільшує на 1 змінну i – номер поточної ланки ланцюга образів – та заносить в i -ту комірку стека $Route_i$ значення коду кінцевого образу маршруту $Bi-I_2$:

$$R_e ::= i + 1 \rightarrow i * Bi - I_2 \rightarrow Route_i. \quad (4.29)$$

Граф-схему оператора R_e представлено на рисунку 4.6.

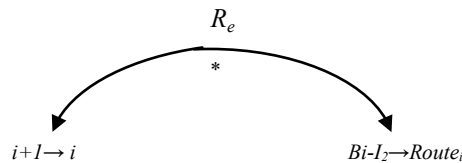


Рисунок 4.6 – Граф-схема оператора R_e

2. Оператор R_i , що збільшує на 1 змінну i та заносить в i -ту комірку стека $Route_i$ значення коду образу-інсайту $Bi-I_x$:

$$R_i ::= i + 1 \rightarrow i * Bi - I_x \rightarrow Route_i. \quad (4.30)$$

Граф-схему оператора R_i представлено на рисунку 4.7.

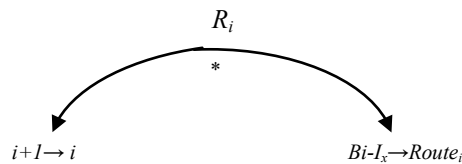


Рисунок 4.7 – Граф-схема оператора R_i

3. Кінцевий оператор R_a , який виводить отриманий маршрут або випадок зациклення у вигляді заповнених комірок стека $Route_i$ від першої по i -ту за допомогою процедури $Output()$:

$$R_a ::= \{ [j = \overline{1, i}] \text{ Output}(Route_j) \}. \quad (4.31)$$

Граф-схему оператора R_a представлено на рисунку 4.8.

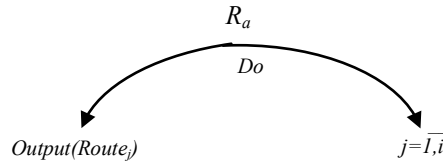


Рисунок 4.8 – Граф-схема оператора R_a

4. Оператор $Find \cdot n'$, що розпочинає алгоритм побудови маршруту обнулінням змінної i , заносить значення коду початкового образу маршруту $Bi-I_1$ у змінну поточного образу $Bi-I_i$ та передає управління оператору $Next$:

$$Find \cdot n' ::= 0 \rightarrow i * Bi - I_1 \rightarrow Bi - I_i * Next. \quad (4.32)$$

Граф-схему оператора $Find \cdot n'$ представлено на рисунку 4.9.

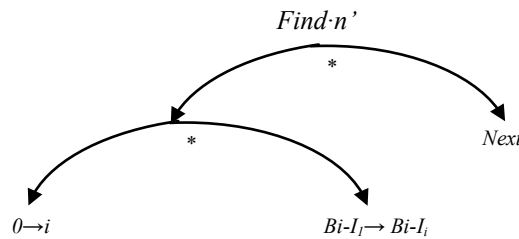


Рисунок 4.9 – Граф-схема оператора $Find \cdot n'$

5. Оператор $Next$, що збільшує на 1 змінну i , заносить в i -ту комірку стека $Route_i$ значення коду поточного образу $Bi-I_i$ та передає управління оператору One :

$$Next ::= i + 1 \rightarrow i * Bi - I_i \rightarrow Route_i * One. \quad (4.33)$$

Граф-схему оператора $Next$ представлено на рисунку 4.10.

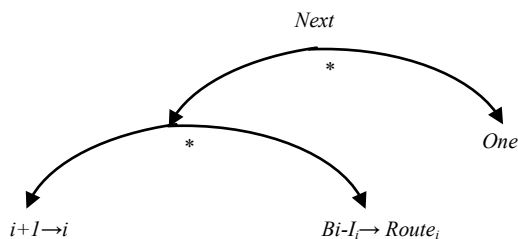


Рисунок 4.10 – Граф-схема оператора *Next*

6. Оператор *One*, що знаходить список образів-асоціатів до образу $Bi-I_i$ за допомогою операції $AList'(Bi - I_i)$ та перевіряє предикат *Stop* – при виявленні кінцевого образу маршруту виконуються операції R_e та R_a , інакше управління передається до оператора *Step*:

$$One ::= AList'(Bi - I_i) * ([Stop] (R_e * R_a), Step). \quad (4.34)$$

Граф-схему оператора *One* представлено на рисунку 4.11.

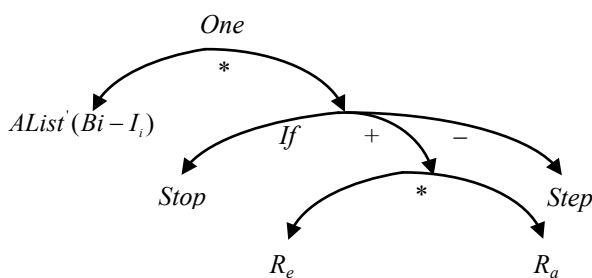


Рисунок 4.11 – Граф-схема оператора *One*

7. Оператор *Step*, що призначає найбільше значення з списку для i -го образу на місці $i+1$ -ї ланки маршруту $Bi-I_{i+1}$ та перевіряє предикат *Cycle* – при виявленні зациклення маршруту виконується операція R_a , інакше управління передається до оператора *Two*:

$$Step ::= Max(List_i) \rightarrow Bi - I_{i+1} * ([Cycle] R_a, Two). \quad (4.35)$$

Граф-схему оператора *Step* представлено на рисунку 4.12.

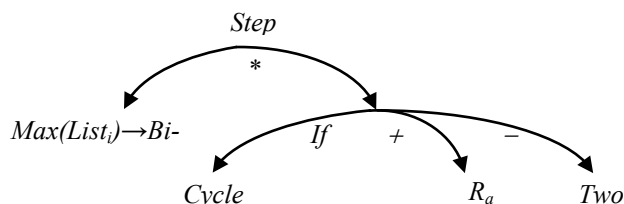


Рисунок 4.12 – Граф-схема оператора *Step*

8. Оператор *Two*, що шукає образ-інсайт між поточним образом $Bi-I_i$ та кінцевим образом маршруту $Bi-I_2$, а потім перевіряє предикат *Insight* – при виявленні образу $Bi-I_x$ послідовно виконуються операції R_i , R_e та R_a , інакше управління рекурсивно перемикається на оператор *Next*:

$$Two ::= Find \cdot 1'(Bi - I_i, Bi - I_2) * ([Insight] (R_i * R_e * R_a), Next). \quad (4.36)$$

Граф-схему оператора *Two* представлено на рисунку 4.13.

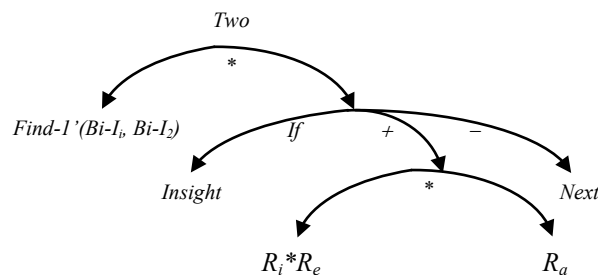


Рисунок 4.13 – Граф-схема оператора *Two*

Відмітимо, що запропонований рекурсивний алгоритм отримує точку зупинки лише за допомогою оператора R_a як для демонстрації знайденого маршруту, так і для визначення ситуації зациклення. Для розмежування цих двох різних випадків у подальшому можна ускладнити алгоритм процедурою подолання зациклення шляхом призначення наступної ланки після проблемного i -го образу другим за вагою образом у списку $List_i$. В силу особливостей графа, що впливають з дослідження моделі образного сенсу, новий скінченний маршрут також існує.

Як ілюстрацію запропонованого підходу розглянемо застосування оператора $Find \cdot n'$ для задачі пошуку маршруту між мовними образами 71 («існує») \rightarrow 17 («мережа») наскрізного прикладу експериментального графа з пп. 3.1, 3.6. «Жадібний» алгоритм операторами *Next*, *One* та *Step* за найбільшими значеннями ваги ребер знаходить перший крок 71–(5)–56, а за допомогою визначення інсайту оператором *Two* – два наступних кроки: 56–(3)–111–(2)–17. Маршрут складається з 3-х ланок загальною вагою 10. Неважко пересвідчитися за допомогою рисунку 3.4, що такий шлях через МО «мова» та «онтологія» є найвагомим з усіх можливих триланкових ланцюгів маршруту 71 \rightarrow 17.

Отриманий за допомогою алгоритму (4.29)–(4.36) «жадібний» опорний маршрут побудовано з розмірністю $o(L)$ за критерієм максимуму сили асоціативного зв'язку між його образами-ланками, що пропорційно кількості сенсу. Проте не можна стверджувати, що цей маршрут є найкоротшим з усіх можливих, оскільки на кожному кроці алгоритму *Find · n'* перевірялося існування інсайтної ланки тільки між поточним образом і кінцевим. Отже, зворотне застосування операції *Find · 1'* та предикату *Insight* до пар образів ланцюга з номерами $(i-1, 1)$, $(i-1, 2)$, ..., $(i-1, i-4)$, $(i-2, 1)$, $(i-2, 2)$, ..., $(i-2, i-5)$, ..., $(4, 1)$ потенційно може скоротити маршрут. На рисунку 4.14 послідовність аналізу таких пар показано пронумерованими дугами для ланцюга з семи вершин графа.

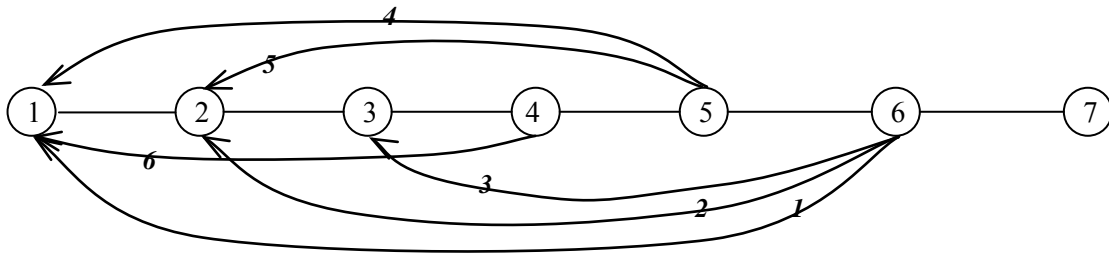


Рисунок 4.14 – Послідовність варіантів оптимізації за критерієм мінімуму довжини маршруту

Знайдений для розглянутого прикладу шлях 71–(5)–56–(3)–111–(2)–17 з 3-х ланок загальною вагою 10 будемо вважати оптимальним, оскільки дволанкових ланцюгів маршруту $71 \rightarrow 17$ не існує, а його вага найбільша з усіх триланкових ланцюгів.

Характерною особливістю запропонованого підходу є суттєве зменшення простору пошуку у графі, оскільки методом математичної індукції легко довести, що найбільша кількість варіантів V скорочення маршруту з L ланок не більша за арифметичну прогресію

$$V = 1 + 2 + \dots + L - 5 + L - 4 = \sum_{i=1}^{L-4} i. \quad (4.37)$$

Таким чином, розмірність алгоритму *Find · n'* зменшується до $o(L + \sum_{i=1}^{L-4} i)$, де $L \leq (n + 1)/2$, проте можна бути впевненим лише в існуванні одного прийняттого розв'язку задачі, а не в знаходженні най-

кращого з них. На відміну від відомих методів розв'язання аналогічних задач пошуку найкоротших шляхів у неорієнтованих зважених графах, запропонований ефективний алгоритм інтерпретує процеси мовленнєвої діяльності людини та зменшує простір пошуку від геометричної до арифметичної прогресії.

4.4 Алгоритм пошуку найвагомішого шляху в орієнтованому графі

Визначений методологією моделювання у розділі 3 орієнтований зважений граф $G_z(V, E)$ застосуємо для традиційної в теорії графів задачі пошуку найкоротшого шляху. Не порушуючи спільності, встановимо $V_z = \{1, 2, \dots, n\}$ та $E_z = \{(i, j)\}$, де i – початок ребра, j – кінець ребра. Для матриці суміжності A_Q з урахуванням (3.1) перевизначимо елементи

$$a_{ij} = \text{Max}_{(i,j) \in E_z} (c_{ij}) - c_{ij} + 1. \quad (4.38)$$

Постановка задачі. Необхідно знайти в графі G найкоротший $(s-t)$ шлях, тобто маршрут $s = v_1, v_2, \dots, v_{k-1}, v_k = t$ з мінімальною вартістю згідно з (4.38). Для комп'ютерного розв'язання задачі вбачається доцільним застосувати оригінальний алгоритм *Find · n* знаходження найкоротшого шляху, оскільки, як було раніше показано, алгоритм Дейкстри [260] в графах з двостороннім рухом типу G_z не працює як внаслідок властивостей АМО, так і у зв'язку з можливими загрозами зациклення.

Введемо такі позначення:

s, t – відповідно початкова та кінцева вершини шляху (задаються);

$v = \{s = v_1, v_2, \dots, v_k\}$, де $v_i \in V$ – маршрут довжиною k , при $1 \leq i \leq k$

$v_k = t$ – $(s-t)$ шлях;

i – вершина пошуку (i -й рядок матриці A_Q);

j_p – номер позиції пошуку в рядку i , $j_p \leq j \leq n$;

kp – номер першої вершини з кінця шляху (або маршруту) с ненульовою пропускнуою здатністю;

S – вартість маршруту, $S = a_{v_1v_2} + a_{v_2v_3} + \dots + a_{v_{k-1}v_k}$;

S_{min} – мінімальна вартість шляху;

V_{min} – шлях з мінімальною вартістю;

Vix – вектор пропускної здатності графа – вектор локальних ступенів (кількість виходів) вершини.

Алгоритм $Find \cdot n$. Знаходить оптимальний $(s-t)$ шлях в орієнтованому зваженому графі $G_z(V, E)$. Алгоритм вичерпної дії, побудований на основі принципу «пошук углиб з поверненням».

Викладемо сутність алгоритму «Вичерпний пошук у вглиб з поверненням»:

1. **[Перевірка існування розв’язку задачі]** – чи є вихід з вершини s та вхід до вершини t ?

Нехай для задачі, що має розв’язок, побудовано маршрут $V = \{s = v_1, v_2, \dots, v_{k-1}\}$ (рисунок 4.15), для якого зафіксуємо $i = v_{k-1}$ та встановимо значення j_p .



Рисунок 4.15 – Маршрут V

2. **[Пошук нової вершини шляху]**. Досліджується i -а вершина. В рядку i матриці A_Q шукаємо новий ненульовий елемент

$$v_k = \min_{j_p \leq j \leq n} \{a_{ij} \neq 0\}.$$

При цьому можливі такі варіанти:

а) v_k не існує (немає виходу з вершини i , тобто $j > n$ та $a_{ij} = 0$).

[Перевірка кінця]. Чи існують ще не побудовані $(s-t)$ шляхи, тобто якщо $Vix = 0$ (всі компоненти вектора Vix дорівнюють 0), то всі $(s-t)$ шляхи побудовані. **[Кінець]**. Інакше:

[Повернення назад] – до пошуку наступної вершини з ненульовою пропускною здатністю. Нехай $kp = k - 2$ (або $k - 1$, якщо $v_{k-2} = n$), $j_p = v_{kp+1} + 1$, $i = v_{kp}$ та $k = kp$ (у маршруті залишити вершини $s = v_1, v_2, \dots, v_{kp=k}$; з вершини v_{kp} – існує вихід). Для нового маршруту V корегувати його вартість S і відновити початкові пропускні здатності

відкинутих вершин з попереднього маршруту $v_m \in V$ та [Пошук нової вершини шляху].

б) $v_k \in V$, тоді прийняти $i = v_{k-1}$, $j_p = j_p + 1$ (виключити виникнення циклів) та [Пошук нової вершини шляху].

в) ($v_k \notin V$) і ($v_k \neq t$), тобто v_k – не остання вершина шляху. Включити v_k до маршруту; $S = S + a_{ij}$, $Vix[k] = Vix[k] - 1$ (зменшити пропускну здатність k -го елемента маршруту, тобто вершини i); $i = v_k$, $j_p = 1$ та [Пошук нової вершини шляху].

г) ($v_k \notin V$) і ($v_k = t$), тобто шлях побудовано. Прийняти $S_{\min} = \min(S, S_{\min})$, $V \rightarrow V_{\min}$. [Перевірка кінця]. Якщо існують ще не побудовані ($s - t$) шляхи, то

[Повернення назад (до витоку)]. Нехай $kp = k - 1$ (або $k - 2$, якщо $v_{k-1} = n$) та шукається перший (з позиції kp вектора V у напрямку до витоку S) елемент з ненульовою пропускну здатністю. Нехай він буде v_{kp} . Тоді прийняти $i = v_{kp}$, $k = kp$, $j_p = v_{kp+1} + 1$ та, у відповідності до побудови маршруту V , корегувати його вартість S та відновити початкові пропускі здатності відкинутих вершин з маршруту $v_m \in V$ та [Пошук нової вершини шляху].

Проілюструємо роботу алгоритму «Вичерпний пошук углиб з поверненням» $Find \cdot n$ на прикладі розглянутої у п. 3.1 АМО у вигляді графа G_z (див. рисунок 3.4). Матрицю суміжності A_o для нього зображено у таблиці 4.5. В додатковому стовпці *Input* цієї таблиці пораховано значення впускної здатності кожної i -ї вершини графа (локальна ступінь входження), а в додатковому стовпці *Output* – значення пропускну здатності (локальна ступінь виходу), що буде використано надалі.

Для проведення програмного експерименту алгоритм $Find \cdot n$ реалізовано у системі *TurboPascal*, текст програми та файли з результатами її роботи представлено у Додатку В. У результаті експерименту з урахуванням нумерації мовних образів наскрізного прикладу (п. 3.6) було одержано такі результати:

І. Найкоротший шлях з вершини $s = 3$ до вершини $t = 8 - 3(35) - 5(56) - 9(88) - 8(87)$, що інтерпретується у МО прикладу як «зміст-мова-XML-RDF».

II. Найкоротший шлях з вершини $s = 2$ до вершини $t = 4 - 2(34) - 3(35) - 5(56)$, що інтерпретується у МО прикладу як «інформація–зміст–мова».

Таблиця 4.5 – Матриця суміжності A_Q для графа G_z з п. 3.1

	17	34	35	37	56	71	73	87	88	93	111	123	Input	Output
17							1						4	1
34	1		2										1	2
35	1	2			2								1	3
37				2							2		3	2
56						2			3	1	3	2	4	5
71					3					1			2	2
73	1				1			2					1	3
87												1	4	1
88								1					1	1
93						1		1					2	2
111	2			1									2	2
123				2	2			1					2	3

Отримані результати програмного експерименту свідчать про високий ступінь адекватності найвагомшого шляху з МО навіть для відносно невеликих обсягів природно-мовного контенту. Отже, для пошуку найкоротшого шляху в орієнтованих зважених графах з двостороннім рухом типу $G_z(V, E)$ запропоновано модифікацію алгоритму вичерпної дії $Find \cdot n$, побудованого на основі принципу «пошук углиб з поверненням». Адекватність алгоритму перевірено на даних наскрізного тестового прикладу графа.

4.5 Метод моделювання механізму оперативної пам'яті СОПМК

Вивчення механізму підсвідомого обміну образів підштовхнуло в свій час З. Фрейда до розроблення методу вільних асоціацій, покладеного згодом в основу психоаналізу [126]. У відповідності до структурно-функціональної моделі ІС (п. 2.7) з механізмом оперативної пам'яті СОПМК мають бути тісно пов'язані такі відомі психологічні феномени, як ансамбль образів, вектор емоцій, орієнтувальний рефлекс, надоперативна, асоціативна і довготривала пам'ять, що і має знайти формальне відображення в моделі оперативної пам'яті [31].

Будемо розрізняти два режими інтелектуальної діяльності – у природних умовах фізіологічний орієнтувальний рефлекс приводить до збудження аферентних виконавчих механізмів і, внаслідок цього, поя-

ви нових образів через надоперативну пам'ять [154]. У складнішому випадку виключно вербального сприйняття інформації поява нових образів в ансамблі оперативної пам'яті можлива тільки через асоціативну і довготривалу пам'ять [31].

Значна частина досліджень щодо вибору і пошуку потрібних образів зосереджена в області фізіології і психології. Практично відсутні публікації щодо механізму взаємодії згаданих психологічних феноменів в контексті оперативної або короткочасної пам'яті людини. Не перевірена у повній мірі гіпотеза про те, що саме образ слід вважати тією одиницею спілкування, в процесі дослідження якої можна буде описати взаємодію всіх чинників успішної комунікації [100, 163]. У зв'язку з цим необхідно відповісти на питання – для чого і як змінюється ансамбль образів оперативної пам'яті, яка при цьому роль фокусу уваги.

Постановка задачі. Виходитимемо з того, що у кожний дискретний проміжок часу оперативна пам'ять об'єднує 7 ± 2 активних образів, на один з яких спрямовано фокус уваги [29, 31]. Принципи побудови механізму оперативної пам'яті повинні показати, як в процесі зміни початкового складу ансамблю образів можна знайти і обрати за певним чисельним критерієм такий образ, який був би найбільш або прийнятно наближеним до ідеального образу-мети. Вихідними даними для пошуку та вибору в загальному випадку можна вважати множину існуючих образів-ознак невідомого образу-мети. Зрештою формально зведемо задачу до алгоритму навігації між вузлами булеану, викладеному у п. 3.2.

Згідно з вихідними положеннями теорії образної обробки текстової інформації (розділ 2) визначимо склад основних задач, розв'язки яких має забезпечувати оперативна пам'ять СОПМК [29]:

- 1) орієнтувальний рефлекс (пошук за типом АО \rightarrow образ) – з початкового стану АО шляхом запуску асоціативного механізму обміну образами знаходиться такий образ-мета пошуку, який найбільш тісно пов'язаний з образами-ознаками пошуку (відповідає типу образного пошуку № 13 узагальненої класифікації *InsZX*);
- 2) конструювання образу розв'язування (пошук за типом стек \rightarrow АО \rightarrow образ, що також відповідає типам образного пошуку № 6 *InsYX* або № 14 *InsZY* узагальненої класифікації) – попередньо відібрані до стеку складові образу розв'язування переносяться в АО, потім образи-складові вектора емоцій міняються місцями з образами АО, а далі за-

пускається алгоритм 1) з можливим залученням інсайтного пошуку № 7 узагальненої класифікації *Find-1*;

3) генерація повідомлення (побудова за типом АО → синтагма), коли з початкового стану АО формується синтагма (відповідає типу образного пошуку № 15 узагальненої класифікації *SynZZ*):

а) у вигляді опису (протоколу) події, що відбувається просто зараз;

б) як запит на основі образів-складових вектора емоцій про те, що потребує в даний час система;

4) відповідь на питання (побудова за типом синтагма → АО → синтагма) – синтагма-питання перетворюється в АО, потім образи-складові вектора емоцій міняються місцями з образами АО, а далі запускається алгоритм 3.б).

Будемо стверджувати, що розв'язок цих задач може бути досягнутий за рахунок введення певних формальних обмежень у концептуальний механізм функціонування ІС (п. 2.6) та на основі модифікації відомих алгоритмів обходу графів.

4.5.1 Визначення основних понять та загальних принципів моделювання

Визначення основних понять методу моделювання механізму оперативної пам'яті ІС обмежує формальні рамки психологічних феноменів, що є значимими з точки зору запропонованого підходу.

Вважатимемо ансамблем образів (АО) сукупність від 5 до 9 образів, що одночасно знаходяться в оперативній пам'яті та доступні для всіх задекларованих операцій з образами. Згідно з (3.37) позначатимемо АО бінарним кодом V_i-OM .

Вектор емоцій (ВЕ) є числовою характеристикою образу з погляду унікальної системи переваг, бажань та цінностей людини. Оскільки природна фізіологія базується на складних гормональних механізмах емоцій, недоступних в принципі для штучного інтелекту, формально обмежимо значення ВЕ множиною інших образів потужністю $n_e \ll n$, заданих як «центри потреб» згідно з рівнем 3 визначення бінарного нечіткого відношення сенсу (п. 2.4).

Ваги образів, що входять до складу АО оперативної пам'яті, визначаються на основі функції від ВЕ. Вага образу, в загальному випадку, є змінною в часі величиною, яка залежить, у тому числі, від емоційного стану системи. В задачах пошуку та вибору образів значенням

ваги будемо вважати сумарну силу асоціативного зв'язку i -го образу з образами-ознаками мети поточного пошуку згідно з (3.5)

$$Se_i^{EV} = \sum_{j=1}^{n_e} c_{ij} . \quad (4.39)$$

Фокус уваги може бути представлений у вигляді вказівника *Focus*, який в кожен момент часу спрямований на образ з найбільшою вагою в АО.

Для продуктивної роботи образного механізму оперативна пам'ять може взаємодіяти з іншими типами пам'яті. Це, у першу чергу, багатифункціональний додатковий стек для планування та регулювання процесу вибору, що може інтерпретуватися таким чином на основі нейропсихологічних даних щодо лобових долей головного мозку [160]. Джерелом нових образів для моделі будемо вважати асоціативну *Assoc–Memory* та довготривалу *Long–Memory* пам'ять, які формалізуються на основі простору асоціативних пар Ω [22]. Підхід, що пропонується, побудований на тому, що всі розглянуті види пам'яті обмінюються образами як неподільними операційними одиницями, а змістовне навантаження образних конструкцій визначається через асоціативні зв'язки між парами образів [31].

Тоді, формально обмежуючи у межах запропонованого підходу відомий психологічний феномен, вважатимемо орієнтувальним рефлексом алгоритм, що дозволяє уникати зациклення в роботі механізму оперативної пам'яті в процесі пошуку потрібного образу.

Визначення загальних принципів моделювання образного механізму оперативної пам'яті дозволяє враховувати особливості природного розв'язання задач, пов'язаних з пошуком та вибором кінцевого образу-мети. Запропоновані в [29] принципи декларують можливість розв'язання поставленої задачі на основі визначених раніше понять.

Принцип 1. АО оперативної пам'яті може змінюватися в кожний такт часу тільки на один образ, а саме:

- а) образ додається до ансамблю оперативної пам'яті, при цьому загальна кількість образів не може перевищувати 9;
- б) образ вилучається з ансамблю оперативної пам'яті, при цьому загальна кількість образів не може бути меншою за 5;

в) новий образ додається до АО, а замість нього вилючається інший образ.

Фізіологічною підставою цього принципу є особливості *надоперативної пам'яті* людини, в яку одночасно і паралельно потрапляють різнорідні сигнали зі всіх органів чуття, проте моменти часу «розпізнавання» образів з безлічі сигналів можна вважати послідовними та кратними дискретному проміжку часу моделювання [155].

Принцип 2. При появі кожного нового образу з інших видів пам'яті для нього в той же такт часу визначається вага. Тому можна вважати, що значення ВЕ у кожний дискретний такт часу відоме для всіх образів оперативної пам'яті, причому окремо фіксуються складові АО з мінімальною і максимальною (інтерпретується як фокус уваги) вагами.

Принцип 3. Власне механізм оновлення АО здійснюється таким чином:

а) якщо вага нового образу більша за максимальну вагу, то склад АО змінюється за принципом 1а) або 1в) (новий образ замінює образ з найменшою вагою);

б) якщо вага нового образу менша за мінімальну вагу, то склад АО не змінюється;

в) проміжне значення ваги нового образу приводить до застосування принципу 1а) або 1в) (так само новий образ замінює в АО образ з найменшою вагою).

Принцип 4. Під час кожного модельного такту образу у фокусі уваги ставиться у відповідність відсортований в порядку зменшення ряд найбільш близьких за силою асоціативного зв'язку образів, для яких послідовно визначається вага. При необхідності значення ВЕ також можуть бути розрахованими для образів певної події, пов'язаної з образом у фокусі уваги.

Принцип 5. Суть обміну образів в оперативній пам'яті полягає у тому, що після кожної чергової ітерації склад АО глобально поліпшується з погляду сумарного значення ВЕ. Образний механізм оперативної пам'яті закінчує свою роботу при досягненні достатнього значення ваги для образу у фокусі уваги.

Принцип 6. Головна проблема образного механізму полягає в зацікленні, яке принципово можливе в таких випадках:

а) ні один з ряду образів, що асоціюються з образом у фокусі уваги не набирає потрібної ваги для входження в АО;

б) образ у фокусі уваги не змінюється, оскільки нові образи-претенденти до АО набирають меншу вагу, ніж у нього;

в) класичне зациклення – фокус уваги переходить по замкнутому колу всередині групи з двох, трьох чи більше образів.

Принцип 7. З метою запобігання зацикленню необхідно відстежувати розглянуті вище випадки і адекватно на них реагувати, а саме:

а) вноситься на розгляд нова підмножина образів за допомогою непрямої асоціації через події, пов'язані з образом у фокусі;

б) образ у фокусі уваги переноситься з АО у стек контролю за принципом 1б) або 1в) (найвагоміший з образів-претендентів замінює в АО образ з найбільшою вагою), а фокус уваги спрямовується на другий за значенням ваги образ АО;

в) n -вимірне зациклення визначається тоді, коли в стеку контролю з'являється дубль вже наявного у ньому образу – у цей момент необхідно закінчити роботу образного механізму з видачею повідомлення про неможливість розв'язання задачі пошуку прийняттого з точки зору ВЕ образу за існуючих даних.

4.5.2 Формалізація образного механізму оперативної пам'яті СОПМК

На основі введених вище принципів можна запропонувати формальну постановку задачі алгебраїчного моделювання механізму оперативної пам'яті СОПМК.

Згідно з п. 3.2, будемо вважати булеаном множину всіх підмножин множини образів *Image*. Якщо поставити у відповідність кожній з цих підмножин бінарний код, то власне образи знаходяться на першому шарі булеану та мають бінарні коди $Bi-I$ тільки з однією одиницею та іншими нулями. Тоді бінарний код $Bi-Sy$ (синтагми) чи $Bi-OM$ (оперативної пам'яті) довільного i -го шару складається з i одиниць і $n - i$ нулів, де n – потужність множини *Image*. В межах запропонованого підходу [31] досягається накопичення сили прямого $Force^+$ та зворотного $Force^-$ асоціативного зв'язку у парах образів з бінарними кодами $Bi-I-1.Assoc-Twice$ та $Bi-I-2.Assoc-Twice$.

Для реалізації методу моделювання будемо використовувати такі змінні:

- *Weight-Current* – вага поточного образу;
 - *Bi-Current* – бінарний код поточного образу;
 - *Focus-Weight* – вага образу у фокусі уваги (найбільша з образів АО);
 - *Focus-Bi* – бінарний код образу у фокусі уваги;
 - *Weight-Min* – найменша вага образу в АО;
 - *Bi-Min* – бінарний код образу АО з найменшою вагою;
 - *Weight-Top* – достатнє (прийнятне) значення ВЕ;
- і такі додаткові конструкції:
- *Vector-Set* – стек з множиною образів-складових вектора емоцій;
 - *Choice-Set* – стек з множиною нових образів-претендентів для вибору;
 - *Check-Set* – стек для контролю за процесом вибору образу.

Розглянемо формальні операції і предикати на булеані, які відповідають вищевикладеним принципам.

1. Операція переходу на вузол вищого шару *Level-Up* (до АО з кодом *Bi-OM* додається образ з кодом *Bi-I*) і предикат *Top-Border*, який істинний тоді, коли не досягнута верхня межа обсягу оперативної пам'яті – принцип 1а):

$$Bi - OM \cup Bi - I \xrightarrow{Level-Up} Bi - OM ; \quad (4.40)$$

$$Level < 9 \rightarrow Top - Border . \quad (4.41)$$

2. Операція переходу на вузол нижчого шару *Level-Down* (від АО з кодом *Bi-OM* віднімається образ з кодом *Bi-I*) і предикат *Bottom-Border*, який істинний тоді, коли не досягнута нижня межа обсягу оперативної пам'яті – принцип 1б):

$$Bi - OM \cap (\neg Bi - I) \xrightarrow{Level-Down} Bi - OM ; \quad (4.42)$$

$$Level > 5 \rightarrow Bottom - Border . \quad (4.43)$$

3. Операція переходу на інший вузол поточного шару *Change-Image* (в АО з кодом *Bi-OM* образ *Bi-Current* замінює образ *Bi-I*) – принцип 1с):

$$(Bi - OM \cap (\neg Bi - I)) \cup Bi - Current \xrightarrow{Change-Image} Bi - OM . \quad (4.44)$$

4. Операція *Evaluate-Vector* визначає ВЕ для образу *Bi-Current* – принцип 2:

$$\begin{aligned} Evaluate-Vector(Bi - Current) ::= \{[Vector - Set] \\ \sum_{i=1}^k (Force- + Force+ | Bi - I - 1.Assoc - Twice = \\ Bi - Current \cup Bi - I - 2.Assoc - Twice = Bi - I_i)\}, \end{aligned} \quad (4.45)$$

де k – загальна кількість образів-складових вектора емоцій *Vector-Set*.

5. Предикат *Weight+* : показує, що вага поточного образу більша за максимальну вагу в АО – принцип 3а):

$$Weight - Current > Focus - Weight \rightarrow Weight + . \quad (4.46)$$

6. Предикат *Weight-* : показує, що вага поточного образу менша за мінімальну вагу АО – принцип 3б):

$$Weight - Current < Weight - Min \rightarrow Weight - . \quad (4.47)$$

7. Операція побудови хеш-таблиці образів *Hash-Table*, що асоціюються з образом *Focus-Bi* – принцип 4:

$$Focus - Bi \xrightarrow{Hash-Table} Choice - Set . \quad (4.48)$$

8. Предикат *Stop-Find*, що показує закінчення успішного пошуку образу, – принцип 5:

$$Focus - Weight \geq Weight - Top \rightarrow Stop - Find . \quad (4.49)$$

9. Предикат *Assoc-False* : показує, що ні один образ-претендент з *Choice-Set* не набрав потрібної ваги для входження в АО – принцип 6а):

$$\{[Choice - Set : n] Weight- = True\} \rightarrow Assoc - False , \quad (4.50)$$

де n – кількість образів в стеку для вибору *Choice-Set*.

10. Предикат *New-False* : показує, що образ у фокусі уваги не змінився, оскільки нові образи-претенденти до АО мають меншу вагу, ніж у нього – принцип 6б):

$$\{[Choice - Set : n] Weight+ = True\} \rightarrow New - False . \quad (4.51)$$

11. Операція *Add-Check*, яка додає образ *Focus-Bi* до стека контролю *Check-Set*:

$$Focus - Bi \xrightarrow{Add-Check} Check - Set . \quad (4.52)$$

12. Операція *Image-to-Event* для пошуку події, найтісніше пов'язаної з образом у фокусі уваги *Focus-Bi* – принцип 7а):

$$Focus - Bi \xrightarrow{Image-to-Event} Bi - Sy . \quad (4.53)$$

13. Операція *Update-Choice*, яка заносить образи події до стека *Choice-Set*:

$$Bi - Sy \xrightarrow{Update-Choice} Choice - Set . \quad (4.54)$$

14. Предикат *Stop-False*, що показує неуспішне закінчення пошуку, оскільки в стеку *Check-Set* з'явився дубль вже існуючого там образу – принципи 6с) і 7с):

$$\{[Check - Set : m] Focus - Bi = Bi - I.Check\} \rightarrow Stop - False, \quad (4.55)$$

де m – кількість образів у стеку контролю *Check-Set*; $Bi - I.Check$ – бінарний код чергового образу в стеку контролю.

Тоді формальною постановкою задачі моделювання орієнтувального рефлексу будемо вважати

$$Bi - OM, Vector - Set \rightarrow Focus - Bi | Focus - Weight \geq Weight - Top. \quad (4.56)$$

4.5.3 Алгебраїчна модель орієнтувального рефлексу

Для розв'язання поставленої задачі (4.56) з урахуванням формальних операцій та предикатів (4.40)–(4.55) поповнимо двоосновну алгебраїчну систему (3.19) за рахунок

$$OP' = \left\{ \begin{array}{l} Level - Up, Level - Down, Change - Image, \\ Evaluate - Vector, Hash - Table, Add - \\ Check, Image - to - Event, Update - Choice \end{array} \right\}, \quad (4.57)$$

$$IF' = \left\{ \begin{array}{l} Top - Border, Bottom - Border, \\ Weight+, Weight-, Stop - Find, \\ Assoc - False, New - False, Stop - False \end{array} \right\}. \quad (4.58)$$

Тоді $OP = OP \cup OP'$, а $IF = IF \cup IF'$. Покажемо, що задача моделювання механізму оперативної пам'яті у межах алгебраїчної системи *BAS* фактично зводиться до алгоритму навігації між вузлами (підмножинами) булеану від 5-го шару до 9-го на основі операцій *OP* та предикатів *IF*. В граф-схемах для алгебраїчних конструктів, що будуть розглядатися з цією метою, використані позначення для операторів структурного програмування з [257].

Достатньою для розв'язання задачі (4.56) є послідовна побудова зі складових сигнатури алгебраїчної системи *BAS* таких операторів.

1. Оператор *Op-Plus*, що додає до АО новий образ (*Bi-Current*) або заміняє образ *Bi-Min* на образ *Bi-Current* – якщо не досягнуто верхнього кордону оперативної пам'яті (предикат *Top-Border* істинний), то виконується оператор *Level-Up*, інакше виконується оператор *Change-Image*:

$$\begin{aligned} Op - Plus ::= ([Top - Border] \quad Level - Up, \\ Change - Image). \end{aligned} \quad (4.59)$$

Граф-схему оператора *Op-Plus* представлено на рисунку 4.16:

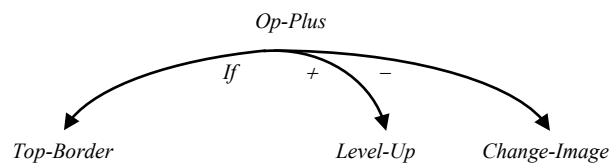


Рисунок 4.16 – Граф-схема оператора *Op-Plus*

2. Оператор *Op-Minus* (рисунок 4.17), що вилучає один образ (*Focus-Bi*) з АО або заміняє образ *Focus-Bi* на образ *Bi-Current* – якщо не досягнуто нижнього кордону оперативної пам'яті (предикат *Bottom-Border* істинний), то виконується оператор *Level-Down*, інакше виконується оператор *Change-Image*:

$$Op - Minus ::= ([Bottom - Border] \\ Level - Down, Change - Image). \quad (4.60)$$

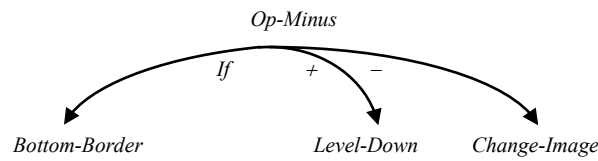


Рисунок 4.17 – Граф-схема оператора *Op-Minus*

3. Оператор *Mechanism* (рисунок 4.18) для моделювання механізму обміну образами в АО – якщо вага поточного образу більша за вагу фокуса уваги (предикат *Weight+* істинний), то виконується оператор *Op-Plus*, а предикат *New-False* отримує значення *False*, інакше, якщо вага поточного образу більша та дорівнює мінімальній вазі АО, то виконується оператор *Change-Image*, що замінює в АО образ *Bi-Min* на образ *Bi-Current*, а предикат *Assoc-False* отримує значення *False*:

$$Mechanism ::= ([Weight+] (Op - Plus * \\ New - False : False), ([Weight-] E, \\ (Change - Image * Assoc - False : False))), \quad (4.61)$$

де *E* – оператор, що не виконує жодних дій.

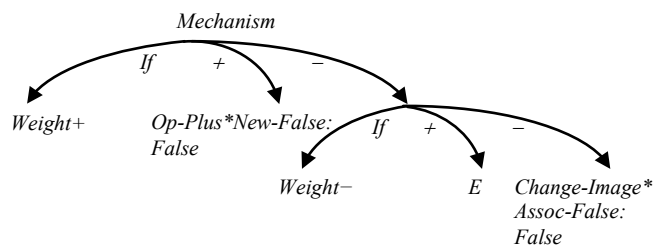


Рисунок 4.18 – Граф-схема оператора *Mechanism*

4. Оператор *Mech-Begin* (рисунок 4.19), що моделює запуск механізму оперативної пам'яті – предикатам *New-False* та *Assoc-False* надається початкове значення *True* та вперше запускається операція *Hash-Table*:

$$Mech - Begin ::= Assoc - False : True * \\ New - False : True * Hash - Table. \quad (4.62)$$

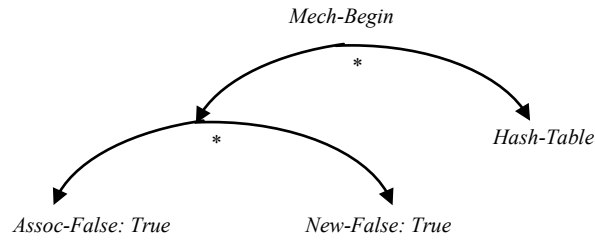


Рисунок 4.19 – Граф-схема оператора *Mech-Begin*

5. Оператор *Focus-Plus* (рисунок 4.20) для зміни фокусу на кращий з точки зору значення ВЕ – виконується операція *Add-Check*, яка додає образ *Bi-Current* до стека контролю *Check-Set* та перевіряється предикат *Stop-False*, якщо такий самий образ в стеку контролю уже присутній, то алгоритм закінчується операцією *Abort*, інакше значення *Bi-Current* присвоюється змінній *Focus-Bi*, а потім виконується операція *Hash-Table*:

$$\begin{aligned}
 \textit{Focus-Plus} ::= & \textit{Add-Check} * ([\textit{Stop-False}] \\
 & \textit{Abort}, \textit{Focus-Bi} = \textit{Bi-Current}) * \textit{Hash-Table}.
 \end{aligned}
 \tag{4.63}$$

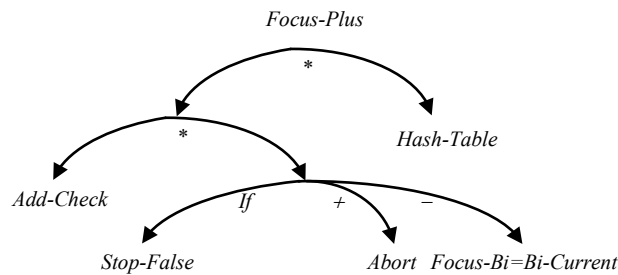


Рисунок 4.20 – Граф-схема оператора *Focus-Plus*

6. Оператор *Focus-Minus* (рисунок 4.21) для зміни фокусу на гірший з точки зору значення ВЕ – виконується оператор *Op-Minus*, потім змінна *Bi-Current* отримує найбільше значення ваги образу з АО та виконується оператор *Focus-Plus*:

$$\begin{aligned}
 \textit{Focus-Minus} ::= & \textit{Op-Minus} * \\
 \textit{Bi-Current} = & \textit{Max}(\textit{Bi}_j) * \textit{Focus-Plus}.
 \end{aligned}
 \tag{4.64}$$

$\textit{Bi}_j \subset \textit{Bi-OM}$

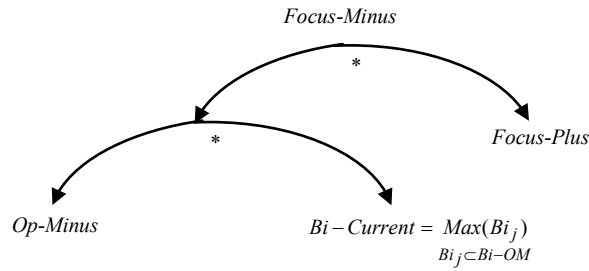


Рисунок 4.21 – Граф-схема оператора *Focus-Minus*

7. Оператор *New-Event* (рисунок 4.22), що моделює згадування події за асоціацією з образом, – послідовно виконуються операції *Image-to-Event* та *Update-Choice*:

$$\begin{aligned}
 & \textit{New-Event} ::= \textit{Image-to-Event} * \\
 & \textit{Update-Choice}.
 \end{aligned}
 \tag{4.65}$$

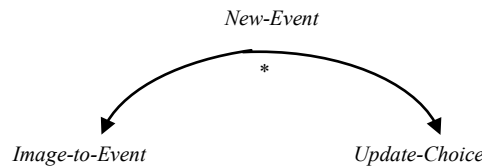


Рисунок 4.22 – Граф-схема оператора *New-Event*

8. Оператор *Operative-Memory* (рисунок 4.23) для моделювання мнемонічних процесів в оперативній пам'яті – перед початком головного циклу виконується оператор *Mech-Begin* та, за кількістю образів n в стеку вибору *Choice-Set*, перевіряється предикат *New-False*, якщо фокус уваги не змінився, то виконується оператор *Mechanism*, в іншому випадку – оператор *Focus-Plus*, після головного циклу перевіряється предикат *Assoc-False*, якщо склад АО не обновлювався, то виконується оператор *New-Event*, інакше – оператор *Focus-Minus*:

$$\begin{aligned}
 & \textit{Operative-Memory} ::= \textit{Mech-Begin} * \\
 & \{ [\textit{Choice-Set} : n] ([\textit{New-False}] \textit{Mechanism}, \\
 & \quad \textit{Focus-Plus}) \} * ([\textit{Assoc-False}] \\
 & \quad \textit{New-Event}, \textit{Focus-Minus}).
 \end{aligned}
 \tag{4.66}$$

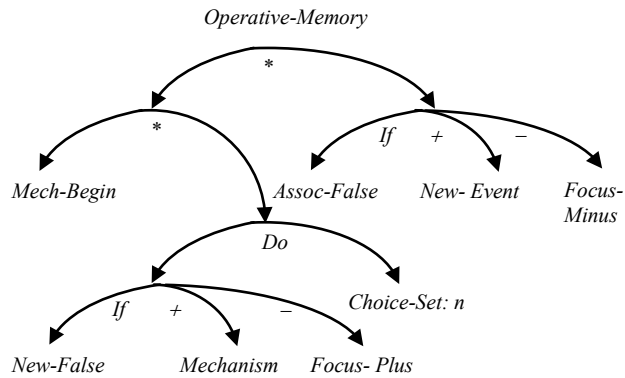


Рисунок 4.23 – Граф-схема оператора *Operative-Memory*

9. Оператор *Orient-Reflect* (рисунок 4.24) для моделювання орієнтувального рефлексу – оператор *Operative-Memory* виконується у циклі до тих пір, доки предикат *Stop-Find* не стане істинним:

$$\begin{aligned}
 & \textit{Orient} - \textit{Re} \textit{flect} ::= \{[\textit{Stop} - \textit{Find}] \\
 & \textit{Operative} - \textit{Memory}\}.
 \end{aligned}
 \tag{4.67}$$

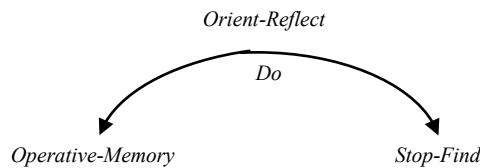


Рисунок 4.24 – Граф-схема оператора *Orient-Reflect*

Розглянемо розв’язок задачі (4.56) на даних наскрізного прикладу п. 3.1, 12-ти мовним образам якого відповідають бінарні коди B_i-I згідно з таблицею 3.3. Задамо початковий стан оперативній пам’яті 6-ма такими МО: «мова (56)», «XML (88)», «існує (71)», «засіб (93)», «ресурс (73)», «RDF (87)», а до складу МО вектора емоцій включимо «інформація (34)» та «мережі (17)». Формально, у відповідності до постановки задачі (4.56), визначимо $Weight-Top=4$, $B_i-OM=001111110000$, $Vector-Set=\{000000000010, 000000000001\}$. На момент початку дії оператора *Orient-Reflect* хоча б з одним образом $Vector-Set$ прями зв’язки має лише МО «ресурс (73)», отже $Focus-Weight=2$.

Оператор *New-Event* знаходить найбільш відповідну до стану B_i-OM синтагму з МО «ресурс (73)» (№ 19 згідно з Додатком Б), звідки до оперативної пам’яті додається оператором *Op-Plus* МО «означення

(37)». Отримуємо нове значення $Bi-OM=001111111000$, але цей крок не змінює значення $Focus-Weight$, оскільки новий МО також не має прямих зв'язків з образами $Vector-Set$. Зрозуміло, що внаслідок аналогічних дій на обмежених даних наскрізного прикладу будуть ще проаналізовані МО «дозвіл (123)», «зміст (35)» та «онтологія 111», причому тільки два останніх мають ненульове значення ваги, але лише МО «зміст (35)» завершить алгоритм з результатом $Focus-Weight=5$. Характер отриманих даних з наскрізного прикладу демонструє, що запропонований підхід до моделювання концептів оперативної пам'яті СОПМК засобами алгебраїчної системи BAS має високу ступінь адекватності навіть на обмеженому експериментальному матеріалі невеликого за обсягом тексту.

Отже, визначено поняття та сформульовано принципи, що дозволяють послідовно побудувати метод моделювання механізму оперативної пам'яті СОПМК за допомогою формальної інтерпретації таких відомих феноменів, як ансамбль образів, вектор емоцій, фокус уваги, орієнтувальний рефлекс. Подальша формалізація запропонованого концептуального підходу призвела до алгебраїчної постановки задачі навігації між вузлами булеану від 5-го шару до 9-го, які відповідають ансамблю образів або подіям.

На основі введених обмежень відомого психологічного феномену орієнтувальний рефлекс представлено як одну з задач оперативної пам'яті, що розв'язується шляхом інтерпретації орієнтувального рефлексу і відповідного механізму оперативної пам'яті СОПМК у вигляді алгоритму навігації у графі булеану алгебраїчної системи BAS . Запропонований алгоритм демонструє високий ступінь адекватності на експериментальних даних наскрізного прикладу.

4.6 Метод самовдосконалення бази знань системи на основі моделювання складових парадигматичного устрою мови

Згідно з концептуальними положеннями (п. 2.2), де визначені 8 типів асоціацій, більшість з них (6 типів) відповідає парадигматичній організації мови, яка є вторинною по відношенню до синтагматичної основи. Маються на увазі:

- 1) P_c – синонімічна асоціація або результат образного пошуку $SynXX$ згідно з (4.3);
- 2) P_a – антонімічна асоціація ($AntXX$);

- 3) P_c – асоціація типу окреме–загальне ($PgXX$);
- 4) P_u – асоціація типу частина–ціле ($CmXX$);
- 5) P_o – омонімічна асоціація (паронімія);
- 6) P_p – асоціація типу «у риму».

Задачею дослідження вважатимемо розробку методу генерації цих типів парадигматичних зв'язків на основі бази знань синтагматичного походження та формальне визначення їх ролі при інтерпретації задач пізнавальної діяльності на основі ІС. З цією метою розглянемо алгоритми утворення всіх видів парадигматичних зв'язків за допомогою графа $G_z(V, E)$. Загальна ідея методу, що пропонується, полягає у визначенні нового парадигматичного зв'язку на основі певного узагальнення вже відомих синтагматичних.

Проінтерпретуємо типову потребу, що виникає в процесі мовної практики – знайти синонім до образу $v_i \in V$. Синонімія, згідно з прийнятими у п. 2.2 мовними концептами образного сенсу, у першу чергу, притаманна якості об'єкта та якості методу. Отже, задача зводиться до визначення двосторонньої синонімічної асоціації Pc_{ij} , $j \neq i$, в якій образи i та j одночасно подаються у вигляді якості об'єкта або якості методу. Ідея відповідного алгоритму полягає у знаходженні такої j -ї вершини графа, яка має множину зв'язків з образами-об'єктами або з образами-методами, подібну до множини зв'язків i -ї вершини графа. Приклад підграфа G_z для визначення синонімії російськомовних образів-понять «быстрота» та «скорость», що грають роль якостей для ряду образів-методів зображено на рисунку 4.25.

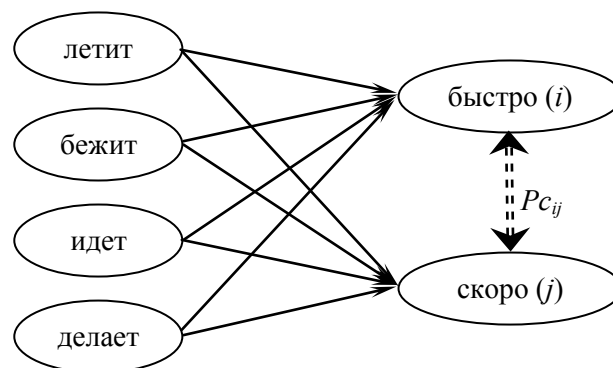


Рисунок 4.25 – Визначення синонімії образів-якостей через образи-методи

Уведемо такі позначення:

j_x – вершина пошуку (синонімічна для i -ї вершини графа);

\bar{x}_{\max} – кількість спільних образів-методів для i -ї та j -ї вершин графа, перед запуском алгоритму $\bar{x}_{\max} = 0$.

Алгоритм *SynXX* визначення синонімічної асоціації $P_{c_{ij}}$ для якості методу складається з таких кроків:

[Крок 1] Досліджується i -та вершина. У стовпці i матриці A_Q визначається множина $\{l\}$ таких елементів, для яких $a_{li} \geq 0$ та кожна l -та вершина має відображення у вигляді методу.

[Крок 2] Організовується цикл для всіх j -х вершин, таких що $j \neq i, j \notin \{l\}$. Для кожної j -ї ітерації циклу:

а) визначається множина $\{m\}$ таких елементів, для яких $a_{mj} \geq 0$ та кожна m -та вершина має відображення у сховищі ІС як метод;

б) визначається потужність \bar{x} множини $X = \{l\} \cap \{m\}$;

с) якщо $\bar{x} > \bar{x}_{\max}$, то $j_x = j$ та $\bar{x}_{\max} = \bar{x}$.

[Крок 3] Вершина j_x виводиться як синонімічна до вершини i .

Примітки:

- якщо необхідно знайти ще один синонім для i -ї вершини, то в умову циклу **[Крок 2]** додатково вноситься попередньо знайдена вершина j_x ;

- у випадку визначення синоніму для якості об'єкта використовується цей же алгоритм *SynXX*, в якому термін «метод» замінюється на термін «об'єкт».

Тестування з деякими обмеженнями алгоритму *SynXX* на експериментальних даних наскрізного прикладу (пп. 3.1, 3.6) дозволило визначити наявність синонімічного зв'язку між МО «онтологія (111)» та «зміст (35)» через МО «мова (56)» та «мережа (17)». Результати тестування представлено у Додатку Д.

З метою визначення антонімічної асоціації P_a скористаємося тим фактом, що заперечення антонімії приводить до синонімії, наприклад «маленький–невеликий», «слабкий–несильний», «низький–невисокий» або «добрий–незлий». Отже, пошук антоніму до образу $v_i \in V$ формально також можна звести до алгоритму *SynXX* з такими змінами:

- умови [Кроку 1] і [Кроку 2а)] необхідно визначати лише для тих зв'язків, які з'єднують відповідний метод з якістю методу через заперечення – це, наприклад, частки не-, без-, анти-, контр- тощо для української мови.

У випадку використання сховища ІС (п. 3.6) для організації бази знань елементів ЕК, таку умову нескладно реалізувати у вигляді реляційного виразу. Приклад підграфа G_z^1 для визначення антонімії російськомовних образів «медленно» та «быстро», що відіграють роль якостей для ряду образів-методів, представлено на рисунку 4.26 (зв'язки з запереченням зображено на рисунку штрих-пунктирними лініями).

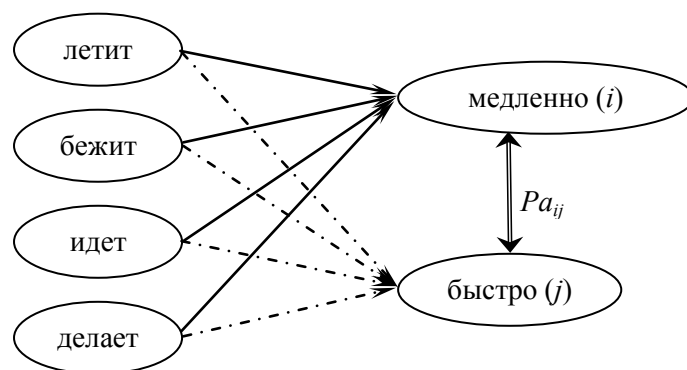


Рисунок 4.26 – Визначення антонімії образів-якостей через образи-методи

Тестування алгоритму *SynXX* для пошуку антонімів на повних даних експериментального прикладу (Додаток Б) дозволило визначити наявність антонімічного зв'язку між МО «довільний (138)» та «обов'язковий (148)» через МО «ресурс (73)» та «вказати (149)», а також між «довільний (138)» та «необхідний (159)» через МО «ресурс (73)» та «поняття (127)». Зрозуміло, що другий варіант не може вважатися антонімією, хоча МО «обов'язковий (148)» та «необхідний (159)» є синонімами. Результати тестування представлено у Додатку Д.

Асоціація типу окреме–загальне P_q характерна для відношень між об'єктами (пізніше в онтогенезі і між поняттями), причому визначається таке відношення, як вже було показано раніше, через методи та якості об'єктів. Саме зв'язок P_q є основою для побудови онтологій, які можна вважати класичним результатом розв'язання окремих задач пізнавальної діяльності [107]. Ідея алгоритму визначення зв'язку P_q полягає у попередній перевірці на синонімію i -ї та j -ї вершин графа як

образів-об'єктів через образи-методи та образи-якості з подальшим додатковим визначенням підлеглого (часткове) та головного (загальне) образів. Ознаками підлеглості у цьому типі зв'язку для синонімічних об'єктів може бути:

- наявність унікальних, на відміну від головного образу, зв'язків з окремими образами-методами та/або образами-якостями об'єкта;
- наявність набагато сильніших за вагою зв'язків з певними образами-методами та/або образами-якостями об'єкта.

Приклад підграфа G'_z для визначення зв'язку типу окреме–загальне між образами «кущ» та «агрус» представлено на рисунку 4.27, де зв'язки з суттєво меншою вагою зображено штриховими лініями.

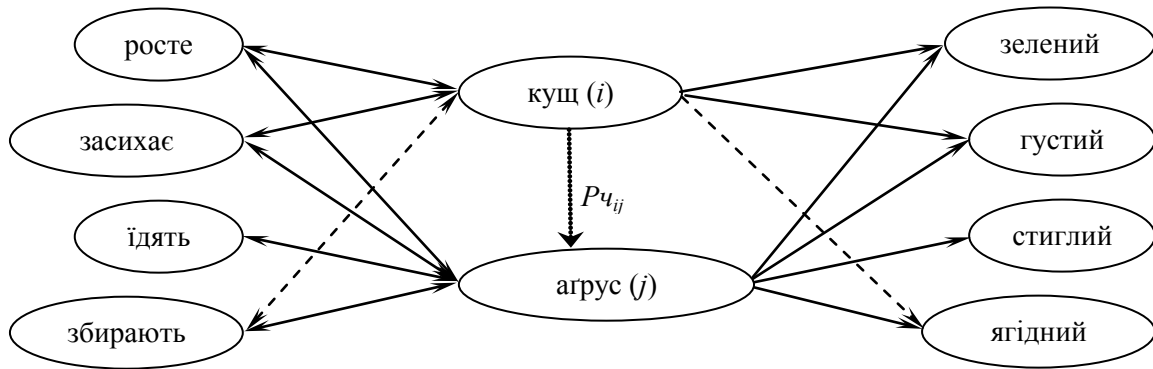


Рисунок 4.27 – Визначення зв'язку типу окреме–загальне для образів-об'єктів

Пошук зв'язків типу окреме–загальне P_u на повних даних експериментального прикладу (Додаток Б) на основі запропонованого підходу показав позитивні результати для часткового МО «XML (88)» та загального «мова (56)» через МО «призначати (99)», «розмітити (89)» та «комп'ютерний (27)». Результати тестування представлено у Додатку Д.

Асоціація типу частина–ціле P_u разом з попереднім парадигматичним зв'язком P_u закладає основу логічного мислення людини. Ідея алгоритму визначення зв'язку P_u полягає в аналізі всієї матриці суміжності A з метою виявлення пар вершин графа виключно з одностороннім рухом у тих синтагмах, де поєднуються між собою два образи-об'єкти або два образи-поняття. У таких випадках головний образ виступає частиною, а підлеглий – цілим, наприклад, «гілка → дерева», «рукав → куртки», «лапа → звіра», «частина → цілого» тощо.

Асоціація за схожістю грає непересічну роль для всієї парадигматичної організації мови і приводить не тільки до появи омонімічної асоціації P_o та асоціації типу «у риму» P_p , але й до побудови всіх можливих морфологічних ознак окремих слів мови. Морфологічний аналіз природно-мовних конструкцій на даний час формалізовано достатньо повно для більшості мов світу [93], а тому для визначення окреслених вище типів асоціацій за схожістю можна скористатися одним з відомих критеріїв, що оцінює довжину спільної для двох слів послідовності символів. Так, для визначення лексичної наближеності слів A та B застосуємо модифікований коефіцієнт Джаккарда [261]:

$$jaccard(A, B) = \frac{|A \cap B|}{\max(|A|, |B|)},$$

де результатом операції перетину \cap є найдовша спільна послідовність символів, що починається з будь-якої позиції у словах A та B . Задача визначення типу асоціації за схожістю формально розв'язується в залежності від значення коефіцієнта лексичної наближеності, додаткових параметрів, закладених у сховище ІС та системи правил морфологічного аналізу певної мови. Окрім цього, актуальною є споріднена задача зменшення розмірності n графа G_z шляхом об'єднання вершин, що належать тим же самим образам.

Пошук зв'язків типу частина–ціле P_u та омонімічної асоціації P_o на повних даних експериментального прикладу (Додаток Б) згідно з вищевикладеним підходом також дав позитивні результати: МО «ресурс (73)» є частиною від цілого «мережа (17)» і «RDF (87)», а МО «призначати (99)» схожий до МО «означати (37)». Результати програмного експерименту представлено у Додатку Д.

Розглянемо тепер інтерпретацію задач пізнавальної діяльності для запропонованого методу самовдосконалення бази знань ІС. До таких задач можна віднести визначення всіх розглянутих вище парадигматичних зв'язків та побудову, зрештою, онтологічних класифікацій (таксонометрій). Оцінимо кількість сенсу-параметра, що змінюється при функціонуванні інфологічної системи згідно з методом побудови нечіткого відношення образного сенсу (п. 2.4).

Запропонований підхід акумулює всю зовнішню образну інформацію про навколишній світ, що надходить на основі синтагматичного

опису подій, у вигляді матриці суміжності A_0 розмірністю $(n \times n)$. Будемо вважати, що загальний сенс синтагматичної інформації про зв'язки між образами або початковий когнітивний потенціал системи є пропорційним кількості виявлених асоціативних зв'язків та загальній кількості повторень цих зв'язків. Якщо (i,j) -м елементом матриці є кількість повторень відповідного асоціативного зв'язку ($a_{ij} = k_{ij}$), то формально введемо загальну функцію сенсу, яка залежить від кількості m ненульових елементів матриці A_0 та суми k_Σ всіх елементів матриці:

$$Se^S = F(m, k_\Sigma), \quad (4.68)$$

$$k_\Sigma = \sum_{i=1}^n \sum_{j=1}^n k_{ij},$$

де Se^S – загальна кількість сенсу інфологічної системи.

В загальному випадку $m \leq n \times (n - 1)$, проте, враховуючи особливості лінгвістичної інтерпретації графа G_z , на практиці обмеження зменшується до $m \leq n \times \bar{\lambda}$, де $\bar{\lambda} \ll n$ – середня кількість вихідних зв'язків кожного з n образів.

Зафіксуємо на певний момент часу стан матриці A_0 , що дозволить стверджувати $n, k_\Sigma, \bar{\lambda} = const$. Тоді, у відповідності до п. 2.4, будемо вважати кількістю нормованого межами відрізка $[0, 1]$ сенсу (i, j) асоціативного зв'язку

$$Se_{ij} = 1 / (1 + e^{-\frac{k_{ij} - \bar{\lambda}}{\mu}}), \quad (4.69)$$

де більший за одиницю коефіцієнт емоційного стану $\mu = const$. Зрештою загальну кількість сенсу (4.68) представимо у вигляді суми

$$Se^S = \sum_{i=1}^n \sum_{j=1}^n Se_{ij}, \quad (4.70)$$

яка так само обмежена значенням $Se^S \leq n \times \bar{\lambda}$ внаслідок присутності у матриці нульових елементів. Враховуючи те, що середня оцінка образного сенсу одинарного синтагматичного зв'язку згідно з (2.18) дорівнює $1/2$, загальна кількість сенсу в системі наближається до $Se^S \approx m/2$.

Відзначимо, що створення будь-якого з парадигматичних зв'язків як задача пізнавальної діяльності потребує узагальнення певної множини синтагматичних асоціацій. Наприклад, згідно з алгоритмом *SynXX* визначення синонімічної асоціації Pc_{ij} між двома існуючими у системі образами базується на використанні сенсу зв'язків з m' спорідненими образами

$$Se^{Pc}_{ij} = \sum_{l=1}^{m'} Se_{li} + \sum_{l=1}^{m'} Se_{lj}. \quad (4.71)$$

Отже, визначення синоніму в середньому приводить до збільшення сенсу в інфологічній системі на m' , оскільки середнє значення сенсу одинарного синтагматичного зв'язку дорівнює $1/2$.

Проведемо порівняння змін загального сенсу системи у 4-х можливих випадках:

1. Підтверджується вже існуючий синтагматичний зв'язок ($k_{ij} = k_{ij} + 1, k_{\Sigma} = k_{\Sigma} + 1$) – сенс збільшується на практично несуттєве

значення $\delta \rightarrow 0$, де $\delta = 1/(1 + e^{\frac{k_{ij}+1-\bar{\lambda}}{\mu}}) - 1/(1 + e^{\frac{k_{ij}-\bar{\lambda}}{\mu}})$. Найбільше значення $\delta \approx 0,231$ досягається за малоймовірних умов $k_{ij} = \bar{\lambda}, \mu = 1$.

2. Створюється новий синтагматичний зв'язок між існуючими образами ($k_{ij} = 1, k_{\Sigma} = k_{\Sigma} + 1, m = m + 1$) – сенс збільшується не більше, ніж на $\delta + 0,5$.

3. До системи вноситься новий образ, який внаслідок орієнтувального рефлексу отримує елементарні зв'язки з 7 ± 2 існуючими образами ($n = n + 1, k_{\Sigma} = k_{\Sigma} + 7, m = m + 7$) – сенс збільшується не більше, ніж на $\delta + 3,5$.

4. Створюється новий парадигматичний зв'язок між існуючими образами – сенс в середньому збільшується на m' : $Se^S \rightarrow Se^S + m'$.

У перших 3-х випадках при збільшенні значення $\bar{\lambda}$ кількість нового образного сенсу наближається, відповідно до 0, 0,5 та 3,5. Отже, розв'язання задач пізнавальної діяльності приводить до помітного збільшення загального образного сенсу інфологічної системи, що можна порівняти лише з дією потужного орієнтувального рефлексу в момент створення нового образу.

РОЗДІЛ 5

РЕАЛІЗАЦІЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ ТА РОЗВ'ЯЗАННЯ СЕМАНТИКО-ЗАЛЕЖНИХ ЗАДАЧ

На основі отриманої методології будується інформаційна технологія та розв'язуються семантико-залежні задачі, а саме – внесення в систему текстової інформації ЕК за допомогою словника образів, формалізація результатів пізнавальної діяльності, конструювання образу розв'язування проблемної ситуації, генерація множини повідомлень СОПМК, побудова відповідей на певні типи питань. Розробляється програмне забезпечення інформаційної технології, проводиться ґрунтовний аналіз впровадження результатів дослідження.

5.1 Побудова інформаційної технології на основі СОПМК

Отримані у попередніх розділах результати моделювання взаємопов'язані між собою не тільки метою дослідження, але й системою формальних обмежень на процеси образної обробки ПМК. У відповідності до запропонованого та розвинутого через методологію підходу побудуємо інформаційну технологію на основі одержаних наукових результатів з урахуванням таких формальних обмежень (рисунок 5.1):

а) на рівні концептуальної ідеї дослідження – формулювання понять онтогенетичного принципу, мовного образу, інфологічної системи та образного сенсу ПМК; розмежування в останньому сенсу-властивості та сенсу-параметра;

б) на рівні формальної метамови – закладення онтогенетичного принципу в основу функціонування інфологічної системи; представлення сенсу-властивості у вигляді семантичної мережі АМО, а сенсу-параметра на основі нечітких відношення та простору образного сенсу;

в) на рівні теорії образного аналізу текстової інформації – формалізація ІС як абстрактної конструктивної моделі об'єкта дослідження; аксіоматика прикладної теорії першого порядку та аксіоматика простору з нечіткою мірою;

г) на рівні синтезу структурно-функціональних моделей СОПМК – механізм онтогенезу ІС на основі кібернетичної інтерпретації нейропсихологічних даних; склад блоків, функцій та образних потоків моделі інфологічної системи; представлення нескінченних множин скінченними; вибір основ та сигнатури алгебраїчної системи; обмеження фу-

нкцій СОПМК класифікацією образного пошуку та функціональними вимогами до ІС;

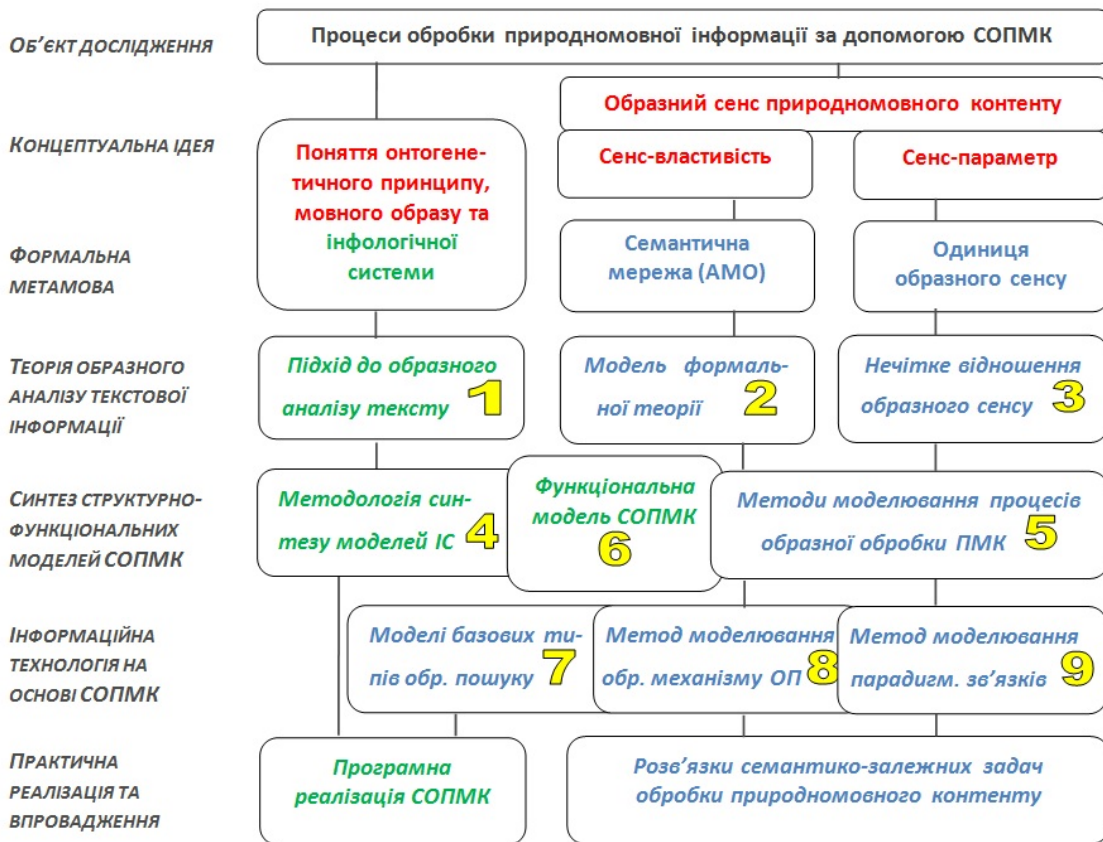


Рисунок 5.1 – Взаємозв'язок системи обмежень та результатів дослідження

д) на рівні інформаційної технології – послідовна побудова графових моделей для базових типів образного пошуку; обмеження методу моделювання образного механізму оперативної пам'яті СОПМК правилами обміну образами та методу моделювання складових парадигматичного устрою мови класифікацією асоціативних відношень.

Розглянемо можливості та особливості інформаційної технології на основі отриманих 9 наукових результатів, позначених на рисунку 5.1 цифрами та вище окресленою п'ятирівневою системою обмежень. Будемо виходити з того, що вхідна інформація для формування бази знань ІС має надходити виключно через ПМК, а саме у вигляді текстів, попередньо відібраних експертом з предметної області згідно з рисунком 1.1.

Розроблено схему організації інформаційних процесів та процедур технології образної обробки ПМК (рисунок 5.2), яка забезпечує розв'язки обумовлені вимогами до ІС семантико-залежних задач як інформаційний продукт на основі отримання інформації з множини текстів та окремих ПМК-запитів.

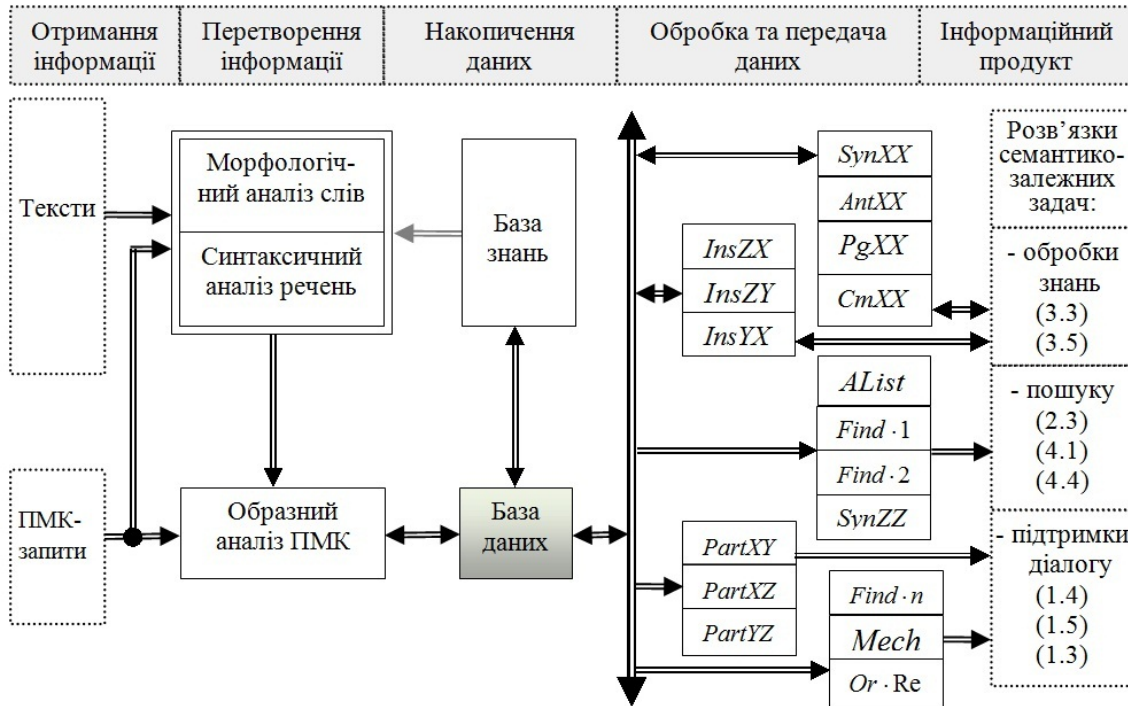


Рисунок 5.2 – Схема організації інформаційних процесів та процедур у межах отриманої інформаційної технології

На рівні процесу перетворення природно-мовної інформації технологія передбачає як відомі процедури морфологічного та синтаксичного аналізу, так і нову процедуру образного аналізу ПМК. Накопичення даних відбувається через загальну базу даних, архітектура якої відповідає п. 3.5 та базу знань образного сенсу у вигляді АМО (п. 3.6). Окрім цього, база знань містить продукційні правила морфологічного та синтаксичного аналізу.

Обґрунтовані у попередніх розділах базові методи, моделі та алгоритми СОПМК виконують роль процедур на рівні процесу обробки та передачі даних інформаційної технології. Частина даних внаслідок розв'язання семантико-залежних задач обробки знань повертається через базу даних з метою самовдосконалення бази знань. Алгоритми

пошуку та підтримки діалогу ініціюються введеними користувачами природно-мовними запитам.

У межах інформаційної технології базовим носієм знань [14] запропоновано вважати модель асоціативної пам'яті СОПМК на основі АМО. Можливими шляхами накопичення такого типу знань принципово можуть бути:

- використання існуючих асоціативних словників, отриманих у результаті широкомасштабного вільного асоціативного експерименту;
- «вирощування» власної АВМ на основі запропонованої онтогенетичної концепції моделювання образного сенсу ПМК.

Обидва шляхи пов'язані з певними труднощами у реалізації, проте для першого з них винятково важливим є коректне проведення вільного асоціативного експерименту зі значною кількістю суб'єктів опитування [121]. З іншого боку, кінцеві результати масштабного експерименту все одно приведуть до отримання, можливо трохи краще узагальнених, ніж у створеному евристичним шляхом гіпертексті, парадигматичних зв'язків дорослих носіїв мови. Проте гіпертекст або онтологія хоча б демонструють зв'язок понять вузької предметної області, а АВМ за визначенням поєднує більш широке коло понять «здорового глузду». Отже, штучне поєднання готового асоціативного словника з навчальним текстом не дає нової якості розуміння сенсу нового матеріалу – це схоже на те, як буде відповідати на питання з абсолютно новою для себе дисципліни загально освічений студент-відмінник.

У межах другого підходу просте розповідне речення (синтагма) тексту розглядається як аналог події, де образи об'єднуються між собою за допомогою асоціативних зв'язків [26]. У такому випадку поступове введення синтагм навчального тексту у систему суттєво доповнює традиційний гіпертекст з парадигматичними зв'язками можливостями синтагматичного знання у вигляді додаткових асоціативних зв'язків між мовними образами. Цей підхід вже нагадує відповіді на питання викладача-виробничника, який на практиці досконало знає предмет, що викладає. Тому відповіді такого викладача мають більш практичну, ніж теоретичну спрямованість.

Зауважимо, що і студент-відмінник і викладач-виробничник алегорично є моделями предметної області знань. Ще однією альтернативною метафорою для популярного підходу до побудови бази знань на основі гіпертексту або онтологій може бути викладач-теоретик. Не

применшуючи роль теоретичного узагальнення, первинним етапом пізнавальної діяльності все ж таки потрібно вважати наочно-дієвий досвід [186]. Тому побудуємо функцію внесення нової синтагми на основі принципів, що закладені до структурно-функціональної моделі ІС (розділ 2). Основні вихідні постулати для реалізації окресленої функції вперше запропоновано в [26], а здатність ІС генерувати відповіді у традиційній формі за такого підходу – в [53].

Передусім необхідно відмітити принцип самонавчання живих носіїв інтелекту. Відомо, що немовля практично з нуля навчається всієї складності світу за простим принципом, сформульованим ще в стародавньому Римі як «поступовість і неперервність». Для сприйняття будь-якого нового знання потрібен міцний фундамент, тому словник образів необхідно «випестити» до необхідного рівня шляхом організаційної опіки від найбільш загальних образів та понять до спеціальних термінів. Інакше парадоксальною виглядає експертна система, що легко вирішує складні, але вузькоспеціалізовані задачі, проте безсила перед «дитячими» питаннями.

Істотною відмінністю природної пам'яті від комп'ютерної є принципова відсутність операції стирання інформації [22]. Якщо людина зустрічає новий образ, то вона ніколи його не проігнорує навіть тому, що в окремих випадках це може призвести до небезпечних наслідків. Так само не можна видалити вже наявний образ, наприклад, негативний – випадки підсвідомого витіснення, як показав ще З. Фрейд, приводять до патології у психіці людини. Психоаналіз як метод лікування пропонує «витягування» образу до свідомості і «нарощування» на нього позитивної мережі зв'язків з іншими образами [126]. Цей принцип організації пам'яті можна, дещо спрощуючи, сформулювати, як «всіх впускати – нікого не випускати».

Третій з основних постулатів підходу до побудови бази знань стосується механізму поповнення словника образів. Відомо, що при сприйнятті нового образу у людини на несвідомому рівні вмикається орієнтувальний рефлекс [29]. Одним із завдань цього рефлексу є побудова мережі зв'язків нового образу з образами, що вже існують в пам'яті. Іншими словами, відбуваються процеси розпізнавання або класифікації образу. Будь-який новий образ принципово не може бути незалежним, оскільки у момент введення в систему поєднується максимально можливою на даний час кількістю зв'язків з іншими образа-

ми. Тим самим виявляється рекурентна сутність визначення мовного образу – «скажи мені, хто твої друзі, і я скажу, хто ти».

З метою реалізації окреслених постулатів пропонується напівавтоматична методика внесення інформації з ЕК в АМО інфологічної системи:

1. Користувачу надається можливість ввести текст речення в систему.

2. Кожному слову речення ставиться у відповідність один з мовних образів:

а) автоматично – у випадку, коли саме таке слово вже існує у пам'яті ІС;

б) напівавтоматично – шляхом вибору користувачем зі списку існуючих образів;

в) ручним введенням користувачем нового образу – у випадку відсутності потрібного образу в системі.

З метою прив'язки вибраного слова пари до образу у випадку 2б) для кожного слова речення складається ранжований список найбільш схожих зовні вербальних позначень з образів *Image*, що вже існують в словнику. При цьому користувач може вибрати в меню існуючу опцію зі списку, при необхідності відкоректувати відповідну статтю словника образів або ввести абсолютно нову.

Для введення асоціативних пар синтагми в систему користувачеві надається можливість задати кількість пар, а потім вибрати кожне слово пари в меню, складеного зі слів поточного речення *Event*.

Питальний займенник між образами пари задається двома способами:

- спочатку можна вибрати тип зв'язку з 7 кортежів відношення *Link* (визначення, присудок, підмет, обставина місця, обставина часу, обставина, доповнення), тоді вибраний тип стає фільтром, і кількість можливих займенників *Inter-Pronoun* зменшується;
- якщо спочатку вибрано питання як кортеж з *Inter-Pronoun*, то для контролю користувачеві автоматично демонструється відповідний йому тип зв'язку з *Link*, оскільки однакові займенники можуть відноситися до різних типів, наприклад питання «що?» ставиться як до підмета, так і до доповнення.

Розглянута методика напівавтоматичного введення природно-мовної інформації позначена як блок 3 на структурній схемі інформаційної технології (рисунок 5.3).

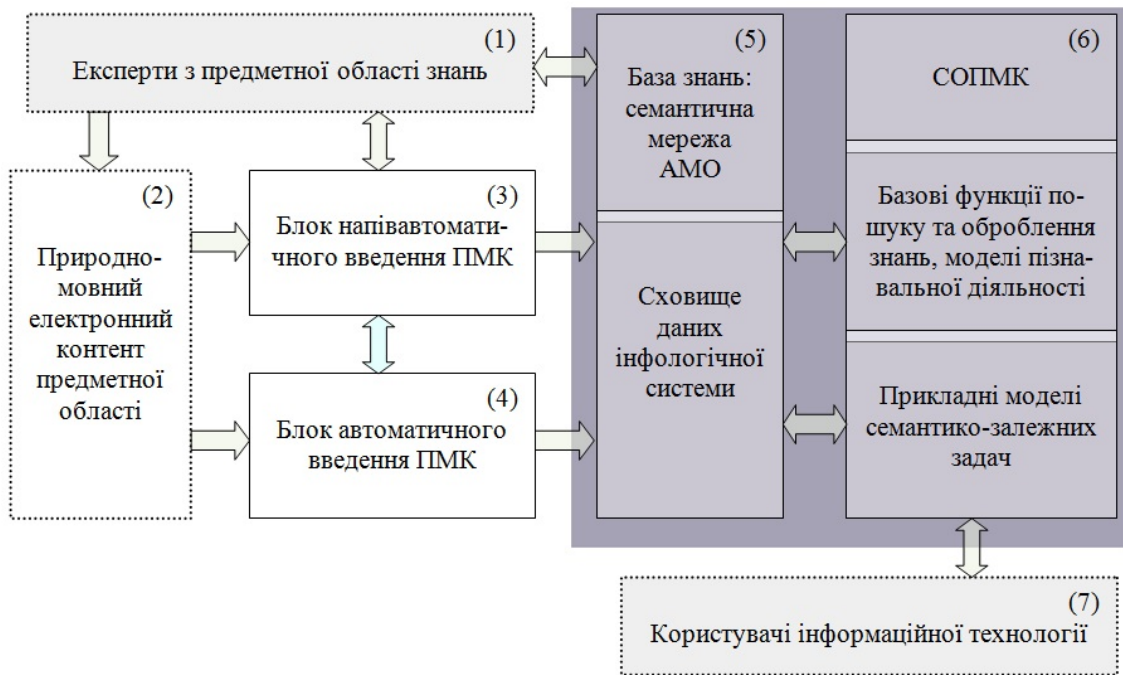


Рисунок 5.3 – Структурна схема інформаційної технології на основі СОПМК

Через блок (3) експерти з предметної області знань (1) вносять до бази знань системи на основі АМО (5) окремі речення з попередньо відібраного природно-мовного ЕК предметної області (2). Прискорення цієї процедури забезпечується блоком автоматичного введення ПМК (4) як альтернативним шляхом досягнення значного обсягу семантичної мережі АМО. Для останнього блока згадане обмеження на рівні практичної реалізації інформаційної технології набуває критичного характеру – якість обслуговування інфологічною системою користувачів технології (7) безпосередньо залежить від якості відібраного природно-мовного контенту.

На відміну від розглянутої методики побудови блока (3) автоматичне введення ЕК до ІС, яка представлена на схемі блоками бази знань (5) та СОПМК (6), у значно більшій мірі залежить від мови контенту. Ключовим елементом у цьому процесі є парсер – синтаксичний аналізатор. Загалом невисока якість роботи парсерів природних мов на даний час помітно вища, наприклад, для англійської мови, ніж для типових афікських мов – української або російської. Для підвищення рівня автоматизації в системі потрібно послідовно застосувати алгоритми, представлені у таблиці 5.1.

Таблиця 5.1 – Склад алгоритмів блока автоматичного введення ПМК до ІС

№ з/п	Назва алгоритму (модуля)	Вхід	Вихід
1.	Виділення синтагми з тексту як речення або автономної частини речення	Текст	Список синтагм
2.	Перетворення синтагми у список слів, що ідентифікуються модулем словника	Окрема синтагма	Список відомих системі слів
3.	Побудова словника мовних образів обраної мови	Окреме слово	Ієрархія: слово → словоформа → лексема → мовний образ
4.	Використання парсеру обраної мови для отримання дерева підлеглості (залежностей)	Окрема синтагма	Дерево орієнтованого графа, вершини якого – МО, а ребра – питальні займенники
5.	Образна індексація предметної області	Корпус відібраних текстів предметної області	Семантична мережа АМО з образним індексом корпусу

Блок бази знань (5) в якості інформаційної основи використовує обґрунтоване у розділі 3 монографії сховище даних ІС. Складова базових функцій пошуку та оброблення знань, моделей пізнавальної діяльності СОПМК (6) повністю відповідає методам і моделям, розробленим у розділі 4. Користувачі (7) інформаційної технології мають можливість отримувати результати як базових, так і прикладних моделей для розв’язання відповідних до п. 1.5 семантико-залежних задач. З метою повної функціональної реалізації задач таблиці 4.2 на основі наукових результатів розділів 3 та 4 потрібно додатково побудувати такі прикладні моделі:

- отримання формальних ознак результатів пізнавальної діяльності,
- конструювання образу розв’язування проблемної ситуації,
- генерація множини повідомлень щодо стану та потреб ІС,
- побудова різних типів відповідей на питання.

Отже, кінцеві результати запропонованої інформаційної технології повністю визначає система формальних обмежень, що сформульована на 5-ти рівнях – від концептуальної ідеї та формальної метамови до практичної реалізації з урахуванням всіх отриманих наукових результатів. Розроблені схеми інформаційної технології образної обробки

ПМК забезпечують внесення попередньо відібраного текстового матеріалу в напівавтоматичному або автоматичному режимах, що імітує природний шлях накопичення знань людиною, побудову бази знань у вигляді семантичної мережі АМО та розв'язання низки актуальних семантико-залежних задач на основі комплексу функцій СОПМК.

5.2 Формалізація результатів пізнавальної діяльності

Значна частина задач електронного навчання передбачає організацію пізнавальної діяльності на основі природно-мовного електронного контенту. Визначимо формальні ознаки такої діяльності як кількісні оцінки образного сенсу складових ПМК у межах запропонованого підходу. Для розв'язання поставленої задачі будемо розглядати типовий навчальний процес як декомпозицію чотирьох основних складових: змістовної (об'єкт та предмет навчання), цільової (мета, цілі та завдання навчального процесу), операційної (етапи пізнавальної діяльності) та технологічної (інструментарій пізнавальної діяльності) [4].

Змістова складова визначає структуроване наповнення навчального матеріалу у вигляді такої ієрархії – навчальний план спеціальності, дисципліни, модулі (кредити), розділи (теми) та навчальні дози. Під останніми будемо розуміти взаємопов'язану сукупність елементів електронного контенту, що пропонується для засвоєння студентам та яку можна вважати змістовною одиницею дисципліни. На рисунку 5.4 представлено приклад змістовної структури типової дисципліни.

Такий чотирирівневий організації навчального матеріалу відповідає нумерація вузлів графа на верхніх рівнях дисциплін–модулів–розділів, яка продовжується на базовому рівні навчальних доз, зображених у вигляді прямокутників. Кількість зв'язків від верхнього вузла ієрархії до підлеглих йому нижніх вузлів може бути довільною, в залежності від визначеного викладачем змісту навчального матеріалу та складу і кількості навчальних задач.

Цільова складова характеризує мету та основні завдання дисципліни, тобто кінцеві результати, що мають бути досягнуті навчальним процесом. Як правило, цільова складова формулюється у вигляді знань, вмінь та навичок, які повинен продемонструвати студент після закінчення навчального процесу. З формальної точки зору цільова складова безпосередньо пов'язана з оцінкою кінцевих результатів функціонування навчальної системи. Тому граф змістовних складових

структури електронного підручника (рисунок 5.4) фактично відповідає ієрархії знань, вмінь та навичок, причому їх кількість на кожному вузлі графа також може бути довільною. Проте знання та навички вищих рівнів, як правило, не є арифметичною сумою ієрархічно підлеглих цільових складових, а являють собою їх певне узагальнення.

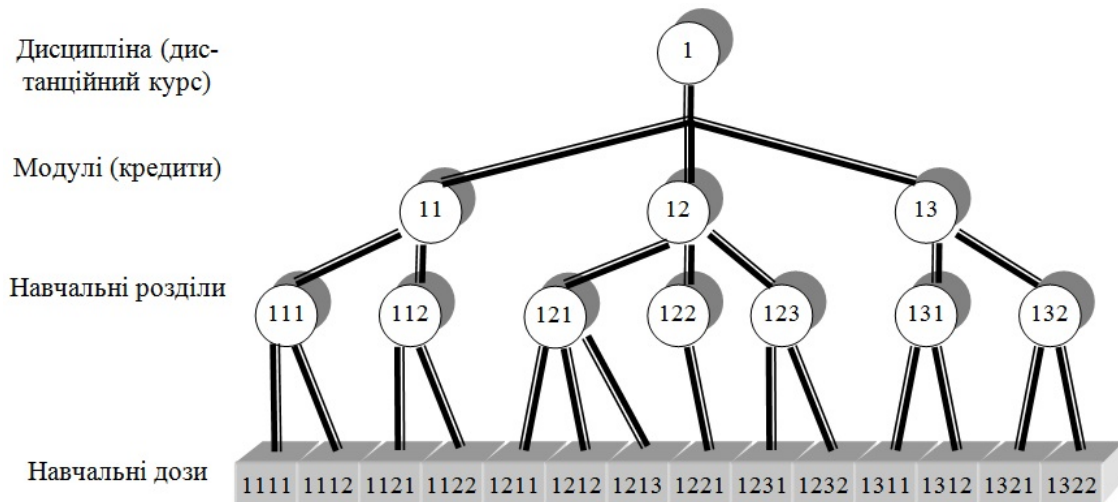


Рисунок 5.4 – Чотирирівнева структура навчального матеріалу дисципліни

Операційна складова безпосередньо пов'язана з дидактикою навчального процесу та відповідає на питання – як знання, вміння та навички передаються від викладача до студентів. Іншими словами, операційна складова у причинно-наслідковому взаємозв'язку розкриває етапи пізнавального процесу, до яких будемо відносити такі групи пізнавальних операцій:

- представлення та усвідомлення нової інформації;
- сприйняття та осмислення нової інформації;
- закріплення та застосування нових знань, вмінь та навичок;
- контроль та оцінка нових знань, вмінь та навичок.

У традиційних технологіях навчального процесу для забезпечення етапів операційної складової застосовуються усталені форми навантаження викладачів. В умовах навчального процесу за відсутності викладача основним засобом впливу на студента стає кадр електронного підручника, який за своєю роллю безпосередньо пов'язаний з поняттями навчальної дози та елементів електронного контенту [7]. Тому в

технологічній складовій у відповідності до окреслених раніше етапів пізнавальної діяльності будемо розрізняти:

- основні інформаційні кадри, додаткові інформаційні кадри, кадри психологічної підтримки студента, структурні (сценарні) кадри;
- кадри закріплення, кадри самоконтролю;
- кадри з тренувальними тестами, консультативні кадри;
- кадри з контрольними тестами, кадри оцінки.

Електронний підручник повинен забезпечувати засвоєння змістовного матеріалу власними специфічними засобами (вище перерахованими кадрами ЕК) в кожній навчальній дозі, причому як кадри, так і результати пізнавальної діяльності (знання/вміння/навички) сформульовані у вигляді ПМК. Оскільки представлена в п. 5.1 інформаційна технологія забезпечує образну індексацію всіх складових контенту, з'являється можливість інтерпретувати засоби та цілі пізнавальної діяльності. Вважатимемо, що знанням відповідає сенс-властивість, тобто підграф загальної АМО, а вмінням – кількісні оцінки певних зв'язків між мовними образами.

Проведемо оцінку образного сенсу КМО та ПМК, враховуючи, що згідно з (3.1), має бути справедливим співвідношення $0 \leq Se_{ij} < 1$. Відповідно до п. 2.4 значення введеної одиниці сенсу Saw характеризує вагу асоціативного зв'язку між парою образів з урахуванням напрямку асоціації. Покажемо, що запропонована функція сенсу Se_{ij} дозволяє оцінити сенс окремого МО чи декількох пов'язаних у КМО образів, наприклад, дати певні кількісні характеристики для таких відомих понять, як «ключовий образ» або «значима подія». Так, будь-який i -й образ АМО характеризується двома головними топологічними параметрами:

- n_i' – кількістю вхідних асоціативних зв'язків;
- n_i'' – кількістю вихідних асоціативних зв'язків.

Тоді, знаючи сенс Se_{ij} для кожного з таких асоціативних зв'язків, можна отримати середню оцінку «сенсу визначення» Se_i' та «вихідного сенсу» Se_i'' i -го образу

$$Se_i' = \sum_{j=1}^{n_i'} Se_{ji}$$

$$Se_i'' = \sum_{j=1}^{n_i} Se_{ij}'' .$$

Зрозуміло, що за визначенням одиниці сенсу вірними є співвідношення

$$0 \leq Se_i' < n_i', \quad 0 \leq Se_i'' < n_i'' .$$

Для моделювання окремих задач мовленнєвої діяльності визначимо оцінку загального сенсу i -го образу у вигляді

$$Se_i = Se_i' + Se_i'' \quad (5.1)$$

та значення загальної кількості асоціативних зв'язків i -го образу $n_i = n_i' + n_i''$. Тоді значення \bar{n}_i топологічного параметра – середньої кількості асоціативних зв'язків мовного образу в АМО – визначається як середнє арифметичне

$$\bar{n}_i = \frac{1}{n} \sum_{i=1}^n n_i .$$

Отже, сенс i -го образу обмежується топологічним параметром $0 \leq Se_i < n_i$, причому середнє значення сенсу \bar{Se}_i по кожному з n образів наближається до $\bar{Se}_i \approx n_i / 2$. Розглянуту сукупність оцінок сенсу і топологічних характеристик АМО застосуємо для побудови лінійки складних якісних критеріїв-ознак. Наприклад, на основі відомих значень Se_i та \bar{n}_i такими ознаками в мовленнєвій діяльності можуть бути:

- маловідомий образ ($0 \leq Se_i < 0,1 \cdot \bar{n}_i$);
- простий образ ($0,1 \cdot \bar{n}_i \leq Se_i < 0,4 \cdot \bar{n}_i$);
- популярний образ ($0,4 \cdot \bar{n}_i \leq Se_i < 0,7 \cdot \bar{n}_i$);
- багатий образ ($0,7 \cdot \bar{n}_i \leq Se_i < \bar{n}_i$);
- базовий (ключовий) образ ($Se_i \geq \bar{n}_i$).

Розрахуємо значення вищенаведених показників для даних наскрізного прикладу з пп. 3.1, 3.6. Згідно з функцією (2.10) і даними таблиці 4.5 (п.4.4), маємо такі значення образного сенсу для $\bar{n}_i = 4,5$:

$k_{ij} = 1 \rightarrow Se_{ij} = 0,35$; $k_{ij} = 2 \rightarrow Se_{ij} = 0,59$; $k_{ij} = 3 \rightarrow Se_{ij} = 0,80$. Для топологічно розвинутого МО «мова (56)» отримуємо формальні параметри $n_i' = 4$, $n_i'' = 5$, $n_i = 9$, $Se_i' = 2,328$, $Se_i'' = 3,126$, $Se_i = 5,454$, які підтверджують його статус ключового образу елемента ЕК.

Подією, згідно з формальною теорією (п. 2.3), можна вважати послідовний або розгалужений ланцюг (орієнтований граф) образів [31]. Основою оцінки сенсу події пропонується визнати сенс асоціативної пари образів, оскільки будь-яка КМО події розкладається на асоціативні пари. Визначення окремої КМО як дерева $T = \langle V_T, E_T \rangle$, $V_T \subseteq V$, $E_T \subseteq E$ на основі рекурсивної процедури через асоціативну пару $(a, b) \in \Omega$ отримуємо згідно з п. 3.1:

а) T є пара $(a, b) \in E$, тобто $V_T = \{a, b\}$, $E_T = (a, b)$;

б) якщо T – це КМО, то $T \oplus (a, b)$ також є КМО тоді і тільки тоді, коли $a, b \in V$, $(a \vee b) \in V_T$, $(a, b) \notin E_T$, де \oplus – операція об'єднання «PLUS OK»;

в) ніяких інших, окрім а) та б) КМО не існує.

Логічно припустити, що загальний образний сенс (i, j) пари образів без урахування напрямку зв'язку визначається як

$$Se_{ij}^p = Se_i - Se_{ij} - Se_{ji} + Se_j. \quad (5.2)$$

Тим самим, у сенс пари МО закладається сума двоспрямованого асоціативного зв'язку з усіма іншими (вхідними та вихідними) зв'язками двох образів, що складають цю пару. Відповідно до цього та з урахуванням деревоподібної структури синтагми образний сенс її визначається як сума сенсу всіх пар, що складають подію за винятком сенсу тих МО, що беруть участь у двох, трьох та більше парах

$$Se_i^e = \sum_{j=1}^{n_i^e} Se_{(i)j}^p - \sum_{j=1}^{n_i^2} Se_j - \sum_{j=1}^{n_i^3} Se_j - \dots, \quad (5.3)$$

де Se_i^e – образний сенс i -ї події; $Se_{(i)j}^p$ – образний сенс j -ї пари i -ї події; n_i^e – кількість асоціативних пар i -ї події; n_i^2 – кількість образів, що беруть участь у двох парах i -ї події; n_i^3 – кількість образів, що беруть участь у трьох парах i -ї події.

За даними наскрізного прикладу для пари «мова – XML» отримуємо значення образного сенсу $Se_{ij}^p = 5,454 - 0,8 + 1,15 = 5,804$. Синтагму «мова XML дозволяє визначити синтаксис [i] структуру документів» згідно з моделлю формальної теорії (п. 2.3) можна представити у вигляді 6-ти пар

мова \ XML \oplus
 мова \ дозволяє \oplus
 дозволяє \ визначити \oplus
 визначити \ синтаксис \oplus
 визначити \ структуру \oplus
 структуру \ документів.

Тоді, з урахуванням всіх даних Додатку Б, значення образного сенсу для цієї синтагми $Se_i^e \geq 5,8 + 5,454 + 3,4 = 14,65$. Як і для попередньо розглянутих критеріїв, кількість образного сенсу асоціативної пари та події має певні межі

$$0 \leq Se_{ij}^p < n_i + n_j,$$

$$0 \leq Se_i^e < \sum_{j=1}^{n_i^e} Se_{(i)j}^p.$$

При інтерпретації пізнавальної, у тому числі і мовленнєвої діяльності оцінка сенсу i -ї події Se_i^e також дозволяє поставити у відповідність чисельному параметру множину якісних критеріїв-ознак, наприклад:

- малоймовірна подія ($0 \leq Se_i^e < 0,2 \cdot n_i^e \cdot \bar{n}_i$);
- незначна подія ($0,2 \cdot n_i^e \cdot \bar{n}_i \leq Se_i^e < 0,4 \cdot n_i^e \cdot \bar{n}_i$);
- типова подія ($0,4 \cdot n_i^e \cdot \bar{n}_i \leq Se_i^e < 0,7 \cdot n_i^e \cdot \bar{n}_i$);
- гучна (відома) подія ($0,7 \cdot n_i^e \cdot \bar{n}_i \leq Se_i^e < n_i^e \cdot \bar{n}_i$);
- значима (ключова) подія ($Se_i^e \geq n_i^e \cdot \bar{n}_i$).

Оцінка образного сенсу $14,65 \geq 0,4 \cdot 6 \cdot 4,5 = 10,8$ синтагми вищеведеного прикладу свідчить, що в ній закладено опис типової події цього елемента ЕК.

Отже, з метою застосування інформаційної технології у семантико-залежних задачах електронного навчання розглянуто змістовну, ці-

льову, операційну та технологічну складові типового навчального процесу. Запропоновано в якості формальних ознак результатів пізнавальної діяльності вважати: для знань – топологічні характеристики фрагментів АМО, а для вмій – чисельні оцінки образного сенсу окремих зв'язків, що забезпечує їх порівняння з елементами ЕК. Отримано та перевірено на даних наскрізного прикладу оцінки образного сенсу для головних конструктів природно-мовного ЕК – образу (МО), асоціативної пари (сенссполучення) та події (синтагми).

5.3 Конструювання образу розв'язування проблемної ситуації

Беззаперечні успіхи в ефективному керуванні надскладними технічними системами не змінюють статус класичної теорії управління як формалізованого розв'язку окремої задачі з природної множини задач асоціативного образного мислення. Парадокс полягає у тому, що принципіві відхилення від формальної схеми відразу приводять до неієздатності моделі управління, а образне мислення людини без жодних складних розрахунків знаходить прийнятний вихід з проблемної ситуації.

Як правило, з метою розв'язання нештатних виробничих ситуацій в контур управління вводять людину (людино-машинні системи [81, 263]) або інтелектуальних агентів (багатоагентні системи [264]). В таких моделях інтелектуальна складова суттєво обмежена з функціональної точки зору, а надійність та безвідмовність системи мають імовірнісний характер [265].

Моделювання логічного мислення людини та природних баз знань забезпечують на основі використання експертних систем, СППР, САПР [84]. Багатокритеріальний вибір рішення за допомогою експертних оцінок можна вважати підкласом такого напряму [266, 267]. Практична реалізація цього підходу показала його доцільність у вузьких предметних областях і значну складність при спробах підвищити ступень універсальності моделі.

Предметом формалізації будемо вважати певні мнемонічні процеси, властиві людині. Хоча моделювання асоціативного образного мислення як природного способу отримання рішень для управління має на даний час, в основному, концептуальний характер [116], у роботі [29] з цією метою запропоновано застосувати алгебраїчні моделі образного механізму оперативної пам'яті СОПМК. Актуальною задачею у

розвиток моделі з п. 4.5 є визначення та деталізація образу-рішення проблемної ситуації шляхом використання групи формальних операцій до ансамблю образів та моделювання орієнтувального рефлексу людини. Вперше алгебраїчний розв'язок цієї задачі було отримано у [24].

До природних задач оперативної пам'яті можна віднести узагальнення певної інформації в єдине ціле. В термінах образного мислення базова операція синтезу забезпечує за Виготським [181] вплив (влиття) сенсу синтагми *Bi-Sy* або АО *Bi-OM* в один образ, наприклад, (рос.) «пшеница, которая в темном чулане хранится в доме, который построил Джек».

Покажемо принципову можливість конструювання образу розв'язування у межах моделі механізму оперативної пам'яті [29]: якщо попередньо *N* образів-претендентів вже відібрано у стек пам'яті *Check-Set*, необхідно остаточно визначити складові вектора емоцій *Vector-Set* та вибрати один образ з *N*. Тоді формальною постановкою задачі конструювання образу розв'язування вважатимемо

$$\begin{aligned} & Check - Set, Vector - Set \rightarrow Focus - Bi | \\ & Focus - Weight = \underset{i \in N}{Max}(Weight_i), \end{aligned} \quad (5.4)$$

де *Focus-Bi* – бінарний код образу у фокусі уваги; *Focus-Weight* – вага асоціативних зв'язків образу у фокусі уваги з образами-складовими вектора емоцій *Vector-Set*; *Weight_i* – вага асоціативних зв'язків *i*-го образу АО з образами *Vector-Set*.

Формально синтез або конструювання образу розв'язування можна представити у вигляді оператора *InsZX*, що відповідає типу № 13 образного пошуку узагальненої класифікації з п. 4.1. Головними джерелами інформації для моделі є такі події, які приводять до виникнення проблемної ситуації та характеризують її. Згідно з п. 5.1, передбачається попереднє внесення в пам'ять системи управління текстів з описом відповідних подій на основі спільного словника МО, що забезпечує накопичення елементарних (синтагматичних) асоціативних зв'язків між образами.

Візьмемо за основу такий підхід пошуку за типом стек-АО-образ (п. 4.5): попередньо відібрані у стек потенційні образи-розв'язки переносяться в АО, потім образи-складові вектора емоцій міняються місцями з образами АО з метою деталізації кінцевого розв'язку, а далі запус-

кається модель механізму оперативної пам'яті (п. 4.5.2) з можливим залученням інсайтного пошуку *Find-I* узагальненої класифікації (п. 4.2).

Використаємо згідно з [24] такі позначення:

$Bi-Vector_i$ – бінарний код i -го образу вектора емоцій *Vector-Set*;

$Bi-OM$ – бінарний код АО оперативної пам'яті.

Додамо до алгебраїчної системи *BAS* (п. 3.3) формальні операції і предикати на булеані, які відповідають поставленій задачі:

1. Операція пошуку i -ї одиниці *Seek“1”(i)* у бінарному коді $Bi-OM$ та предикат *End-Bi*, який істинний тоді, коли ще не переглянуто всі розряди бінарного коду

$$Bi - OM \xrightarrow{Seek^{1''(i)}} k, \quad (5.5)$$

$$i \leq n \rightarrow End - Bi, \quad (5.6)$$

де i – номер образу в АО оперативної пам'яті; k – номер одиничного розряду бінарного коду, що відповідає i -му образу АО; n – кількість одиниць у бінарному коді (поточний обсяг оперативної пам'яті).

2. Операція *New-OM* перенесення образів-складових вектора емоцій *Vector-Set* в АО оперативної пам'яті:

$$\bigcup_{i=1}^n Bi - Vector_i \xrightarrow{New-OM} Bi - OM. \quad (5.7)$$

3. Операція *New-Vector* занесення до вектора емоцій *Vector-Set* значень зі стеку *Choice-Set*:

$$Choice - Set \xrightarrow{New-Vector} Vector - Set. \quad (5.8)$$

4. Операція *New-Choice* занесення до стеку *Choice-Set* образів з АО оперативної пам'яті:

$$Bi - OM \xrightarrow{New-Choice} Choice - Set. \quad (5.9)$$

У граф-схемах для алгебраїчних конструктів задіємо такі позначення операторів структурного програмування [257]:

Do – цикл за параметром або за умовою;

*** – композиція.

Розглянемо розв'язання задачі (5.4) на основі таких складових операторів.

1. Оператор *Pyramid-Images* декомпозиції бінарного коду *Bi-OM* на коди образів-складових. У результаті дії оператора, який представлено на рисунку 5.5, за відомим кодом *Bi-OM* визначаються коди образів, що входять до складу АО.

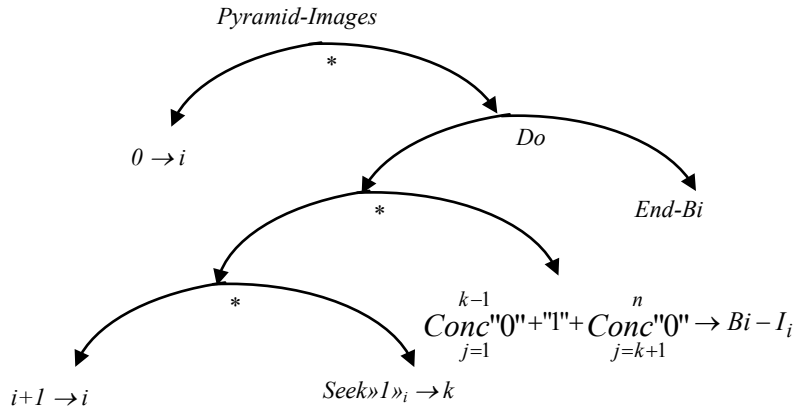


Рисунок 5.5 – Граф-схема оператора *Pyramid-Images*

З цією метою організовується цикл по довжині коду *Bi-OM*, якщо знаходиться одиниця, то до списку образів додається ще один елемент з одиницею саме на цій позиції та нулями на всіх інших (цикл обмежується кількістю одиниць *n* в бінарному коді *Bi-OM*):

$$\begin{aligned}
 \text{Pyramid - Images} ::= & 0 \rightarrow i * \{ [\text{End} - \text{Bi}] \} (i + 1 \rightarrow i * \\
 & \text{Seek} "1" _i \rightarrow k * \text{Conc} "0" + "1" + \text{Conc} "0" \rightarrow \text{Bi} - I_i \} , \quad (5.10)
 \end{aligned}$$

де $Bi - I_i$ – бінарний код *i*-го образу в АО оперативної пам'яті; *Conc "0"* – конкатенація символів «0» в розряди бінарного коду.

2. Оператор *OM-Change-Vector* (рисунку 5.6) моделює запуск механізму конструювання образу розв'язування – послідовно виконуються операції *New-Choice*, *New-OM* та *New-Vector*:

$$\begin{aligned}
 \text{New - Change - Vector} ::= & \text{New - Choice} * \\
 & \text{New - OM} * \text{New - Vector} . \quad (5.11)
 \end{aligned}$$

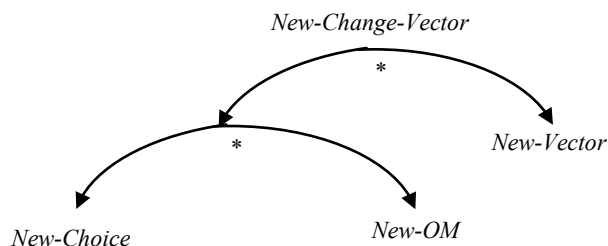


Рисунок 5.6 – Граф-схема оператора *OM-Change-Vector*

3. Оператор конструювання образу розв'язування *InsZX* (рисунок 5.7) запускає відповідний механізм *OM-Change-Vector* та подає його результати на вхід моделі орієнтувального рефлексу *Orient-Reflect* з п. 4.5.3:

$$InsertZX ::= OM - Change - Vector * Orient - Reflect . \quad (5.12)$$

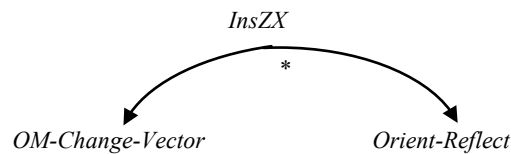


Рисунок 5.7 – Граф-схема оператора *InsZX*

Алгоритм конструювання образу розв'язування проблемної ситуації реалізовано у межах моделі орієнтувального рефлексу [29] за допомогою алгебраїчних операторів (5.5)–(5.12), а також програмно реалізовано та апробовано на основі технології Python + SQLite.

Проілюструємо можливості запропонованого підходу шляхом розв'язання задачі (5.4) для експериментальних даних наскрізного тестового прикладу пп. 3.1, 3.6. На відміну від прикладу п. 4.5.3, відберемо до стеку пам'яті *Check-Set* 5 мовних образів – «онтологія (111)», «XML (88)», «засіб (93)», «ресурс (73)», «RDF (87)», а складовими вектора емоцій *Vector-Set* встановимо МО «мова (56)», «інформація (34)» та «мережа (17)».

Формально, у відповідності до постановки задачі 5.4, визначаємо $Check-Set=011111000000$, $Vector-Set=\{000000010000, 000000000010, 000000000001\}$. В результаті дії оператора *OM-Change-Vector* мовні образи *Vector-Set* переносяться в АО, а на їх місце потрапляють МО з *Check-Set*. Оператор орієнтувального рефлексу *Orient-Reflect* послідовно визначає для всіх МО з оперативної пам'яті ($Bi-OM=000000010110$) значення ваги $Weight_i$, що зрештою приводить до знаходження найбільшого з них – $Focus-Weight=8$ для МО «мова (56)». Проте МО «мережа (17)» навіть на обмеженому природномовному матеріалі елемента ЕК наскрізного прикладу набирає вагу 4, що робить його конкурентом МО «мова (56)» при узагальненні значимих понять теми «Стандарти Semantic WEB».

Отже, на основі формального погляду на феномени ансамблю образів, вектора емоцій, фокусу уваги та орієнтувального рефлексу по-

казано можливість конструювання образу розв'язування проблемної ситуації засобами алгебраїчної системи *BAS*. Основним джерелом інформації для моделі служить синтагматичний опис подій, які, власне, приводять до проблемної ситуації та визначають її особливості. Запропонований підхід до побудови оператора *InsZX* базується на вже відомих складових моделі оперативної пам'яті СОПМК (п. 4.5) та забезпечує інваріантне представлення змісту КМО, що продемонстровано на даних наскрізного тестового прикладу.

5.4 Генерація повідомлень щодо стану та потреб системи

Одним з проявів інтелектуальної поведінки вважається здатність системи передати назовні інформацію про свій внутрішній стан та актуальні на даний час потреби. Такою здатністю володіє не тільки людина та створені нею штучні системи, але й представники вищого тваринного світу, що ведуть, переважно, колективний спосіб життя [254]. Докорінною відмінністю між означеними двома класами систем виступає характер передачі повідомлення – штучний або символно-сигнальний у першому випадку та інстинктивний або природно-сигнальний у другому. Відомим ще за часів Ч. Дарвіна доказом тваринного походження людини є підсвідоме володіння та розуміння нею «мови рухів тіла», яка дуже часто дозволяє набагато глибше зрозуміти стан та бажання співбесідника, ніж зміст сказаних ним слів [268].

Відомі штучні системи генерують виключно символні повідомлення щодо стану власного функціонування переважно у таких випадках [91, 269]:

- при невірних діях користувача (у більшості систем);
- для контролю значимих параметрів системи (операційні системи, бази даних та знань, охоронні системи);
- з метою пояснення отриманих результатів (експертні системи, автоматичне доведення теорем);
- для отримання проміжних даних у процесі налагодження алгоритмів та програмних продуктів (системи програмування);
- з метою сигналізації про небезпечні ситуації (банківські та охоронні системи, підтримка послуг зв'язку, комп'ютерні ігри тощо).

Зрозуміло, що тільки в першому та останньому випадку мова йде про обмежену внутрішню потребу системи повідомити назовні інформацію щодо появи проблем при виконанні чітко визначених задале-

гідь функцій. У літературі відсутні формальні постановки задач про генерацію таких повідомлень відносно однієї з можливих потреб відкритої системи, які відповідають комплексним критеріям функціонування цієї системи. Підхід до формалізації образного сенсу ПМК, що пропонується, дозволяє розв'язати таку задачу і, тим самим, отримати модель адекватного функціонування СОПМК.

Розглянемо ІС, яка обмінюється інформацією з зовнішнім світом за допомогою потоків образів. Як було показано у п. 2.1 внутрішні критерії функціонування системи, що характеризують принцип мінімізації енергетичних витрат теж мають образну природу, оскільки залежать від попередньої історії накопичення сполучень зовнішніх образів у вигляді ланцюга подій. Якщо представити емоції згідно з п. 4.5 у вигляді найбільш близьких до них образів, то постановка задачі полягає у формалізації процедур для таких повідомлень системи:

- 1) визначення «емоційного стану» ІС у вигляді ансамблю образів;
- 2) необхідність зовнішньої верифікації отриманого системою розв'язку однієї з семантико-залежних задач, що належить до моделі пізнавальної діяльності;
- 3) отримання вимог системи до зовнішнього оточення, наприклад у випадку застосування когнітивного критерію система має вказати на те, що незрозуміло, які образи ще вимагають додаткового визначення;
- 4) сигналізація про «нештатні ситуації» у функціонуванні системи.

Вирішимо поставлені задачі генерації 4-х типів повідомлень у розвиток запропонованого підходу до визначення сенсу КМО та ПМК шляхом моделювання образної обробки природно-мовного контенту. З цією метою будемо використовувати можливості математичного апарата, який розроблено для реалізації концептуальних основ образного сенсу з урахуванням особливостей архітектури бази знань СОПМК.

1. Для визначення «емоційного стану» системи будемо спиратися на модель механізму оперативної пам'яті (п. 4.5), згідно з якою в інфологічній системі у кожний такт часу в блоці 5 структурної моделі (п. 2.6) відбувається оновлення ансамблю образів. В цей же такт часу образи з оперативної пам'яті потрапляють в блок емоцій 1б), задача якого полягає у спрямуванні фокусу уваги на один з цих образів на основі оцінки зв'язків образів-складових вектора емоцій. Формально це означає визначення бінарного коду образу у фокусі уваги *Focus–Vi*

та вагу образу у фокусі уваги *Focus-Weight*, яка є найбільшою з усіх образів АО.

Для розв'язання поставленої задачі до складу вектора емоцій *Vector-Set* включимо найбільш емоційно забарвлені образи. Формально це забезпечується відбором 5–9 вершин графа АМО з найбільшою кількістю сенсу Se_i згідно з (5.1) – для даних обмеженої кількості МО наскрізного прикладу (таблиця 4.5) такими є «мова (56)» та «дозвіл (123)». Тоді вектор емоцій кожного образу АО визначається на основі операції *Evaluate-Vector* (4.45), внаслідок чого отримуємо формальні ознаки фокуса уваги *Focus-Bi* та *Focus-Weight*. Зрештою, повідомленням про «емоційний стан» системи будемо вважати функцію F_e – формальну залежність зміни фокуса уваги *Focus-Bi* від модельного часу t

$$Focus - Bi = F_e(t).$$

Запропонований підхід не виключає неперервне відображення назовні всього складу ансамблю образів оперативної пам'яті, яке буде мати вигляд опису (протоколу) внутрішніх подій, що відбувається в інфологічній системі просто зараз.

2. Поняття інфологічної системи (п. 2.1) передбачає необхідність зовнішньої верифікації або «підкріплення» за І. П. Павловим отриманого системою розв'язку однієї з задач пізнавальної діяльності (п. 4.6). Будемо вважати, що результатом такого розв'язку є поява одного нового парадигматичного зв'язку з множини (3.6), а формальною ознакою – збільшення кількості сенсу в системі, наприклад, для синонімічного зв'язку згідно з (4.71). Зрозуміло, що задачі пізнавальної діяльності можуть мати складну ієрархічну структуру [20], проте за елементарну складову пізнавальної діяльності приймемо створення нового парадигматичного зв'язку (п. 4.6).

Отже, мета генерації повідомлення про верифікацію розв'язку задачі пізнавальної діяльності полягає у підтвердженні істинності нового парадигматичного зв'язку. Лишивши осторонь планування та ініціалізацію задач пізнавальної діяльності формально таке повідомлення для синонімії представимо за допомогою

$$\exists Pc_{ij} \mid Pc_{ij} \in Pc \rightarrow If - Synonim ,$$

$$Se^s + Se_{ij}^{Pc} \xrightarrow{Synonim} Se^s ,$$

де *If-Synonym* – предикат, що вимагає зовнішнього підтвердження; *Synonym* – функція збільшення загального сенсу системи внаслідок додавання синонімічного зв'язку.

Задача зводиться до виконання оператора *Synonym-Plus*, що збільшує сенс інфологічної системи внаслідок верифікації синонімічного зв'язку – якщо отримано зовнішнє підтвердження синонімії (предикат *If-Synonym* істинний), то виконується оператор *Synonym* :

$$Synonym - Plus ::= ([If - Synonym] Synonym, E), \quad (5.13)$$

де *E* – пустий оператор. Граф-схему оператора *Synonym-Plus* представлено на рисунку 5.8:

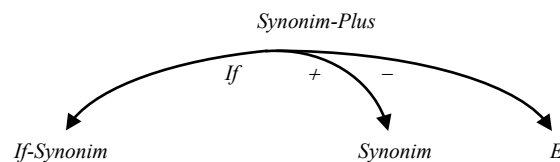


Рисунок 5.8 – Граф-схема оператора *Synonym-Plus*

Запропонований підхід може бути застосований в процесі діалогу з ІС (п. 2.1), під час якого відбувається природна ініціалізація задач пізнавальної діяльності та образний пошук виду *InsZY*. В такому випадку предикати, подібні до *If-Synonym*, можна вважати істинними за замовчуванням до тих пір, доки «вчитель», як інша сторона діалогу, не «виправить» помилку системи у розв'язку відповідної часткової задачі пізнавальної діяльності. Так, у прикладі знайденої антонімії для МО «довільно (138)» та «необхідно (159)» з п. 4.6, у результаті діалогу аналогічний предикат *If-Antonim* має отримати значення *False*.

3. На відміну від повідомлень про «емоційний стан» інфологічна система здатна генерувати вимоги до зовнішнього оточення внаслідок неперервного аналізу стану модулі внутрішнього когнітивного простору асоціативних пар у вигляді графа G_z (п. 3.5). При цьому може бути застосовано один з критеріїв функціонування системи (п. 2.1), наприклад, максимальне збільшення загальної кількості сенсу Se^s інфологічної системи як модель домінуючого мотиву «чистої» інтелектуальної діяльності.

Будемо в подальшому вважати когнітивним критерієм загальну кількість сенсу Se^s інфологічної системи. Тоді загальну задачу пізнавальної діяльності інтерпретуємо у вигляді

$$Se^s \rightarrow \max |G_z|. \quad (5.14)$$

Для розв'язання задачі (5.14) інфологічна система може ініціювати діалог з зовнішнім «вчителем», наприклад, вказати на те, які образи (вершини графа G_z) вимагають додаткового визначення. З цією метою пропонується провести кластеризацію вершин графа G_z за топологічними параметрами n'_i і n''_i та значенням сенсу Se_i згідно з (5.1). Розбиття вершин на кластери можна провести на основі одного з відомих методів [82, 267, 270], або за допомогою спеціалізованого алгоритму, який додатково ініціює діалог з зовнішнім «вчителем».

Розглянемо алгоритм *Question–Out*, для роботи якого попередньо визначаються середні значення \overline{n}_i та \overline{Se}_i :

[Крок 1] $n'_{\min} = \overline{n}_i / 2$, $n''_{\min} = \overline{n}_i / 2$, $Se_{\min} = \overline{Se}_i$, $Minimum - Bi = 1$.

[Крок 2] Відкрито цикл по кожній i -й вершині графа G_z .

[Крок 3] Якщо $(n'_i < n'_{\min}) \vee (n''_i < n''_{\min}) \vee (Se_i < Se_{\min})$, то $n'_{\min} = n'_i$, $n''_{\min} = n''_i$, $Se_{\min} = Se_i$, $Minimum - Bi = i$.

[Крок 4] Цикл за індексом i закрито.

[Крок 5] Виведення (як об'єкта запитання) образу з номером *Minimum-Bi*.

Для даних наскрізного прикладу пп. 3.1, 3.6 було визначено, що $\overline{n}_i = 4,5$ та $\overline{Se}_i = 0,497$, а застосування алгоритму *Question–Out* дозволило отримати МО з $Se_{\min} = 1,15$ та $Minimum - Bi = 9$, а саме – «XML (88)». Реакція зовнішнього «вчителя» па подібні запити передбачається у вигляді додаткового внесення в ІС множини нових синтагм, до складу яких входить образ *Minimum–Bi*.

Будемо виходити з припущення про те, що кількість нових зв'язків для підтримки «найслабшого» образу буде максимальною. У відповідності до п. 4.6 внесення в систему кожного нового синтагматичного зв'язку збільшує сенс-параметр системи

$$Se^s \approx Se^s + \delta + 0,5, \quad (5.15)$$

що і є кінцевою метою алгоритму *Question–Out* та образного пошуку виду *InsZX*. Отже, поява нових синтагм з МО «XML (88)» у елементі ЕК з розглянутого прикладу призведе до найшвидшого зростання загального сенсу Se^s ІС.

4. Розбиття вершин графа G_z на кластери забезпечує можливість генерування сигналів назовні щодо різного роду «нештатних ситуацій» в роботі, що залежать від специфіки та предметної області функціонування інфологічної системи. Якщо не враховувати динамічні характеристики роботи системи в режимі реального часу, то під «нештатною ситуацією» у вузькому сенсі будемо розуміти випадок визначення нового кластера на основі сукупності інформації про граф G_z , що змінилася за певний проміжок часу.

У результаті виникнення «нештатної ситуації», наприклад, в мовленнєвій діяльності з'являється можливість надання імені новому кластеру, що приведе до появи нового парадигматичного зв'язку типу окреме–загальне P_c між всіма образами кластера та збільшення загального сенсу системи аналогічно (4.71) згідно з (5.14). Отже, задача формально зводиться до:

- виявлення факту появи нового кластера образів;
- повідомлення зовнішнього «вчителя» щодо складу нового кластера;
- отримання у відповідь номера вершини графа G_z , відповідний до якої образ визначається назвою нового кластера.

Незалежно від того, чи входить образ-назва кластера до складу кластера чи ні, при такому підході кількість доданого сенсу інфологічної системи залежить від кількості об'єднаних таким чином синтагматичних зв'язків. Характерним з цієї точки зору є приклад п. 4.6 щодо визначення зв'язку типу окреме–загальне, результати якого наведені у Додатку Д. Підтвердження появи нового кластера під назвою «мова» дозволить об'єднати сенс 12 синтагматичних зв'язків МО «XML», «RDF», «онтологія», «розмітка» та «метадані» згідно з (4.71), що відразу збільшить загальний сенс ІС до $Se^s \approx Se^s + 6$.

Отже, на основі запропонованого підходу до оцінки образного сенсу ПМК отримано формальні методи генерації 4-х типів повідомлень інфологічної системи за видами образного пошуку *InsZX* та *InsZY*.

Повідомлення про «емоційний стан» системи інтерпретовано як відображення функціональної залежності зміни фокусу уваги *Focus–Bi* від модельного часу. Повідомлення про розв’язок однієї з задач пізнавальної діяльності зводиться до оператора, що збільшує загальний сенс інфологічної системи внаслідок зовнішньої верифікації відповідного парадигматичного зв’язку. Генерування вимог до зовнішнього оточення реалізовано у вигляді алгоритму *Question–Out*, який на основі перебору вершин графа G_z знаходить «найслабший» образ *Minimum–Bi*, що відрізняється найменшими значеннями топологічних параметрів n_i' і n_i'' та значенням сенсу Se_i . Сигнал про «нештатну ситуацію» зведено до появи за участю «вчителя» парадигматичного зв’язку типу окреме–загальне між образами виявленого нового кластера та збільшення загальної кількості Se^s сенсу системи.

5.5 Побудова відповіді на питання в процесі діалогу

Підтримка повноцінного діалогу штучної (інфологічної) системи з людиною передбачає генерацію відповідей системою на довільні питання людини. Умови тесту Тьюрінга вважаються універсальним критерієм інтелектуальності штучних систем, принаймні, на сучасному етапі [165, 274]. Під час генерації розглянутих у попередньому параграфі 4-х типів повідомлень щодо стану та/або мотивуючих факторів інфологічна система за власною ініціативою розпочинає діалог з зовнішнім «вчителем», причому з цілком визначеною кожний раз метою. Тому у межах підходу до підтримки розуміння образного сенсу ПМК, що пропонується, будемо розглядати такі можливі типи відповідей на питання в процесі діалогу:

I. Побудова відповіді на питання, якщо такою є слово (*PartXZ*) або частина синтагми (*PartYZ*), що вже внесені у сховище ІС:

а) пошук синтагм, в яких певні образи поєднані логічними операторами «АБО» чи «І» як удосконалений пошук за маскою слів у відповідності до критерію 4.1 (п. 1.6);

б) визначення слова або частини синтагми, що відповідають питальному займеннику речення-питання (*PartXY*) згідно з моделлю формальної теорії у вигляді комутативної напівгрупи ОК (п. 2.3).

II. Побудова відповіді шляхом моделювання механізму маніпулювання ансамблем образів оперативної пам’яті СОПМК (побудова за ти-

пом синтагма \rightarrow АО \rightarrow синтагма) засобами алгебраїчної системи *BAS* (*Mech*, *Or* · *Re*).

Розглянемо формальні методи реалізації діалогу такого типу без додаткового залучення модулів морфологічного та синтаксичного аналізу, які, у принципі, не грають вирішальної ролі для розуміння сенсу природно-мовних складових електронного контенту [31, 53]. Джерелом інформації для підтримки діалогу будемо вважати дані сховища інфологічної системи (п. 3.6).

I. Задамо питання системі у вигляді послідовності з k слів $\{Word_1, \dots, Word_k\}$. Без порушення спільності будемо вважати, що кожному i -му з цих k слів ставиться у відповідність код образу $Bi-I_i$ та всі вони потенційно можуть входити до складу кортежів відношення *Words* сховища ІС згідно з (3.58). Розглянемо два можливих випадки:

а) $\{Word_1, \dots, Word_k\} \cap Words = \emptyset$ – жодне з слів питання не було у складі синтагм, що внесені до сховища ІС. У цьому випадку за кодом відповідного образу $Bi-I_i$ перетворюємо послідовність слів у множину образів, для якої справедливо $\{Bi-I_1, \dots, Bi-I_k\} \subset Image$, де *Image* визначається згідно з (3.51). Оскільки всі синтагми сховища ІС також можна перетворити у множини образів, результат відповіді на питання зводиться до пошуку одного з двох типів:

- «АБО» – відбираються всі синтагми сховища, в яких присутній хоча б один образ $Bi-I_i$ з питання;
- «І» – відбираються тільки ті синтагми сховища, в яких присутні всі образи з питання.

Формально алгоритм пошуку «І» складається з таких кроків:

[Крок 1] Визначити коди образів для кожного слова питання та скласти з них множину *Images-List-X*.

[Крок 2] Аналогічно визначити множину *Images-List-i* для кожної i -ї синтагми з n синтагм, що зберігаються у сховищі ІС.

[Крок 3] Відбирати у список відповідей *Syntagma-List* ті i -ті синтагми, для яких

$$Images - List - X \subseteq Images - List - i. \quad (5.16)$$

Результати відповідей ІС на основі пошуку типу «АБО» та «І» для експериментальних даних наскрізного прикладу представлено у Додатку Д.

б) $\{Word_1, \dots, Word_k\} \cap Words \neq \emptyset$ – хоча б одне зі слів питання вже зустрічалось у складі синтагм, що внесені до сховища ІС. У цьому випадку відбирається множина потрібних синтагм, яку можна відсортувати за кількістю спільних з питанням слів, а відповідь зводиться до алгоритму обходу дерева відібраної синтагми.

Результатами цього типу пошуку принципово можуть бути: а) пустий список; б) слово; в) частина речення; г) множина синтагм *Syntagma-List*. Використовуючи особливості запропонованого підходу, алгоритм відповіді на питання згідно з правилом 2 формальної теорії (п. 2.3) представимо у вигляді таких кроків:

1. Якщо результат пошуку типу «І» *Syntagma-List* пустий, то виходом стає варіант а).

2. Інакше організувати цикл за елементами множини *Syntagma-List*:

2.1. Розкласти чергову синтагму на список пар *Pair-List*, які складають її.

2.2. Перевірити у циклі кожну пару з *Pair-List* з метою знаходження такої з них, в якій головний образ збігається з образом першого слова, а питальний займенник пари співпадає з питальним займенником *Pronoun*; якщо виконується тільки перша умова і тип зв'язку *link_id* у обох питальних займенників збігається, також вважати, що пару знайдено.

2.2.1. Якщо пару знайдено, то виконується перевірка: чи є підлеглий образ пари головним для будь-яких інших пар цієї ж синтагми:

2.2.1.1. Якщо ні, то підлеглий образ пари закласти як відповідь за варіантом б).

2.2.1.2. Якщо так, то відповідь за варіантом в) формувати за допомогою рекурсивного алгоритму, що «витягає» з синтагми піддерево образів, яке є підлеглим підлеглому образу пари.

3. Якщо образ першого слова жодного разу не був головним в жодній з пар, то відповідь будується за варіантом г).

Результати відповідей ІС на основі пошуку з питальним займенником для експериментальних даних наскрізного прикладу представлено у Додатку Д.

II. Побудова відповіді також можлива шляхом моделювання ансамблю образів оперативної пам'яті людини, при цьому неможливість чітко структурувати відповідь робить її схожою на пророцтва дельфійського оракула, які ще потрібно тлумачити. Для цього випадку

пропонується алгоритм, що використовує підхід до моделювання механізму оперативної пам'яті з п. 4.5:

[Крок 1] Синтагма-питання перетворюється в ансамбль образів з $\{Bi - I_1, \dots, Bi - I_k\} \subset Image$.

[Крок 2] Відбувається обмін образів-складових вектора емоцій з образами АО,

[Крок 3] Запускається алгоритм 1) з п. 5.3, а повідомлення системи вважається її відповіддю на поставлене питання.

Результати відповідей інфологічної системи на основі цього алгоритму для окремих ПМК представлено у п. 5.6, а для експериментальних даних наскрізного прикладу – у Додатку Д.

Отже, в межах запропонованого підходу створено формальні процедури для побудови таких типів відповідей на питання до ІС у відповідності до *PartXY*, *PartXZ*, *PartYZ*, *Mech* та *Or · Re*:

- відповіддю вважається синтагма, в яких певні образи речення-питання поєднані логічними операторами «АБО» чи «І»;
- відповіддю є слово або частина синтагми, що відповідають питальному займеннику речення-питання;
- відповіддю вважається ансамбль образів оперативної пам'яті (за типом пророцтва дельфійського оракула).

5.6 Програмна реалізація інформаційної технології

У процесі програмної реалізації функцій системи обробки природно-мовного контенту за базову було обрано технологію *Python* + *SQLite*, яка поєднує можливості мови запитів SQL до реляційної бази даних з парадигмами об'єктно-орієнтованого і функціонального програмування. Інструментарій цієї технології було також збагачено можливостями пакетів утиліт *GraphViz* та *Networkx*. На даний час створене програмне забезпечення інформаційної технології реалізовано на рівні діючого прототипу СОПМК для всіх засобів підтримки розуміння образного сенсу ЕК, які розроблено в межах цієї роботи.

Окрім цього, часткові задачі моделювання в межах запропонованого концептуального підходу з метою відлагодження відповідних алгоритмів у різний час було реалізовано таким чином:

1. На мові програмування *Pascal* – алгоритм пошуку найкоротшого шляху в орієнтованому графі (п. 4.4) – текст програми для моделю-

вання обходу графа АМО (алгоритм *OPTWAY*) і файли з тестовими результатами її виконання представлено у Додатку В.

2. На мові програмування Лісп (оболонка *DrScheme*) – алгоритм навігації в графі булеану для реалізації «піраміди сенсу» та моделювання механізму оперативної пам’яті (пп. 3.3, 4.5, 5.3). На рисунках 5.9 і 5.10 представлено результат роботи програмного комплексу для тестових прикладів за мотивами Б. Окуджави і довідника з штучного інтелекту [78], які вміщено у Додатку В.

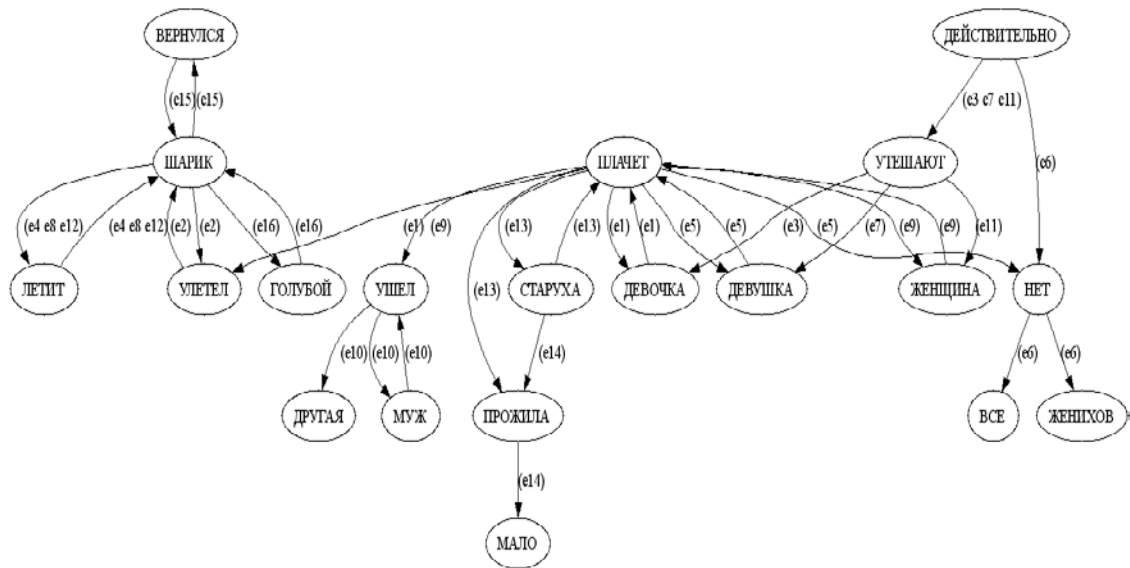


Рисунок 5.9 – Результат побудови булеану для тестового прикладу 1 на мові Лісп

За прикладом, що зображено на рисунку 5.9, було реалізовано такий діалог з оболонкою *DrScheme* :

- > (?Сколько 'СТАРУХА 'ПРОЖИЛА)
- (МАЛО)
- > (?Куда 'МУЖ 'УШЕЛ)
- (|ДРУГАЯ|)
- > (?Кого 'НЕТ 'ВСЕ)
- (ЖЕНИХОВ)
- > (?Почему 'ДЕВОЧКА 'ПЛАЧЕТ)
- ((6 ((УЛЕТЕЛ) (ШАРИК))))

> (?Почему 'ДЕВУШКА 'ПЛАЧЕТ)
 ((6 ((НЕТ) (ЖЕНИХОВ ВСЕ))))
 > (?Почему 'ЖЕНЩИНА 'ПЛАЧЕТ)
 ((6 ((УШЕЛ) (МУЖ |ДРУГАЯ))))
 > (?Почему 'СТАРУХА 'ПЛАЧЕТ)
 ((6 ((ПРОЖИЛА) (МАЛО))))

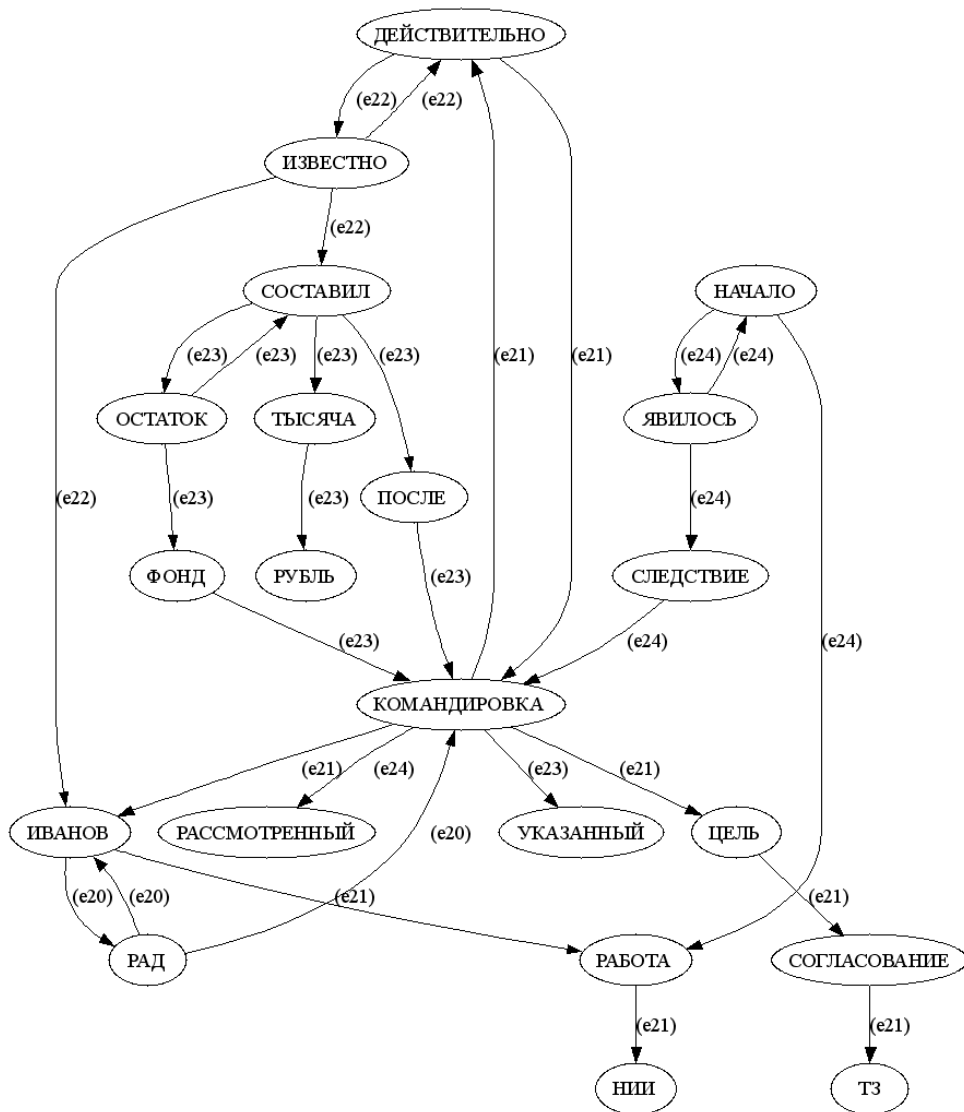


Рисунок 5.10 – Результат побудови булеану для тестового прикладу 2 на мові Лісп

За прикладом, зображеним на рисунку 5.10, оболонкою *DrScheme* було побудовано і застосовано для підтримки діалогу з ІС таку асоціативну таблицю *AList* :

ИВАНОВ -> РАД[e20] РАБОТА[e21]
ФОНД -> КОМАНДИРОВКА[e23]
КОМАНДИРОВКА -> РАССМОТРЕНИЕ[e24] УКАЗАНИЕ[e23]
ДЕЙСТВИТЕЛЬНО[e21] ИВАНОВ[e21] ЦЕЛЬ[e21]
РАД -> ИВАНОВ[e20] КОМАНДИРОВКА[e20]
ПОСЛЕ -> КОМАНДИРОВКА[e23]
ЦЕЛЬ -> СОГЛАСОВАНИЕ[e21]
РАБОТА -> НИИ[e21]
СОСТАВИЛ -> ОСТАТОК[e23] ПОСЛЕ[e23] ТЫСЯЧА[e23]
СОГЛАСОВАНИЕ -> ТЗ[e21]
НАЧАЛО -> ЯВИЛОСЬ[e24] РАБОТА[e24]
ТЫСЯЧА -> РУБЛЬ[e23]
ЯВИЛОСЬ -> НАЧАЛО[e24] СЛЕДСТВИЕ[e24]
ДЕЙСТВИТЕЛЬНО -> ИЗВЕСТНО[e22] КОМАНДИРОВКА[e21]
ИЗВЕСТНО -> ДЕЙСТВИТЕЛЬНО[e22] ИВАНОВ[e22]
СОСТАВИЛ[e22]
СЛЕДСТВИЕ -> КОМАНДИРОВКА[e24]
ОСТАТОК -> СОСТАВИЛ[e23] ФОНД[e23]

Текст основних програмних модулів комплексу на мові Лісп, якими забезпечено реалізацію тестових прикладів, також представлено в Додатку В.

3. На мові *SQL* (СУБД *Access*) – алгоритм для реалізації кредитно-модульної системи для діючого прототипу електронного підручника (п. 5.2) у складі АРМ викладача.

4. За допомогою технології *PHP-MySQL* – система пошуку зображень на основі образної індексації фолксонометричного простору [63], що побудовано з назв зображень – на рисунку 5.11 відображено

фрагмент роботи програмного комплексу, а результати пошуку зображень представлено у Додатку В.

Особливості та основні результати програмної реалізації СОПМК в базовій технології *Python + SQLite* докладно перевірено на даних наскрізного тестового прикладу ППЗ «WEB-технології: стандарти Semantic WEB» (Додаток Б). Програмну реалізацію словника образів було здійснено у відповідності зі схемою інформаційної технології п. 5.1 та з урахуванням ключової ролі цього програмного модуля в СОПМК [26]. Вперше результати цієї роботи опубліковано в [27]. З метою поліпшення одночасного доступу багатьох користувачів до словника образів забезпечено зберігання в одній базі даних не менше двох таблиць *Image i Assoc-Twice* як ядра системи, окреме адміністрування бази і організація до неї паралельного доступу в режимі реального часу.

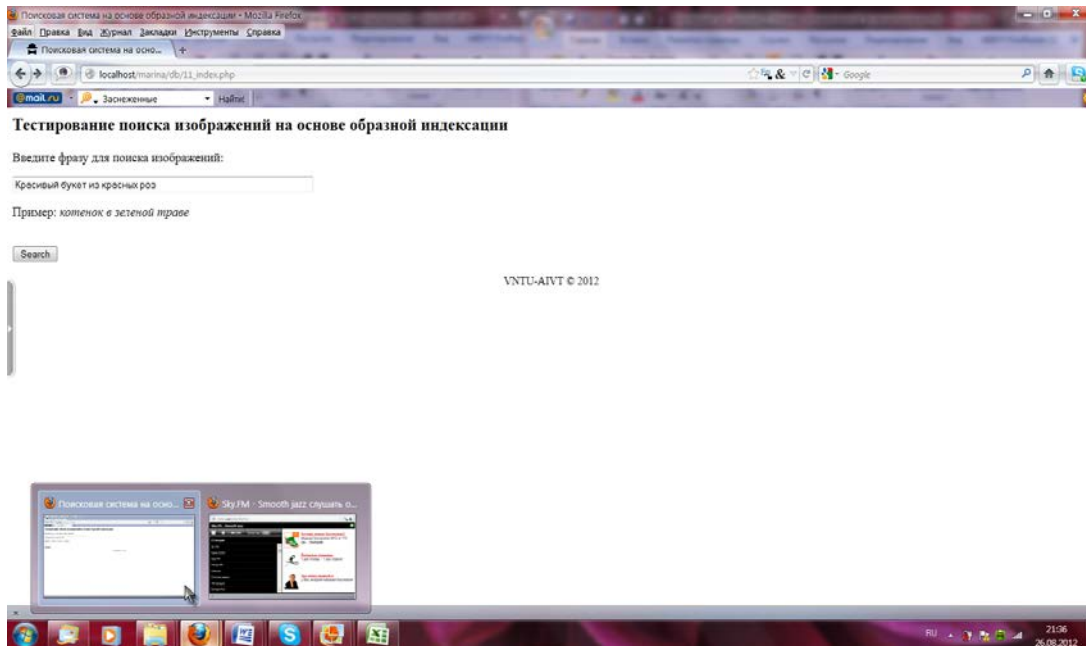


Рисунок 5.11 – Фрагмент роботи пошукової системи на основі образної індексації фолксонометричного простору

В *SQLite* використовується база даних, схема якої зображена на рисунку 3.11 та відповідає реляційній моделі (3.51)–(3.59) БАС. У Додатку Д представлено зовнішній вигляд оболонки навчальної системи, що заповнена тестовими синтагмами прикладу. Верхнє вікно головної WEB-сторінки (форма 1) цієї оболонки призначено для запису тексту нової синтагми, а кнопка «+», що знаходиться праворуч від вікна, вносить цей текст до бази даних системи.

В Додатку Д також представлено режим введення синтагми як множини асоціативних пар до навчальної системи (форма 2). Перехід до цього режиму забезпечується натисканням на відповідне гіперпосилання з форми 1. Згідно з розглянутою методикою, кожному слову речення ставиться у відповідність один з образів *Image*, а три вікна форми та кнопка «+» дозволяють вносити асоціативні пари. Список введених пар розташовано у нижній частині форми 2.

На рисунку 5.12 показано режим графічної інтерпретації зв'язків синтагми, який викликається за допомогою відповідної іконки з форми 2. Режим служить для візуального контролю правильності побудови АВМ окремого речення.

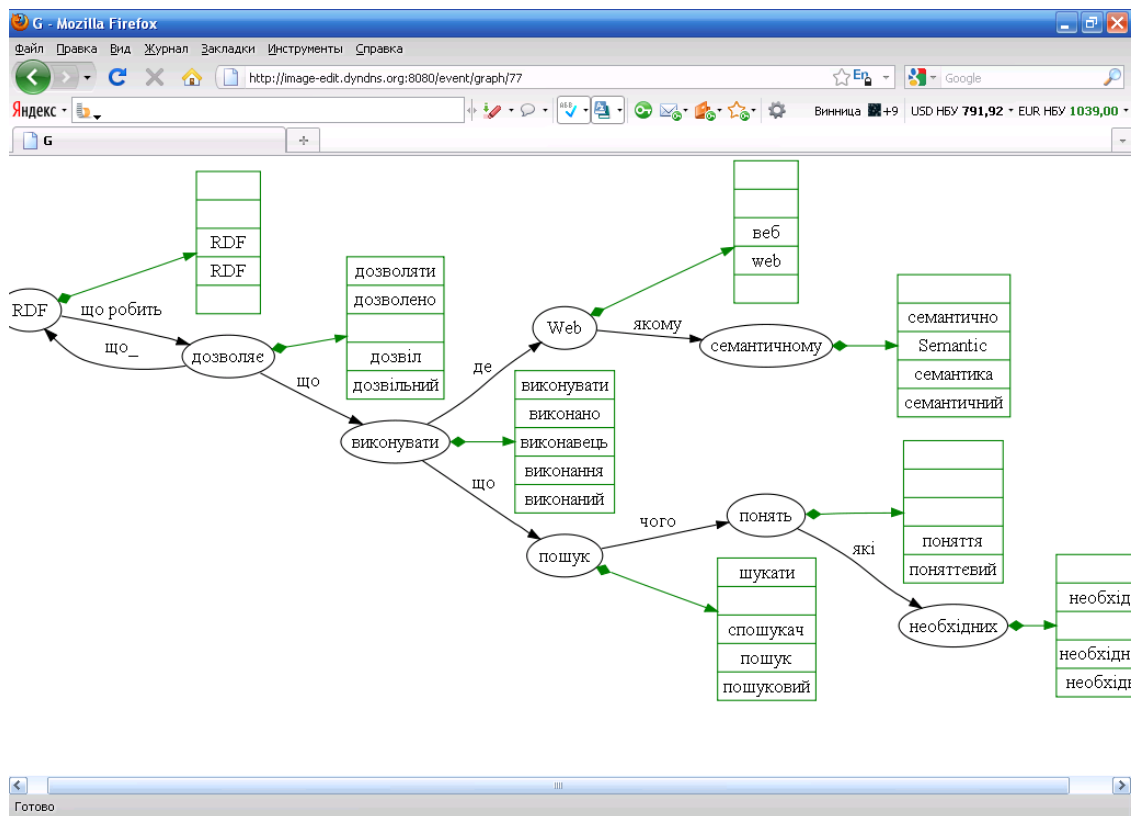


Рисунок 5.12 – Режим графічної інтерпретації зв'язків синтагми

У таблиці 5.2 показано найбільш близькі асоціативні зв'язки до образу «мова», отримані у результаті такого запиту в *SQLite* до бази даних тестового прикладу:

```
select total(force) as f, * from
(select force, b1 as b from a_t where b2=56
union all
```

select force, b2 as b from a_t where b1=56), image
 where bi_i=b group by b order by f desc,

що використовує таке представлення *a_t*:

```
CREATE VIEW a_t as select count(*) as force, b1, b2 from
(select construct_id, bi_i as b1, word_id as w1 from construct as c,
words as w where w.word_id = word_id_1) as a,
(select construct_id, bi_i as b2, word_id as w2 from construct as c,
words as w where w.word_id = word_id_2) as b
where a.construct_id=b.construct_id group by b1,b2 order by force
desc .
```

Таблиця 5.2 – Образи, що асоціативно наближені до образу «мова»

Кількість зв'язків	Код образу	ОQ	О	N	М	Н
5	71	існуючий	{null}	існування	є	{null}
4	123	дозвільний	{null}	дозвіл	дозволяти	дозволено
3	88	{null}	XML	XML	{null}	{null}
3	111	онтологічний	{null}	онтологія	{null}	{null}

Аналіз отриманих результатів програмного експерименту показує, що асоціативні зв'язки між образами «є», «дозволяти», «XML», «онтологія» та образом «мова» за своєю природою є синтагматичними, а тому пошук по кожній з цих 4-х асоціативних пар відокремить всі речення, де така пара присутня. Нова якість пошуку проявляється у тому, що його результати залежать виключно від змістовних характеристик синтагм, причому на пошук не впливають синтаксичні особливості речення та морфологічні форми застосованих в ньому слів.

Такий підхід також забезпечує можливість автоматичного визначення як найбільш значимих МО та сенсосполучень (пар МО). Наприклад, 12 МО з найбільшим значенням сенсу було отримано за допомогою таких *Python*-команд:

```
>>> import m
>>> a = m.create_adjacency_matrix()
>>> g = m.create_graph(a)
>>> all_nodes = [(i,g.degree(i)) for i in g.nodes()]
>>> sorted_top_nodes = sorted(all_nodes, key=lambda n: -n[1])[12]
>>> sorted_top_nodes
```

[(37, 25), (56, 24), (73, 16), (71, 12), (17, 11), (19, 11), (34, 11), (76, 11), (30, 10), (35, 10), (49, 10), (93, 10)]

Сортування результатів представлення *a_t* дозволило отримати ключові словосполучення елемента ЕК, що позначені кольором у таблиці 5.3.

Таблиця 5.3 – Аналіз результатів представлення *a_t*

<i>F+</i>	<i>F-</i>	<i>Sum</i>	<i>Bi_1</i>		<i>Bi_2</i>	
5		5	79	веб web	10	Семантичний Semantic семантика семантично
4		4	135	ідентифікований ідентифікатор ідентифікація ідентифікувати	134	URI
3		3	17	мережевий мережво мережа мережити	18	Інтернет
3		3	38	розумний розум зрозуміти розумно	27	комп'ютерний комп'ютер
3		3	56	мовний мовити мова	88	XML
3		3	56	мовний мовити мова	111	онтологічний онтологія
3	2	5	71	існуючий існування є	56	мовний мовити мова
3		3	76	описаний опис писати	73	ресурсний ресурс

Слід зазначити, що нові можливості запропонованого підходу [15, 17], в тому числі нова якість пошуку, на даний час досягаються за рахунок більшої трудомісткості внесення текстів навчального змісту в систему внаслідок розширеного синтаксичного аналізу речень. Проте порівняно невеликі обсяги навчального контенту в системах такого класу [4] не дозволяють вважати це обмеження критичним.

Розглянемо реалізацію типів асоціативного пошуку як відповідей на питання (п. 5.5) в умовах застосування словника образів. Першим з них апробовано відомий пошук типу «АБО» – запропонований підхід забезпечує пошук всіх словоформ без додаткового морфологічного аналізатора.

Реалізовано пошук типу «АБО» на основі такого *SQL*-запиту (коду 56 в тестовому прикладі відповідає МО «мова»):

```
select syntagma from event where event_id in (select distinct event_id from construct where word_id_1 and word_id_2 in (select word_id from words where bi_i=56)).
```

У Додатку Д представлено результат пошуку «АБО» всіх синтагм тестового прикладу, в яких бере участь образ «мова». Запропонована оболонка системи автоматично вкладає у текст *SQL*-запиту код того образу, який обирається користувачем в одному з полів вікна пошуку (форма 3). При необхідності пошуку типу «АБО» декількох образів текст запиту несуттєво змінюється (код 88 у тестовому прикладі відповідає образу «XML»):

```
select syntagma from event where event_id in (select distinct event_id from construct where word_id_1 and word_id_2 in (select word_id from words where bi_i=56 or bi_i=88)).
```

У додатку Д представлено тестовий результат пошуку типу «АБО» для двох МО («мова» та «XML»).

Згідно з алгоритмом п. 5.5, тип пошуку «І» також зберігає незалежність результатів від словоформ тих образів, що входять до його складу. Цей пошук забезпечує перемикач *Search2(and)* форми 3. Не дивлячись на формальну наближеність до першого типу пошуку («АБО»), пошук «І» набагато ближче відповідає мисленню людини. Так, на рисунку 5.13 зображено тестовий результат пошуку типу «І» для тих же двох образів «мова» та «XML», який значно менший попереднього результату.

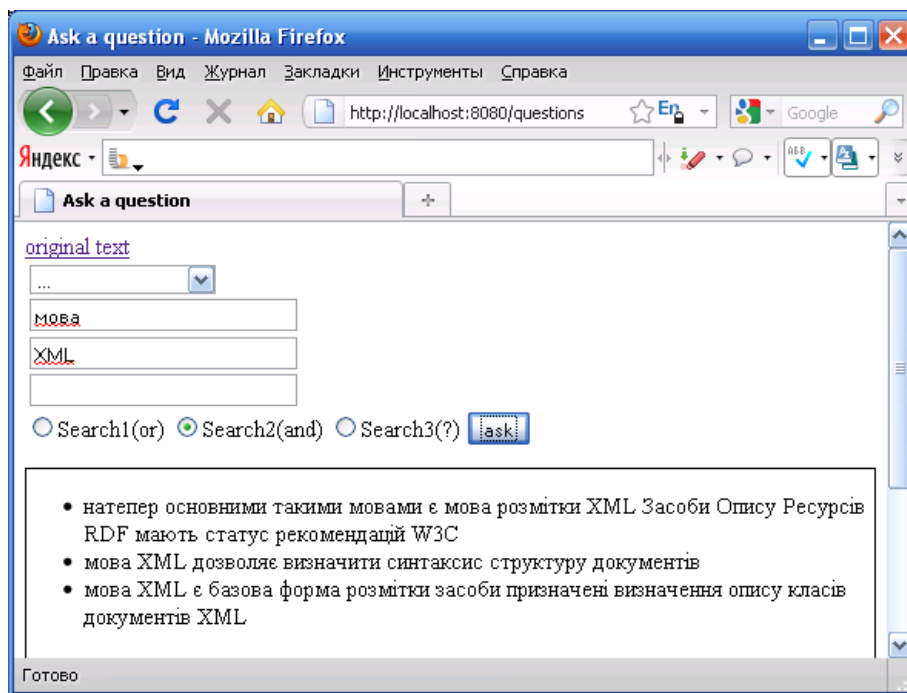


Рисунок 5.13 – Пошук всіх речень, в яких присутні образи «мова» та «XML»

Такий тип пошуку моделює ситуацію, коли людина, маючи на ду-мці щось своє, неухважно слухає співбесідника. Або людина слухає текст на малознайомій мові та сприймає значення лише окремих слів. І в тому і в іншому випадку на основі декількох сприйнятих слів у го-лові цієї людини підсвідомо виникає згадка про подію, де ці образи суміщені одночасно. Використання запропонованого підходу забезпе-чує появу на виході системи цитати з тексту у вигляді тієї синтагми, в якій беруть участь вхідні образи.

Третім з розглянутих типів пошуку *Search3(?)* будемо вважати ві-дповідь на формальне питання, яке починається питальним займенни-ком згідно з п. 5.5. У вікні форми 3, яку зображено на рисунку 5.13, займенник можна обрати з верхнього списку, що виключає викорис-тання невідомих системі питань. На рисунку 5.14 представлено прик-лад тестового результату пошуку за типом *Search3(?)*, що реалізує ал-горитм п. 5.4 за варіантом в).

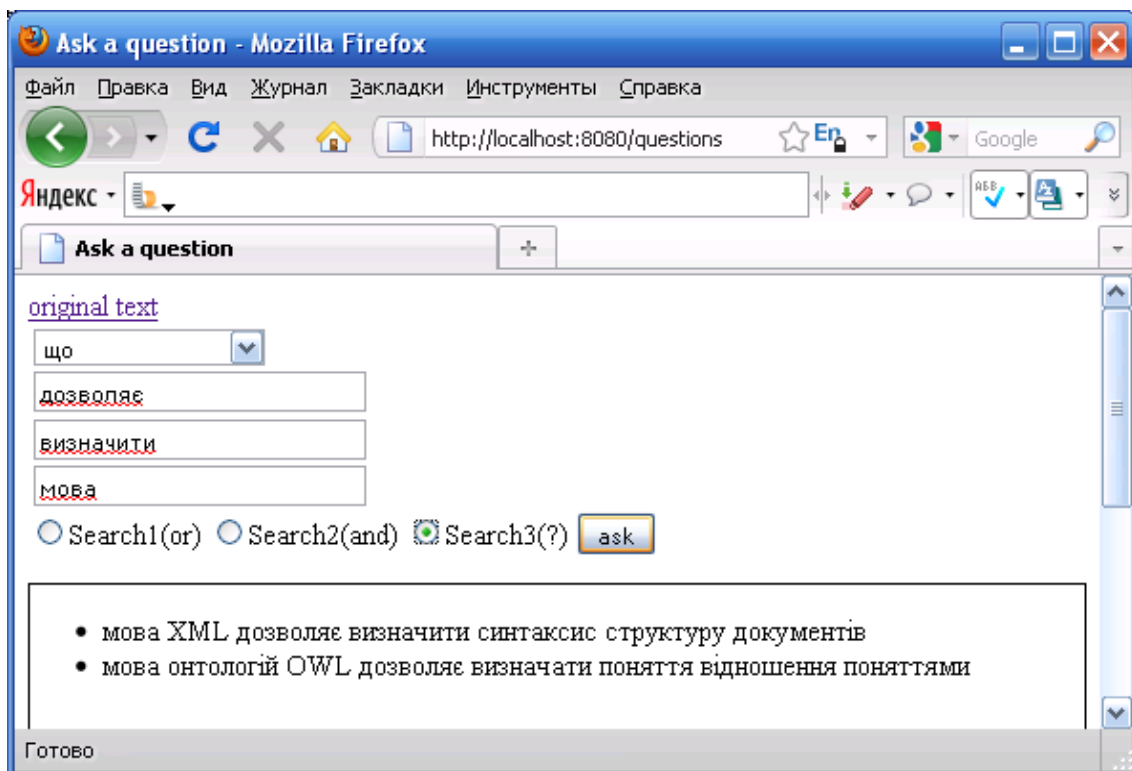


Рисунок 5.14 – Пошук типу *Search3(?)* як відповідь на питання

У Додатку Д також проілюстровано додаткові можливості (фор-ма 4) інфологічної системи з природно-мовним елементом ЕК на-

скрізного текстового прикладу, що відповідають розробленим у пп. 4.5, 4.6 та 5.4 моделям СОПМК.

Отже, основні результати дослідження програмно реалізовано за допомогою технології *Python + SQLite*. Внаслідок цього до отриманої програмної оболонки як основного діючого прототипу СОПМК забезпечено доступ через WEB. Для апробації окремих задач моделювання і впровадження отриманих результатів дослідження в навчальний процес було використано також мову програмування Паскаль, мову програмування Лісп (оболонка *DrScheme*), мову *SQL* в СУБД *Access* та, разом з мовою *PHP*, в технології *PHP-MySQL*. В результаті апробації програмного комплексу підтверджено адекватність результатів застосування інформаційної технології, на основі якої реалізовано базові та прикладні функції розробленої СОПМК.

5.7 Аналіз результатів впровадження інформаційної технології

Основні результати, отримані у цій роботі, впроваджено на підприємствах «AETHER. Inc» (США), «Smart.exe Ltd» (Ізраїль), «ІнноВінн» (м. Вінниця), в Донецькому інституті проблем штучного інтелекту НАН та МОНМС України, Державній науковій сільськогосподарській бібліотеці УААН, у Всеукраїнській громадській організації «Національна асоціація сільськогосподарських дорадчих служб України», в управлінні освіти і науки Вінницької обласної державної адміністрації, Центрі інформаційних технологій Вінницького державного аграрного університету, а також у навчальний процес фізико-математичної гімназії № 17 (м. Вінниця), на кафедрах економічної кібернетики Харківського національного аграрного університету, інформатики та кібернетики Сумського національного аграрного університету, інформаційних технологій Подільського державного аграрно-технічного університету, інформаційних систем і технологій Полтавської державної аграрної академії, автоматичної та інформаційно-вимірювальної техніки Вінницького національного технічного університету.

У відповідності до Субдоговору підряду для підприємства «AETHER. Inc» (США) розроблено комплекс моделей, алгоритмів і програм для створення системи повнотекстового пошуку з удосконаленими функціями семантичного пошуку. Для тестування запропонованих моделей і алгоритмів проведено програмний експеримент з такими характеристиками:

1) Здійснення програмного експерименту було забезпечено діючим прототипом системи повнотекстового пошуку на основі базової технології PHP-MySQL з використанням окремих модулів бібліотек nltk та csv на мові програмування Python.

2) Для створення тестової корпусної бази програмою dreamstime.py заповнено 2 поля (URL, Title) для 8660 записів та вміщено до файла dreamstime.csv, джерело інформації – відкритий ресурс Dreamstime©2012 (<http://www.dreamstime.com>);

3) Технологія образної індексації передбачає послідовний запуск 11-ти PHP-скриптів з метою створення та заповнення 6-ти таблиць бази даних MySQL: Pictures (8660 записів), Words (18794), Images (5045), Stems (4231), Construct (10146), Twice (7768).

4) Узагальнення англійських слів у мовні образи було забезпечено програмою stem_eng.py з використанням модуля SnowballStemmer бібліотеки nltk.

Результати запропонованого методу пошуку порівнювалися за стандартною методикою [281] з методами Гаусової, Жаккардової і косинусної подібності. Альтернативний пошук забезпечено програмою search.py з використанням бібліотеки csv. Результати програмного експерименту для 20 запитів узагальнено та в графічному вигляді показано на рисунку 5.15. Загальне зі всіма результатами поліпшення точності пошуку склало 11,3 %, а збільшення повноти пошуку – 17,5 %.

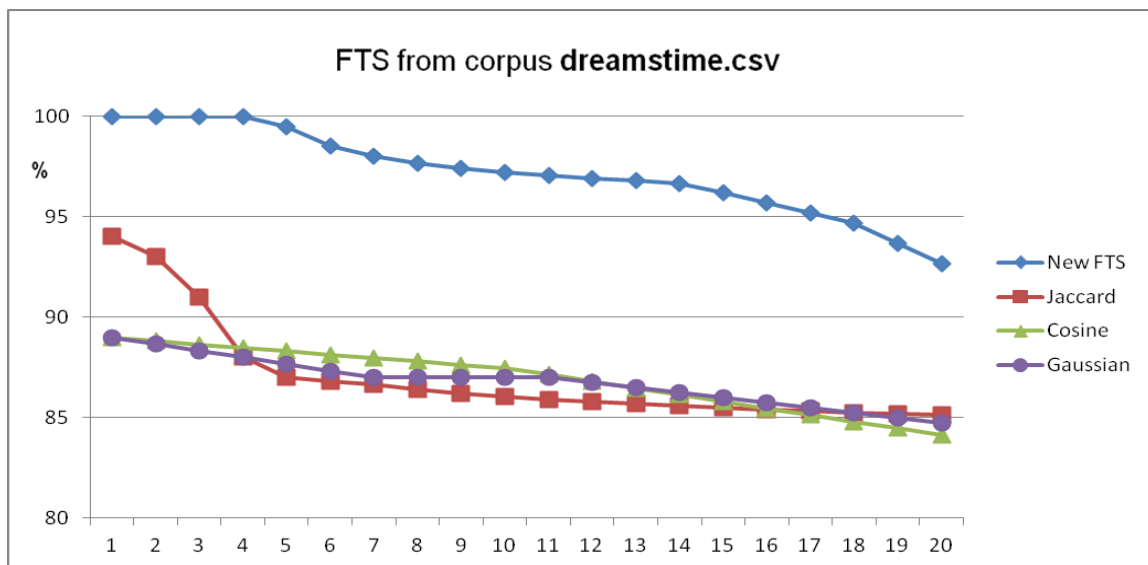


Рисунок 5.15 – Результати програмного експерименту в «AETHER. Inc»

У відповідності до Угоди про співпрацю для підприємства «Smart.exe Ltd» (Ізраїль) розроблено комплекс моделей та алгоритмів для підтримки системи автоматичного розпізнавання та пошуку образів (APRS) в репозиторії мультимедійних ресурсів. Система APRS класифікує за сенсом множину мультимедійних об'єктів, які зібрані та описані користувачами у відповідності до сховища багатомовних зображень. Функціонально система забезпечує:

- вибір мультимедійного об'єкта та аналіз його вербального опису;
- відокремлення ключових слів опису і визначення відповідних до них статей словника мовних образів;
- підтримку фолксономії та спільного маркування;
- визначення інваріантів сенсу в багатомовному середовищі;
- можливості самонавчання та навчання з вчителем.

Отримане програмне забезпечення методу розуміння сенсу мультимедійних об'єктів з підтримкою багатомовності протестовано та впроваджено в навчальних курсах «Android Development Class» та «Intro to Objective C & iPhone Training» підприємства «Smart.exe Ltd». Окрім цього, було проведено програмний експеримент для визначення повноти та точності інтелектуалізованого пошуку зображень на основі образної індексації фолксометричного простору. Оцінка результатів експерименту проводилася у відповідності до відомих методів, викладених у [281].

На основі методології образного аналізу та синтезу структурно-функціональних моделей систем для впровадження на підприємстві «Інновінн» (м. Вінниця) розроблено концептуальну модель системи для боротьби з шахрайством в мережах зв'язку, що забезпечує такі функції:

- автоматична кластеризація всіх абонентів;
- інтерактивна взаємодія з особою, що приймає рішення з метою інтерпретації та визначення потенційного ступеня небезпеки кожного нового кластера;
- контроль у режимі реального часу за «історією з'єднань» кожного абонента та оперативна сигналізація про випадки його виходу за межі власного кластера;
- можливість автоматичного перенавчання системи шляхом введення штучного кластера на основі аналізу «історій з'єднань» абонентів;

- автоматичне створення бази знань для боротьби з шахрайством з параметрами, що відповідають базі даних оператора зв'язку;
- виявлення фактів шахрайства у режимі реального часу;
- визначення «топологічних» параметрів абонентської мережі у вигляді кількості та динаміки зростання вхідних і вихідних зв'язків;
- автоматизоване поповнення бази знань невідомими раніше видами шахрайства.

У порівнянні з існуючими методами нові можливості впровадження результатів цієї роботи на практиці потребують лінійного збільшення часу обчислювальних процедур (в обсязі $O(k \cdot n)$, де n – загальна кількість абонентів, k – середня кількість зв'язків абонента), що не є критичним для систем комп'ютерної підтримки операторів зв'язку.

Дослідження базових характеристик моделі здійснено на прикладі технології інтелектуалізованого пошуку зображень на основі формалізації фолксонометричного простору. При цьому виявлення вже відомих типів шахрайства інтерпретовано як пошук на орієнтованому зваженому графі, що дозволяє порівняти отримані результати структурно-функціонального підходу з існуючими методами пошуку.

Програмний експеримент у «ІнноВінн» проведено для графічних ресурсів, вільно доступних на WEB-сайті «Банк обоев – бесплатные обои для рабочего стола © 2012» (<http://www.bankoboev.ru/>). Фолксонометричний простір репозиторію отримано на основі вербального опису зображень на російській мові як обмежений варіант простору образного сенсу з нечіткою мірою (п. 2.5). Основні обмеження, закладені в інтелектуалізовану пошукову систему, такі:

1. У зв'язку із загальноновживаним лаконізмом опису зображень парами вважаються сусідні слова назви.
2. Займенники, прийменники та службові слова не враховуються у пошуку.
3. Кількість зображень з відповідним описом обмежена 47754 файлами з <http://www.bankoboev.ru/>, що належать до 71 категорії.
4. Приналежність зображення до категорії не впливає на пошук, тому що слова з назв категорій не включено до фолксонометричного простору.

5. Інформаційна технологія узагальнює російські слова у мовні образи на основі стандартних можливостей лінгвістичного пакета NLTK Stemmers (<http://nltk.org/api/nltk.stem.html#module-nltk.stem.snowball>).

Первинна оцінка релевантності результатів пошуку $O_{\alpha}P_{e_i}$ здійснювалася незалежно 3-ма експертами за 4-рівневою шкалою: 3 – відмінно, 2 – добре, 1 – задовільно, 0 – незадовільно. Оцінку було отримано для 50-ти перших результатів 10-х запитів, половина яких складалася з 3-х значимих слів, інша половина – з 4-х.

Приклади порівняння інтелектуалізованого та стандартного пошуків зображень за 4-ма запитами з найгіршими/найкращими результатами для 3-х та 4-х слів проілюстровано у Додатку В. У таблицях Додатку В наведено усереднені оцінки релевантності результатів пошуку для обраних запитів: 1) «Лесная река осенью», 2) «Заснеженные горные вершины», 3) «Красивый букет из красных роз» та 4) «Рыжий котенок на зеленой траве». Додатково визначалися бінарна оцінка $B_iO_{\alpha_i}$ та оцінка точності $O_{\alpha}T_{o_i}$ кожного i -го результату за формулами:

$$B_iO_{\alpha_i} = \begin{cases} 1, & O_{\alpha}P_{e_i} > 1 \\ 0, & O_{\alpha}P_{e_i} \leq 1 \end{cases}$$

$$O_{\alpha}T_{o_i} = \frac{1}{i} \sum_{j=1}^i B_iO_{\alpha_j},$$

а також загальна точність $T_{o_{\Sigma}} = \frac{1}{N} \sum_{i=1}^N O_{\alpha}T_{o_i}$, де N – загальна кількість оцінок.

Порівняння отриманих результатів у графічному вигляді зображено на рисунках 5.16–5.19. За результатами програмного експерименту виявлено, що покращення точності складає від 5,9 % до 11,7 % для запитів з 3-х слів та від 25,8 % до 29,7 % для запитів з 4-х слів. Відповідно підвищується повнота пошуку від 14,3 % до 28,2 % для запитів з 3-х слів та від 17,4 % до 95,5 % для запитів з 4-х слів. Отже, збільшення кількості значимих слів у запиті покращує результати пошуку зображень як за точністю, так і за повнотою. Середнє за всіма результатами покращення точності пошуку склало 18,3 %, а збільшення повноти пошуку – 38,8 %.

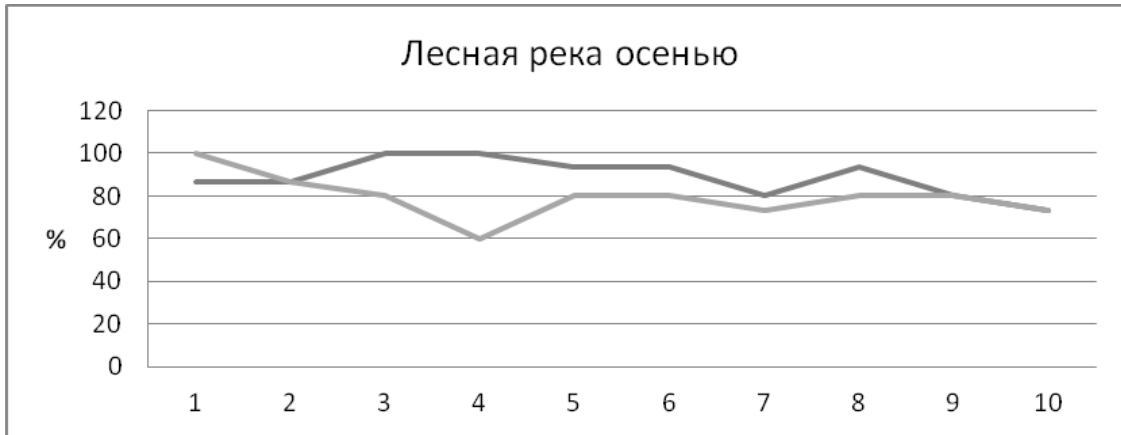


Рисунок 5.16 – Запит 1) з 3-х слів з найменшим покращенням точності пошуку



Рисунок 5.17 – Запит 2) з 3-х слів з найбільшим покращенням точності пошуку

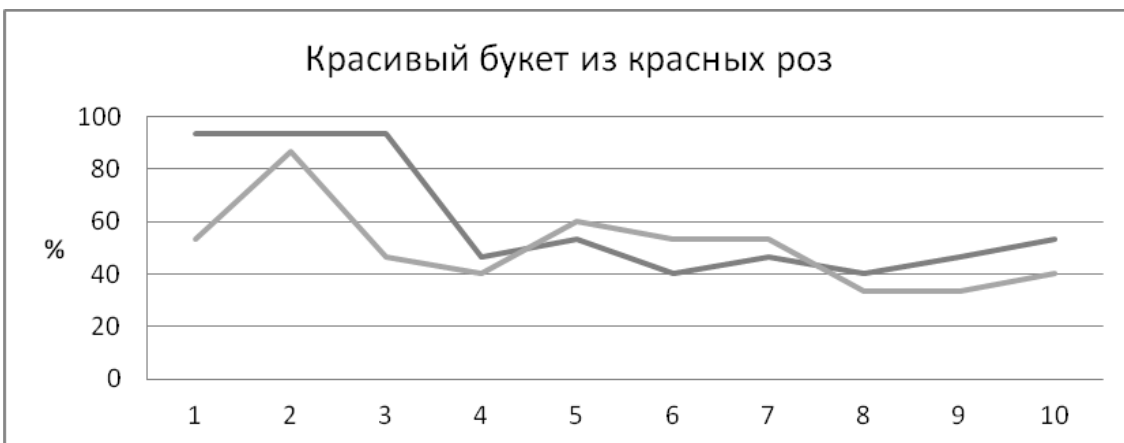


Рисунок 5.18 – Запит 3) з 4-х слів з найменшим покращенням точності пошуку



Рисунок 5.19 – Запит 4) з 4-х слів з найбільшим покращенням точності пошуку

В управлінні освіти і науки Вінницької ОДА внаслідок виконання НДР (№ державної реєстрації 0110U004301), в якому автор брав участь як науковий керівник впроваджено:

1. Комплекс методів та засобів підтримки розуміння сенсу ЕК на основі моделювання образного мислення людини.

2. Програмне забезпечення, яке забезпечує створення електронних підручників з розширеними дидактичними і лінгвістичними можливостями, підтримкою обмежених задачами навчального контенту типів діалогу та семантичного пошуку.

В результаті експлуатації вищезгаданого програмного забезпечення встановлено його високу ефективність при створенні «розумного» навчального контенту та навчальних об'єктів у форматі SCORM як окремих складових електронного підручника. Зокрема, досягнення мети впровадження наукових результатів у напрямку інтелектуалізації дидактичних функцій електронного підручника характеризується такими середніми показниками по 9-ти навчальних закладах, що підпорядковані управлінню освіти і науки Вінницької обласної державної адміністрації:

– зменшення на 45 % робочого часу викладача для створення електронного підручника (ЕП);

– підвищення ефективності застосування ЕП на 24 % в режимі самонавчання.

У Донецькому інституті проблем штучного інтелекту НАН та МОНМС України, Державній науковій сільськогосподарській бібліо-

теці УААН, Всеукраїнській громадській організації «Національна асоціація сільськогосподарських дорадчих служб України», Центрі інформаційних технологій Вінницького державного аграрного університету та в аграрних вузах Харкова, Сум, Полтави і Кам'янця-Подільського впроваджено комплекс інфологічного моделювання образного мислення (КІМОМ) людини та інструментальні засоби його підтримки у вигляді інтелектуальних інформаційних технологій. Означені результати досліджень було використано:

- для створення електронних посібників, електронного журналу викладача та інших продуктів і послуг в галузі дистанційної освіти згідно з вимогами кредитно-модульної та модульно-рейтингової систем;
- для забезпечення низки операцій асоціативного пошуку інформації в базі даних електронного підручника, що будується на основі спеціального аналізу текстових документів навчального контенту;
- для підтримки бази знань курсу дистанційного навчання «Сільськогосподарська дорадча діяльність» та WEB-сайту асоціації «Дорада».

Внаслідок впровадження КІМОМ та інструментальних засобів його підтримки у вигляді інтелектуальних інформаційних технологій було забезпечено:

- суттєве скорочення термінів (звичайно з 6–12 місяців до 2–4 місяців) розробки електронних підручників та іншої програмно-інформаційної продукції навчального призначення за рахунок застосування технології автоматизованої побудови бази знань для підтримки процесів пізнавальної діяльності;
- підвищення ефективності нових продуктів і послуг для дистанційної освіти та електронного навчання за рахунок більшої якості формального розуміння природно-мовних конструкцій віртуального навчального контенту.

У Центрі дистанційної освіти ВНТУ впроваджено та інтегровано в систему eLearning Server програмну оболонку для проектування електронних навчальних об'єктів. В результаті застосування програмної оболонки згідно з навчальною програмою спеціальності 7.091401 було створено програмно-педагогічний засіб (ППЗ) «WEB-технології: стандарти Semantic WEB». Апробацію ППЗ проведено в період з 1.12.2009 по 12.12.2009 на базі 3-х груп 5-го курсу ФАКСУ (2АВ-09сп, 4АС-09сп, 5АС-09сп) для загальної кількості 59 студентів з метою вивчення теми «Перспективи розвитку WEB-технологій» в режи-

мі самопідготовки. Контрольну групу зазначеного педагогічного експерименту склали 19 студентів групи ЗАВ-09сп (5-й курс ФАКСУ), заняття для яких з цієї ж теми проводились також в режимі самопідготовки на основі традиційного ЕК у вигляді гіпертексту з аналогічним змістом.

За результатами навчального процесу було визначено оцінку математичного сподівання середнього балу за максимальним балом 16 для цієї теми, яка склала 14,3 для експериментальної групи з 59 студентів та 12,9 для контрольної групи з 19 студентів. Різниця складає 8,27 % від максимальної оцінки. Також було визначено 95 % довірчі інтервали для значень оцінки знань студентів експериментальної та контрольної груп. Це дозволяє стверджувати, що із надійністю 95 % мінімальне покращення результатів навчання внаслідок використання ППЗ «WEB-технології: стандарти Semantic WEB» становить 0,62 бали (4 % від 16), а максимальне покращення становить 2,02 бали (13 % від 16).

У термін з 1.03.2010 по 10.03.2010 додатково проводився експеримент з експертної оцінки можливостей електронних навчальних об'єктів, створених за допомогою програмної оболонки. В експерименті брали участь члени комісії та 17 студентів експериментальної групи, що отримали оцінку «А» за шкалою ECTS з дисципліни «WEB-технології» у 1-му семестрі 2009–2010 навчального року. В результаті експерименту отримано представлені у таблиці 5.4 оцінки за п'ятибальною системою щодо можливостей навчального контенту ППЗ «WEB-технології: стандарти Semantic WEB»:

У Вінницькій ФМГ № 17 в результаті застосування програмної оболонки для проектування електронних навчальних об'єктів було створено програмно-педагогічний засіб (ППЗ) «Фізика 7-й клас: базові фізичні поняття». Апробацію ППЗ проведено згідно з навчальною програмою в період з 15.09.2009 по 25.09.2009 на базі 7-А та 7-Б класу для загальної кількості 59 учнів з метою вивчення теми «Спостереження та експеримент. Базові фізичні поняття». З контрольної групи зазначеного педагогічного експерименту, яку склали 32 учня 7-В класу, отримано результати навчального процесу з цієї ж теми. Заняття в контрольній групі проводились на основі апробованого у попередні періоди традиційного електронного підручника у вигляді гіпертексту

[282]. В усіх класах учням ставилося завдання самостійно вивчити матеріал на уроці у присутності вчителя.

Таблиця 5.4 – Експертна оцінка можливостей ЕК ППЗ (ВНТУ)

№ з/п	Семантичні можливості ППЗ	Експертні оцінки								Середній бал
		«2»		«3»		«4»		«5»		
		К-сть	%	К-сть	%	К-сть	%	К-сть	%	
1	Асоціативний пошук	–	–	–	–	2	10,0	18	90,0	4,90
2	Інсайтний пошук	–	–	2	10,0	6	30,0	12	60,0	4,50
3	Визначення синонімії	–	–	2	10,0	4	20,0	14	70,0	4,60
4	Визначення антонімії	–	–	2	10,0	5	25,0	13	65,0	4,55
5	Визначення зв'язку окреме–загальне	–	–	6	31,6	6	31,6	7	36,8	4,05
6	Визначення зв'язку частина–ціле	–	–	5	25,0	10	50,0	5	25,0	4,00
7	Інваріантність мовних конструкцій	–	–	5	25,0	8	40,0	7	35,0	4,10
8	Генерація повідомлень	–	–	3	15,0	6	30,0	11	55,0	4,40
9	Підтримка спрощеного діалогу	–	–	7	35,0	7	35,0	6	30,0	3,95

За результатами навчання було визначено оцінку математичного сподівання середнього бала учнів за 12-бальною системою, яка склала 9,3 для групи учнів з 7-А та 7-Б класів і 8,4 для 7-В класу. Різниця складає 7,5 % від максимальної оцінки. Також було визначено 95 % довірчі інтервали для значень оцінки знань учнів експериментальної та контрольної груп (таблиця 5.5).

Таблиця 5.5 – Оцінка математичного сподівання та довірчі інтервали успішності 91 учня з теми «Спостереження та експеримент. Базові фізичні поняття»

Класи	Математичне сподівання	Довірчий інтервал (95 %)
7-А та 7-Б (59 учнів)	9,3	(8,9; 9,7)
7-В (32 учня)	8,4	(8,1; 8,7)

Результати педагогічного експерименту у вигляді відсотків учнів експериментальних та контрольних груп, що отримали оцінки за 16-бальною шкалою (ВНТУ) та 12-бальною (ФМГ № 17) узагальнено на рисунку 5.20.

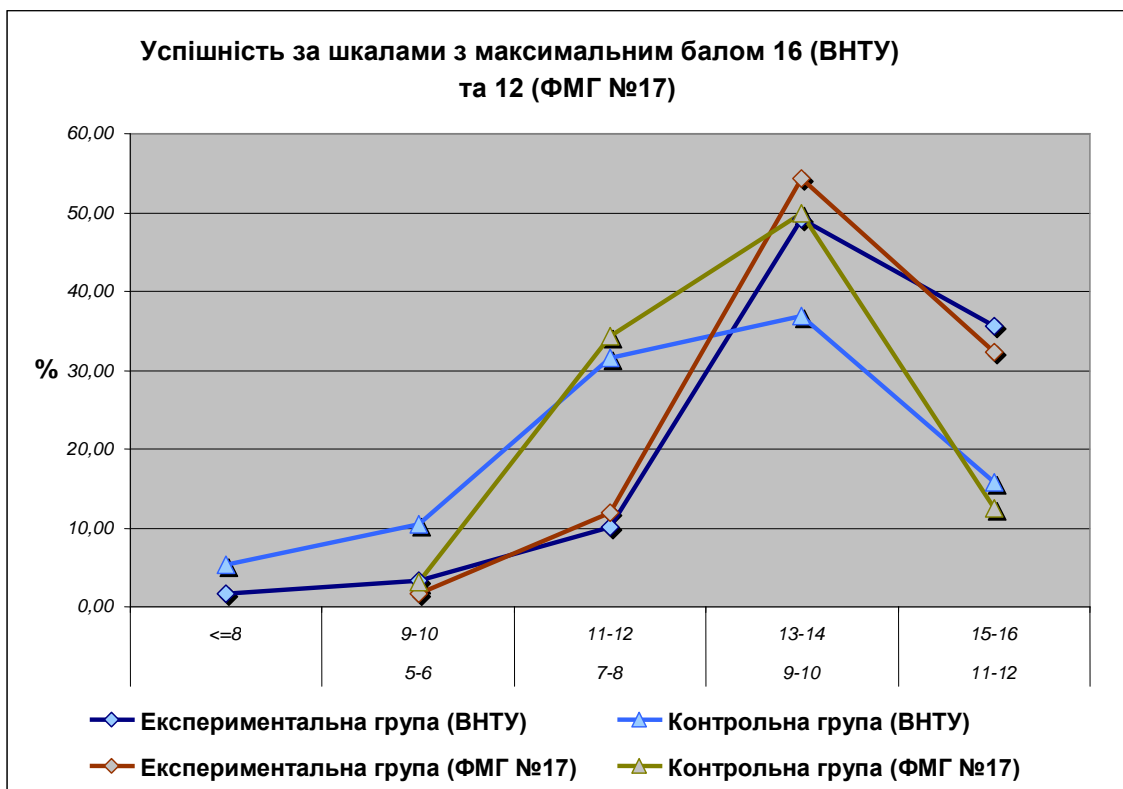


Рисунок 5.20 – Результати педагогічного експерименту

Отримані дані дозволяють стверджувати, що із надійністю 95 % мінімальне покращення результатів навчання внаслідок використання ППЗ «Фізика 7-й клас: базові фізичні поняття» становить 0,2 бали (2 % від 12), а максимальне покращення становить 1,6 бали (13 % від 12). Виявлене покращення успішності учнів було досягнуто за рахунок принципів побудови, структури та нових можливостей ППЗ, отриманих на основі функцій СОПМК.

З 25.09.2009 по 5.02.2010 додатково проводився експеримент з метою експертної оцінки функціональних можливостей електронних навчальних об'єктів, створених за допомогою програмної оболонки. В експерименті брали участь члени комісії та 9 учнів 7-А та 7-Б класів, що отримали оцінку «12 балів» з фізики за 1-й семестр 2009–2010 навчального року. Отримані у результаті експерименту оцінки за 5-бальною системою щодо можливостей навчального контенту ППЗ «Фізика 7-й клас: базові фізичні поняття» представлено в таблиці 5.6.

Таблиця 5.6 – Експертна оцінка можливостей ЕК ППЗ (ФМГ № 17)

№ з/п	Семантичні можливості ППЗ	Експертні оцінки								Середній бал
		«2»		«3»		«4»		«5»		
		К-сть	%	К-сть	%	К-сть	%	К-сть	%	
1	Асоціативний пошук	–	–	–	–	1	8,3	11	91,7	4,9
2	Інсайтний пошук	–	–	2	16,7	3	25,0	7	58,3	4,4
3	Визначення синонімії	–	–	1	8,3	3	25,0	8	66,7	4,6
4	Визначення антонімії	–	–	1	8,3	4	33,3	7	58,3	4,5
5	Визначення зв'язку окреме–загальне	–	–	4	33,3	4	33,3	4	33,3	4,0
6	Визначення зв'язку частина–ціле	–	–	4	33,3	5	41,7	3	25,0	3,9
7	Інваріантність мовних конструкцій	–	–	3	25,0	5	41,7	4	33,3	4,1
8	Генерація повідомлень	–	–	2	16,7	4	33,3	6	50,0	4,3
9	Підтримка спрощеного діалогу	–	–	5	41,7	4	33,3	3	25,0	3,8

Узагальнюючи результати проведених педагогічних експериментів можна стверджувати, що покращення успішності студентів та учнів досягнуто за рахунок нових засобів образної обробки природно-мовного контенту.

Аналіз всіх результатів дослідження та впровадження узагальнено на основі сукупності проведених експериментів та оцінки теоретичних моделей згідно з функціональними вимогами, обґрунтованими у п. 1.5. На рисунку 5.21 вміщено у жирну рамку такі функції з рисунку 1.2, для яких було побудовано відповідні математичні моделі та алгоритми. Пунктирною рамкою відмічено такі функції, прогрес для яких було досягнуто за рахунок концептуальних та конструктивних особливостей створеної СОПМК.

Зрештою, за межами дослідження залишилися критерії, пов'язані з морфологією і синтаксисом ПМК та пошуком за маскою слів, тобто задачі з відомими практичними розв'язками.

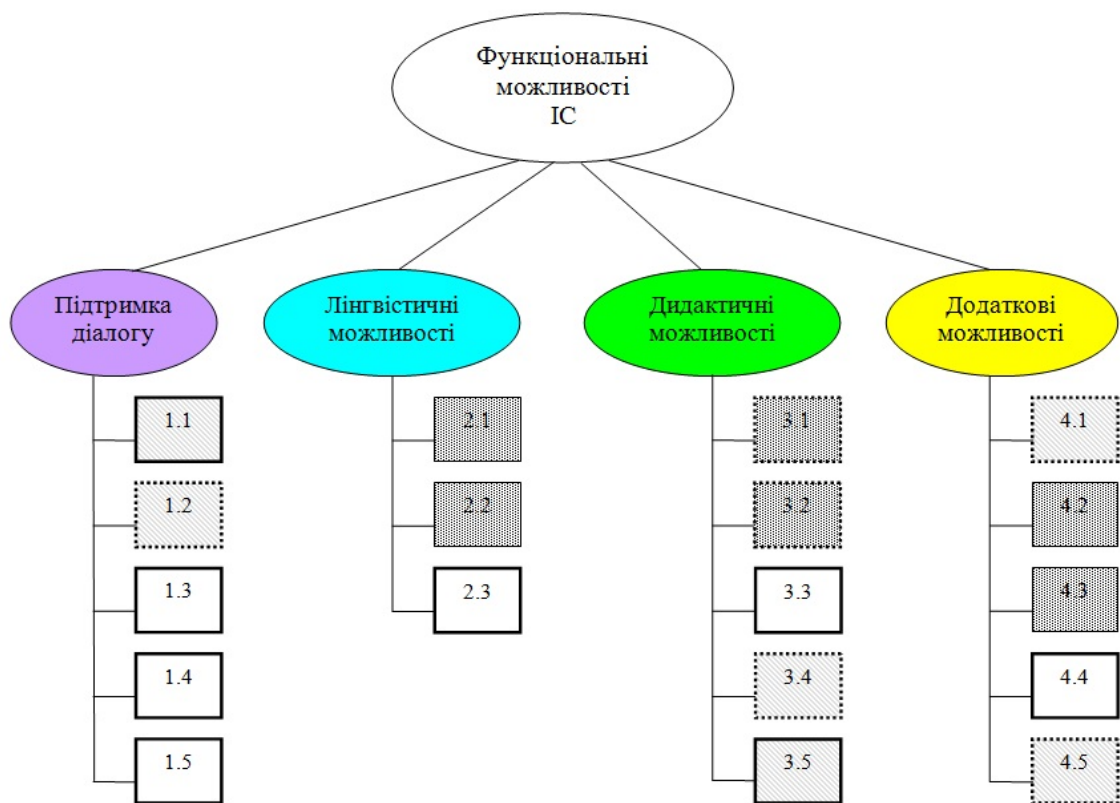


Рисунок 5.21 – Узагальнення можливостей інфологічної системи згідно з вимогами п. 1.6

Отже, основні результати проведеного дослідження підтверджено актами впровадження на 5-х підприємствах та установах різної форми власності, в т. ч. за кордоном, в адміністративній установі обласного рівня, громадській організації, 6-ти вищих та одному спеціалізованому загальноосвітньому навчальних закладах України. Об'єктами впровадження стали складові СОПМК, АРМ викладача, програмна оболонка для проектування електронних навчальних об'єктів та ППЗ, створені на її основі, концептуальна модель системи для боротьби з шахрайством в мережах зв'язку, які на практиці чисельно підтвердили достовірність викладених у монографії результатів дослідження.

ПІСЛЯМОВА

Внаслідок проведеного дослідження отримано нову інформаційну технологію, що забезпечує значний прогрес у розв'язанні актуальної проблеми створення науково-методологічних засад і відповідних технологічних засобів побудови СОПМК, здатних розв'язувати семантико-залежні задачі на основі накопичування загальних знань образного сенсу. Формалізовано нове поняття образного аналізу текстової інформації, що виявилось корисним для розв'язання такого класу задач у ролі проміжної ланки між синтаксичним та семантичним аналізом. Отримане теоретичне підґрунтя забезпечило синтез природно-мовних конструкцій на рівні алгебраїчних операцій з мовними образами та удосконалення структурно-функціональних моделей СОПМК.

Порівняльний огляд існуючих методів семантичного аналізу текстової інформації та мультидисциплінарних результатів дослідження когнітивної сфери людини виявив їх особливості та протиріччя, що обумовило доцільність структурно-функціонального підходу до побудови СОПМК. В основі підходу закладено онтогенетичне моделювання процесів обробки природно-мовного контенту з метою визначення образного сенсу як прагматичної першооснови змісту ПМК. На основі дослідження актуальних семантико-залежних задач визначено 18 функціональних вимог до інформаційної системи, яка у ролі абстрактної моделі СОПМК отримала назву інфологічної. Цим самим множини обраних семантико-залежних задач було закладено до таких груп функцій системи, як підтримка діалогу, лінгвістичні, дидактичні та додаткові (пошук і переклад) можливості.

Побудовано основи теорії образного аналізу текстової інформації у відповідності до ідеї формалізації поняття образного сенсу через визначення його властивості та параметра. Вибір базових концептуальних понять – інфологічної системи, онтогенетичного принципу та мовного образу – дозволив представити на рівні метамови сенс-властивість як семантичну мережу у вигляді АМО і дати чисельну оцінку сенсу-параметра на основі введеної одиниці сенсу. Зокрема:

- Уперше запропонований підхід до образного аналізу текстової інформації, що базується на обраних концептуальних поняттях, дозволив отримати чисельні ознаки загального образного сенсу природно-мовної конструкції.

- У комутативній напівгрупі образних конструкцій набула подальшого розвитку модель формальної теорії за рахунок застосування бінарного оператора спрямованого асоціативного зв'язку та поняття асоціативної нормальної форми, що забезпечило представлення образної конструкції простого речення у вигляді 3-х складових питальної конструкції мовних образів.

- Удосконалення методу створення нечіткого відношення образного сенсу завдяки зведення аксіоматики до онтогенетичного принципу, формалізації понять мотиваційної мети та емоційного стану ІС, введення одиниці сенсу на основі сили елементарного асоціативного зв'язку дозволило побудувати топологічний і квазіметричний простір упорядкованих пар образів.

Синтез структурно-функціональних моделей систем на основі природно-мовної інформації досягнуто шляхом обґрунтування інтерпретацій відомих психофізіологічних феноменів інтелектуальної діяльності з визначенням архітектури бази знань та структури СОПМК. У тому числі:

- Уперше розроблена методологія синтезу структурно-функціональних моделей інфологічної системи, в якій інтелектуальну діяльність людини інтерпретовано як взаємопов'язану обробку даних 9-ма функціональними блоками-моделями головного мозку у межах 19 інформаційних потоків образної природи, забезпечила формалізацію механізму обробки природно-мовного контенту.

- Удосконалення методів моделювання процесів образної обробки ПМК на основі орієнтованих і неорієнтованих зважених графів та графа булеану через алгебраїчну систему БАС дозволило представити базові функції ІС у вигляді алгоритмів на графах і формалізувати архітектуру бази знань СОПМК.

- Уперше запропонована функціональна модель СОПМК на основі побудови нової класифікації типів образного пошуку згідно з функціональними вимогами до ІС забезпечила уніфікацію процесів обробки даних і знань для розв'язання базових та прикладних семантико-залежних задач.

У межах загальної методології побудовано інформаційну технологію образної обробки природно-мовного контенту на основі СОПМК,

склад моделей для функцій якої відповідає функціональним вимогам до ІС. Зокрема:

- Набули подальшого розвитку математичні моделі базових типів образного пошуку з включенням інсайту як окремого випадку, що дозволило зменшити розмірність пошуку на орієнтованих та неорієнтованих зважених графах в процесі розв'язання семантико-залежних задач.

- Уперше розроблений метод моделювання образного механізму оперативної пам'яті СОПМК засобами двоосновної алгебраїчної системи БАС забезпечив отримання розв'язків семантико-залежних задач у вигляді алгоритмів навігації на графі булеану.

- Уперше запропонований метод моделювання складових парадигматичного устрою мови на основі підходу до образного аналізу тексту дозволив отримати рекурсивні алгоритми самовдосконалення бази знань СОПМК.

Розроблена інформаційна технологія образної обробки природно-мовного контенту забезпечує внесення попередньо відібраного текстового матеріалу в напівавтоматичному або автоматичному режимах, що імітує природний шлях накопичення знань людиною, побудову бази знань СОПМК у вигляді семантичної мережі АМО та розв'язання комплексу актуальних семантико-залежних задач на основі базових і прикладних функцій СОПМК. Результати дослідження програмно реалізовано за допомогою сучасних інструментальних засобів.

Отримані результати мають цілу низку перспективних напрямків подальшого розвитку, до яких можна віднести:

- Формальне визначення та дослідження класу семантико-залежних задач як підкласу NP-повних задач на основі моделі розповсюдження обмежень.

- Розвиток формальної теорії – доведення корисних з точки зору обробки природно-мовного контенту теорем, наприклад, мінімізація термів у АНФ.

- Удосконалення критеріального апарата, у т. ч. через функції передбачення та довіри, постановка та розв'язання відповідних задач оптимізації з метою синтезу довірчих систем.

- Дослідження самоподібності графових і алгебраїчних моделей ПМК, застосування фракталів для розв'язання семантико-залежних задач.

- Розвиток семантичної теорії інформації, теоретичне та експериментальне дослідження понять «бідного» та «багатого» тезаурусу на образному рівні.

- Доведення теорії до множини простих інженерних методик та співвідношень, які дають можливість отримати практично корисні результати.

- Автоматизація синтаксичного аналізу з метою побудови масштабного тезаурусу мовних образів для різних природних мов на основі існуючих електронних лінгвістичних та лексикографічних ресурсів.

- Збільшення ступеня когнітивності інфологічної системи за рахунок поєднання природно-мовних та формальних графічних ознак.

- Використання художніх (поетичних) контентів як об'єкта дослідження.

Отже, одержані при проведенні досліджень теоретичні результати відкривають перспективу науково обґрунтованого розв'язання актуального кола прикладних задач, пов'язаних з автоматизацією обробки електронного контенту на основі методології образного аналізу й синтезу його природно-мовних складових, а практичні результати створюють нове експериментальне середовище для подальших досліджень у широкій предметній галузі інтелектуальних інформаційних технологій.

ЛІТЕРАТУРА

1. Бисикало О. В. Игровой тренажер «Эффект» / О. В. Бисикало, А. Н. Мудрый, М. Л. Литвинов // Проблемы создания и применения автоматизированных обучающих комплексов в курсах высшей и прикладной математики : тез. докл. на межвуз. научно-метод. конф. – Винница : ВПИ, 1989. – С. 20.
2. Адаптивная информационно-обучающая система / О. В. Бисикало, А. С. Васюра, В. М. Дубовой, Ю. А. Пасихов // Приборостроение–93 и новые информационные технологии : тез. докл. на научно-техн. конф. с международным участием. – Николаев : НКИ, 1993. – С. 60.
3. Волков О. А. Підхід до створення інформаційно-аналітичної системи міністерства аграрної політики України / Олександр Андрійович Волков, Олег Володимирович Бісікало // Зб. наук. праць Вінницького державного аграрного університету (Серія «Технічні науки»). – 2004. – № 17. – С. 234–239.
4. Методичні рекомендації щодо створення електронних посібників: з досвіду роботи / В. В. Ільїн, М. З. Швиденко, М. М. Пастушенко, О. В. Бісікало. – К. : Наукметодцентр аграрної освіти Мінагрополітики України, 2004. – 35 с.
5. Ільїн В. В. Дидактичні та технологічні вимоги до оболонки для підготовки та використання електронних навчальних посібників / Ільїн В. В., Теплюк В. М., Бісікало О. В. – К. : Аграрна освіта, 2004. – 20 с.
6. Бісікало О. В. Концепція проектування електронного навчального посібника / О. В. Бісікало // Наука і методика. – 2005. – № 3. – С. 73–78.
7. Підготовка змісту електронних посібників: методичний посібник для науково-педагогічних працівників та викладачів аграрних вищих навчальних закладів / В. В. Ільїн, М. З. Швиденко, М. М. Пастушенко, О. В. Бісікало. – К. : Наукметодцентр аграрної освіти, 2005. – 40 с.
8. Бісікало О. В. Дистанційний навчальний процес як об'єкт системного аналізу / О. В. Бісікало // Інформаційні технології та комп'ютерна інженерія. – 2005. – № 2. – С. 78–83.

9. Бісікало О. В. Проектування процесів дистанційного навчання на основі формалізації пізнавальної діяльності людини / О. В. Бісікало // Інформаційні технології та комп'ютерна інженерія. – 2005. – № 3. – С. 274–280.

10. Бісікало О. В. Характеристика і формальна оцінка переваг дистанційної форми навчання / Олег Володимирович Бісікало, Ірина Іванівна Бурденюк // Оптико-електронні інформаційно-енергетичні технології. – 2005. – № 2 (10). – С. 5–12.

11. Бісікало О. В. Можливості комп'ютерних комунікацій при дистанційній формі навчання на основі Internet-технологій / Олег Володимирович Бісікало, Ірина Іванівна Бурденюк // Зб. наук. праць Вінницького державного аграрного університету (Серія «Технічні науки»). – 2006. – № 24. – С. 147–157.

12. Bisikalo O. V. Principles of concept model developing of image thinking / O. V. Bisikalo // First International Conference «New Information Technologies in Education for All», Extended Conference Proceedings, (Kiev, 29-31 May 2006). – K. : Akadempriodika, 2006. – P. 25–34.

13. Бісікало О. В. Визначення результатів тестового контролю знань засобами нелінійних спискових структур / О. В. Бісікало // Інформаційні технології та комп'ютерна інженерія. – 2006. – № 2 (6). – С. 152–155.

14. Бісікало О. В. Принципи побудови бази знань експертної системи в галузі приладобудування / О. В. Бісікало // Вісник Черкаського державного технологічного університету. – Черкаси, 2006. – СПЕЦВИПУСК-2006. – С. 12–14.

15. Бісікало О. В. Асоціативний пошук інформації в межах моделі образного мислення людини / О. В. Бісікало // Інформаційні технології та комп'ютерна інженерія. – 2007. – № 1 (8). – С. 248–252.

16. Bisikalo O. V. The Approach to Automation of Designing Knowledge Base in the Device-Making Industry / O. V. Bisikalo // Proc. of IEEE East-West Design & Test Workshop. EWDTW'06, (Sochi, 15–19 September 2006). – Sochi, Russia, 2006. – P. 440–442.

17. Бісікало О. В. Система «питання–відповідь» у межах моделі образного мислення / О. В. Бісікало // Матеріали XIII міжнар. конф. з автоматичного управління. Автоматика-2006, (Вінниця, 25–28 вересня 2006 р.). – Вінниця : УНІВЕРСУМ-Вінниця, 2006. – С. 537–542.

18. Бісікало О. В. Дослідження простору асоціативних пар в контексті бази знань електронного підручника / О. В. Бісікало // Вимірвальна та обчислювальна техніка в технологічних процесах. – 2006. – № 2 (28). – С. 109–113.

19. Бісікало О. В. Класифікація образного пошуку / О. В. Бісікало // Інтелектуальні системи в промисловості і освіті – 2007 : тези доповідей Першої міжнар. наук.-техн. конф. (Суми, 7–9 листопада 2007 р.). – Суми, 2007. – С. 14–15.

20. Бісікало О. В. Проектування електронного підручника на основі формалізації пізнавальної діяльності людини / О. В. Бісікало // Перспективні технології навчання та освітні простори : зб. наук. праць. – К. : МННЦ ІТiС, 2007. – Вип. 1. – С. 179–190.

21. Bisikalo O. Approach to the modeling of imaging mechanism of operative memory / O. Bisikalo // Second International Conference «New Information Technologies in Education for All», Conference Proc., (Kiev, 21–23 November 2007). – K. : Akadempriodika, 2007. – P. 336–344.

22. Бісікало О. В. Структура блоку пам'яті на основі моделі образного мислення людини / О. В. Бісікало // Искусственный интеллект. – 2007. – № 3. – С. 461–468.

23. Bisikalo O. V. The Determination of Results of Knowledge Test Control by Means of the Hierarchic Structures / Oleg Vladimirovich Bisikalo, Robert Gevorgovich Tadevosyan // Proc. of IEEE East-West Design & Test Symposium. EWDTs'07, (Yerevan, 7–10 September 2007). – Yerevan, Armenia, 2007 – P. 585–588.

24. Бісікало О. В. Конструювання образу-рішення для моделі інтелектуального управління / О. В. Бісікало // Системні технології: регіональний міжвуз. зб. наук. праць. – Дніпропетровськ, 2008. – Вип. 3(56). – Т. 2. – С. 123–128.

25. Bisikalo O. Knowledge base of teaching system construction supported by creative thinking model / O. Bisikalo // Third International Conference «New Information Technologies in Education for All: e-education», Proc., (Kiev, 1–3 October 2008). – Kiev: Akadempriodika, 2008. – P. 413–421.

26. Бісікало О. В. Методика побудови тезауруса навчальної системи на основі моделі образного мислення / О. В. Бісікало // Искусственный интеллект. – 2008. – № 4. – С. 730–735.

27. Bisikalo O. V. Model of image thinking based development of glossary / O. V. Bisikalo // Искусственный интеллект. Интеллектуальные системы: материалы IX междунар. научн.-техн. конф. – Донецк : ИПИИ «Наука і освіта», 2008. – Т. 2. – С. 87–91.

28. Бісікало О. В. Реалізація модульно-рейтингової системи у вигляді додатку до бази даних / О. В. Бісікало // Наука і методика. – 2008. – № 14. – С. 42–49.

29. Бісікало О. В. Орієнтувальний рефлекс як задача моделювання образного механізму оперативної пам'яті людини / О. В. Бісікало // Бионика интеллекта. – 2008. – № 2 (69). – С. 89–94.

30. Бісікало О. В. Класифікація образного пошуку та моделювання інсайту / О. В. Бісікало // Вісник СумДУ (Серія «Технічні науки»). – 2008. – № 2. – С. 53–59.

31. Бісікало О. В. Концептуальні основи моделювання образного мислення людини / О. В. Бісікало. – Вінниця : ПП Балюк І.Б., ВДАУ, 2009. – 163 с.

32. Бісікало О. В. Побудова ланцюга образів у межах моделі асоціативного образного мислення [Електронний ресурс] / О. В. Бісікало // Наукові праці Вінницького національного технічного університету: електронне наукове фахове видання. – 2009. – № 2. – С. 1–8. – Режим доступу: http://www.nbu.gov.ua/e-journals/VNTU/2009_2/2009-2.files/uk/09ovbapt_ua.pdf.

33. Бисикало О. В. Ассоциативный поиск для задач обучения на основе электронного тезауруса образов / О. В. Бисикало // Управляющие системы и машины. – 2009. – № 2. – С. 28–33. – ISSN 0130-5395.

34. Бісікало О. В. Інфологічний підхід до моделювання образного мислення людини [Електронний ресурс] / О. В. Бісікало // Вісник СумДУ (Серія «Технічні науки»). – 2009. – № 2. – С. 15–20. – Режим доступу: http://visnyk.sumdu.edu.ua/arhiv/2009/Tech_2_09/09bovoml.pdf.

35. Бісікало О. В. Формалізація образної пам'яті людини у вигляді графа / О. В. Бісікало // Бионика интеллекта. – 2009. – № 1 (70). – С. 127–131.

36. Бісікало О. В. Комплекс інфологічного моделювання образного мислення людини / О. В. Бісікало // Восточно-Европейский журнал передовых технологий. – 2009. – № 2/2 (38). – С. 14–17.

37. Бісікало О. В. Представлення пізнавальної діяльності на основі інфологічної моделі образного мислення людини / О. В. Бісікало //

Оптико-електронні інформаційно-енергетичні технології. – 2009. – № 1 (17). – С. 90-97.

38. Бисикало О. В. Принципы построения тезауруса на основе модели образного мышления / О. В. Бисикало // Междунар. науч.-практ. конф. «Е-обучение в высшей школе – проблемы и перспективы» (INCEL–08). – Одесса, 2008. – 5 с.

39. Бісікало О. В. Представлення асоціативної мережі образів за допомогою графів / Олег Володимирович Бісікало, Роберт Геворкович Тадевосян // Вісник Національного ун-ту «Львівська політехніка»: Комп'ютерні науки та інформаційні технології. – 2009. – № 650. – С. 73–80.

40. Бісікало О. В. Моделювання складових парадигматичної будови мови / О. В. Бісікало // Искусственный интеллект. – 2009. – № 4. – С. 188–194.

41. Бисикало О. В. Субъективная единица смысла образных конструкций / О. В. Бисикало // Nauka: teoria i praktyka – 2009 : materialy V miedzynar. naukowii-praktycznej konf., (Przemysl, 7–15 sierpnia 2009). – Przemysl : Nauka i studia, 2009. – Vol. 6. – P. 9–12.

42. Бісікало О. В. Модель визначення синонімічного зв'язку у просторі асоціативних пар / О. В. Бісікало // Системи та засоби штучного інтелекту: тези доповідей міжнар. наук. молодіжної школи. – Донецьк: ППШ «Наука і освіта», 2009. – С. 33–35.

43. Бісікало О. В. Аксиоматизація простору сенсу образних конструкцій / Олег Володимирович Бісікало, Роман Наумович Кветний // Зб. наук. праць Військового інституту Київського національного університету імені Тараса Шевченка. – К. : ВІКНУ, 2009. – Вип. 20. – С. 121–127.

44. Бисикало О. В. Об одном подходе к созданию обучающих систем в области оперативного управления приборостроительным производством / О. В. Бисикало, Ю. Я. Пасихов // Приборостроение-93 и новые информационные технологи : тез. докл. на научно-техн. конф. с международным участием. – Николаев : НКИ, 1993. – С. 59.

45. Бісікало О. В. Концептуальні основи корпоративної бази даних з WEB-технологіями доступу / О. В. Бісікало // Прикладні комп'ютерні програми для навчальної, методичної та організаційної роботи у вищих аграрних навчальних закладах III-IV рівнів акредита-

ції. Доповіді, виступи та повідомлення семінару, (Суми, 1–3 листопада 2000 р.). – К. : Аграрна освіта, 2001. – С. 86–90.

46. Бісікало О. В. Організація самостійної роботи студентів в умовах дистанційного навчання / Олег Володимирович Бісікало, Ірина Іванівна Бурденюк // Контроль та управління в складних системах. (КУСС-2003) : матеріали VII міжнар. наук.-техн. конф., (Вінниця, 8–11 жовтня 2003 р.). – Вінниця : УНІВЕРСУМ-Вінниця, 2003. – С. 224–227.

47. Бісікало О. В. Підхід до створення електронних підручників з тестуючими компонентами на основі моделі адаптивного навчання / О. В. Бісікало // Контроль та управління в складних системах. (КУСС-2003) : матеріали VII міжнар. наук.-техн. конф., (Вінниця, 8–11 жовтня 2003 р.). – Вінниця : УНІВЕРСУМ-Вінниця, 2003. – С. 227–234.

48. Бісікало О. В. Методика проектування процесів дистанційного навчання на основі формалізації пізнавальної діяльності людини / О. В. Бісікало // Контроль та управління в складних системах. (КУСС-2005) : тези доповідей VIII міжнар. наук.-техн. конф., (Вінниця, 24-27 жовтня 2005 р.). – Вінниця : УНІВЕРСУМ-Вінниця, 2005. – С. 287.

49. Бісікало О. В. Методика підготовки електронного посібника на прикладі дисципліни «Моделі та структури даних» / О. В. Бісікало // Сучасні технології підготовки фахівців в умовах подальшого розвитку вищої освіти України : матеріали міжнар. наук.-метод. конф., (Харків, 27–28 жовтня 2005 р.). – Харків : ХНАДУ, 2005. – С. 41–43.

50. Бісікало О. В. Проектування архітектури бази знань експертної системи для створення електронних посібників / О. В. Бісікало // Інформаційні технології в освіті, науці і техніці (ІТОНТ-2006) : матеріали V Всеукраїнської конф. молодих науковців, (Черкаси, 3-5 травня 2006 р.). – Черкаси : ЧНУ, 2006. – С. 98.

51. Бісікало О. В. Архітектура електронного підручника на основі бази знань навчальної експертної системи / О. В. Бісікало // Інформаційні технології в освіті: матеріали Всеукр. наук.-практ. конф., (Мелітополь, 24–26 травня 2006 р.). – Мелітополь : МДПУ, 2006. – С. 7–9.

52. Бісікало О. В. Алгоритм пошуку інформації на основі моделі асоціативної пам'яті людини / О. В. Бісікало // ІНТЕРНЕТ–ОСВІТА–НАУКА–2006 (ІОН–2006) : зб. матеріалів V міжнар. конф., (Вінниця,

10–14 жовтня 2006 р.). – Вінниця : УНІВЕРСУМ-Вінниця, 2006. – Т. 2. – С. 560–565.

53. Бисикало О. В. Принципы построения лингвистической системы на основе модели образного мышления / О. В. Бисикало // Труды восьмой международной научно-практ. конф. «Современные информационные и электронные технологии», (Одесса, 21–24 мая 2007 г.). – Одесса : СИЭТ–2007, 2007. – С. 63.

54. Бисикало О. В. Подход к построению ассоциативной памяти на основе модели образного мышления / О. В. Бисикало // Искусственный интеллект. Интеллектуальные системы (ИИ–2007) : материалы междунар. научн.-техн. конф., (Дивноморское, 24–29 сентября 2007 г.). – Донецк–Таганрог–Минск, 2007. – С. 6–10.

55. Бисикало О. В. Использование понятия «пирамида смысла» для построения лингвистического процессора / О. В. Бисикало // Труды девятой международной научно-практ. конф. «Современные информационные и электронные технологии», (Одесса, 19–23 мая 2008 г.). – Одесса : СИЭТ–2008, 2008. – Т. 1. – С. 35.

56. Бісікало О. В. Алгебраїчна модель лінгвістичного процесора / О. В. Бісікало // Информационные технологии в управлении сложными системами : сб. докладов и тезисов междунар. научн.-практ. конф., (Днепропетровск, 22–23 мая 2008 г.). – Днепропетровск : ИТМ НАНУ и НКАУ, 2008. – С. 23–24.

57. Об одной модели взаимодействия «пользователь–ПК» / Р. Г. Тадевосян, О. В. Бисикало, Н. П. Ильницкий [и др.] // Нові технології навчання. Розвиток духовності та професіоналізму в умовах глобалізації : зб. наук. праць V міжнар. наук.-метод. конф. – Київ–Вінниця, 2008. – Спецвип. 55. – Ч. 2. – С. 257–259.

58. Бісікало О. В. Розуміння сенсу навчального контенту на основі моделювання образного мислення людини / О. В. Бісікало // Тези доповідей П'ятої наук.-практ. конф. з міжнар. участю «Математичне та імітаційне моделювання систем. МОДС'2010», (Київ, 21–25 червня 2010 р.). – К., 2010. – С. 183–185.

59. Бісікало О. В. Концептуальне поєднання понять образного мислення та мовленнєвої діяльності / О. В. Бісікало // Інформаційні технології та комп'ютерна інженерія. – 2010. – № 1(17). – С. 72–77.

60. Бісікало О. В. Онтогенетичний підхід до розуміння сенсу навчальних об'єктів / О. В. Бісікало // Теоретичні та прикладні аспекти

побудови програмних систем (ТААПСД'2010) : тези доповідей 7-ї міжнар. конф., (Київ, 4–8 жовтня 2010 р.). – Київ, 2010. – С. 85–94.

61. Бісікало О. В. Когнітивний простір образних конструкцій / О. В. Бісікало // Искусственный интеллект. Интеллектуальные системы. (ИИ–2010) : материалы международной научно-техн. конф., (Казивели, 20–24 сентября 2010 г.). – Донецк : ИПИИ «Наука і освіта». – 2010. – Т. 1. – С. 17–21.

62. Кветний Р. Н. Морфологічний аналіз слова на основі асоціативно-статистичного підходу / Р. Н. Кветний, О. В. Бісікало, І. А. Кравчук // Вісник Черкаського державного технологічного університету. – 2010. – № 3. – С. 132–135.

63. Бісікало О. В. Інтелектуальний пошук графічних даних на основі фолксонометричних даних / Олег Володимирович Бісікало, Марина Володимирівна Савелова // Veda a technologie: krok do budoucnosti – 2011 : materialy VII mezinar. vedecko-prakticka konf., (Praha, 27 unora – 05 brezen 2011). – Praha : Publishing House «Education and Science», 2011. – Dil. 16. – P. 6–8.

64. Бісікало О. В. Використання морфологічного аналізу в задачах дистанційної освіти / Олег Володимирович Бісікало, Ірина Анатоліївна Кравчук // Системні технології : регіональний міжвуз. зб. наук. праць. – Дніпропетровськ, 2011. – Вип. 4 (75). – С. 41–47.

65. Бісікало О. В. Побудова нечітких відношення і простору сенсу образних конструкцій / О. В. Бісікало // Вісник Київського національного університету імені Тараса Шевченка. Серія: фізико-математичні науки. – 2011. – Вип. № 1. – С. 70–73.

66. Бісікало О. В. Концептуальні алгоритми виокремлення морфем для реалізації інформаційної технології обробки природномовних текстів / О. В. Бісікало, І. А. Кравчук // Системи обробки інформації. – Харків : ХУПС, 2011. – Вип. 3 (93). – С. 7–9.

67. Бісікало О. В. Розв'язання задач дистанційного навчання на основі формалізації психодидактичного середовища / Олег Володимирович Бісікало, Тетяна Василівна Вінійчук // Europejska nauka XXI rowieka – 2011 : materialy VII miedzynar. naukowi-praktycznej konf., (Przemysl, 7–15 maja 2011). – Przemysl : Nauka i studia, 2011. – Vol. 19. – P. 6–9.

68. Бісікало О. В. Аналіз морфологічної структури слова на основі асоціативно-статистичного підходу / О. В. Бісікало, І. А. Кравчук //

Вісник Вінницького політехнічного інституту. – 2011. – № 4. – С. 134–136.

69. Бісікало О. В. Формальна теорія та модель комутативної напівгрупи образних конструкцій / О. В. Бісікало, Р. Г. Тадевосян // Математичні машини і системи. – 2011. – № 4. – С. 39–47.

70. Бісікало О. В. Онтогенетичний метод побудови нечіткого відношення сенсу / О. В. Бісікало // Штучний інтелект. – 2011. – № 1. – С. 134–140.

71. Bisikalo O. Formalization of semantic network of image constructions in electronic content [Електронний ресурс] / O. Bisikalo, I. Kravchuk // Cornell University Library (Computer Science, Computation and Language), arXiv: 1201.1192v1. – January 2012. – 4 с. – Режим доступу: <http://arxiv.org/abs/1201.1192v1>.

72. Бісікало О. В. Застосування поняття ентропії для чисельної оцінки образного сенсу вербальних конструкцій / О. В. Бісікало, Н. В. Кондратюк // Сучасні інформаційні системи і технології (AIST-2012): матеріали I міжнар. наук.-практ. конф. (Суми, 15-18 травня 2012 р.). – Суми: СДУ. – 2012. – С. 213–214.

73. Кветний Р. Н. Визначення сенсу текстової інформації на основі моделі розповсюдження обмежень [Електронний ресурс] / Р. Н. Кветний, О. В. Бісікало, І. О. Назаров // Вимірювальна та обчислювальна техніка в технологічних процесах. – 2012. – № 1. – С. 93–96. – ISSN 2219-9365. – Режим доступу:

http://www.nbu.gov.ua/portal/Natural/Vtot/2012_1/46kve.pdf.

74. Бісікало О. В. Формалізація понять мовного образу та образного сенсу природно-мовних конструкцій / О. В. Бісікало // Математичні машини і системи. – 2012. – № 2. – С. 70–73.

75. Бісікало О. В. Образний аналіз текстової інформації з Wikipedia / О. В. Бісікало, І. О. Назаров // Вісник інженерної академії України. – 2012. – № 2. – С. 52–56.

76. Широков В. А. Комп'ютерна лексикографія / В. А. Широков. – К.: Наукова думка, 2011. – 351 с.

77. Искусственный интеллект: в 3 кн. – Кн. 1: Системы общения и экспертные системы: справочник / Под. ред. Э. В. Попова. – М.: Радио и связь, 1990. – 464 с.

78. Козеренко Е. Б. Формальные модели анализа и распознавания языковых структур [Електронний ресурс] / Е. Б. Козеренко // Диалог:

междунар. конф. по компьютерной лингвистике : форум. – Режим доступа: http://www.dialog-21.ru/trends/?id=2026&forum_id=17&f=1.

79. Нильсон Н. Принципы искусственного интеллекта / Н. Нильсон. – М. : Радио и связь, 1985. – 376 с.

80. Люггер Дж. Искусственный интеллект: стратегии и методы решения сложных проблем / Дж. Люггер – 4-е изд. – М. : Вильямс, 2003. – 864 с.

81. Стюарт Р. Искусственный интеллект: современный подход / Р. Стюарт, Н. Питер – 2-е изд. – М. : Вильямс, 2006. – 1408 с.

82. Шлезингер М. И. Математические средства обработки изображений / М. И. Шлезингер. – К. : Наукова думка, 1989. – 198 с.

83. Джонс М. Т. Программирование искусственного интеллекта в приложениях / М. Т. Джонс. – М. : ДМК Пресс, 2004. – 312 с.

84. Представление и использование знаний / под ред. Х. Уэно, М. Исидзука. – М. : Мир, 1989. – 220 с.

85. Schank R. Scripts, plans, goals and understanding: An inquiry into human knowledge structures / R. Schank, R. Abelson. – Hillsdale, N.J : L. Erlbaum, 1977. – 248 p.

86. Минский М. Фреймы для представления знаний / М. Минский; пер с англ. под ред. Ф. М. Кулакова. – М. : Энергия, 1979. – 151 с.

87. Грэхем И. Объектно-ориентированные методы. Принципы и практика = Object-Oriented Methods: Principles & Practice / Грэхем И – 3-е изд. – М. : Вильямс, 2004. – 880 с.

88. Codd E. F. Extending the database relational model to capture more meaning / E. F. Codd // ACM Trans. on Database Systems. – 1979. – № 4. – P. 397–434.

89. Конноли Т. Базы данных: проектирование, реализация и сопровождение. Теория и практика / Т. Конноли, К. Бегг, А. Страчан – 3-е изд. – М. : Вильямс, 2003. – 1440 с. – ISBN 5-8459-0527-3.

90. Цаленко М. Ш. Моделирование семантики в базах данных / М. Ш. Цаленко. – М. : Наука, Гл. ред. физ.-мат. лит., 1989. – 288 с.

91. Рубашкин В. Ш. Представление и анализ смысла в интеллектуальных информационных системах / В. Ш. Рубашкин. – М. : Наука, Гл. ред. физ.-мат. лит., 1989. – 192 с.

92. Попов Э. В. Общение с ЭВМ на естественном языке / Э. В. Попов. – М. : Наука, 1982. – 360 с.

93. Gelbukh A. Use of a weighted topic hierarchy for text retrieval and classification / A. Gelbukh, G. Sidorov, A. Guzman-Arenas // Vaclav Matoušek et al. (eds.). Text, Speech and Dialogue. Proc. TSD-99. Lecture Notes in Artificial Intelligence. – No. 1692, Springer-Verlag, 1999. – P. 130–135.

94. Попов Э. В. Общение с базами данных на ограниченном языке: прошлое, настоящее, будущее / Э. В. Попов // Новости искусственного интеллекта. – 2002. – № 1. – С. 21–26.

95. Осипов Г. С. Реляционно-ситуационный метод поиска и анализа текстов и его приложения / Г. С. Осипов, И. В. Смирнов, И. А. Тихомиров // Искусственный интеллект и принятие решений. – 2008. – № 2. – С. 3–10.

96. Виноград Т. Программа, понимающая естественный язык / Т. Виноград. – М. : Мир, 1976. – 296 с.

97. Арбиб М. Мозг, машина и математика / М. Арбиб. – М. : Наука, Гл. ред. физ.-мат. лит., 1968. – 224 с.

98. Аткинсон Р. Человеческая память и процесс обучения / Р. Аткинсон. – М. : Прогресс, 1980. – 528 с.

99. Temple J. How Google understands language like a 10-year-old [Электронный ресурс] / James Temple // Business Report, The Chronicle with Bloomberg. – 18.10.2010. – Режим доступа: [www/URL:](http://www.sfgate.com/cgi-bin/article.cgi?f=/c/a/2010/10/18/BUJ61FTF9I.DTL)

<http://www.sfgate.com/cgi-bin/article.cgi?f=/c/a/2010/10/18/BUJ61FTF9I.DTL>.

100. Поспелов Г. С. Искусственный интеллект – основа новой информационной технологии / Г. С. Поспелов. – М. : Наука, 1988. – 280 с.

101. Широков В. А. Элементы лексикографії / В. А. Широков. – К. : Довіра, 2005. – 304 с.

102. Apresyan J. ETAP-3 Linguistic Pricessor: a full-fledged NLP implementation of the MTT / J. Apresyan, I. Boguslavsky, L. Iomdin etc. // MTT 2003, First International Conference on Meaning – Text theory (Paris, June 16-18, 2003). – Paris, ENS, 2003. – P. 279–288.

103. Большаков И. А. Модель «Смысл ↔ Текст»: тридцать лет спустя [Электронный ресурс] / И. А. Большаков, А. Ф. Гельбух // J. International Forum on Information and Documentation. – 2000. – № 1. – Режим доступа: [www/URL:](http://gelbukh.com/CV/Publications/2000/Forum-MTM-rus.htm)

<http://gelbukh.com/CV/Publications/2000/Forum-MTM-rus.htm>.

104. Кобозева И. М. Лингвистическая семантика : учебное пособие / И. М. Кобозева. – М. : Эдиториал УРСС, 2000. – 352 с.
105. Харкевич А. А. О ценности информации / А. А. Харкевич // Проблемы кибернетики. – М. : Физматгиз, 1960. – Вып. 4. – С. 53–57.
106. Заде Л. А. Размытые множества и их применение в распознавании образов и кластер-анализе / Л. А. Заде // Вопросы анализа и процедуры принятия решений. – М. : Мир, 1976. – С. 172–215.
107. Гаврилова Т. А. Базы знаний интеллектуальных систем / Т. А. Гаврилова, В. Ф. Хорошевский. – СПб. : Питер, 2000. – 384 с.
108. Кофман А. Введение в теорию нечетких множеств / А. Кофман. – М. : Радио и связь, 1982. – 432 с.
109. Широков В. А. Информационный подход в моделировании лексикографических систем [Электронный ресурс] / В. А. Широков // Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества : материалы 4-й междунар. конф. «Крым-97». – 1997. – Т. 2. – Режим доступа:
<http://www.nbu.gov.ua/articles/crimea/1997/DOC/sem11/s11doc8.html>.
110. Garlick D. Intelligence and the Brain: Solving the Mystery of Why People Differ in IQ and How a Child Can Be a Genius / D. Garlick. – Burbank : Aesop Press, 2010. – 241 p.
111. Дударь З. В. Семантическая нейронная сеть, как формальный язык описания и обработки смысла текстов на естественном языке / Зоя Владимировна Дударь, Дмитрий Евгениевич Шуклин // Радиоэлектроника и информатика. – 2000. – № 3. – С. 72–76.
112. Попов Э. В. «Естественный» подход к обучению модели ребенка естественному языку [Электронный ресурс] / Э. В. Попов // Десятая национальная конф. по искусственному интеллекту с междунар. участием КИИ–2006, (Обнинск, 25–28 сентября 2006 г.). – Режим доступа: [www/URL: www.raai.org/resurs/papers/kii-2006/doklad/Popov.doc](http://www.raai.org/resurs/papers/kii-2006/doklad/Popov.doc).
113. Пальчунов Д. Е. Моделирование мышления и формализация рефлексии. Теоретико-модельная формализация онтологии и рефлексии / Д. Е. Пальчунов // Философия науки. – 2006. – № 4 (31). – С. 86–113.
114. Гладун В. П. Партнерство с компьютером / В. П. Гладун. – К. : Port-Royal, 2000. – 128 с.
115. Куссуль Э. М. Ассоциативные нейроподобные структуры / Э. М. Куссуль. – К. : Наукова думка, 1992. – 140 с.

116. Валькман Ю. Р. О языке образного мышления / Ю. Р. Валькман, Л. Р. Исмагилова // Доклады международной конференции «Диалог 2004». – Верхневолжский, 2004. – С. 90–97.
117. Talbot D. Extracting Meaning from Millions of Pages [Электронный ресурс] / David Talbot // Technology Review, MIT. – 10.03.2009. – Режим доступа: www/URL:
<http://www.technologyreview.com/computing/22773/>.
118. The DeepQA Project [Электронный ресурс] / The IBM Jeopardy! Challenge, IBM. – Режим доступа:
<http://www.research.ibm.com/deepqa/deepqa.shtml>.
119. Голубицкий С. Чудо Compreno [Электронный ресурс] / С. Голубицкий // Компьютерра-Онлайн. – 28.02.2012. – Режим доступа: [www/URL: http://www.computerra.ru/sgolub/663954/](http://www.computerra.ru/sgolub/663954/).
120. Schneider David. Google Brings New Meaning to the Web [Электронный ресурс] / D. Schneider // IEEE Spectrum inside technology. – June 2012. – Режим доступа:
http://spectrum.ieee.org/telecom/internet/google-brings-new-meaning-to-the-web/?utm_source=techalert&utm_medium=email&utm_campaign=061412.
121. Горошко Е. И. Интегративная модель свободного ассоциативного эксперимента / Е. И. Горошко. – М. : Харьков : ИЯ РАН ; Каравелла, 2001. – 318 с.
122. Ron S. The Cambridge Handbook of Computational Psychology / S. Ron. – New York : Cambridge University Press, 2008. – 432 p.
123. Спенсер Г. Основания психологии; Циген Т. Физиологическая психология в 14 лекциях // Основные направления психологии в классических трудах: Ассоциативная психология. – М. : АСТ-ЛМД, 1998. – 560 с.
124. Эббингауз Г. Очерк психологии; Бэн А. Психология // Основные направления психологии в классических трудах: Ассоциативная психология. – М. : АСТ-ЛМД, 1998. – 544 с.
125. Глейтман Г. Основы психологии / Г. Глейтман, А. Фридлунд, Д. Райсберг. – СПб. : Речь, 2001. – 1247 с.
126. Фрейд З. Психология бессознательного : сб. произведений / Фрейд З. ; сост. М. Г. Ярошевский. – М. : Просвещение, 1990. – 448 с.
127. Рубинштейн С. Л. Основы общей психологии / С. Л. Рубинштейн. – СПб. : Питер, 2009. – 712 с.

128. Шульц Д. П. История современной психологии / Д. П. Шульц, С. Э. Шульц. – СПб. : Евразия, 1998. – 528 с.
129. Белянин В. П. Основы психолингвистической диагностики (Модели мира в литературе) / В. П. Белянин. – М. : Тривола, 2000. – 248 с.
130. Лурия А. Р. Язык и сознание / Лурия А.Р. ; под ред. Е. Д. Хомской. – М. : Издательство Московского университета, 1979. – 320 с.
131. Зейгарник Б. В. Введение в патопсихологию / Б. В. Зейгарник. – М. : Издательство Московского университета, 1969. – 172 с.
132. Залевская А. А. Значение слова и возможности его описания / А. А. Залевская // Языковое сознание: формирование и функционирование : сб. статей / Отв. ред. Н. В. Уфимцева. – М., 1998. – С. 35–54.
133. Овчинникова И. Г. Ассоциации и высказывание: Структура и семантика : учебное пособие по спецкурсу / И. Г. Овчинникова. – Пермь : Перм. гос. ун-т, 1994. – 124 с.
134. Золотова Н. О. Ядро лексикона человека как феномен языкового сознания / Н. О. Золотова // Языковое сознание: содержание и функционирование : тезисы докладов XIII Международного симпозиума по психолингвистике и теории коммуникации (Москва, 1–3 июня 2000 г.) / Под ред. Е. Ф. Тарасова, И. Г. Овчинниковой. – М., 2000. – С. 94.
135. Караулов Ю. Н. Активная грамматика и ассоциативно-вербальная сеть / Караулов Ю. Н. – М. : ИРЯ РАН, 1999. – 180 с.
136. Караулов Ю. Н. Семантический гештальт ассоциативного поля и образы сознания / Ю. Н. Караулов // Языковое сознание: содержание и функционирование : тезисы докладов XIII Международного симпозиума по психолингвистике и теории коммуникации (Москва, 1–3 июня 2000 г.) / Под ред. Е. Ф. Тарасова. – М., 2000. – С. 107–108.
137. Прист С. Теории сознания / Прист С. ; пер. с англ. и предисл. А. Ф. Грязнова. – М. : Идея-пресс, 2000. – 287 с.
138. Гуссерль Э. Идеи к чистой феноменологии и феноменологической философии / Э. Гуссерль. – М., 1999. – Т. 1. – 336 с.
139. Портнов А. И. Язык и сознание: основные парадигмы исследования проблемы в философии XIX – XX вв. / А. И. Портнов. – Ивово, 1994. – 370 с.
140. Metzinger T. Being No One – The Self Model Theory of Subjectivity / T. Metzinger. – Cambridge : MIT Press, 2003. – P. 349–366.

141. Тихомиров О. К. Структура мыслительной деятельности человека / О. К. Тихомиров. – М. : МГУ, 1969. – 350 с.
142. Сержантов В. Ф. Человек как предмет философского и естественнонаучного познания / В. Ф. Сержантов, В. В. Гречаный. – Л., 1980. – 216 с.
143. Мамардашвили М. К. Стрела познания: набросок естественноисторической гносеологии / М. К. Мамардашвили ; под. ред. Ю. П. Сенокосова. – М. : Языки русской культуры, 1997. – 304 с.
144. Теплов Б. М. Психология и психофизиология индивидуальных различий : избранные психологические труды / Теплов Б. М. ; под ред. М. Г. Ярошевского. – М. : Институт практической психологии, Воронеж : МОДЭК, 1998. – 544 с.
145. Павилёнис Р. И. Проблема смысла : современный логико-фило-софский анализ языка / Р. И. Павилёнис. – М., 1983. – 286 с.
146. Хабаров И. А. Философские проблемы семиотики / И. А. Хабаров. – М., 1978. – 160 с.
147. Лотман Ю. М. Внутри мыслящих миров: Человек–текст–семиосфера–история / Ю. М. Лотман. – М., 1999. – 447 с.
148. Пономарев Я. А. Психология творения / Я. А. Пономарев. – М. : Московский психолого-социальный институт, Воронеж : МОДЭК, 1999. – 480 с.
149. Абушенко В. Л. Структурно-функциональный анализ [Электронный ресурс] / В. Л. Абушенко // Новейший философский словарь. – Режим доступа: www/URL:
http://www.gumer.info/bogoslov_Buks/Philos/fil_dict/767.php.
150. Jurafsky D. Speech and Language Processing [Электронный ресурс] / D. Jurafsky, J. Martin // An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition (2nd edition). – Prentice-Hall, 2009. – Режим доступа: www/URL:
<http://www.cs.colorado.edu/~martin/slp.html#top>.
151. Физиология человека : в 3 т. / пер. с англ. под ред. Р. Шмидта, Г. Тевса. – М. : Мир, 1996. – 880 с.
152. Павлов І. П. Двадцятирічний досвід об'єктивного вивчення вищої нервової діяльності (поведінки) тварин / Павлов І. П. – К., 1953. – 531 с.
153. Общий курс физиологии человека и животных : в 2 кн. / под ред. А. Д. Ноздрачева. – М. : Высшая школа, 1991. – 1040 с.

154. Соколов Е. Н. Механизм памяти / Е. Н. Соколов. – М. : Издательство Московского университета, 1969. – 175 с.
155. Ратанова Т. А. Субъективное шкалирование и объективные физиологические реакции человека / Т. А. Ратанова. – М. : Педагогика, 1990. – 216 с.
156. Хомская Е. Д. Нейропсихология / Е. Д. Хомская. – М. : Издательство Московского университета, 1987. – 288 с.
157. Анохин П. К. Биология и нейрофизиология условного рефлекса / П. К. Анохин. – М. : Медицина, 1968. – 548 с.
158. Черниговская Т. В. Билингвизм и функциональная асимметрия мозга / Т. В. Черниговская, Л. Я. Баллонов, В. Л. Деглин // Учен. зап. ТГУ : труды по знаковым системам. – 1983. – XVI. – С. 62–83.
159. Данилова Н. Н. Физиология высшей нервной деятельности / Н. Н. Данилова, А. Л. Крылова. – Ростов-на-Дону : Феникс, 2005. – 478 с.
160. Лурия А. Р. Основные проблемы нейролингвистики / Лурия А. Р. – М.: Издательство Московского университета, 1975. – 253 с.
161. Heilbrunner R. Forensic Neuropsychology Casebook / R. Heilbrunner. – New York, London : The Guilford Press, 2005. – 362 p.
162. Зейгарник Б. В. Патопсихология / Б. В. Зейгарник. – 2-е изд. – М. : Издательство Московского университета, 1986. – 287 с.
163. Урусов В. В. Инфология (R). Тема: метапсихология. Экибана [Электронный ресурс] / В. В. Урусов. – Режим доступа:
http://read.bookam.net/read/urusov_v_/page0/metapsihologija__yekibana.html.
164. Лохоня В. И. Моделирование поведения человека при помощи новых технологий Искусственного Интеллекта [Электронный ресурс] / В. И. Лохоня. – Режим доступа: <http://melos.info/rus/index.shtml>.
165. Saygin A. P. Turing Test: 50 Years Later / A. P. Saygin, I. Cicekli, V. Akman // Minds and Machines. – 2000. – Vol. 10, № 4. – P. 463–518.
166. Зейгарник Б. В. Теория личности в зарубежной психологии / Б. В. Зейгарник. – М. : Издательство Московского университета, 1982. – 128 с.
167. Блюм Г. Психоаналитические теории личности / Блюм Г. ; пер. с англ. и вступ. ст. А. Б. Хавина. – М. : КСП, 1996. – 247с.
168. Психология личности : словарь–справочник / под ред. П. П. Горностая и Т. М. Титаренко. – К. : Рута, 2001. – 320 с.

169. Олпорт Г. Становление личности / Г. Олпорт // Избранные труды. – М. : Смысл, 2002. – С. 166–216.
170. Маслоу А. Мотивация и личность / Маслоу А. – СПб : Питер, 2003. – 352 с. – (Серия: Мастера психологии).
171. Роджерс К. Теория личности / К. Роджерс // Клиентоцентрированная терапия. – К. : Ваклер, 1997. – С. 28–83.
172. Леонтьев А. Н. Деятельность, сознание, личность / А. Н. Леонтьев. – М. : Политиздат, 1975. – 304 с.
173. Сарджвеладзе Н. И. Личность и ее взаимодействие с социальной средой / Н. И. Сарджвеладзе. – Тбилиси : Мецниереба, 1998. – 204 с.
174. Хьелл Л. Теории личности / Л. Хьелл, Д. Зиглер. – СПб : Питер, 2000. – 608 с.
175. Холл К. С. Теории личности / К. С. Холл, Г. Линдсей. – М. : КСП+, 1997. – 720 с.
176. Мадди С. Р. Теория личности: сравнительный анализ / Мадди С. Р. ; Д. А. Леонтьев (науч. ред.), И. Ю. Авидон (пер.). – СПб. : Речь, 2002. – 539 с.
177. Якобсон П. М. Психологические проблемы мотивации поведения человека / П. М. Якобсон. – М. : Просвещение, 1969. – 317 с.
178. Изард К. Э. Психология эмоций / К. Э. Изард. – СПб. : Питер, 2000. – 464 с.
179. Симонов П. В. Высшая нервная деятельность человека. Мотивационно-эмоциональные аспекты / П. В. Симонов. – М. : Наука, 1975. – 173 с.
180. Ахутина Т. В. Порождение речи: Нейролингвистический анализ синтаксиса / Т. В. Ахутина. – 2-е изд. – М. : Изд-во ЛКИ, 2007. – 216 с.
181. Выготский Л. С. Мышление и речь / Л. С. Выготский. – 5-е изд., испр. – М. : Лабиринт, 1999. – 352 с.
182. Выготский Л. С. Развитие высших психических функций / Л. С. Выготский. – М. : Изд-во Ак. пед. наук, 1960. – 500 с.
183. Лурия А. Р. Об историческом развитии познавательных процессов / А. Р. Лурия. – М. : Наука, 1974. – 172 с.
184. Лурия А. Р. Поражения мозга и мозговая локализация высших психических функций / А. Р. Лурия // Этапы пройденного пути : науч-

ная автобиография. – М. : Издательство Московского университета, 1982. – С. 110–121, 130–138.

185. Лурия А. Р. Мозг человека и психические процессы / А. Р. Лурия. – М. : Педагогика, 1970. – Т. 2. – 495 с.

186. Лурия А. Р. Маленькая книжка о большой памяти / А. Р. Лурия. – М. : Издательство Московского университета, 1968. – 87 с.

187. Broks P. Into the Silent Land: Travels in Neuropsychology / P. Broks. – London : Atlantic Books, 2003. – 245 p.

188. Адрианов О. С. Узловые вопросы локализации и организации церебральных функций / О. С. Адрианов // Современные аспекты учения о локализации и организации церебральных функций. – М. : Медицина, 1980. – С. 200–215.

189. Вертгеймер М. Продуктивное мышление / Вертгеймер М. ; общ. ред. С. Ф. Горбова, В. П. Зинченко ; вступ. ст. В. П. Зинченко. – М. : Прогресс, 1987. – 336 с.

190. Аладжалова Н. А. Психофизиологические аспекты сверхмедленной ритмической активности головного мозга / Н. А. Аладжалова. – М. : Медицина, 1979. – 216 с.

191. Соколов Е. Н. Цветовое зрение / Е. Н. Соколов, Ч. А. Измайлов. – М., 1984. – 175 с.

192. Соколов Е. Н. Геометрическая модель структуры знания / Е. Н. Соколов, А. Ю. Терехина, С. Б. Ребрик // Вопросы психологии. – 1986. – № 6. – С. 130–138.

193. Айзенк Г. Исследования человеческой психики: почему мы ведем себя именно так / Г. Айзенк, М. Айзенк. – М. : ЭКСМО-Пресс, 2002. – 480 с.

194. Якобсон Р. О. Мозг и язык / Р. О. Якобсон // Избранные работы. – М. : Прогресс, 1985. – С. 270–286.

195. Соссюр Ф. Труды по языкознанию / Ф. Соссюр. – М. : Прогресс, 1977. – 696 с.

196. Блумфилд Л. Язык / Л. Блумфилд. – М. : Прогресс, 1968. – 606 с.

197. Бодуэн де Куртенэ И. А. Опыт теории фонетических альтернатив / И. А. Бодуэн де Куртенэ // Избр. труды по общему языкознанию. – М. : Изд-во АН СССР, 1963. – Т. 1. – С. 265–347.

198. Лингвистический процессор для сложных информационных систем / Ю. Д. Апресян, И. М. Богуславский, Л. Л. Иомдин [и др.]. – М. : Наука, 1992. – 256 с.
199. Леонтьев А. А. Язык, речь, речевая деятельность / А. А. Леонтьев. – М. : Просвещение, 1969. – 211 с.
200. Німчук В. В. Мовознавство на Україні в XIV—XVII ст. / В. В. Німчук. – К. : Наукова думка, 1985. – 222 с.
201. Колобаев В. К. Подвижность границ синтагмы: последствия для лингвистического и методического знания [Электронный ресурс] / В. К. Колобаев. – СПб. : ГМА им. И. И. Мечникова. – Режим доступа: http://old.mechnik.spb.ru/lat_pb3.htm.
202. Хомский Н. Аспекты теории синтаксиса / Н. Хомский. – М. : Издательство Московского университета, 1972. – 259 с.
203. Chomsky N. Syntactic Structures / N. Chomsky // Walter de Gruyter. – 2nd edition, 2002. – 117 p. – ISBN 3110172798.
204. Демьянков В. З. Доминирующие лингвистические теории в конце XX века / В. З. Демьянков // Язык и наука конца 20 века. – М. : Институт языкознания РАН, 1995. – С. 239–320.
205. Лакофф Дж. Когнитивная семантика / Дж. Лакофф // Язык и интеллект. – М. : Прогресс, 1995. – С. 143–184.
206. Kövecses Z. Metaphor: a practical introduction / Z. Kövecses. – New York : Oxford University Press, 2002. – 285 p.
207. Chu Yun. Human Performance on Insight Problem Solving: A Review [Электронный ресурс] / Yun Chu, J. MacGregor // The Journal of Problem Solving. – 2011. – Vol. 3, Iss. 2, Article 6. – Режим доступа: <http://docs.lib.purdue.edu/jps/vol3/iss2/6>.
208. Демьянков В. З. Интерпретация, понимание и лингвистические аспекты их моделирования на ЭВМ / Демьянков В. З. – М. : Изд-во Моск. ун-та, 1989. – 172 с.
209. Лурия А. Р. Высшие корковые функции человека / А. Р. Лурия. – М. : Академический проект, 2000. – 505 с.
210. Рубинштейн С. Л. Проблема сознания и деятельности в советской психологии / С. Л. Рубинштейн // Принципы и пути развития психологии. – М., 1959. – С. 249–255.
211. Леонтьев А. А. Психолингвистические единицы и порождение речевого высказывания / А. А. Леонтьев. – М. : Наука, 1969. – 307 с.

212. Обухова Л. Ф. Неоконченные споры: П. Я. Гальперин и Ж. Пиаже / Л. Ф. Обухова // Психологическая наука и образование. – 1996. – № 1. – С. 32–39.
213. Леонтьев А. А. Основы психолингвистики / А. А. Леонтьев. – М. : Смысл, 1999. – 287 с.
214. Мельчук И. А. Русский язык в модели «Смысл–Текст» / И. А. Мельчук. – М. : Языки русской культуры, 1995. – 682 с.
215. Allen J. F. Natural Language Understanding / J. F. Allen. – 2nd ed. – Benjamin-Cummings Publishing, 1994. – 654 p.
216. Михайлов М. И. Эпос, лирика, драмы как роды литературы [Электронный ресурс] / М. И. Михайлов // Библиотечный каталог российских и украинских диссертаций. – Режим доступа: <http://www.lib.ua-ru.net/diss/cont/183464.html>
217. Демьянков В. З. «Теория речевых актов» в контексте современной лингвистической литературы: (Обзор направлений) / В. З. Демьянков // Новое в зарубежной лингвистике. – М. : Прогресс, 1986. – Вып. 17: Теория речевых актов. – С. 223–235.
218. Hausser R. Foundations of computation linguistics / R. Hausser. – Springer, 1999. – 534 p.
219. Роммейтвейт Р. Слова значения и сообщения / Р. Роммейтвейт // Психолингвистика за рубежом: сб. / Отв. ред. А. А. Леонтьев, Л. В. Сахарный. – М. : Наука, 1972. – С. 53–56, 65–66.
220. Щерба Л. В. Языковая система и речевая деятельность / Л. В. Щерба. – Л. : Наука, 1974. – 428 с.
221. Osgood C. E. Lectures on language performance (Springer Series in Language and Communication) / C. E. Osgood. – New-York, Heidelberg, Berlin: Springer-Verlag, 1980. – Vol. 7. – 276 p.
222. Солсо Р. Когнитивная психология / Р. Солсо. – 6-е изд. – СПб. : Питер, 2006. – 589 с.
223. Ждан А. Н. История психологии. От Античности до наших дней : уч. для вузов / А. Н. Ждан. – 5-е изд., перераб. и доп. – М. : Академический проект, 2004. – 576 с.
224. Пиаже Ж. Речь и мышление ребенка / Ж. Пиаже. – М. : Педагогика-Пресс, 1994. – 528 с.
225. Панов В. И. Психодидактика образовательных систем: теория и практика / В. И. Панов. – СПб. : Питер, 2007. – 352 с.

226. Turing A. Computing Machinery and Intelligence / A. Turing // Mind. – 1950. – Vol. LIX, No. 236. – P. 433–460.
227. Образный відео-комп'ютер око-процесорного типу / В. П. Кожем'яко, Г. Л. Лисенко, А. А. Яровий, А. В. Кожем'яко – Вінниця : Універсум-Вінниця, 2008. – 215 с.
228. Глушков В. М. Введение в кибернетику / Глушков В. М. – К. : Изд-во АНУССР, 1964. – 324 с.
229. Глушков В. М. Алгебра. Языки. Программирование / В. М. Глушков, Г. Е. Цейтлин, В. М. Ющенко ; АН УССР, Ин-т киберн. им. В. М. Глушкова. – К. : Наукова думка, 1989. – 376 с.
230. Pickering A. The cybernetic brain : sketches of another future / A. Pickering. – Chicago : University of Chicago Press, 2010. – 526 p.
231. Широчин В. П. Архитектоника мышления и нейроинтеллект : в 2 кн. – Кн. 1 : Архитектоника мышления и нейроинтеллект; Кн. 2 : Программирование доверия в эволюции интеллекта / под ред. Ю. С. Ковтанюка. – К. : Юниор, 2004. – 560 с.
232. Колмогоров А. Н. Основные понятия теории вероятностей / А. Н. Колмогоров. – 2-е изд. – М. : Наука, 1974. – 120 с.
233. WordNet® [Электронный ресурс] / Princeton University. – Режим доступа: <http://wordnet.princeton.edu>.
234. Крылов С. А. Некоторые уточнения к определениям понятий словоформы и лексемы / С. А. Крылов // Семиотика и информатика. – 1982. – Вып. 19. – С. 118–136.
235. Анисимов А. В. Компьютерная лингвистика для всех: Мифы. Алгоритмы. Язык / А. В. Анисимов. – К. : Наукова думка, 1991. – 208 с.
236. Бардачов Ю. М. Дискретна математика : підручник / Ю. М. Бардачов, Н. А. Соколова, В. Є. Ходаков ; за ред. В. Є. Ходакова. – К. : Вища школа, 2002. – 287 с.
237. Валькман Ю. Р. Образы и образное мышление: некоторые отношения и структуры [Электронный ресурс] / Ю. Р. Валькман // V Междунар. науч.-практ. конф. «Интегрированные модели и мягкие вычисления в искусственном интеллекте», (Коломна, 28–30 мая 2009 г.). – Режим доступа:
<http://raai.org/resurs/papers/kolomna2009/doklad/Valkman.doc>.

238. Жилиякова Л. Ю. Структурирование знаний в ассоциативной модели / Л. Ю. Жилиякова // КИИ-2008, (Дубна, 29 сентября – 3 октября 2008 г.). – 2008. – Т. 1. – С. 104–111.
239. Раскин Л. Г. Нечеткая математика. Основы теории. Приложения / Л. Г. Раскин, О. В. Серая. – Х. : Парус, 2008. – 352 с.
240. Cox E. The fussy systems handbook / E. Cox. – 2nd ed. – Academic Press Professional, 1999. – 716 p.
241. Штовба С. Д. Введение в теорию нечетких множеств и нечеткую логику / [Электронный ресурс] / MATLAB. Exponenta, Fuzzy Logic Toolbox. – Режим доступа:
<http://matlab.exponenta.ru/fuzzylogic/book1/index.php>.
242. Коллатц Л. Функциональный анализ и вычислительная математика / Коллатц Л.; пер. с нем. – М. : Мир, 1969. – 447 с.
243. Кузьмин И. В. Основы теории информации и кодирования / И. В. Кузьмин, В. А. Кедрус. – К. : Вища школа, 1986. – 238 с.
244. Трифонов Е. В. Психофизиология человека [Электронный ресурс] / Е. В. Трифонов. – Режим доступа:
<http://www.tryphonov.narod.ru/tryphonov2/terms2/hemcer.htm>.
245. Краткая анатомия нервной системы. Большие полушария головного мозга [Электронный ресурс]. – Режим доступа:
<http://www.neuro.net.ru/educ/100/ana/an1.html>.
246. Professor Mark Dubin - University of Colorado. Brodmann Areas in the Human Brain with an Emphasis on Vision and Language [Электронный ресурс]. – Режим доступа:
<http://spot.colorado.edu/~dubin/talks/brodmann/brodmann.html>.
247. Автоматизированная система научных исследований ассоциативных экспериментов (АСНИ АЭ) [Электронный ресурс] / CLAIM – научно-образовательный кластер. – Режим доступа: www/URL:
<http://www.philipovich.ru/Projects/ASIS/index.htm>.
248. Столл Р. Множества. Логика. Аксиоматические теории / Р. Столл. – М. : Просвещение, 1968. – 231 с.
249. Беран Л. Упорядоченные множества / Л. Беран. – М. : Наука, 1981. – 64 с.
250. Новиков Ф. А. Дискретная математика для программистов / Ф. А. Новиков. – СПб : Питер, 2001. – 304 с.

251. Черняков А. Г. Онтология времени. Бытие и время в философии Аристотеля, Гуссерля, Хайдеггера / А. Г. Черняков. – СПб. : Высшая религиозно-философская школа, 2001. – 460 с.

252. Успенский В. А. Теорема Гёделя о неполноте / В. А. Успенский. – М. : Наука, 1982. – 110 с.

253. Широков В. А. Феноменологія лексикографічних систем / В. А. Широков ; НАН України, Укр. мов.-інформ. фонд. – К. : Наукова думка, 2004. – 327 с.

254. Нойманн Э. Происхождение и развитие сознания / Э. Нойманн. – К. : Ваклер ; М. : Реал-бук, 1998. – 462 с.

255. Рейнгольд Э. Комбинаторные алгоритмы. Теория и практика / Э. Рейнгольд, Ю. Нивергельд, Н. Део. – М. : Мир, 1980. – 476 с.

256. Грей П. Логика, алгебра и базы данных / П. Грей. – М. : Машиностроение, 1989. – 360 с.

257. Кук Д. Компьютерная математика / Д. Кук, Г. Бейз. – М. : Наука, Гл. ред. физ.-мат. лит., 1990. – 384 с.

258. Цейтлин Г. Е. Введение в алгоритмику / Г. Е. Цейтлин. – К. : Сфера, 1998. – 310 с.

259. Основи дискретної математики : підручник / Ю. В. Капітонова, С. Л. Кривий, О. А. Летичевський [та ін.]. – К. : Наукова думка, 2002. – 579 с.

260. Ахо А. Структуры данных и алгоритмы / А. Ахо, Д. Хопкрофт, Д. Ульман. – М. : Вильямс, 2001. – 384 с.

261. Список функций Statistics Toolbox: расчет парных расстояний между объектами исходного множества данных [Электронный ресурс] / Консультационный центр Matlab компании Softline. – Режим доступа: <http://matlab.exponenta.ru/statist/book2/14/pdist.php>.

262. Заде Л. А. Понятие лингвистической переменной и его применение к принятию приближенных решений / Л. А. Заде. – М. : Мир, 1976. – 165 с.

263. Глонь О. В. Моделювання систем керування в умовах невідзначеності : монографія / О. В. Глонь, В. М. Дубовой. – Вінниця : УНІВЕРСУМ-Вінниця, 2004. – 170 с.

264. Капитонова Ю. В. О некоторых тенденциях развития и проблемах искусственного интеллекта / Ю. В. Капитонова, В. И. Скурихин // Кибернетика и системный анализ. – 1999. – № 1. – С. 43–50.

265. Штовба С. Д. Разработка метода моделирования надежности алгоритмических процессов при неопределенных исходных данных : дис... канд. техн. наук : 05.13.02 / Штовба Сергей Дмитриевич ; Винницкий гос. технический ун-т. – Винница, 1996. – 200 с.

266. Ротштейн О. П. Інтелектуальні технології ідентифікації: нечіткі множини, генетичні алгоритми, нейронні мережі / О. П. Ротштейн. – Вінниця : УНІВЕРСУМ–Вінниця, 1999. – 320 с. – ISBN 966-7199-49-5.

267. Николаев В. И. Системотехника: методы и приложения / В. И. Николаев, В. М. Брук. – Л. : Машиностроение, Ленингр. отд-е, 1985. – 199 с.

268. Пиз А. Язык телодвижений / А. Пиз. – М. : Изд-во Эксмо, 2006. – 272 с.

269. Ефимов В. И. Решатели интеллектуальных задач / В. И. Ефимов. – М. : Наука, Гл. ред. физ.-мат. лит., 1982. – 320 с.

270. Заманский М. Введение в современную алгебру и анализ / М. Заманский. – М. : Наука, 1974. – 488 с.

271. Gritsenko V. Towards the global e-Learning: approaches and solutions / V. Gritsenko // Second International Conference «New Information Technologies in Education for All», Conference Proceedings, (Kiev, 21–23 November 2007). – К. : Akadempriodika, 2007. – P. 10–17.

272. Программа ЮНЕСКО «Информация для всех» в России : руководящие документы [Электронный ресурс]. – Режим доступа:

<http://www.ifap.ru/ofdocs/index.htm>.

273. Вінцюк Т. К. Образний комп'ютер: концепції, методологія, підходи / Т.К. Вінцюк // Праці п'ятої міжн. конф. «УкрОБРАЗ'2000», (Київ, 27 листопада – 1 грудня 2000 р.). – К., 2000. – С. 87–94.

274. The Turing Test Page [Електронний ресурс]. – Режим доступу:

<http://www.fil.ion.ucl.ac.uk/~asaygin/tt/ttest.html#top>.

275. Кобозева И. М. Лингвистическая семантика [Електронний ресурс] / И. М. Кобозева // Диалог. Междунар. конф. по компьютерной лингвистике. Форум. – Режим доступу:

<http://www.dialog-21.ru/forum/actualtopics.aspx?bid=16>.

276. Russian WordNet [Електронний ресурс] / Петербургский государственный университет путей сообщения. – СПб. – Режим доступу: <http://www.pgups.ru/WebWN/wordnet.uix>.

277. Комп'ютерно-орієнтовані системи навчання : зб. наук. праць. – К. : НПУ ім. Драгоманова. – 2001. – Вип. 4. – 230 с.
278. Дистанційне навчання / В. М. Кухаренко, О. В. Рибалко, Т. О. Олійник [та ін.] ; за ред. проф. Кухаренка В. М. – Харків : ХДПУ, 1999. – 216 с.
279. Атанов Г. А. Деятельностный подход в обучении / Г. А. Атанов. – Донецк : «ЕАИ-пресс», 2001. – 160 с. – ISBN 966-7200-13-2.
280. Хювенен Э. Мир Лиспа : в 2 т. – Т. 1: Введение в язык Лисп и функциональное программирование / Э. Хювенен, Й. Сеппянен. – М. : Мир, 1990. – 447 с.
281. Manning C. Introduction to Information retrieval [Електронний ресурс] / Christopher D. Manning, Prabhakar Raghavan, Hinrich Schutse. – Cambridge University Press, 2008. – ISBN 0521865719. – Режим доступу: <http://nlp.stanford.edu/IR-book/>.
282. Спостереження та експеримент. Фізика 7-й клас, 12-ти річна школа, 1 семестр [Електронний ресурс] / ФМГ № 17. – Вінниця. – Режим доступу: <http://disted.edu.vn.ua/courses/learn/1663>.

Додаток А
Приклади застосування формальної теорії Th для
російськомовних речень

1. Забытую песню несет ветерок (в) задумчивых травах звеня
 $(x_1x_2x_3x_4x_5x_6x_7)$.

Згідно з правилом 1 будуємо терм

$$x_1 2x_2 3x_3 4x_4 3x_5 6x_6 7x_7 3;$$

продукція 3.12 до підрядка $x_1 2x_2$ веде до

$$x_2 \setminus x_1 \oplus x_2 3x_3 4x_4 3x_5 6x_6 7x_7 3;$$

продукція 3.12 до підрядка $x_2 3x_3$ веде до

$$x_2 \setminus x_1 \oplus x_3 \setminus x_2 \oplus x_3 4x_4 3x_5 6x_6 7x_7 3;$$

продукція 3.12 до підрядка $x_3 4x_4$ веде до

$$x_2 \setminus x_1 \oplus x_3 \setminus x_2 \oplus x_4 \setminus x_3 \oplus x_4 3x_5 6x_6 7x_7 3;$$

продукція 3.12 до підрядка $x_4 3x_5$ веде до

$$x_2 \setminus x_1 \oplus x_3 \setminus x_2 \oplus x_4 \setminus x_3 \oplus x_3 \setminus x_4 \oplus x_5 6x_6 7x_7 3;$$

продукція 3.12 до підрядка $x_5 6x_6$ веде до

$$x_2 \setminus x_1 \oplus x_3 \setminus x_2 \oplus x_4 \setminus x_3 \oplus x_3 \setminus x_4 \oplus x_6 \setminus x_5 \oplus x_6 7x_7 3;$$

продукція 3.12 до підрядка $x_6 7x_7$ веде до

$$x_2 \setminus x_1 \oplus x_3 \setminus x_2 \oplus x_4 \setminus x_3 \oplus x_3 \setminus x_4 \oplus x_6 \setminus x_5 \oplus x_7 \setminus x_6 \oplus x_7 3;$$

продукція 3.13 до підрядка $x_7 3$ веде до

$$x_2 \setminus x_1 \oplus x_3 \setminus x_2 \oplus x_4 \setminus x_3 \oplus x_3 \setminus x_4 \oplus x_6 \setminus x_5 \oplus x_7 \setminus x_6 \oplus x_3 \setminus x_7 \text{ — маємо терм в АНФ.}$$

Отже, початкова природно-мовна конструкція в АНФ має такий вигляд:

$$\begin{aligned} & \text{песню} \setminus \text{забытую} \oplus \\ & \text{несет} \setminus \text{песню} \oplus \\ & \text{ветерок} \setminus \text{несет} \oplus \\ & \text{несет} \setminus \text{ветерок} \oplus \\ & \text{травах} \setminus \text{задумчивых} \oplus \\ & \text{звеня} \setminus \text{травах} \oplus \\ & \text{несет} \setminus \text{звеня} . \end{aligned}$$

На рис. А1 представлено дерево графа речення з виділенням двох асоціативних пар.

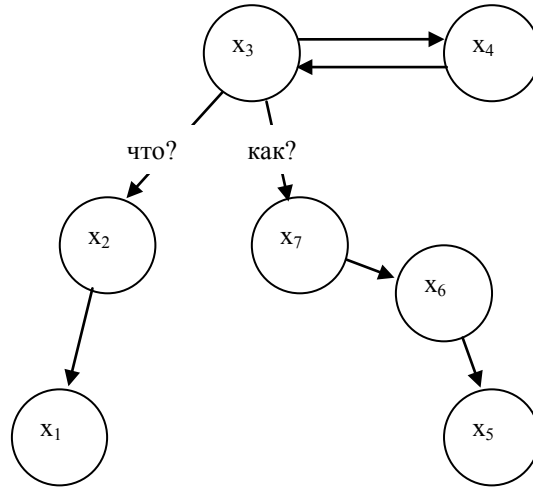


Рисунок А1 – Дерево графа терма $x_1 2x_2 3x_3 4x_4 3x_5 6x_6 7x_7 3$ з виділенням асоціативних пар $x_3 \setminus x_2$ та $x_3 \setminus x_7$

Позначимо $\langle AN\Phi? \rangle := x_3 \setminus x_2$ словом $\langle \text{что?} \rangle$. За теоремою 2 $\langle AN\Phi a \rangle \rightarrow x_2 \setminus x_1$, $\langle AN\Phi q \rangle \rightarrow x_4 \setminus x_3 \oplus x_3 \setminus x_4 \oplus x_6 \setminus x_5 \oplus x_7 \setminus x_6 \oplus x_3 \setminus x_7$.

Тоді, згідно з правилом 2, $\langle tA \rangle \rightarrow x_2 x_1$, а $\langle tQ \rangle \rightarrow x_3 x_4 x_7 x_6 x_5$. Отже, маємо такий результат:

что? несет ветерок звеня задумчивых травах ? песню забытую.

Тепер позначимо $\langle AN\Phi? \rangle := x_3 \setminus x_7$ словом $\langle \text{как?} \rangle$. За теоремою 2 $\langle AN\Phi a \rangle \rightarrow x_7 \setminus x_6 \oplus x_6 \setminus x_5$,

$\langle AN\Phi q \rangle \rightarrow x_4 \setminus x_3 \oplus x_3 \setminus x_4 \oplus x_3 \setminus x_2 \oplus x_2 \setminus x_1$.

Тоді, згідно з правилом 2, $\langle tA \rangle \rightarrow x_7 x_6 x_5$, а $\langle tQ \rangle \rightarrow x_3 x_4 x_2 x_1$. Отже, маємо такий результат:

как? несет ветерок песню забытую ? звеня задумчивых травах.

2. Царь небесный пошлет мне прощенье (за) прегрешенья $(x_1 x_2 x_3 x_4 x_5 x_6)$.

Згідно з правилом 1 будуємо терм

$x_1 3x_2 1x_3 1x_4 3x_5 3x_6 5$;

5 разів продукція 3.12 до підрядка $x_1 3x_2 1x_3 1x_4 3x_5 3x_6$

$x_3 \setminus x_1 \oplus x_1 \setminus x_2 \oplus x_1 \setminus x_3 \oplus x_3 \setminus x_4 \oplus x_3 \setminus x_5 \oplus x_6 5$;

продукція 3.13 до підрядка $x_6 5$

$x_3 \setminus x_1 \oplus x_1 \setminus x_2 \oplus x_1 \setminus x_3 \oplus x_3 \setminus x_4 \oplus x_3 \setminus x_5 \oplus x_5 \setminus x_6$ – маємо терм в АНФ.

Отже, початкова природно-мовна конструкція в АНФ має такий вигляд:

пошлет \ царь \oplus
 царь \ небесный \oplus
 царь \ пошлет \oplus
 пошлет \ мне \oplus
 пошлет \ прощенье \oplus
 прощенье \ прегрешенья.

На рис. А2 представлено дерево графа речення з виділенням двох асоціативних пар.

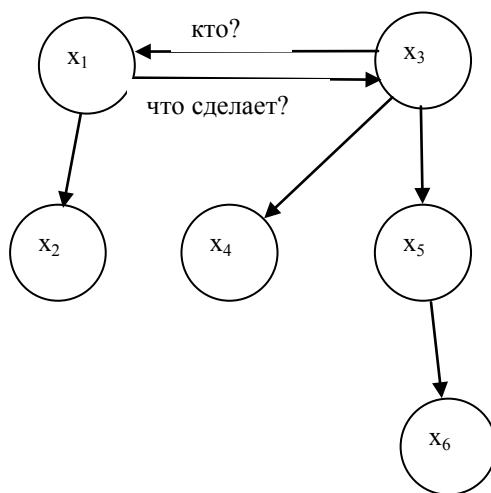


Рисунок А2 – Дерево графа терма $x_1 \setminus x_3 x_2 \setminus x_4 x_1 \setminus x_3 x_5 \setminus x_6$ з виділенням асоціативних пар $x_3 \setminus x_1$ та $x_1 \setminus x_3$

Позначимо $\langle АНФ? \rangle ::= x_3 \setminus x_1$ словом $\langle \underline{\text{кто?}} \rangle$. За теоремою 2 $\langle АНФа \rangle \rightarrow x_3 \setminus x_4 \oplus x_3 \setminus x_5 \oplus x_5 \setminus x_6$, $\langle АНФq \rangle \rightarrow x_1 \setminus x_2$.

Тоді, згідно з правилом 2, $\langle tA \rangle \rightarrow x_3 x_4 x_5 x_6$, а $\langle tQ \rangle \rightarrow x_1 x_2$. Отже, маємо такий результат:

кто? пошлет мне прощенье прегрешенья ? царь небесный.

Тепер позначимо $\langle АНФ? \rangle ::= x_1 \setminus x_3$ словом $\langle \underline{\text{что сделает?}} \rangle$. За теоремою 2 $\langle АНФа \rangle \rightarrow x_1 \setminus x_2$, $\langle АНФq \rangle \rightarrow x_3 \setminus x_4 \oplus x_3 \setminus x_5 \oplus x_5 \setminus x_6$.

Тоді, згідно з правилом 2, $\langle tA \rangle \rightarrow x_1 x_2$, а $\langle tQ \rangle \rightarrow x_3 x_4 x_5 x_6$. Отже, маємо такий результат:

что сделает? царь небесный ? пошлет мне прощенье прегрешенья.

Додаток Б

Наскрізний тестовий приклад даних з тематики «WEB-технології: стандарти Semantic WEB»

*/ Первинний текст елемента ЕК – ч. I ППЗ «WEB–технології: стандарти Semantic WEB»

Семантичною павутиною називають нову концепцію розвитку всесвітньої мережі Інтернет. Концептуальна надбудова над існуючою Всесвітньою павутиною створена і впроваджується консорціумом W3C під головуванням Тіма Бернерса-Лі. Комп'ютери поки що не можуть сприйняти та осмислити текстову інформацію, розміщену в Інтернеті. Слово «семантична» у визначенні павутини означає «осмислена», «зрозуміла» для комп'ютерів щодо змісту розміщеної в мережі інформації.

В даний час комп'ютери беруть досить обмежену участь у формуванні й обробці інформації в мережі Інтернет. Людина, як і раніше, створює інформацію, дає її оцінку, будує класифікацію. Вся інформація в Інтернеті повинна розміщуватись двома мовами: людською мовою для людини і комп'ютерною мовою для розуміння комп'ютера. Семантична павутина – це концепція мережі, у якій кожен ресурс людською мовою був би доповнений описом, зрозумілим комп'ютеру.

Semantic web можна представити як два напрями, перше з яких охоплює мови представлення даних. Натепер основними такими мовами є мова розмітки XML і Засоби Опису Ресурсів RDF, що мають статус рекомендацій W3C. RDF був затверджений як стандарт W3C у лютому 2004 року. Формат RDF призначений для збереження метаданих. Метадані – це дані про дані. Опис у форматі RDF повинен прикріплюватися до кожного ресурсу мережі Інтернет. Другий концептуальний напрямок несе в собі теоретичне уявлення про моделі предметних областей або онтології. 10 лютого 2004 року консорціумом W3C була затверджена й опублікована специфікація мови мережевих онтологій OWL.

Дві гілки Семантичного Web використовують три ключові мови:

- мова XML дозволяє визначити синтаксис і структуру документів;
- механізм опису ресурсів RDF забезпечує модель кодування для значень, визначених в онтології;

▪ мова онтологій OWL дозволяє визначати поняття і відносини між ними.

Технологія Semantic web базується на чотирьох логічних рівнях:

1. Уніфікований ідентифікатор URI визначає спосіб запису адреси довільного ресурсу. За допомогою URI-ідентифікатора семантичний Web дає можливість кожному висловлювати ті поняття, якими він користується. Типовими прикладами URI-ідентифікаторів є URL-адреси. URI-ідентифікатор задаючи або посилаючись на деякий ресурс, не обов'язково при цьому вказує на його місцезнаходження в Internet.

2. Мова XML як базова форма розмітки та засоби, що призначені для визначення й опису класів XML-документів.

3. На базі XML розгортаються засоби опису ресурсів RDF. RDF-схеми пояснюють, як узгодити XML-дані в мережі, будувати каталоги і словники понять. RDF дозволяє виконувати пошук необхідних понять в семантичному Web.

4. Мова мережевих онтологій OWL призначена для опису класів, об'єктів і відносин між ними. OWL забезпечує більш повну автоматичну обробку мережевого контенту у порівнянні з XML і RDF. Окрема онтологія визначає семантику конкретної предметної області і сприяє встановленню зв'язків між значеннями її елементів.

*/ Адаптований текст елемента ЕК

- 1 семантичною павутиною називають нову концепцію розвитку всесвітньої мережі Інтернет
- 2 концептуальна надбудова існуючої Всесвітньої павутиною створена впроваджується консорціумом W3C головуванням Тім_Бернерс_Лі
- 3 комп'ютери поки не можуть сприйняти осмислити текстову інформацію розміщену Інтернеті
- 4 слово «семантична» визначенні павутини означає «осмислена» «зрозуміла» комп'ютерів змісту розміщеної мережі інформації
- 5 даний час комп'ютери беруть досить обмежену участь формуванні обробці інформації мережі Інтернет
- 6 людина як раніше створює інформацію дає оцінку будує класифікацію

- 7 вся інформація Інтернеті повинна розміщуватись двома мовами людською мовою людини комп'ютерною мовою розуміння комп'ютера
- 8 семантична павутина є концепція мережі кожен ресурс людською мовою був_би доповнений описом зрозумілим комп'ютеру
- 9 Semantic web можна представити два напрями перше охоплює мови представлення даних
- 10 натепер основними такими мовами є мова розмітки XML Засоби Опису Ресурсів RDF мають статус рекомендацій W3C
- 11 RDF був затверджений стандарт W3C лютому 2004 року
- 12 формат RDF призначений збереження метаданих
- 13 метадані означають дані відносно даних
- 14 опис форматі RDF повинен прикріплюватися кожного ресурсу мережі Інтернет
- 15 другий концептуальний напрямок несе теоретичне уявлення моделі предметних областей онтології
- 16 10 лютого 2004 року консорціумом W3C була затверджена опублікована специфікація мови мережевих онтологій OWL
- 17 дві гілки Семантичного Web використовують три ключові мови
- 18 мова XML дозволяє визначити синтаксис структуру документів
- 19 механізм опису ресурсів RDF забезпечує модель кодування значень визначених онтологією
- 20 мова онтологій OWL дозволяє визначати поняття відношення поняттями
- 21 технологія Semantic web базується чотирьох логічних рівнях
- 22 уніфікований ідентифікатор URI визначає спосіб запису адреси довільного ресурсу
- 23 допомогою URI ідентифікатора семантичний Web дає можливість кожному Y висловлювати поняття користується
- 24 типовими прикладами URI ідентифікаторів є URL адреси
- 25 URI ідентифікатор задаючи посилаючись деякий ресурс не обов'язково вказує місцезнаходження даного Internet ресурсу
- 26 мова XML є базова форма розмітки засоби призначені визначення опису класів документів XML

- 27 бази XML розгортаються засоби опису ресурсів RDF
- 28 схеми RDF пояснюють узгодити дані XML мережі будувати каталоги словники понять
- 29 RDF дозволяє виконувати пошук необхідних понять семантичному Web
- 30 мова мережових онтологій OWL призначена опису класів об'єктів відношень об'єктами класами
- 31 OWL забезпечує більш повну автоматичну обробку мережевого контенту порівнянні XML RDF
- 32 окрема онтологія визначає семантику конкретної предметної області сприяє встановленню зв'язків значеннями елементів онтології

*/ Дані таблиць бази даних SQLite – записи таблиці *Event* (всього 32)

event_id	i_s	text_id	syntagma
19	65	{...}	1 механізм опису ресурсів XML забезпечує модель кодування значень визначених онтологією
20	66	{...}	1 мова онтологій OWL дозволяє визначити поняття відношення поняттями
21	67	{...}	1 технологія Semantic web базується чотирьох логічних рівнях
22	68	{...}	1 уніфікований ідентифікатор URI визначає спосіб запису адреси довільного ресурсу
23	69	{...}	1 допомогою URI ідентифікатора семантичний Web дає можливість кожному Y висловлювати поняття користуватися
24	70	{...}	1 типовими прикладами URI ідентифікаторів є URL адреси
25	71	{...}	1 URI ідентифікатор задаючи посилання на деякий ресурс не обов'язково вказує місцезнаходження даного Internet ресурсу
26	73	{...}	1 мова XML є базова форма розмітки засоби призначені визначення опису класів документів XML
27	75	{...}	1 бази XML розгортаються засоби опису ресурсів RDF
28	76	{...}	1 схеми RDF пояснюють узгодити дані XML мережі будувати каталоги словники понять
29	77	{...}	1 RDF дозволяє виконувати пошук необхідних понять семантичному Web
30	79	{...}	1 мова мережових онтологій OWL призначена опису класів об'єктів відношень об'єктами класами
31	80	{...}	1 OWL забезпечує більш повну автоматичну обробку мережевого контенту порівнянні XML RDF
32	81	{...}	1 окрема онтологія визначає семантику конкретної предметної області сприяє встановленню зв'язків значеннями елементів онтології

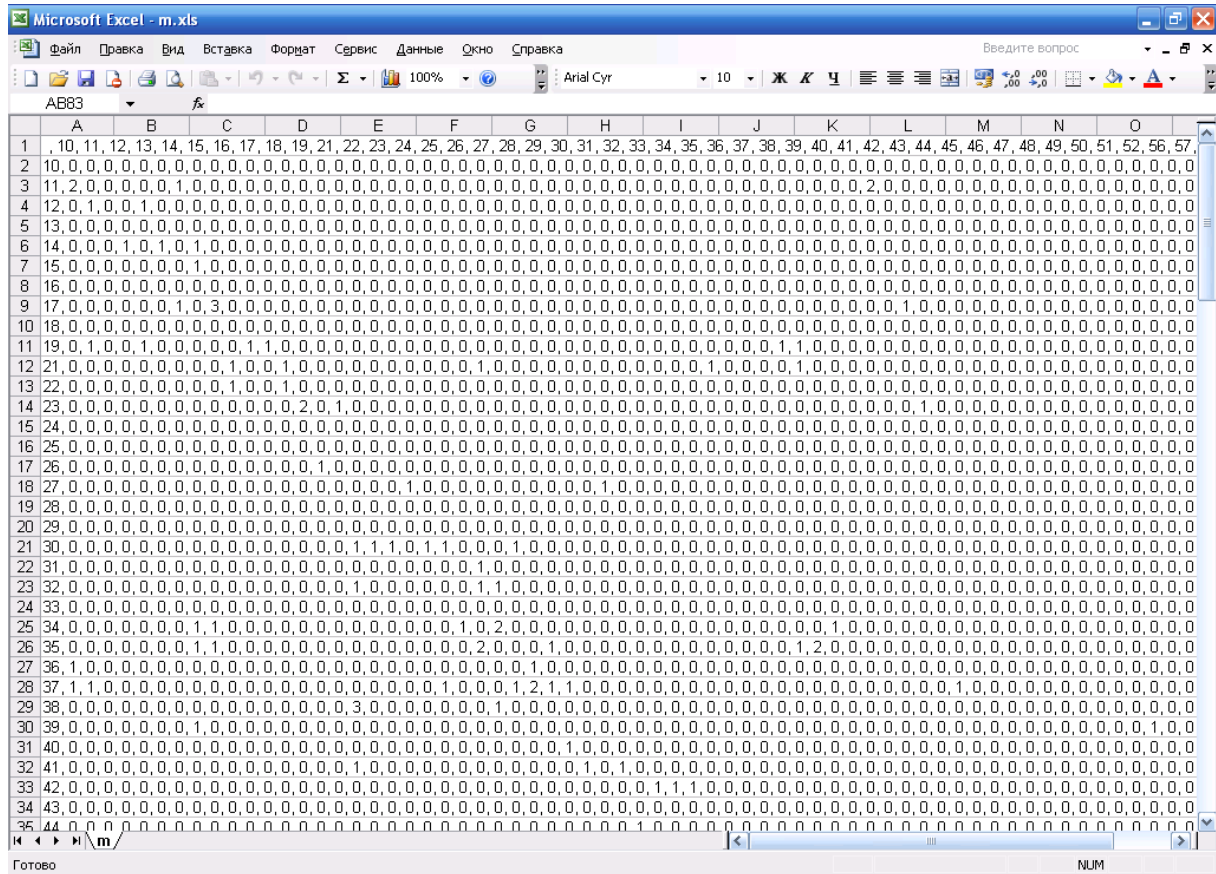
*/ Записи таблиці *Image* (всього 143)

	bi_i	object_quality	object	notion	method	method_quality
1	10	семантичний	Semantic	семантика	{null}	семантично
2	11	{null}	павутина	{null}	{null}	{null}
3	12	названий	{null}	назва	називати	{null}
4	13	новий	{null}	новина	{null}	ново
5	14	концептуальний	{null}	концепція	{null}	{null}
6	15	розвинутий	{null}	розвиток	розвивати	{null}
7	16	всесвітній	всесвіт	{null}	{null}	{null}
8	17	мережевий	мереживо	мережа	мережити	{null}
9	18	{null}	Інтернет	Internet	{null}	{null}
10	19	збудований	будова	надбудова	будувати	{null}
11	21	створений	{null}	створіння	створити	{null}
12	22	{null}	{null}	впровадження	впровадити	{null}
13	23	{null}	{null}	консорціум	{null}	{null}
14	24	{null}	W3C	W3C	{null}	{null}
15	25	{null}	Тім Берн ерс Лі	{null}	{null}	{null}
16	26	головний	голова	{null}	головувати	{null}
17	27	комп'ютерний	комп'ютер	{null}	{null}	{null}
18	28	{null}	{null}	{null}	{null}	поки
19	29	ніякий	ні	заперечення	заперечувати	ніяк
20	30	можливий	{null}	можливість	могти	можна
21	31	{null}	{null}	сприйняття	сприймати	{null}
22	32	{null}	{null}	осмислення	осмислювати	{null}
23	33	текстовий	текст	{null}	{null}	{null}
24	34	інформаційний	{null}	інформація	інформувати	{null}
25	35	змістовний	зміст	розміщення	розміщувати	змістовно
26	36	словесний	слово	слово	висловлювати	{null}
27	37	визначений	{null}	означення	означати	{null}
28	38	розумний	{null}	розум	зрозуміти	розумно
29	39	даний	дані	даність	{null}	{null}
30	40	часовий	{null}	час	{null}	{null}
31	41	відібраний	{null}	{null}	брати	{null}
32	42	{null}	учасник	участь	{null}	{null}
33	43	{null}	{null}	{null}	{null}	досить
34	44	обмежений	{null}	обмеження	обмежувати	{null}
35	45	сформований	форма	формування	формувати	{null}
36	46	{null}	робота	оброблення	обробляти	{null}
37	47	який	{null}	якість	{null}	як
38	48	ранішній	{null}	ранок	{null}	раніше
39	49	{null}	{null}	{null}	давати	{null}
40	50	цінний	{null}	оцінка	оцінювати	{null}
41	51	класифікаційний	клас	класифікація	класифікувати	{null}
42	52	людський	людина	{null}	{null}	по людськи
43	56	мовний	{null}	мова	мовити	{null}
44	57	весь	{null}	все	{null}	{null}
45	59	повинний	{null}	повинність	повинен	{null}
46	71	існуючий	{null}	існування	є	{null}
47	72	кожний	{null}	{null}	{null}	{null}

48	73	ресурсний	{null}	ресурс	{null}	{null}
49	74	{null}	{null}	бутність	бути	{null}
50	75	повний	{null}	доповнення	доповнити	повно
51	76	описаний	{null}	опис	писати	{null}
52	77	{null}	{null}	представлення	представляти	{null}
53	79	{null}	веб	web	{null}	{null}
54	80	{null}	напрямок	направлення	направляти	{null}
55	81	перший	один	першість	{null}	1
56	82	другий	два	2	{null}	{null}
57	83	{null}	{null}	охоплення	охопити	{null}
58	84	теперішній	тепер	натепер	{null}	{null}
59	85	{null}	маєток	маєтність	мати	{null}
60	86	{null}	статут	статус	стати	{null}
61	87	{null}	RDF	RDF	{null}	{null}
62	88	{null}	XML	XML	{null}	{null}
63	89	розмічений	{null}	розмітка	розмітити	{null}
64	90	{null}	{null}	рекомендація	рекомендувати	{null}
65	91	основний	{null}	основа	основувати	{null}
66	92	такий	{null}	так	такати	{null}
67	93	{null}	{null}	засіб	{null}	{null}
68	94	стандартний	{null}	стандарт	стандартизувати	стандартно
69	95	річний	{null}	рік	{null}	щорічно
70	96	{null}	{null}	2004	{null}	{null}
71	97	затверджений	{null}	затвердження	затвердити	затверджено
72	98	{null}	лютий	лютий	{null}	{null}
73	99	призначений	{null}	призначення	призначати	{null}
74	100	збережений	{null}	збереження	зберігати	{null}
75	101	{null}	метадані	метадані	{null}	{null}
76	102	{null}	{null}	відношення	відносити	відносно
77	103	прикріплений	{null}	прикріплення	прикріплювати	{null}
78	105	{null}	{null}	{null}	нести	{null}
79	106	теоретичний	{null}	теорія	теоретизувати	теоретично
80	107	модельний	{null}	модель	моделювати	{null}
81	108	обласний	{null}	область	{null}	{null}
82	109	{null}	{null}	уявлення	уявляти	{null}
83	110	предметний	{null}	предмет	{null}	{null}
84	111	онтологічний	{null}	онтологія	{null}	{null}
85	112	{null}	OWL	OWL	{null}	{null}
86	113	{null}	10	десять	{null}	{null}
87	114	специфікований	{null}	специфікація	{null}	{null}
88	115	опублікований	{null}	публікація	опублікувати	{null}
89	116	{null}	3	три	{null}	{null}
90	117	ключовий	ключ	{null}	включати	{null}
91	118	гілковий	гілка	{null}	{null}	{null}
92	119	корисний	{null}	користь	використовувати	{null}
93	120	синтаксичний	{null}	синтаксис	{null}	синтаксично
94	121	структурни	{null}	структура	структурувати	структурно
95	122	документальний	{null}	документ	документувати	документально
96	123	дозвільний	{null}	дозвіл	дозволяти	дозволено
97	124	механічний	{null}	механізм	механізувати	механічно
98	125	закодований	{null}	кодування	кодувати	{null}

99	126	забезпечений	{null}	забезпечення	забезпечувати	{null}
100	127	поняттєвий	{null}	поняття	{null}	{null}
101	128	базовий	база	базування	базуватися	базово
102	129	{null}	4	чотири	{null}	{null}
103	130	логічний	логік	логіка	{null}	логічно
104	131	рівний	{null}	рівень	рівняти	рівно
105	132	технологічний	{null}	технологія	{null}	технологічно
106	133	{null}	{null}	спосіб	{null}	{null}
107	134	{null}	URI	URI	{null}	{null}
108	135	ідентифікований	ідентифікатор	ідентифікація	ідентифікувати	{null}
109	136	уніфікований	{null}	уніфікація	уніфікувати	{null}
110	137	записаний	{null}	запис	записувати	{null}
111	138	довільний	{null}	довільність	{null}	довільно
112	139	адресний	адреса	{null}	адресувати	адресно
113	140	{null}	невідомий object	Y	{null}	{null}
114	141	{null}	{null}	Z	невідомий-method	{null}
115	142	{null}	помічник	допомога	допомогти	{null}
116	143	типовий	{null}	тип	{null}	типово
117	144	прикладний	{null}	приклад	прикладати	{null}
118	145	{null}	URL	URL	{null}	{null}
119	146	{null}	{null}	посилання	посилати	{null}
120	147	деякий	{null}	{null}	{null}	{null}
121	148	обовязковий	{null}	обовязок	зобовязати	обовязково
122	149	вказаний	вказівник	{null}	вказати	вказано
123	150	{null}	{null}	місцезнаходження	{null}	{null}
124	151	розгорнутий	розгортка	розгортання	розгортати	розгорнуто
125	152	словниковий	словник	{null}	{null}	{null}
126	153	{null}	{null}	каталог	каталогизувати	{null}
127	154	схематичний	{null}	схема	схематизувати	схемно
128	155	узгоджений	{null}	узгодження	узгодити	погоджено
129	156	ясний	{null}	пояснення	пояснювати	ясно
130	157	пошуковий	спошукач	пошук	шукати	{null}
131	158	виконаний	виконавець	виконання	виконувати	виконано
132	159	необхідний	{null}	необхідність	{null}	необхідно
133	160	об'єктивний	{null}	об'єкт	{null}	об'єктивно
134	161	{null}	{null}	порівняння	порівнювати	{null}
135	162	більший	{null}	більшість	збільшити	Більше
136	163	автоматичний	автомат	{null}	автоматизувати	автоматично
137	164	контентний	{null}	контент	{null}	Контентно
138	165	встановлений	{null}	встановлення	встановити	встановлено
139	166	конкретний	{null}	конкретність	конкретизувати	Конкретно
140	167	окремий	{null}	{null}	відокремити	Окремо
141	168	зв'язний	{null}	зв'язок	зв'язувати	Зв'язно
142	169	сприятливий	{null}	сприяння	сприяти	{null}
143	170	елементарний	{null}	елемент	{null}	елементарно

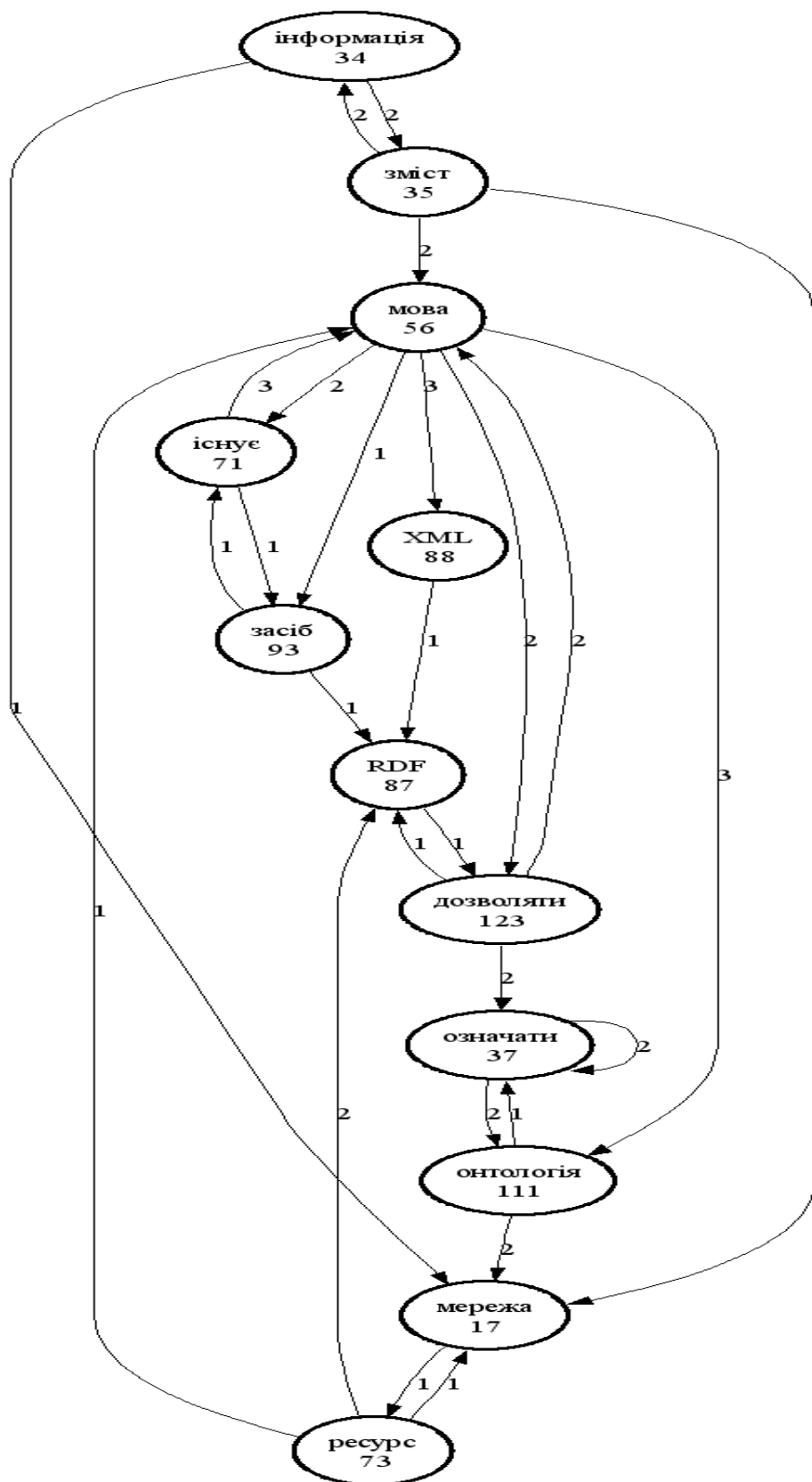
*/ Фрагмент матриці суміжності (143×143) для графа АМО



*/ Отримання значень функції сенсу-параметра на основі даних прикладу графа (п.3.1)

	17	34	35	37	56	71	73	87	88	93	111	123	
17	0,16	0,16	0,16	0,16	0,16	0,16	0,35	0,16	0,16	0,16	0,16	0,16	0,35
34	0,35	0,16	0,59	0,16	0,16	0,16	0,16	0,16	0,16	0,16	0,16	0,16	0,94
35	0,35	0,59	0,16	0,16	0,59	0,16	0,16	0,16	0,16	0,16	0,16	0,16	1,53
37	0,16	0,16	0,16	0,59	0,16	0,16	0,16	0,16	0,16	0,16	0,59	0,16	1,18
56	0,16	0,16	0,16	0,16	0,16	0,59	0,16	0,16	0,80	0,35	0,80	0,59	3,126
71	0,16	0,16	0,16	0,16	0,80	0,16	0,16	0,16	0,16	0,35	0,16	0,16	1,15
73	0,35	0,16	0,16	0,16	0,35	0,16	0,16	0,59	0,16	0,16	0,16	0,16	1,29
87	0,16	0,16	0,16	0,16	0,16	0,16	0,16	0,16	0,16	0,16	0,16	0,35	0,35
88	0,16	0,16	0,16	0,16	0,16	0,16	0,16	0,35	0,16	0,16	0,16	0,16	0,35
93	0,16	0,16	0,16	0,16	0,16	0,35	0,16	0,35	0,16	0,16	0,16	0,16	0,70
111	0,59	0,16	0,16	0,35	0,16	0,16	0,16	0,16	0,16	0,16	0,16	0,16	0,94
123	0,16	0,16	0,16	0,59	0,59	0,16	0,16	0,35	0,16	0,16	0,16	0,16	1,53
	1,63	0,59	0,59	1,53	2,33	0,94	0,35	1,63	0,80	0,70	1,39	0,94	13,42

*/ Інтерпретація прикладу графа (п. 3.1) як підграфа загального графа АМО



Додаток В

Підтримка базових функцій пошуку та генерації знань СОПМК: тексти програм і результати тестування

*/ Програма для моделювання обходу графа АМО (алгоритм *Find · n*)

```
uses crt;
label 1,2,3,4,5;
const n=12; { kol. verшин }
var
  F : Text;
  A:array[1..n,1..n] of byte; { graf }
  V:array[1..n] of byte;
  Vmin:array[1..n] of byte;
  vix,vz:array[1..n] of byte; { kol. vixodov iz verшин }
  vx:array[1..n] of byte; { kol. vxodov v verшин }
  i, { nomer verшинi poiska (ctroka) }
  j, { nomer ctroki, chto a[i,j]<>0 }
  priz,
  k,kn,kp,kmin:byte; { dlja verшин puti }
  s, { nachalo puti }
  t, { konetch puti }
  vp, { verшина poiska }
  Sp, { ctويمoct puti }
  Smin:byte;
begin
  Assign(F, 'Output.txt');
  ReWrite(F);
  for i:=1 to n do
  for j:=1 to n do
  a[i,j]:=0;
  a[1,7]:=4-1; { graf c pomotchju spicok reber}
  a[2,1]:=4-1; a[2,3]:=4-2;
  a[3,1]:=4-1; a[3,2]:=4-2; a[3,5]:=4-2;
  a[4,4]:=4-2; a[4,11]:=4-2;
  a[5,6]:=4-2; a[5,9]:=4-3; a[5,10]:=4-1; a[5,11]:=4-3; a[5,12]:=4-2;
  a[6,5]:=4-3; a[6,10]:=4-1;
  a[7,1]:=4-1; a[7,5]:=4-1; a[7,8]:=4-2;
  a[8,12]:=4-1;
  a[9,8]:=4-1;
  a[10,6]:=4-1; a[10,8]:=4-1;
  a[11,1]:=4-2; a[11,4]:=4-1;
  a[12,4]:=4-2; a[12,5]:=4-2; a[12,8]:=4-1;
  for i:=1 to n do begin vix[i]:=0; vx[i]:=0; vz[i]:=0; end;
  for i:=1 to n do
  for j:=1 to n do
  if a[i,j]<>0 then begin vix[i]:=vix[i]+1; vz[i]:=vz[i]+1; end;
  for j:=1 to n do
  for i:=1 to n do
  if a[i,j]<>0 then vx[j]:=vx[j]+1;
  Write(' s='); readln(s);
  Write(' t='); readln(t);
  Smin:=100; k:=1; v[1]:=s;
```

```

Sp:=0; vp:=s; j:=1;
if vx[t]=0
  then begin writeln(F,' net puti v ',t:3); goto 4; end;
if vix[s]=0
  then begin writeln(F,' net puti ot ',s:3); goto 4; end;
l:i:=vp;
while a[i,j]=0 do { poisk nenulevogo elementa }
  j:=j+1;
if j>n then
  begin
    { net vixoda -vozvrat nazad }
    kn:=1;
    while (kn<=k-1) and (vix[kn]=0) do
      kn:=kn+1;
    if kn>k-1 then goto 3;
    kp:=k-1; sp:=sp-a[v[k-1],v[k]];
    if v[k]+1>n then
      begin kp:=k-2; sp:=sp-a[v[k-2],v[k-1]]; end;
    vp:=v[kp]; j:=v[kp+1]+1;
    for kn:=kp to k-1 do
      vix[v[kn]]:=vz[v[kn]];
    k:=kp;
    goto 1; { prodolgat poick }
  end;
{ a[i,j]<>0 }
kn:=1;
while v[kn]<>j do { vt imeetcja v spicke V ? }
  kn:=kn+1;
if kn>k
  then { vt ne imeetcja v V }
    begin
      vix[i]:=vix[i]-1;
      sp:=sp+a[i,j];
      k:=k+1;
      v[k]:=j; { vkljuchit vershinu v epicok V }
      if v[k]=t { eto konec puti ? }
        then { da }
          begin
            for kp:=1 to k do { pechat put i ego stoimost }
              write(F,V[kp]:3); writeln(F,' ',sp:3);
            if sp<smin then
              begin
                smin:=sp; kmin:=k;
                for kp:=1 to k do
                  vmin[kp]:=v[kp];
                end;
            kn:=1;
            while (kn<=k-1) and (vix[kn]=0) do
              kn:=kn+1;
            if kn>k-1 then goto 3;
            kp:=k-1; sp:=sp-a[v[k-1],v[k]];
            if v[k]+1>n then
              begin kp:=k-2; sp:=sp-a[v[k-2],v[k-1]]; end;
              while (kp>1) and (vix[kp]=0) do

```

```

        begin
            sp:=sp-a[v[kp-1],v[kp]];
            kp:=kp-1;
        end;
        vp:=v[kp]; j:=v[kp+1]+1;
        for kn:=kp to k-1 do
            vix[v[kn]]:=vz[v[kn]];
            k:=kp;
            goto 1; { prodolgat poick }
        end
    else { vt ne konec puti }
        begin
            vp:=v[k]; j:=1;
            goto 1; { prodolgat poick }
        end;
end
else { vt imeetja v spicke V }
begin
    vp:=v[k];
    j:=j+1;
    goto 1; { prodolgat poick }
end;
3: writeln(F,'----- opt -----');
for kp:=1 to kmin do
    write(F,vmin[kp]:3); writeln(F,' Smin=',smin:2);
Close(F);
4: readln
end.

```

*/ Результати виконання програми за алгоритмом *Find · n*.

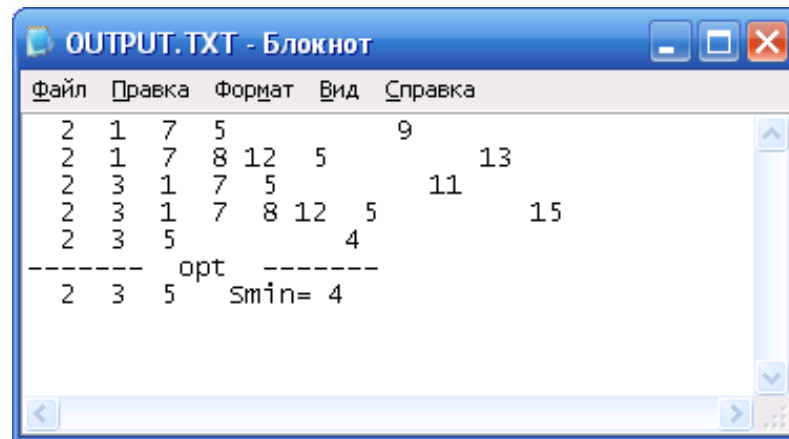
Результат визначення найкоротшого шляху з вершини $s=3$ до вершини $t=8$:

```

OUTPUT.TXT - Блокнот
Файл  Правка  Формат  Вид  Справка
3 1 7 5 6 10 8      17
3 1 7 5 9 8        13
3 1 7 5 10 8       15
3 1 7 5 12 8       14
3 1 7 8            8
3 2 1 7 5 6 10 8   19
3 2 1 7 5 9 8     15
3 2 1 7 5 10 8    17
3 2 1 7 5 12 8    16
3 2 1 7 8         10
3 5 6 10 8        10
3 5 9 8           6
3 5 10 8          8
3 5 11 1 7 8     10
3 5 12 4 11 1 7 8 15
3 5 12 8          7
----- opt -----
3 5 9 8      Smin= 6

```

*/ Результат визначення найкоротшого шляху з вершини $s=2$ до вершини $t=4$:



*/ Текст основних програмних модулів комплексу в оболонці DrScheme мови Лісп для побудови «піраміди сенсу» та реалізації діалогу

```
(define *links-syntagma* (make-immutable-hash-table '(
(Какой? 1) (Какая? 1) (Какое? 1) (Какие? 1) (Каким? 1)
(Что_делать? 2) (Что_делает? 2) (Что_делают? 2) (Что_делаешь? 2) (Что_сделал? 2)
(Что_будешь_делать? 2) (Что_делал? 2) (Кто? 3) (Что_? 3) (Где? 4) (Откуда? 4) (Куда? 4)
(Когда? 5) (В_какое_время? 5) (Сколько_времени? 5) (Как_долго? 5)
(Как? 6) (Каким_образом? 6) (Сколько? 6) (Отчего? 6) (Почему? 6)
(Кого? 7) (Что? 7) (Чего? 7) (Кому? 7) (Чему? 7) (Кем? 7) (Чем? 7) (О_ком? 7) (О_чем? 7)
)))
```

```
(define (answer-1 i1 i2 q a-m l-m)
  (let ((assoc-list (find-associations (normalize-image-id i1) a-m))
        (result '()))
    (for-each
     (lambda (a)
       (when (eq? (first a) (normalize-image-id i2))
         (for-each
          (lambda (e)
            (for-each
             (lambda (p)
               (for-each
                (lambda (i)
                  (for-each
                   (lambda (s)
                     (when (= (normalize-link (first s)) (normalize-link q))
                       (set! result (append result (list (find-word (second s) (normalize-link q))))))
                   (rest i)))
                  p)))
              (hash-table-get l-m e (lambda () '()))))
          (second a))))
    assoc-list)
```

```

result))

(define (make-answerer answer-function q a-m l-m)
  (lambda (i1 i2)
    (answer-function i1 i2 q a-m l-m)))

(define ?Кого (make-answerer answer-1 'Кого? test-a-m *long-memory*))
(define ?Сколько (make-answerer answer-1 'Сколько? test-a-m *long-memory*))
(define ?Почему (make-answerer answer-1 'Почему? test-a-m *long-memory*))
(define ?Куда (make-answerer answer-1 'Куда? test-a-m *long-memory*))

(define (find-word img link-type)
  (if (syntagma? img)
      (list link-type (syn->sentence img)
              ((link->find (normalize-link link-type)) (find-sublist img *image*))))
      (find-word (normalize-link link-type) (find-sublist img *image*)))

(define (print-assoc l)
  (let ((assoc-s «"))
    (for-each
      (lambda (assoc)
        (set! assoc-s
              (string-append assoc-s (format «~a[~a] «
                                              (find-word (first assoc) 3)
                                              (second assoc)))))) l)assoc-s))

(define (print-assoc-memory a-m)
  (hash-table-for-each a-m
    (lambda (k v)
      (display (format «~a -> ~a~%» (find-word k 3) (print-assoc v))))))

```

***/ Тестовий приклад за мотивами Б. Окуджави**

```

(define *image* '(
i1 (  ()  ()  ВЕРНУЛСЯ ВЕРНУЛСЯ () )
i2 (  ()  ()  ВСЕ () ВСЕ )
i3 (  ГОЛУБОЙ () () () )
i4 (  ()  ДЕВОЧКА () () () )
i5 (  ()  ДЕВУШКА () () () )
i6 (  ()  ДРУГАЯ () () () )
i7 (  ()  ЖЕНИХОВ () () () )
i8 (  ()  ЖЕНЩИНА () () () )
i9 (  ()  ()  ЛЕТИТ ЛЕТИТ () )
i10 (  ()  ()  МАЛО () МАЛО )
i11 (  ()  МУЖ МУЖЕСТВО () () )
i12 (  ()  ()  НЕТ НЕТ () )
i13 (  ()  ()  ПЛАЧЕТ ПЛАЧЕТ () )
i14 (  ()  ()  ПРОЖИЛА ПРОЖИЛА ПРОЖИТО )
i15 (  ()  СТАРУХА () СТАРИТЬ () )
i16 (  ()  ()  УЛЕТЕЛ УЛЕТЕЛ () )
i17 (  ()  ()  УТЕШАЮТ УТЕШАЮТ () )
i18 (  ()  ()  УШЕЛ УШЕЛ () )
i19 (  ()  ШАРИК () () () )
True ( ()  ()  ДЕЙСТВИТЕЛЬНО () ДЕЙСТВИТЕЛЬНО ))

```

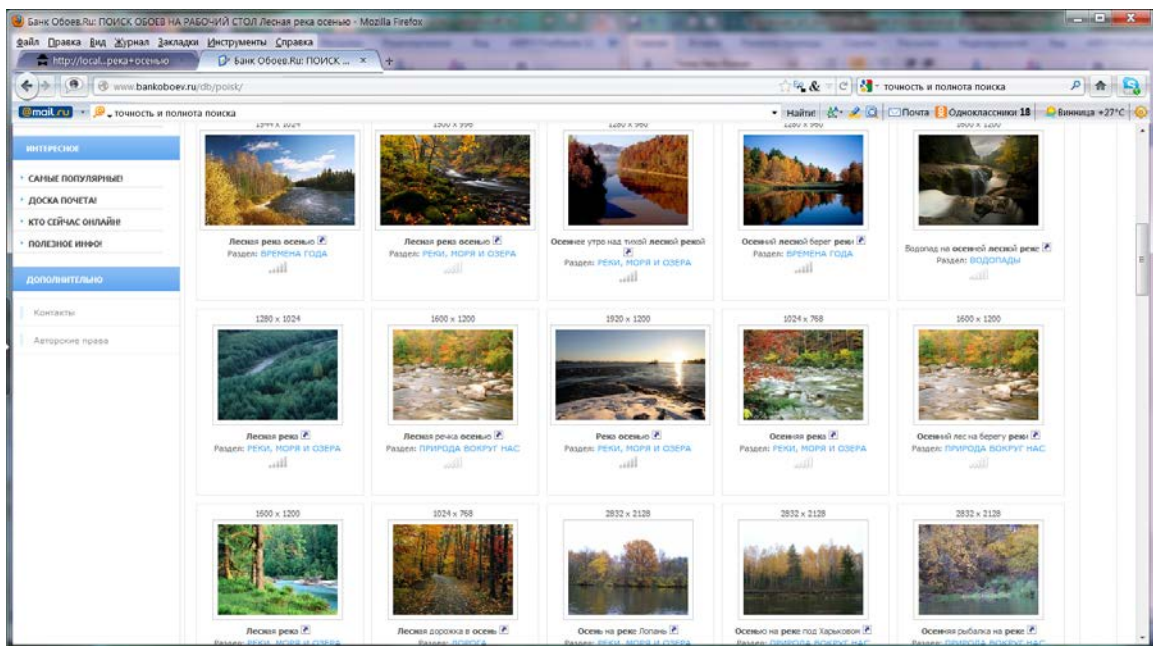
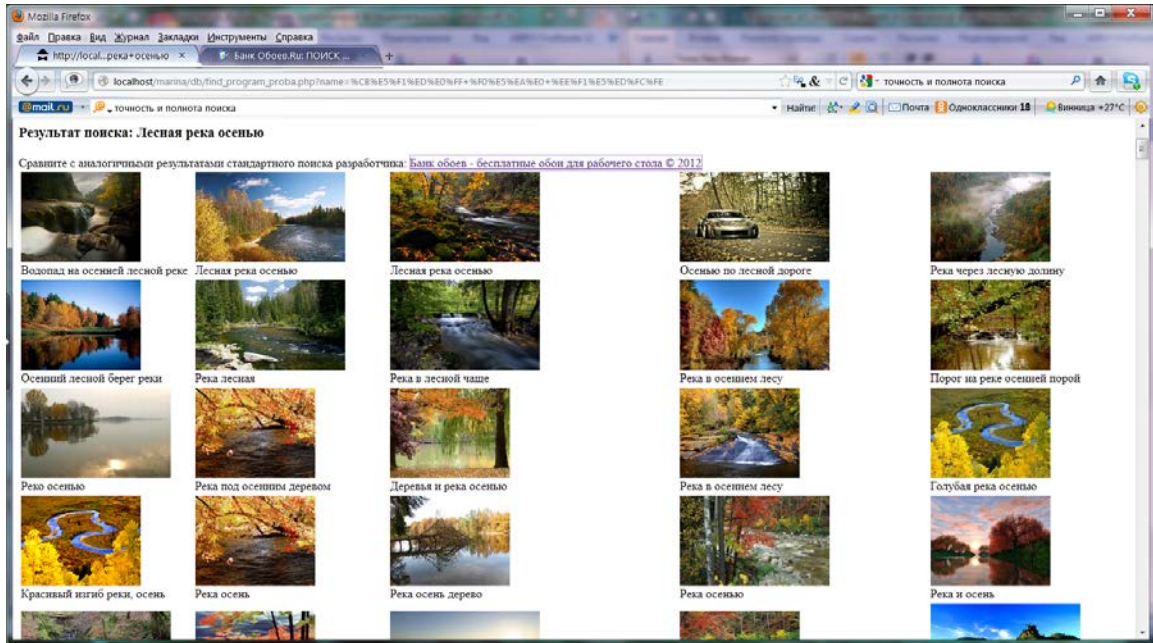
```
(define *long-memory* (make-immutable-hash-table '(
  (e1 ((i4 (2 i13)) (i13 (3 i4) (Почему? e2))))
  (e2 ((i19 (2 i16)) (i16 (3 i19))))
  (e3 ((True (2 i17)) (i17 (7 i4))))
  (e4 ((i19 (2 i9)) (i9 (3 i19))))
  (e5 ((i5 (2 i13)) (i13 (3 i5) (Почему? e6))))
  (e6 ((True (2 i12)) (i12 (Кого? i7) (6 i2))))
  (e7 ((True (2 i17)) (i17 (7 i5))))
  (e8 ((i19 (2 i9)) (i9 (3 i19))))
  (e9 ((i8 (2 i13)) (i13 (3 i8) (Почему? e10))))
  (e10 ((i11 (2 i18)) (i18 (3 i11) (4 i6))))
  (e11 ((True (2 i17)) (i17 (7 i8))))
  (e12 ((i19 (2 i9)) (i9 (3 i19))))
  (e13 ((i15 (2 i13)) (i13 (3 i15) (Почему? e14))))))
  (e14 ((i15 (2 i14)) (i14 (6 i10))))
  (e15 ((i19 (2 i1)) (i1 (3 i19))))
  (e16 ((i19 (2 i3)) (i3 (3 i19))))))
```

*/ Тестовий приклад з довідника: Искусственный интеллект: в 3 кн. Кн. 1. Системы общения и экспертные системы: Справочник / Под. ред. Э. В. Попова. – М. : Радио и связь, 1990. – 464 с.

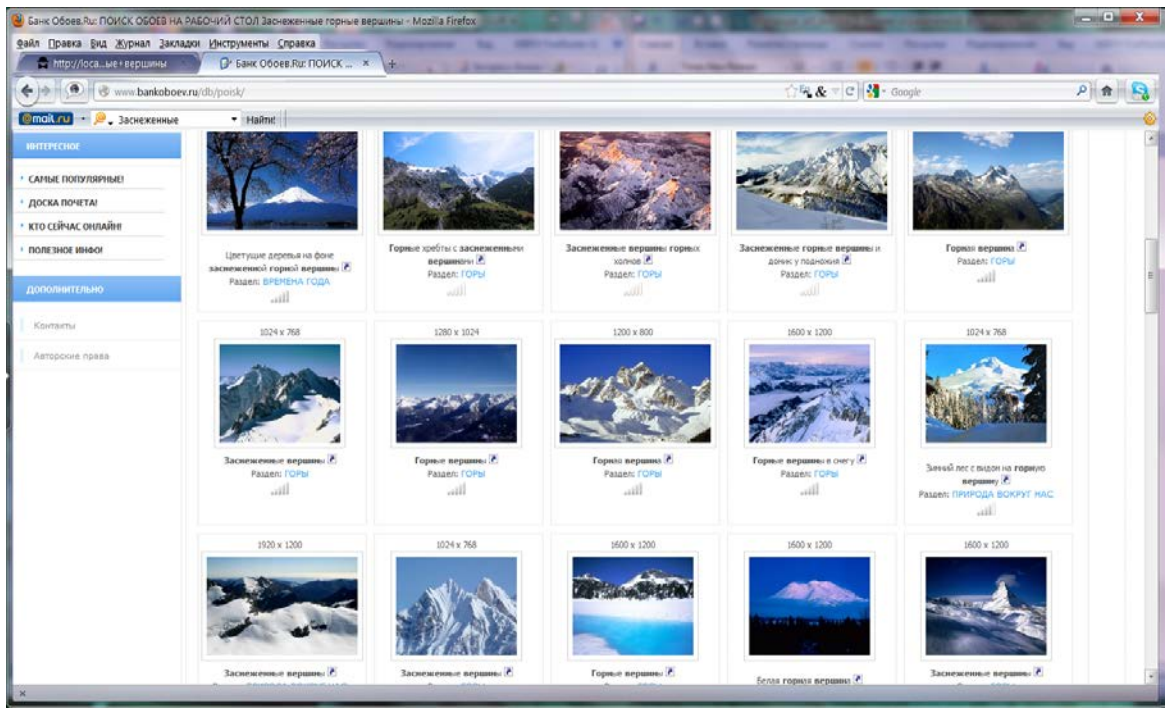
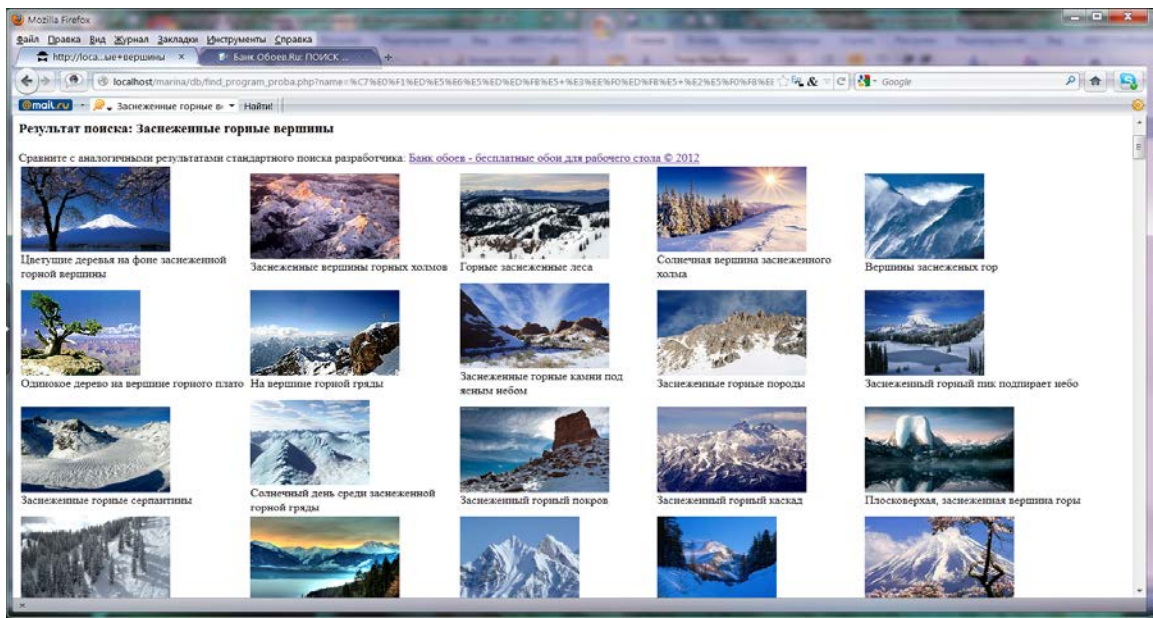
```
(module complex-sentences-data mzscheme
(define *image* '(
i50 () ИВАНОВ () ()
i51 () РАД РАДОСТЬ РАДОВАТЬСЯ РАДОСТНО)
i52 (ЦЕЛЕВОЙ () ЦЕЛЬ ЦЕЛИТЬ ЦЕЛЬНО)
i53 () () СОГЛАСОВАНИЕ СОГЛАСОВЫВАТЬ СОГЛАСОВАННО)
i54 () () РАБОТА РАБОТАТЬ ОТРАБОТАННО)
i55 () НИИ () ()
i56 (ИЗВЕСТНЫЙ () ИЗВЕСТИЕ ИЗВЕСТИТЬ ИЗВЕСТНО)
i57 (ОСТАТОЧНЫЙ () ОСТАТОК ОСТАВИТЬ ОСТАТОЧНО)
i58 (ФОНДОВЫЙ () ФОНД () ())
i59 () () ПОСЛЕ () ПОСЛЕ)
i60 (СОСТАВЛЕННЫЙ () СОСТАВИЛ СОСТАВИЛ СОСТАВЛЕННО)
i61 (УКАЗАННЫЙ () УКАЗАНИЕ УКАЗАТЬ УКАЗАННО)
i62 () () НАЧАЛО НАЧИНАТЬ НАЧАЛЬНО)
i63 () () ЯВИЛЕНИЕ ЯВИЛОСЬ ЯВНО)
i64 (СЛЕДСТВЕННЫЙ () СЛЕДСТВИЕ СЛЕДИТЬ СЛЕДСТВЕННО)
i65 (РАССМОТРЕННЫЙ () РАССМОТРЕНИЕ РАССМАТРИВАТЬ РАССМОТРЕННО)
i66 (КОМАНДИРОВОЧНЫЙ () КОМАНДИРОВКА КОМАНДИРОВАТЬ ())
i67 () ТЗ () ()
i68 () РУБЛЬ () ()
i69 () ТЫСЯЧА ТЫСЯЧА () ())
True (ДЕЙСТВИТЕЛЬНЫЙ () ДЕЙСТВИТЕЛЬНОСТЬ () ДЕЙСТВИТЕЛЬНО)
))
(define *long-memory*
(make-immutable-hash-table '(
(e20((i50 (2 i51))(i51 (3 i50) (Чему? e21))))
(e21((True (2 i66))(i66 (3 True) (7 i50) (7 i52)) (i52 (7 i53)) (i53 (7 i67)) (i50 (1 i54)) (i54
(4 i55))))
(e22 ((True (2 i56))(i56 (3 True) (Кому? i50) (Что? e23))))))
```

(e23 ((i57 (2 i60) (Чего? i58)) (i58 (1 i66)) (i60 (7 i57) (Когда? i59) (Сколько? i69)) (i69 (Чего? i68)) (i59 (Чего? i66)) (i66 (1 i61))))
 (e24((i62 (2 i63) (Чего? i54)) (i63 (3 i62) (7 i64)) (i64 (Чего? i66)) (i66 (Какой? i65))))
)))
 (provide *image* *long-memory*)

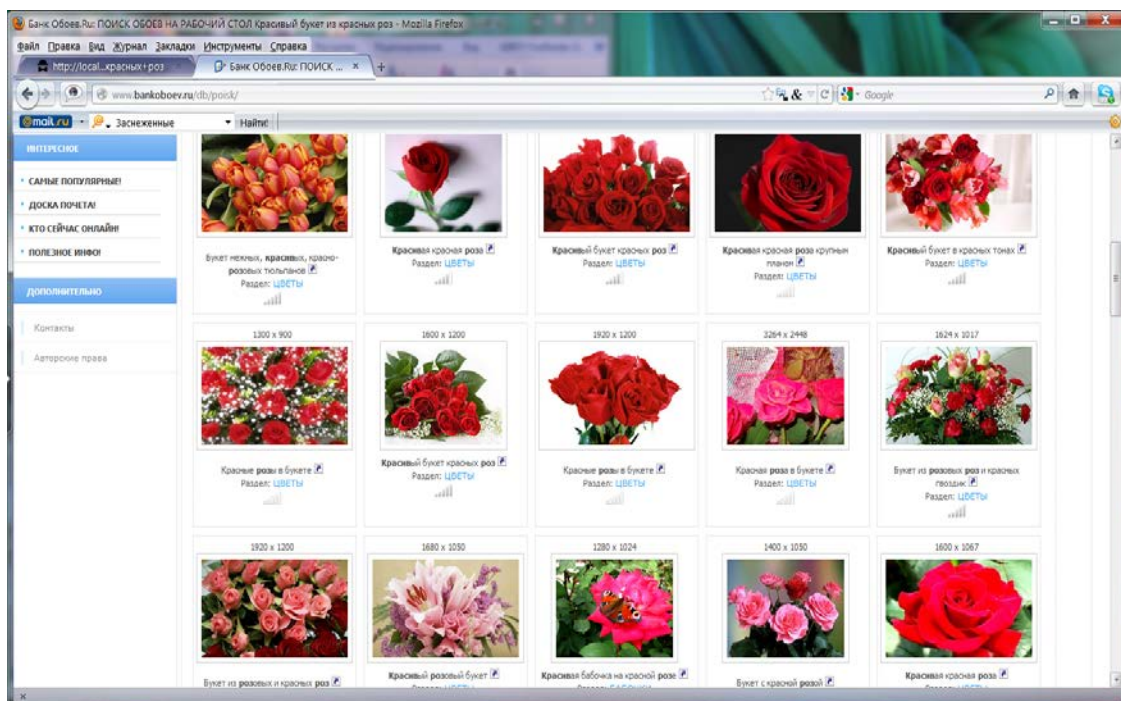
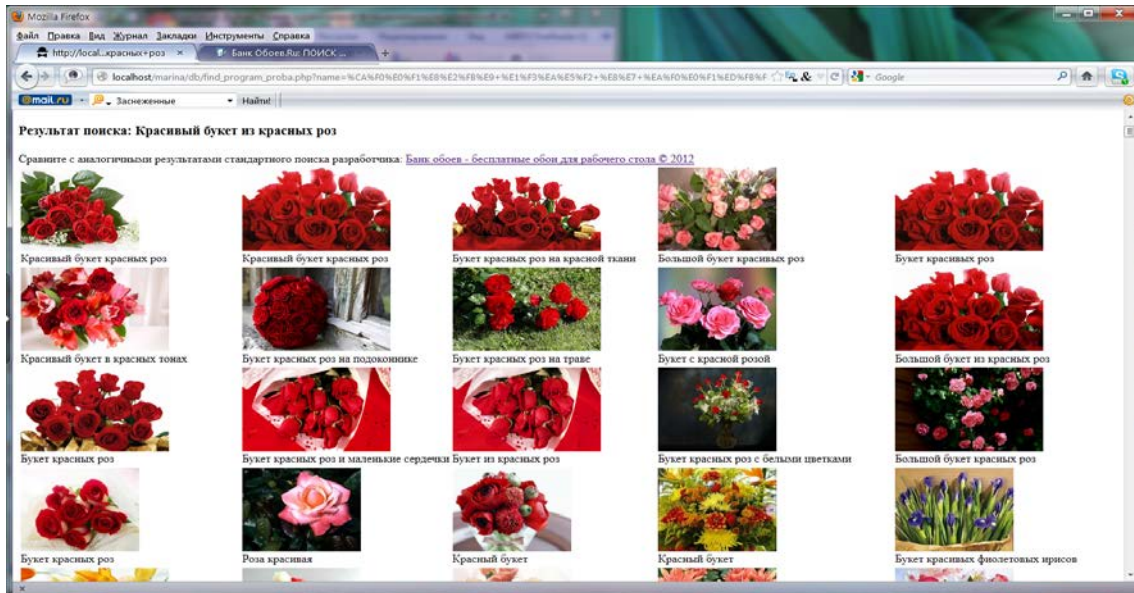
*/ Результаты поиска зображень для запиту «Лесная река осенью»



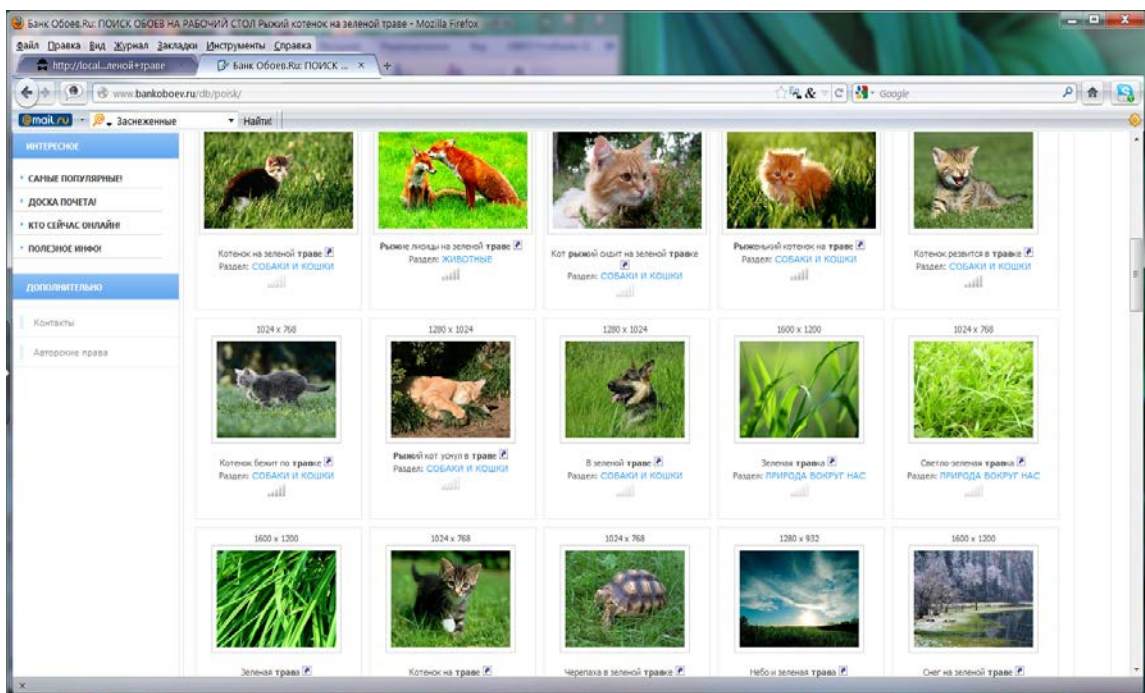
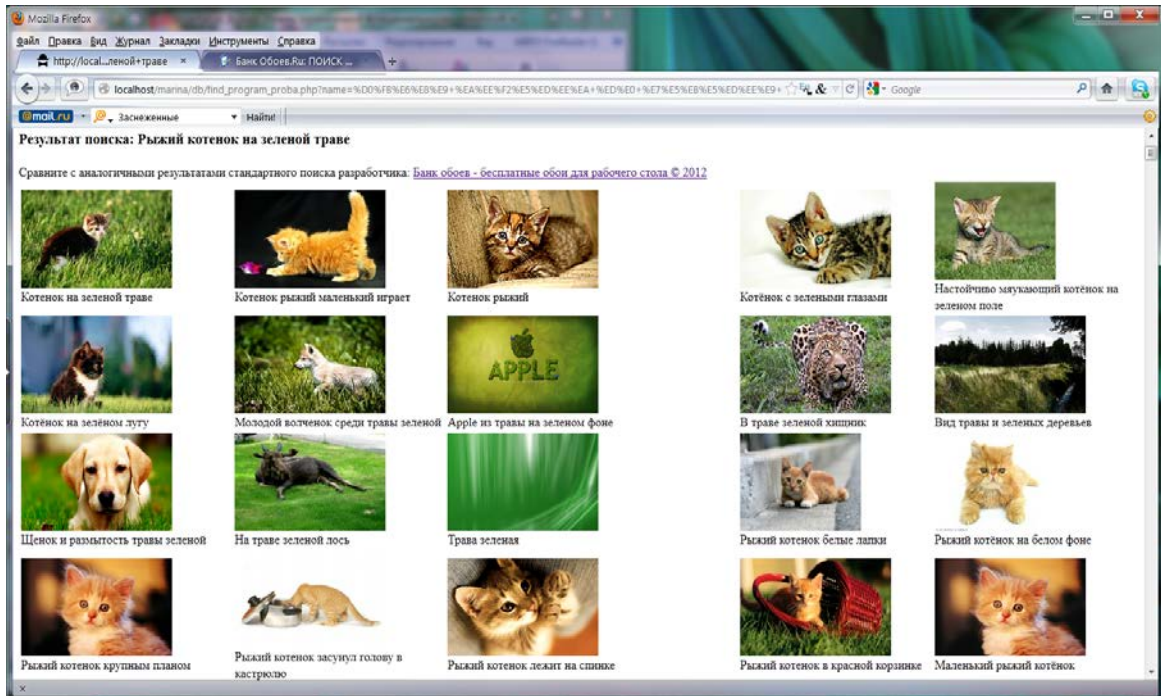
*/ Результаты поиска изображений для запросу «Заснеженные горные вершины»



*/ Результаты пошуку зображень для запиту «Красивый букет из красных роз»



*/ Результаты поиска изображений для записи «Рыжий котенок на зеленой траве»



*/ Усереднена оцінка релевантності запитів з 3-х слів

№ з/п	Образний пошук 1)			Стандартний пошук 1)			Образний пошук 2)			Стандартний пошук 2)		
	ОцРе	БіОц	ОцТо	ОцРе	БіОц	ОцТо	ОцРе	БіОц	ОцТо	ОцРе	БіОц	ОцТо
1	3	1	1,00	3	1	1,00	3	1	1,00	3	1	1,00
2	3	1	1,00	3	1	1,00	3	1	1,00	3	1	1,00
3	3	1	1,00	3	1	1,00	3	1	1,00	3	1	1,00
4	1	0	0,75	3	1	1,00	3	1	1,00	3	1	1,00
5	3	1	0,80	3	1	1,00	3	1	1,00	2	1	1,00
6	3	1	0,83	2	1	1,00	2	1	1,00	3	1	1,00
7	2	1	0,86	3	1	1,00	3	1	1,00	2	1	1,00
8	2	1	0,88	2	1	1,00	3	1	1,00	3	1	1,00
9	3	1	0,89	3	1	1,00	3	1	1,00	3	1	1,00
10	3	1	0,90	3	1	1,00	3	1	1,00	3	1	1,00
11	3	1	0,91	2	1	1,00	3	1	1,00	3	1	1,00
12	3	1	0,92	1	0	0,92	3	1	1,00	3	1	1,00
13	3	1	0,92	3	1	0,92	3	1	1,00	3	1	1,00
14	3	1	0,93	3	1	0,93	3	1	1,00	3	1	1,00
15	3	1	0,93	3	1	0,93	3	1	1,00	3	1	1,00
16	3	1	0,94	3	1	0,94	3	1	1,00	2	1	1,00
17	3	1	0,94	1	0	0,88	3	1	1,00	3	1	1,00
18	3	1	0,94	1	0	0,83	3	1	1,00	3	1	1,00
19	3	1	0,95	3	1	0,84	3	1	1,00	1	0	0,95
20	3	1	0,95	1	0	0,80	3	1	1,00	3	1	0,95
21	3	1	0,95	2	1	0,81	3	1	1,00	1	0	0,90
22	3	1	0,95	2	1	0,82	3	1	1,00	1	0	0,86
23	2	1	0,96	2	1	0,83	3	1	1,00	1	0	0,83
24	3	1	0,96	3	1	0,83	3	1	1,00	3	1	0,83
25	3	1	0,96	3	1	0,84	3	1	1,00	3	1	0,84
26	2	1	0,96	1	0	0,81	3	1	1,00	3	1	0,85
27	3	1	0,96	3	1	0,81	3	1	1,00	3	1	0,85
28	3	1	0,96	2	1	0,82	3	1	1,00	1	0	0,82
29	3	1	0,97	3	1	0,83	2	1	1,00	3	1	0,83
30	3	1	0,97	3	1	0,83	3	1	1,00	3	1	0,83
31	1	0	0,94	2	1	0,84	3	1	1,00	1	0	0,81
32	3	1	0,94	2	1	0,84	3	1	1,00	3	1	0,81
33	3	1	0,94	3	1	0,85	2	1	1,00	3	1	0,82
34	3	1	0,94	1	0	0,82	3	1	1,00	3	1	0,82
35	2	1	0,94	3	1	0,83	3	1	1,00	1	0	0,80
36	3	1	0,94	3	1	0,83	3	1	1,00	2	1	0,81
37	3	1	0,95	3	1	0,84	3	1	1,00	3	1	0,81
38	2	1	0,95	1	0	0,82	3	1	1,00	3	1	0,82
39	3	1	0,95	2	1	0,82	3	1	1,00	1	0	0,79
40	3	1	0,95	3	1	0,83	3	1	1,00	1	0	0,78
41	3	1	0,95	3	1	0,83	3	1	1,00	3	1	0,78
42	3	1	0,95	3	1	0,83	3	1	1,00	3	1	0,79

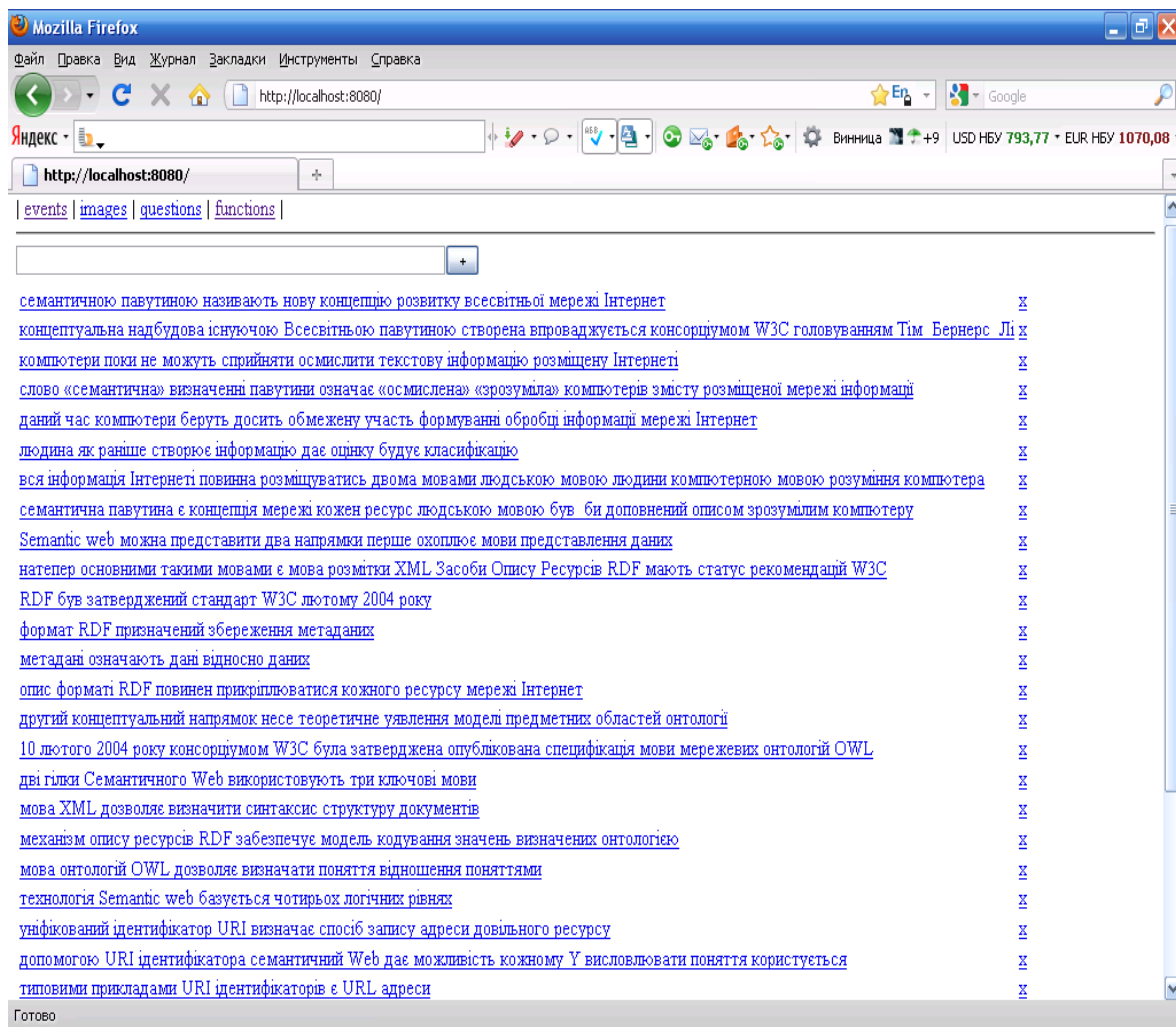
*/ Усереднена оцінка релевантності запитів з 4-х слів

№ з/п	Образний пошук 1)			Стандартний пошук 1)			Образний пошук 2)			Стандартний пошук 2)		
	ОцРе	БіОц	ОцТо	ОцРе	БіОц	ОцТо	ОцРе	БіОц	ОцТо	ОцРе	БіОц	ОцТо
1	3	1	1,00	1	0	0,00	2	1	1,00	2	1	1,00
2	3	1	1,00	1	0	0,00	2	1	1,00	1	0	0,50
3	3	1	1,00	3	1	0,33	2	1	1,00	3	1	0,67
4	2	1	1,00	1	0	0,25	2	1	1,00	3	1	0,75
5	3	1	1,00	2	1	0,40	2	1	1,00	2	1	0,80
6	3	1	1,00	3	1	0,50	2	1	1,00	2	1	0,83
7	3	1	1,00	3	1	0,57	1	0	0,86	3	1	0,86
8	3	1	1,00	3	1	0,63	1	0	0,75	1	0	0,75
9	3	1	1,00	2	1	0,67	0	0	0,67	0	0	0,67
10	2	1	1,00	2	1	0,70	1	0	0,60	0	0	0,60
11	3	1	1,00	2	1	0,73	0	0	0,55	0	0	0,55
12	3	1	1,00	1	0	0,67	1	0	0,50	2	1	0,58
13	3	1	1,00	1	0	0,62	0	0	0,46	1	0	0,54
14	3	1	1,00	2	1	0,64	3	1	0,50	0	0	0,50
15	2	1	1,00	1	0	0,60	2	1	0,53	0	0	0,47
16	2	1	1,00	2	1	0,63	3	1	0,56	0	0	0,44
17	1	0	0,94	1	0	0,59	3	1	0,59	0	0	0,41
18	1	0	0,89	0	0	0,56	2	1	0,61	0	0	0,39
19	2	1	0,89	1	0	0,53	2	1	0,63	0	0	0,37
20	1	0	0,85	2	1	0,55	3	1	0,65	1	0	0,35
21	1	0	0,81	3	1	0,57	3	1	0,67	1	0	0,33
22	1	0	0,77	3	1	0,59	3	1	0,68	1	0	0,32
23	2	1	0,78	1	0	0,57	3	1	0,70	1	0	0,30
24	3	1	0,79	1	0	0,54	3	1	0,71	2	1	0,33
25	1	0	0,76	1	0	0,52	2	1	0,72	3	1	0,36
26	1	0	0,73	2	1	0,54	2	1	0,73	0	0	0,35
27	1	0	0,70	1	0	0,52	2	1	0,74	2	1	0,37
28	1	0	0,68	2	1	0,54	2	1	0,75	0	0	0,36
29	1	0	0,66	2	1	0,55	2	1	0,76	0	0	0,34
30	2	1	0,67	1	0	0,53	2	1	0,77	0	0	0,33
31	2	1	0,68	2	1	0,55	2	1	0,77	0	0	0,32
32	1	0	0,66	2	1	0,56	3	1	0,78	0	0	0,31
33	2	1	0,67	1	0	0,55	2	1	0,79	3	1	0,33
34	1	0	0,65	2	1	0,56	2	1	0,79	2	1	0,35
35	1	0	0,63	1	0	0,54	2	1	0,80	2	1	0,37
36	1	0	0,61	1	0	0,53	2	1	0,81	0	0	0,36
37	1	0	0,59	2	1	0,54	3	1	0,81	0	0	0,35
38	1	0	0,58	0	0	0,53	2	1	0,82	3	1	0,37
39	1	0	0,56	1	0	0,51	2	1	0,82	3	1	0,38
40	2	1	0,58	1	0	0,50	2	1	0,83	2	1	0,40
41	2	1	0,59	0	0	0,49	2	1	0,83	2	1	0,41
42	1	0	0,57	3	1	0,50	3	1	0,83	2	1	0,43

Додаток Д

Діючий прототип інформаційної технології: результати тестування

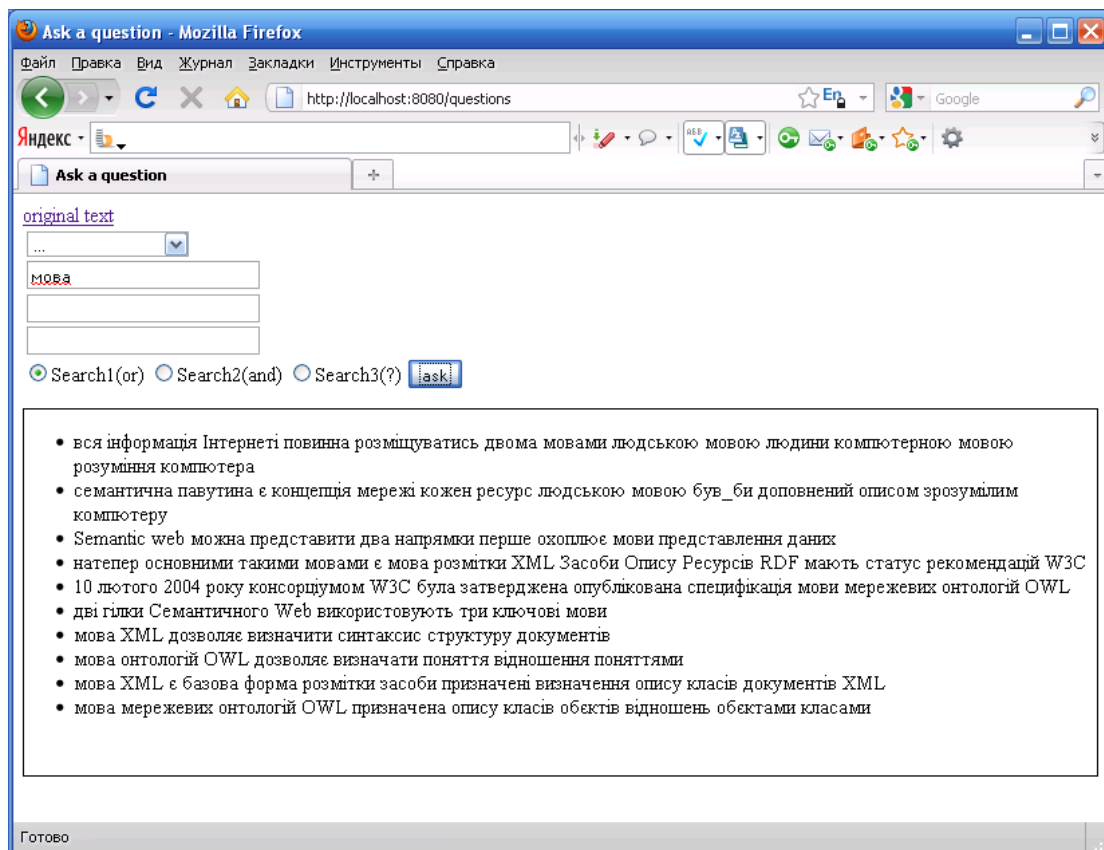
*/ Зовнішній вигляд оболонки навчальної системи (форма 1) з синтагмами елемента ЕК наскрізного прикладу згідно з Додатком Б



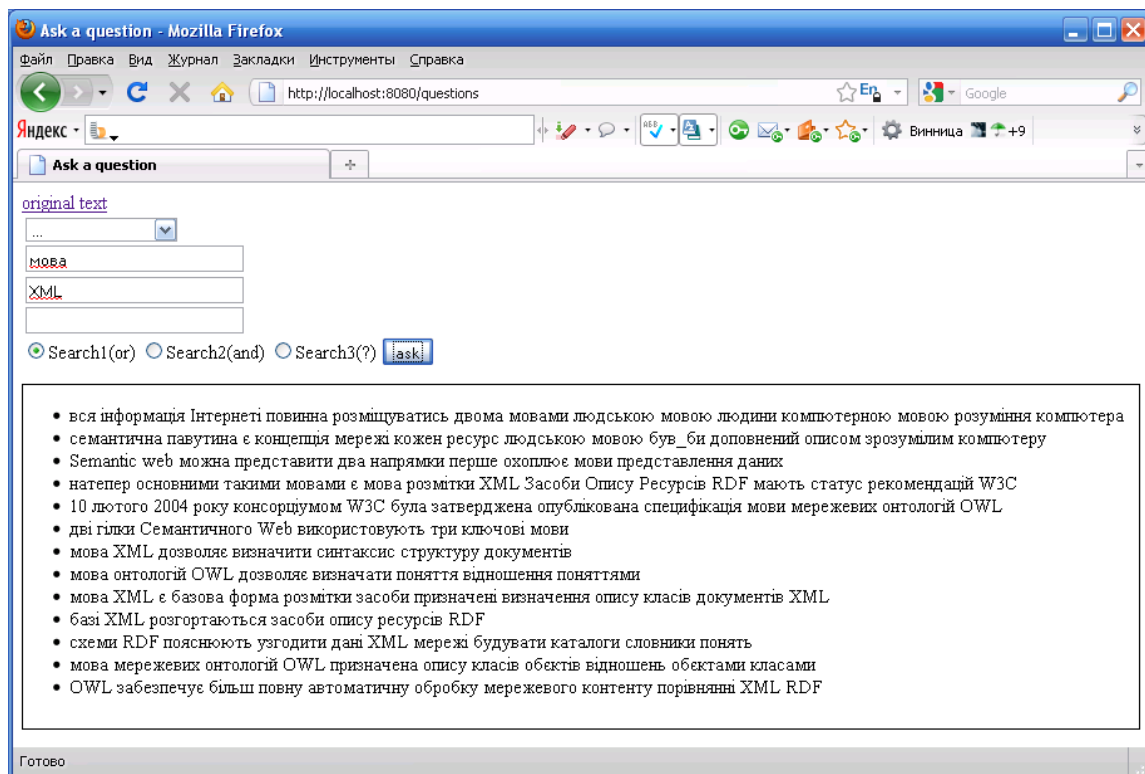
*/ Режим введення синтагми в оболонку навчальної системи (форма 2)



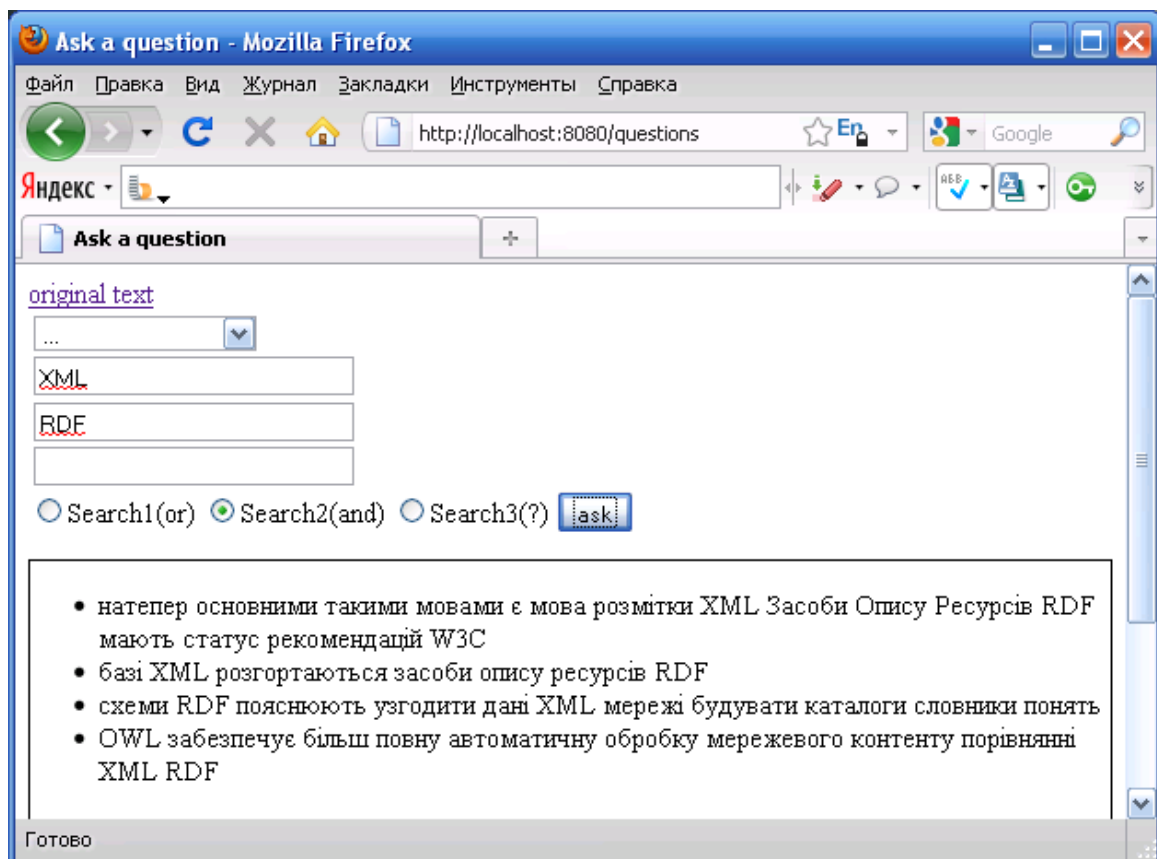
*/ Пошук всіх речень з образом «мова» у будь-якій словоформі (форма 3)



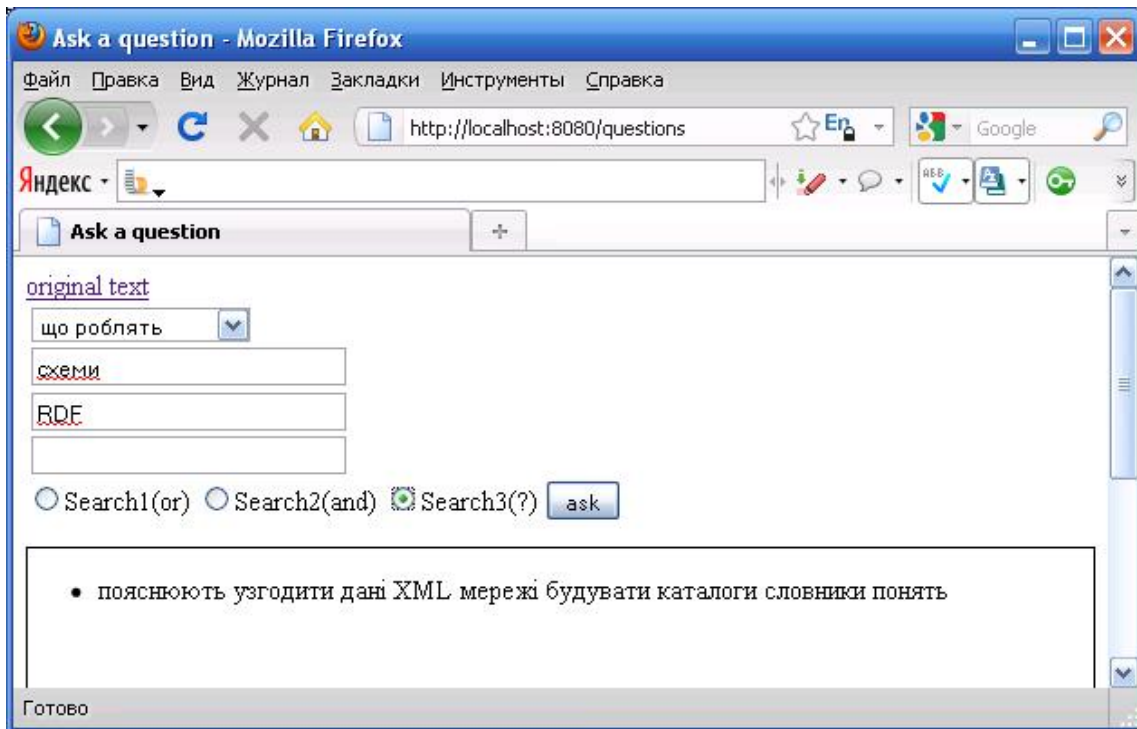
* / Пошук всіх речень з образами «мова» або «XML»



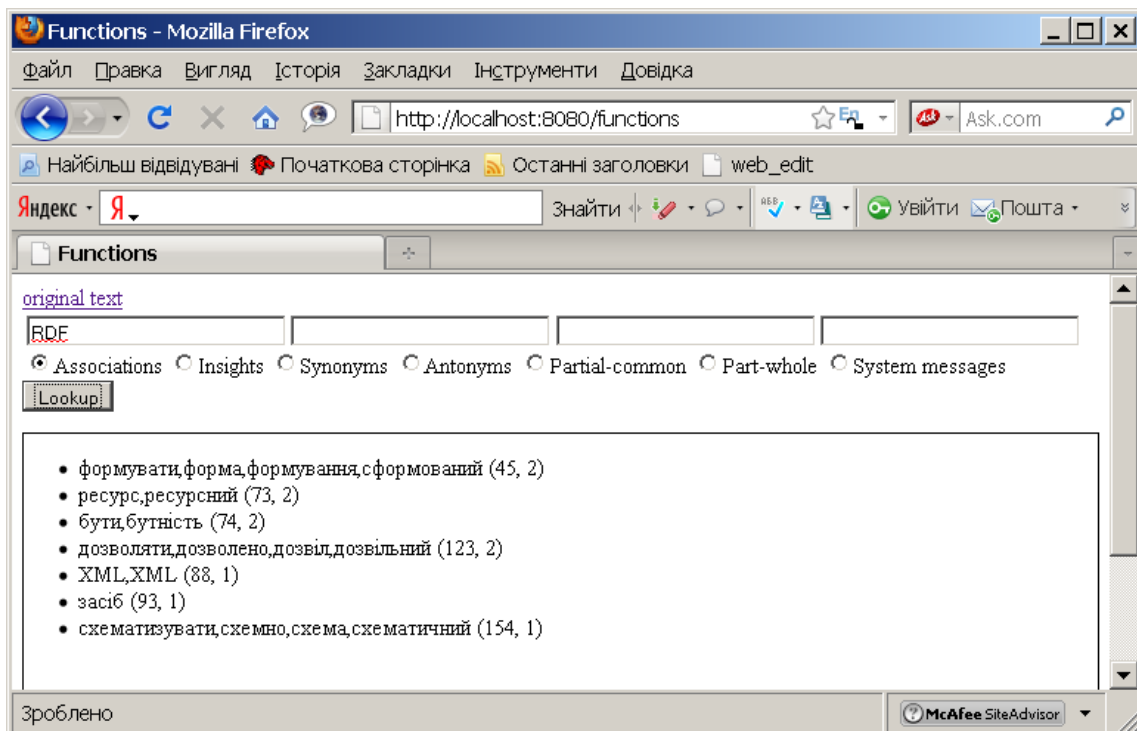
* / Пошук всіх речень з образами «XML» та «RDF»



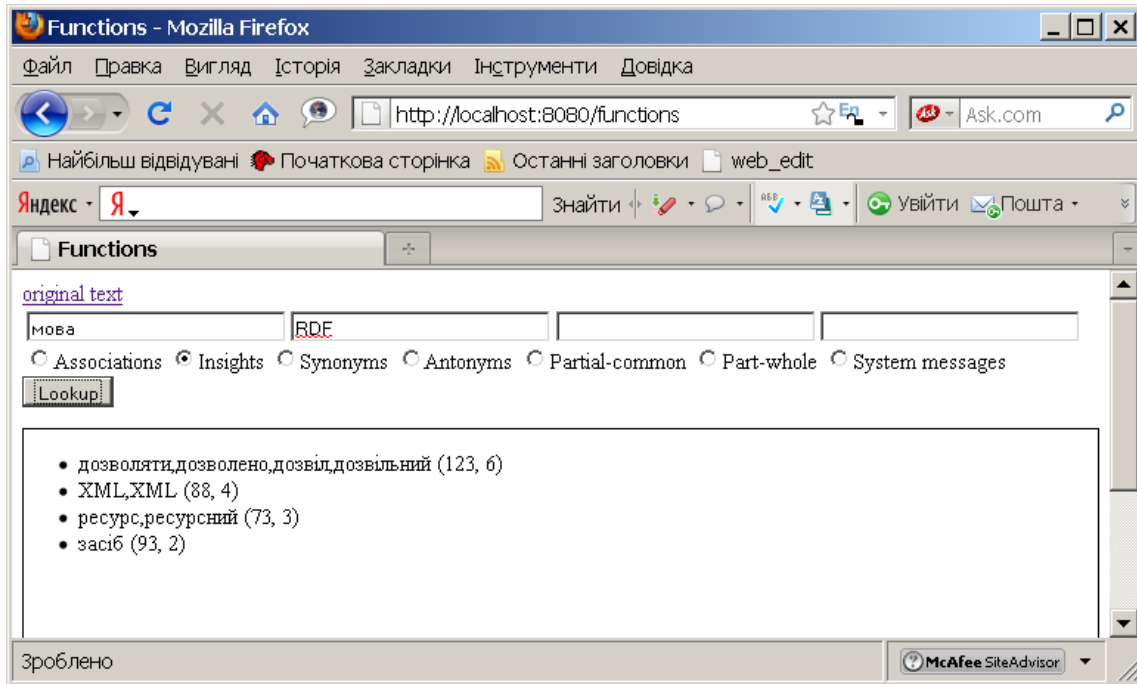
*/ Результат відповіді ІС на основі пошуку з питальним займенником



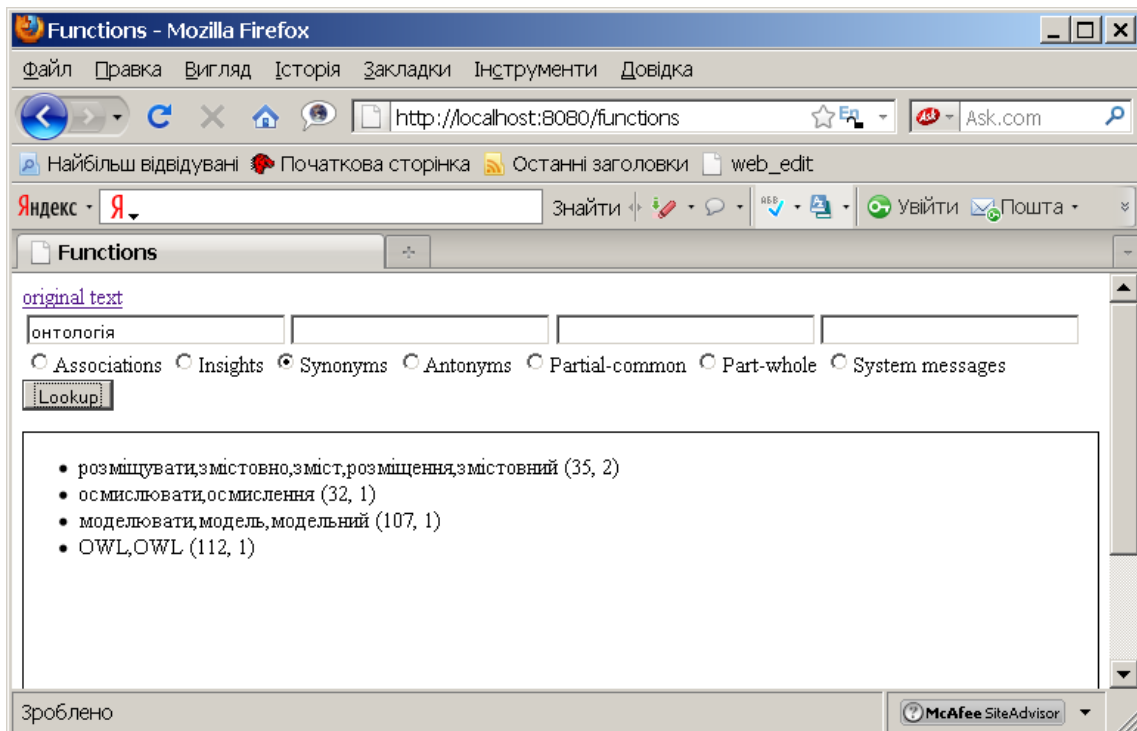
*/ Визначення списку асоціативних образів до МО «RDF» (форма 4)



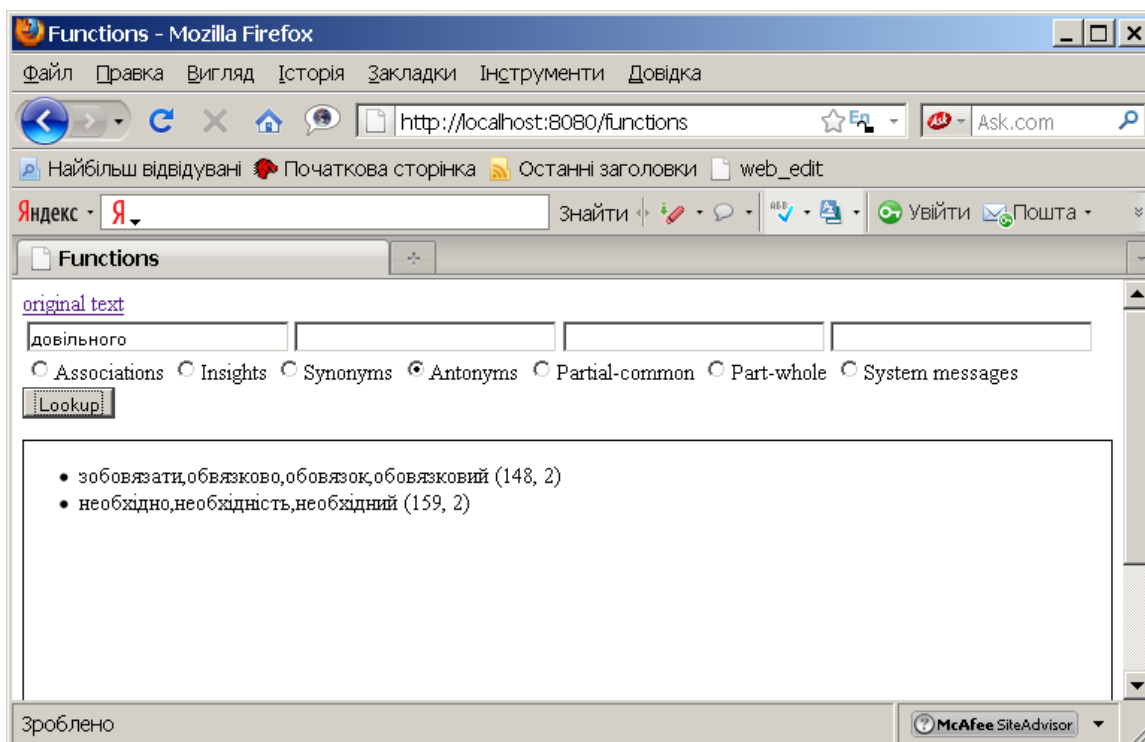
*/ Пошук інсайтних образів між МО «мова» та «RDF»



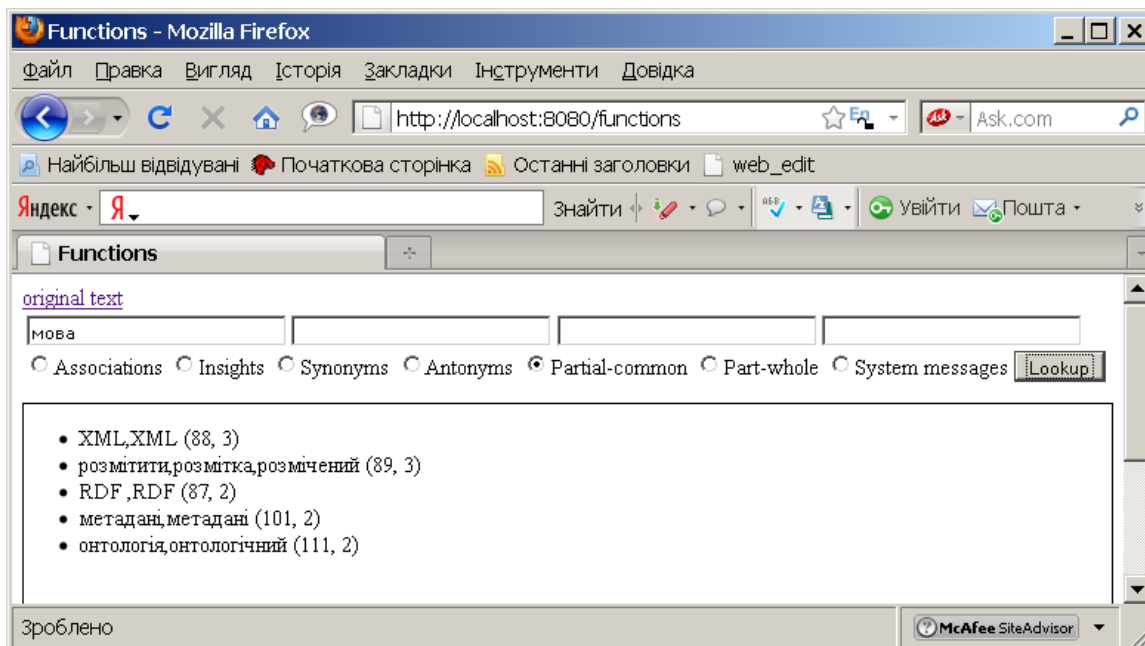
*/ Визначення синонімічних зв'язків для МО «онтологія (111)»



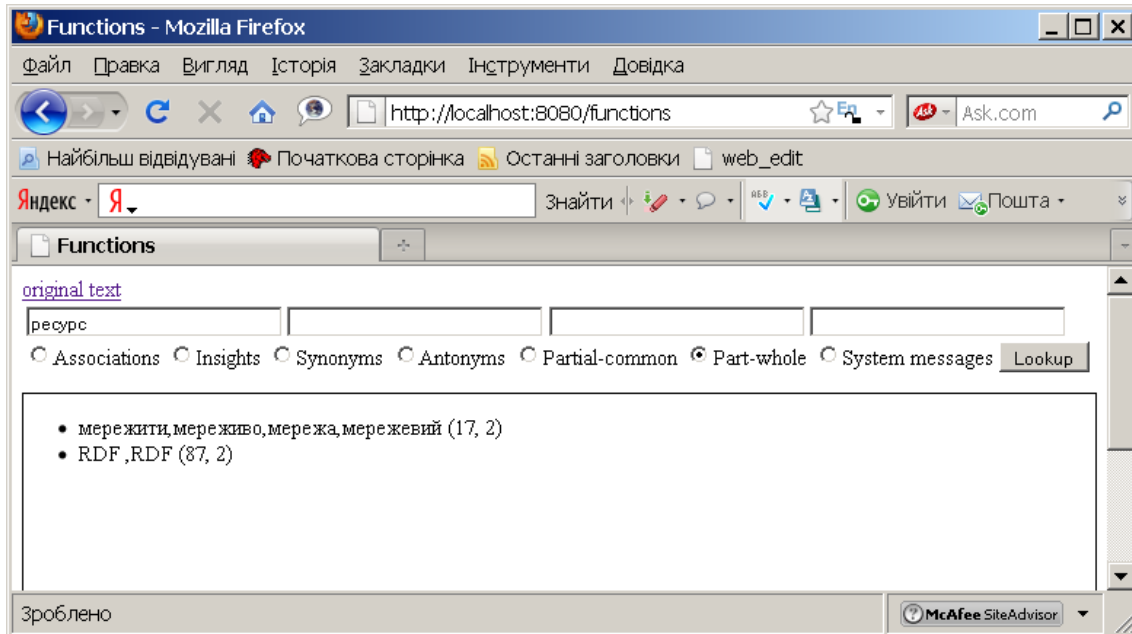
*/ Визначення антонімічних зв'язків для МО «довільно (138)»



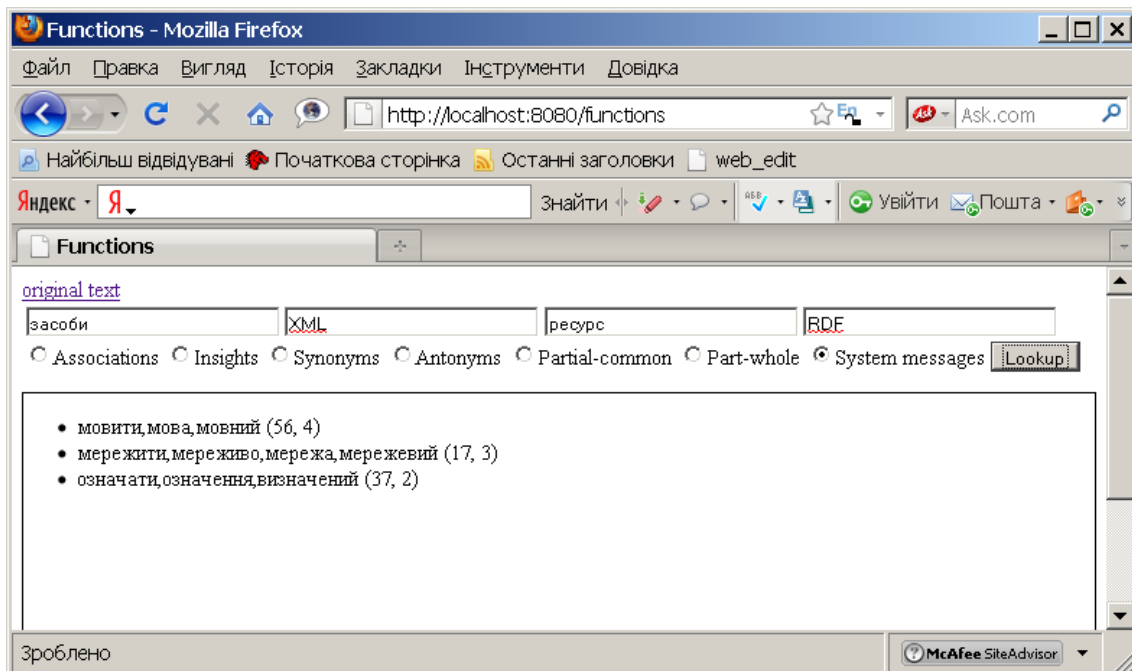
*/ Визначення часткових МО до загального МО «мова (56)»



*/ Пошук зв'язків типу частина–ціле для даних наскрізного прикладу



*/ Повідомлення системи як відповідь на запит користувача за типом пророцтва дельфійського оракула



Наукове видання

Бісікало Олег Володимирович

**ФОРМАЛЬНІ МЕТОДИ
ОБРАЗНОГО АНАЛІЗУ ТА СИНТЕЗУ
ПРИРОДНО-МОВНИХ КОНСТРУКЦІЙ**

Монографія

Редактор С. А. Малішевська

Оригінал-макет підготовлено О. В. Бісікало

Підписано до друку 29.05.2013 р.
Формат 29,7×42¼. Папір офсетний.
Гарнітура Times New Roman.
Друк різнографічний. Ум. др. арк. 18,25
Наклад 300 (1-й запуск 1–75) Зам № 2013-107

Вінницький національний технічний університет,
КІВЦ ВНТУ,
21021, м. Вінниця, Хмельницьке шосе, 95,
ВНТУ, ГНК, к. 114.
Тел. (0432) 59-85-32.
Свідоцтво суб'єкта видавничої справи
серія ДК № 3516 від 01.07.2009 р.

Віддруковано у Вінницькому національному технічному університеті,
в комп'ютерному інформаційно-видавничому центрі,
21021, м. Вінниця, Хмельницьке шосе, 95,
ВНТУ, ГНК, к. 114.
Тел. (0432) 59-81-59
Свідоцтво суб'єкта видавничої справи
серія ДК № 3516 від 01.07.2009 р.