

МІНІСТЕРСТВО ОСВІТИ І НАУКИ, МОЛОДІ ТА СПОРТУ УКРАЇНИ  
ВІННИЦЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ

Бісікало Олег Володимирович

УДК 004.94:159.95

**ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ОБРАЗНОГО АНАЛІЗУ ТЕКСТУ ТА  
СИНТЕЗУ СТРУКТУРНО-ФУНКЦІОНАЛЬНИХ МОДЕЛЕЙ  
СИСТЕМИ ОБРОБКИ ПРИРОДНОМОВНОГО КОНТЕНТУ**

Спеціальність 05.13.06 – інформаційні технології

**АВТОРЕФЕРАТ**

дисертації на здобуття наукового ступеня  
доктора технічних наук

Вінниця – 2012

Дисертацією є рукопис.

Робота виконана у Вінницькому національному технічному університеті, Міністерства освіти і науки, молоді та спорту України.

**Науковий консультант**

доктор технічних наук, професор  
**Квстний Роман Наумович**,  
Вінницький національний технічний університет,  
завідувач кафедри автоматичної та інформаційно-вимірної техніки.

**Офіційні опоненти:**

доктор технічних наук, професор  
**Бушуєв Сергій Дмитрович**,  
Київський національний університет будівництва і архітектури,  
завідувач кафедри управління проектами

доктор технічних наук, професор  
**Петров Едуард Георгійович**,  
Харківський національний університет радіоелектроніки,  
завідувач кафедри системотехніки

доктор технічних наук, професор  
**Петух Анатолій Михайлович**,  
Вінницький національний технічний університет,  
завідувач кафедри програмного забезпечення.

Захист відбудеться « 21 » грудня 2012 р. о 9-30 годині на засіданні спеціалізованої вченої ради Д 05.052.01 у Вінницькому національному технічному університеті за адресою: 21021, м. Вінниця, Хмельницьке шосе, 95, ГНК, ауд.210.

З дисертацією можна ознайомитись у бібліотеці Вінницького національного технічного університету за адресою: 21021, м. Вінниця, Хмельницьке шосе, 95, ГНК.

Автореферат розісланий « 13 » листопада 2012 р.

Учений секретар  
спеціалізованої вченої ради

С.М. Захарченко

## ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

**Актуальність теми.** 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, Колмогорова А.М., Глушкова В.М., Поспелова Д.А., Гаврилової Т.А., Т. Бернерс-Лі тощо. Проте можливості сучасних технологічних засобів оброблення контенту не відповідають вимогам міжнародних стандартів до його семантичних властивостей. Всупереч зростаючим очікуванням Інтернет-спільноти 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 Андон П.І., Палагін О.В., Широков В.А., Шабанов-Кушнарєнко Ю.П., Кузьмін І.В., Бондаренко М.Ф., Шлезінгер М.І., Анісімов А.В., Бушуєв С.Д., Петров Е.Г., Валькман Ю.Р. та ін. Проте відчутна парадоксальність проблеми розв'язання семантико-залежних задач обробки природномовної інформації вимагає застосування 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 ре-

зультатами наук когнітивного напрямку гіпотеза про витoki прагматичних і семантичних аспектів сенсу з понять образу, асоціації, потреб і емоцій.

Отже, актуальною *науково-прикладною проблемою*, що підлягає вирішенню, є відсутність теоретико-методологічних основ і технологічних засобів побудови систем обробки природномовного контенту, здатних розв'язувати семантико-залежні задачі на основі накопичування загальних знань образного сенсу. Необхідність проведення досліджень з метою вирішення науково-прикладної проблеми обумовлена розвитком сучасних інформаційно-комунікаційних технологій.

**Зв'язок роботи з науковими програмами, планами, темами.** Дисертаційне дослідження проводилось згідно з планами науково-дослідних робіт кафедри економічної кібернетики Вінницького національного аграрного університету і кафедри автоматизації та інформаційно-вимірювальної техніки Вінницького національного технічного університету, в тому числі в межах договору «Розробка програмної оболонки для створення електронних підручників з розширеними дидактичними можливостями» (номер державної реєстрації 0110U004301, 2010 р.) та держбюджетної теми «Методологія моделювання процесів образного мислення у відповідності до задач пізнавальної та мовленнєвої діяльності людини» (номер державної реєстрації 0111U001115, 2011-2013 рр.), при виконанні яких автор брав участь як науковий керівник та відповідальний виконавець. Тематика роботи пов'язана з Державною науково-технічною програмою "Образний комп'ютер" на 2000-2010 рр., затвердженою постановою Кабінету Міністрів України №1652 від 08.11.2000 р., Державною програмою розвитку і функціонування української мови на 2004-2010 рр., затвердженою постановою Кабінету Міністрів України №1546 від 02.10.2003 р. та відповідає Основним науковим напрямкам та найважливішим проблемам фундаментальних досліджень у галузі природничих, технічних і гуманітарних наук на 2009–2013 рр., затвердженим постановою Президії НАН України № 55 від 25.02.2009 р.

**Мета і задачі дослідження.** Мета дослідження полягає у підвищенні якості розв'язання комплексу семантико-залежних задач обробки інформації шляхом створення методології образного аналізу і синтезу структурно-функціональних моделей систем на основі природномовної інформації. Досягнення мети має забезпечити побудована на основі методології інформаційна технологія образного аналізу тексту та синтезу структурно-функціональних моделей системи обробки природномовного контенту, що орієнтована на загальні знання образного сенсу, має здатність до їх самовдосконалення та складом власних функцій забезпечує розв'язання комплексу семантико-залежних задач.

Відповідно до мети мають бути вирішені такі **основні задачі** дослідження:

1. Аналіз методів моделювання процесів обробки природномовної інформації та обґрунтування структурно-функціонального підходу.
2. Обґрунтування напрямку досліджень та визначення вимог до семантичних функцій інформаційної системи обробки природномовного контенту.
3. Побудова теоретичних основ образного аналізу текстової інформації через:

3.1. Визначення концептуальних основ представлення образного сенсу природномовних конструкцій.

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

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

4. Синтез структурно-функціональних моделей системи обробки природномовного контенту (СОПМК), а саме:

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

4.2. Побудова методів моделювання процесів образної обробки природномовного контенту з обґрунтуванням архітектури бази знань СОПМК.

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

5. Побудова інформаційної технології образної обробки природномовного контенту на основі СОПМК, зокрема:

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

5.2. Побудова методу моделювання механізму оперативної пам'яті СОПМК.

5.3. Створення методу моделювання складових парадигматичного устрою мови.

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

7. Практична реалізація й аналіз результатів впровадження отриманих методологічних і технологічних засобів.

*Об'єктом дослідження* обрано процеси обробки тексту, образного аналізу та синтезу структурно-функціональних моделей систем на основі природномовної інформації.

*Предметом дослідження* є методологія, що складається з моделей, методів аналізу текстової інформації й синтезу інтерпретацій інтелектуальної діяльності на основі структурно-функціонального підходу до побудови СОПМК.

**Методи дослідження.** Сукупність обраних методів визначена сутністю задач дослідження. Для розробки теорії образного аналізу текстової інформації використано методи системного та функціонального аналізу, теорії множин, теорії формальних систем, нечіткої математики та теорії інформації. Синтез структурно-функціональних моделей СОПМК здійснювався за допомогою методів системного аналізу, кібернетичної інтерпретації процесів інтелектуальної діяльності, теорії множин, алгебри логіки та теорії графів. Для розв'язання прикладних задач інформаційної технології образної обробки природномовного контенту застосовувалися методи теорії графів, теорії множин, алгебри логіки та теорії алгоритмів.

**Наукова новизна одержаних результатів** полягає в обґрунтуванні теорії образного аналізу текстової інформації, синтезі структурно-функціональних

моделей СОПМК та розробці інформаційної технології образної обробки природномовного контенту, зокрема:

*Вперше:*

- запропоновано підхід до образного аналізу текстової інформації, який, на відміну від існуючих, базується на формально введених поняттях інфологічної системи, онтогенетичного принципу та мовного образу, що дозволяє отримати чисельні ознаки загального образного сенсу природномовної конструкції;
- розроблено методологію синтезу структурно-функціональних моделей інфологічної системи, в якій, на відміну від існуючих, інтерпретовано інтелектуальну діяльність людини як взаємопов'язану обробку даних 9-ма функціональними блоками-моделями головного мозку у межах 19 інформаційних потоків образної природи, що забезпечує формалізацію механізму обробки природномовного контенту;
- запропоновано функціональну модель СОПМК на основі побудови нової класифікації синтагматичного та парадигматичного типів образного пошуку, що дозволяє уніфікувати процеси обробки даних і знань для розв'язання базових та прикладних семантико-залежних задач;
- розроблено метод моделювання образного механізму оперативної пам'яті СОПМК засобами двохосновної алгебраїчної системи БАС, що дозволяє отримати розв'язки семантико-залежних задач у вигляді алгоритмів навігації на графі булеану;
- запропоновано метод моделювання складових парадигматичного устрою мови на основі підходу до образного аналізу текстової інформації, що дозволяє отримати рекурсивні алгоритми самовдосконалення бази знань СОПМК.

*Удосконалено:*

- метод створення нечіткого відношення образного сенсу шляхом приведення аксіоматики до онтогенетичного принципу, формалізації понять мотиваційної мети та емоційного стану інфологічної системи, введення одиниці сенсу на основі сили елементарного асоціативного зв'язку, що дозволяє побудувати топологічний і квазіметричний простір упорядкованих пар образів;
- методи моделювання процесів образної обробки природномовного контенту на основі орієнтованих і неорієнтованих зважених графів і графа булеану через алгебраїчну систему БАС, що дозволяє представити базові функції інфологічної системи у вигляді алгоритмів на графах і формалізувати архітектуру бази знань СОПМК.

*Набули подальшого розвитку:*

- модель формальної теорії у вигляді комутативної напівгрупи образних конструкцій, в якій, на відміну від відомих, застосовано бінарний оператор спрямованого асоціативного зв'язку та поняття асоціативної нормальної форми, що забезпечує представлення образної конструкції простого речення у вигляді 3-х складових питальної конструкції мовних образів;
- математичні моделі базових типів образного пошуку з включенням інсайту як часткового випадку, що відповідають новій класифікації та дозволяють

зменшити розмірність пошуку на орієнтованих та неорієнтованих зважених графах в процесі розв'язання семантико-залежних задач.

**Практичне значення одержаних результатів** в галузі обробки природномовного контенту полягає в створенні методології побудови СОПМК та відповідної інформаційної технології, які у сукупності вирішують важливу науково-прикладну проблему розв'язання семантико-залежних задач на основі накопичування загальних знань образного сенсу. Зокрема:

1. Запропоновані підходи, методи формалізації та моделі процесів обробки природномовної інформації як складові отриманої методології дозволили представити розв'язки комплексу актуальних семантико-залежних задач у вигляді функцій єдиної СОПМК, що забезпечило покращення якості WEB-сайтів, баз даних і знань, електронного журналу викладача та інших прикладних продуктів. Вказані результати впроваджені у Донецькому інституті проблем штучного інтелекту НАН та МОНМС України – акт від 25.02.2009 р., Державній науковій сільськогосподарській бібліотеці УААН – акт від 15.04.2009 р., Всеукраїнській громадській організації «Національна асоціація сільськогосподарських дорадчих служб України» – акт від 29.05.2009 р., Центрі інформаційних технологій Вінницького державного аграрного університету – акт від 09.12.2008 р.

2. Створені формальні методи та технологічні засоби, на відміну від існуючих, за рахунок додаткових можливостей оброблення текстової інформації шляхом її образного індексування підтримують декілька типів семантичного пошуку та діалогу, що забезпечує підвищення релевантності пошуку за показниками точності на 11,3% та повноти на 17,5% для англійської та, відповідно, 18,3% та 38,8% для російської мов. Вказані результати впроваджені на підприємствах Aether. Inc. (США) – акт від 28.08.2012 р., Smart.exe (Ізраїль) – акт № 513604439, Інновін (Вінниця) – акт №01 від 23.08.2012 р.

3. На основі інформаційної технології розроблено методи й алгоритми для підтримки прикладних функцій електронного підручника, які дозволяють скоротити на 45% витрати часу на створення програмно-педагогічних засобів, а також підвищити успішність на 7,5% для учнів і 8,27% для студентів. Перераховані результати впроваджені в управлінні освіти і науки Вінницької обласної державної адміністрації (ДР 0110U004301) – акт від 26.09.2010 р., Центрі дистанційної освіти Вінницького національного технічного університету (ДР 0111U001115) – акт від 27.05.2010 р., фізико-математичній гімназії №17 (м. Вінниця) – акт від 11.03.2010 р.

4. Розроблено практичні методи підтримки функцій СОПМК, у т.ч. модифікованого пошуку найкоротшого шляху у зваженому графі, що зменшує розмірність пошуку від геометричної до арифметичної прогресії, а також організації бази знань природномовного контенту, обсяг якої не перевищує поліноміальної складності з порядком не більше третього. Методи використано у навчальному процесі Харківського національного аграрного університету – довідка №1654 від 30.12.2008 р., Сумського національного аграрного університету – акт №4-кїї від 9.02.2009 р., Полтавської державної аграрної академії – акт від 15.01.2009 р. і Подільського державного аграрно-технічного університету – акт

від 20.12.2008 р. при викладанні 6-ти навчальних дисциплін протягом 2007/2008, 2008/2009 навчальних років.

**Особистий внесок здобувача.** Усі основні результати дисертаційної роботи отримані автором особисто та опубліковані в [1, 3-4, 7-21, 23, 25, 28, 32, 34, 36-46, 48-57]. У роботах, опублікованих у співавторстві, здобувачеві належать такі положення: [2] – системний аналіз і декомпозиція складових інформаційно-аналітичної системи; [5, 6, 59] – постановка і розв’язання семантико-залежних задач електронної освіти; [22, 33, 35] – ідея прикладної інтерпретації моделей образного аналізу і синтезу та постановка наукових задач; [24, 31, 60] – підхід до аксіоматизації простору сенсу образних конструкцій, отримання математичних співвідношень; [26, 27, 29, 30] – ідея застосування структурно-функціонального підходу для морфологічного аналізу мовних образів та постановка наукових задач; [47] – підхід до формалізації результатів тестового контролю за допомогою ієрархічних структур; [58] – постановка і розв’язання задачі інтелектуального пошуку графічних зображень на основі фолксонометричних даних; [61] – підхід до застосування поняття ентропії для чисельної оцінки образного сенсу вербальних конструкцій; [62, 63] – підхід до формалізації методів адаптивного навчання, моделі та алгоритми базових типів образного пошуку.

**Апробація результатів дисертації.** Результати досліджень, що викладені в дисертаційній роботі, пройшли апробацію на більше ніж 40 наукових конференціях та семінарах, зокрема на 7–10-й міжнародних науково-технічних конференціях «Контроль і управління в складних системах (КУСС)» (м. Вінниця, 2003, 2005, 2008, 2010 рр.); міжнародній науково-методичній конференції «Сучасні технології підготовки фахівців в умовах подальшого розвитку вищої освіти України» (м. Харків, 2005 р.); Всеукраїнській науково-практичній конференції «Інформаційні технології в освіті» (м. Мелітополь, 2006 р.); 1-й, 2-й та 3-й міжнародних конференціях «Нові інформаційні технології в освіті для всіх» (м. Київ, 2006, 2007, 2008 рр.); міжнародній науково-технічній конференції «Датчики, приборы и системы – 2006» (м. Гурзуф, 2006 р.); 13-й міжнародній науково-технічній конференції з автоматичного управління «Автоматика-2006» (м. Вінниця, 2006 р.); 5-й та 7-й міжнародних науково-технічних конференціях «ІНТЕРНЕТ–ОСВІТА–НАУКА (ІОН)» (м. Вінниця, 2006, 2010 рр.); 4<sup>th</sup> & 5<sup>th</sup> «IEEE East-West Design & Test Workshop (EWDTW)» (м. Сочі, Росія, 2006 р.; м. Єреван, Вірменія, 2007 р.); 8-й та 9-й міжнародних науково-технічних конференціях «Современные информационные и электронные технологии (СИЭТ)» (м. Одеса, 2007, 2008 рр.); 8–11-й міжнародних науково-технічних конференціях «Искусственный интеллект. Интеллектуальные системы» (м. Дивноморське, Росія, 2007, 2009 рр.; м. Кацивелі, 2008, 2010 рр.); 1-й та 3-й міжнародній науково-технічній конференції «Інтелектуальні системи в промисловості і освіті» (м. Суми, 2007, 2011 рр.); міжнародній науково-практичній конференції «Информационные технологии в управлении сложными системами» (м. Дніпропетровськ, 2008 р.); міжнародній науково-практичній конференції «Е-обучение в высшей школе – проблемы и перспективы (INCEL – 08)» (м. Одеса, 2008 р.); науково-технічній конференції «Информационные технологии в металлургии и машиностроении (ИТММ'2008)» (м. Дніпропетровськ, 2008 р.); 5-й



міжнародній науково-практичній конференції «Наука: теорія і практика» (м. Перемишль, Польща, 2009 р.); 5-й міжнародній науково-практичній конференції «Військова освіта і наука: сьогодення та майбутнє» (м. Київ, 2009 р.); 1-й і 2-й міжнародній науково-практичній конференції «Інформаційні технології та комп'ютерна інженерія» (м. Вінниця, 2010 р.; м. Харків, 2011 р.); 5-й та 7-й міжнародних науково-практичних конференціях «Математичне та імітаційне моделювання систем (МОДС)» (м. Київ-Жукін, 2010 р.; м. Чернігів-Жукін, 2012 р.); 7-й міжнародній конференції «Нові інформаційні та комп'ютерні технології в освіті та науці (IES-2010)» (м. Вінниця, 2010 р.); 7-й міжнародній конференції «Теоретичні та прикладні аспекти побудови програмних систем (TAAPSD'2010)» (м. Київ, 2010 р.); 7-й міжнародній науково-практичній конференції «Наука і технології: крок у майбутнє» (м. Прага, Чехія, 2011 р.); науковому семінарі «Перспективні технології навчання та освітні простори» та науковому семінарі з проблеми «Образний комп'ютер» (м. Київ, МННЦІТС 2006, 2007 рр.); наукових семінарах інституту проблем штучного інтелекту НАН та МОНМС України (м. Донецьк, 2009 р.) та Інституту проблем математичних машин і систем (м. Київ, 2010 р.).

**Публікації.** Результати дисертаційної роботи представлено у 63 працях, зокрема в 1 монографії, 34 наукових статтях у фахових виданнях (з них 22 одноосібних), 26 публікаціях у збірках матеріалів міжнародних конференцій та наукових журналах опубліковані основні наукові результати. Додатково наукові результати дослідження відображені у 2 навчальних посібниках.

**Структура та обсяг дисертації.** Дисертація складається зі вступу, 5 розділів, висновків, списку використаних джерел із 282 найменувань та 5 додатків. Повний обсяг друкованого тексту становить 416 сторінок, з них 28 сторінок списку використаних джерел та 81 сторінка додатків. Робота містить 76 рисунків і 22 таблиці по тексту та 8 рисунків і 6 таблиць на окремих 9 сторінках.

## ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** обґрунтовано актуальність теми дисертації, визначено науково-прикладну проблему, що вимагає розв'язання, показано зв'язок роботи з науковими програмами, планами, темами. Сформульовано мету і основні задачі досліджень, охарактеризовано наукову новизну та практичну цінність одержаних результатів. Наведено дані про особистий внесок здобувача, апробацію результатів, основні праці, опубліковані за темою дисертації, структуру та обсяг роботи.

У **розділі 1** проведено аналіз існуючих методів моделювання процесів обробки природномовної інформації та обґрунтовано структурно-функціональний підхід до розв'язання науково-прикладної проблеми.

Розглянуто існуючі підходи до моделювання семантики, що історично склалися у штучному інтелекті. На основі досягнень фізіології і генетики виник так званий висхідний або м'який напрям, представлений нейронними мережами, методами розпізнавання образів, генетичними алгоритмами тощо (У. Маккалок, Ф. Розенблат, А. Фразер, Д. Кросби). Дослідники низхідного або жорст-

кого напряму запропонували математичну інтерпретацію усвідомлених процесів абстрактного мислення – алгебра логіки та нечітка логіка (Д. Буль, Л. Заде), продукційні моделі, семантичні мережі (Р. Річенс), фрейми (М. Мінський), онтології, об'єктно-орієнтоване програмування (А. Кей), евристичні моделі, пошук семантичної інформації у великих масивах даних Data mining тощо. Проте 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». В той же час мовленнєва діяльність людини природним чином забезпечує навчання та отримання нових знань на основі вербальної інформації з тексту.

У комп'ютерній лінгвістиці більшість моделей належить до традиційної тріади морфологічного-синтаксичного-семантичного аналізу природномовного контенту (ПМК), в якій результати попереднього аналізу використовуються у наступному. Типовим прикладом такого підходу є модель «Зміст – текст» Московської семантичної школи (І.О. Мельчук, Ю.Д. Апресян), витоки якої можна зустріти в сучасних лексикології та лексикографії (В.А. Широков). Але, якщо морфологічний аналіз для природних мов можна вважати задовільним, то труднощі синтаксичного аналізу контекстно-залежних граматик та врахування значної кількості семантичних функцій призводять до загально низької якості розв'язання семантико-залежних задач.

У результаті аналізу відомих методів та моделей обробки природномовної інформації та побудови баз знань виявлено, що:

- відсутні цілісні концептуальні підходи та відповідні теорії, за головну мету яких покладено розуміння сенсу як прагматичної першооснови змісту;
- відсутнє поняття одиниці сенсу, не розроблено критерії для кількісної оцінки сенсу в ПМК та інших продуктах когнітивної діяльності людини;
- моделювання образного мислення на відміну від логічних та евристичних моделей має на даний час, переважно, концептуальний характер;
- ідею Тьюринга про онтогенетичний характер моделювання продуктів інтелектуальної діяльності практично не розвинуто у низхідному (жорсткому) напряму штучного інтелекту;
- відомі формальні методи не забезпечують прийняттого рівня розв'язку широкого кола актуальних семантико-залежних задач, а тому потребують розробки нові комплексні підходи та методи.

Проведено аналіз поглядів на об'єкт дослідження цілого ряду наук когнітивного спрямування, які, зауважимо, і до сих пір не дійшли до спільної думки щодо механізму образного мислення людини. Як основи моделювання обиралися лише загально визнані результати, а саме:

- з асоціативної психології: вільний асоціативний експеримент (К. Юнг) та асоціативно-вербальна мережа;
- з філософії: феноменологія (Е. Гуссерль, М. Хайдеггер), рефлексія (В.О. Лефевр, Г.П. Щедровицький);
- з фізіології вищої нервової діяльності: теорія І.П. Павлова;
- з психології: психоаналіз (З. Фрейд), теорія особистості;

– з лінгвістики: трансформаційний генеративізм (Н. Хомський), формування та розуміння мовних висловлювань.

Інтегральну роль для всіх цих досліджень грають лише поняття образу (гештальту) та асоціації, що відповідає зв'язку певної сили між двома образами.

На основі синдромного аналізу мозкових механізмів вищих психічних функцій з нейропсихології запропоновано структурно-функціональний підхід до побудови СОПМК. Результати нейропсихологічних досліджень онтогенезу вищих психічних функцій (О.Р. Лурія) демонструють єдину образну основу всіх типів мислення. Кожна вища психічна функція жорстко не прив'язана до якого-небудь одного мозкового центру, а є результатом системної діяльності мозку. Різні структури (аферентні та еферентні ланки) здійснюють власний специфічний внесок в побудову кожної такої функції.

Згідно з Лурією, головний мозок людини за своїми макрофункціями складається з трьох основних блоків: підтримка робочого тону кори, забезпечення тривалих селективних форм діяльності; отримання, переробка і зберігання екстероцептивної інформації; програмування, регуляція і контроль діяльності мислення. Проведено кібернетичну інтерпретацію нейропсихологічних даних, що дозволило виокремити з основних блоків такі функціональні блоки-моделі:

1. Блок емоцій як ефективна й афективна складові нервової системи (стовбурові відділи мозку).

2. Виконавчі механізми (моторні відділи мозку).

3. Надоперативна пам'ять або буфер для побудови складних образів (скронева область).

4. Блок кінестетичних мелодій або буфер для створення складних рухів (передмоторна область).

5. Оперативна пам'ять як ансамбль з  $7 \pm 2$  образів (так звана зона *ТРО* або перекриття скроневої, тім'яної і потиличної кори).

6. Довготривала пам'ять (потилична область).

7. Асоціативна пам'ять (тім'яна область).

8. Блок моделювання і вибору стратегій (лобово-скроневі відділи).

9. Блок програмування, регулювання і контролю (лобові доли).

Обґрунтовано характерні особливості структурно-функціонального підходу, що пропонується та висунуто провідну ідею дослідження – отримати розв'язки актуальних семантико-залежних задач у вигляді комплексу функцій інформаційної системи обробки природномовного контенту. Показано хибність погляду на те, що розуміння мовних висловлювань можна досягти лише на основі використання словника як системи понять у поєднанні з системою граматичних правил. Проблема розуміння полягає у складному характері взаємодії психічних процесів, що одночасно відбуваються у напрямках пошуку загальної думки та декодування на лексико-фонемному і синтаксичному рівнях.

З метою врахування екстралінгвістичних знань окреслено множину значимих понять образного аналізу – образ, асоціацію, подію, фокус уваги, потреби та емоції, орієнтувальний рефлекс, несвідоме, підсвідоме та свідоме. Значенням слова, що є формальною вербальною ознакою образу, прийнято 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 індивідуальна прагматика слова, наслідок відокремлення таких зв'язків із загальної системи асоціацій, що мають відношення до даного моменту і даної ситуації.

Для розв'язання науково-прикладної проблеми пропонується під сенсом у широкому значенні вважати визначену за мету пізнавальної діяльності людини деяку властивість вербальної інформації. Тоді образний сенс – 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 ПМК на основі формальної метамови у вигляді семантичної мережі як моделі сенсу-властивості, простору образного сенсу на основі відповідної одиниці сенсу та 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.

Вибір критеріального апарата для оцінки досягнення мети дослідження побудовано на основі аналізу широкого кола семантико-залежних задач, актуальних для комп'ютерної лінгвістики, електронного навчання та інших предметних областей. В табл. 1 найбільш значимі функціональні вимоги до інфологічної системи (ІС) відображено у відтинках штриховки від відсутніх до частково реалізованих у існуючих системах та узагальнено за чотирма категоріями.

Таблиця 1 – Функціональні вимоги до інфологічної системи

<i>Підтримка деяких типів діалогу</i>	<i>Лінгвістичні можливості</i>	<i>Дидактичні можливості</i>	<i>Мультимедіа, пошук, переклад</i>
У предметній області (1.1)	Морфологічний аналіз ПМК (2.1)	Презентація знань (3.1)	Сенс мультимедіа (4.1)
Сценарний варіант (1.2)	Синтаксичний аналіз ПМК (2.2)	Оцінювання знань (3.2)	Пошук за маскою (4.2)
«Дельфійський оракул» (1.3)	Інваріантність сенсу (2.3)	Самовдосконалення бази знань (3.3)	Пошук за правилами (4.3)
«Магістр Йо-да» (1.4)		Адаптивне навчання (3.4)	Семантичний пошук (4.4)
«Basic English» (1.5)		Генерація повідомлень (3.5)	Багатомовність (4.5)

Діалог на природній мові (тест Тьюрінга) як універсальний засіб перевірки інформаційної системи на інтелектуальність відносять до *AI*-повних задач. Для задач дослідження достатньо, щоб ІС підтримувала обмежені типи діалогу – у межах лексики предметної області (1.1), за певним тестовим сценарієм (1.2), відповідь надається як: множина слів, асоціативно пов'язаних з питанням (1.3),

цитата з вихідних текстових джерел (1.4), каркас речення з відхиленнями від синтаксису (1.5).

Результат підвищення якості розв'язання комплексу семантико-залежних задач обробки інформації як мети дослідження можна оцінити шляхом порівняння розв'язків, отриманих запропонованими у межах створеної методології та відомими методами. Якість побудови інформаційної технології на основі СОПМК також характеризується обсягом комплексу семантико-залежних задач. На основі мети та визначених функціональних вимог до ІС сформульовано основні задачі дослідження.

У **розділі 2** закладено основи теорії образного аналізу текстової інформації та формалізовано ключове поняття інфологічної системи.

Визначено клас інфологічних систем як частковий випадок більш загального класу кібернетичних систем, що орієнтований на моделювання образної основи психічної діяльності людини. В ІС принципово немає зовнішнього об'єкта управління, натомість:

- а) «чорним ящиком» служить весь навколишній світ;
- б) як критерії можуть застосовуватися 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;
- в) діалог та інші форми мовленнєвої діяльності можуть розглядатися як результат взаємодії двох чи декількох схожих інфологічних систем.

Нехай інфологічна система  $S$  здатна розпізнавати окремі образи з нескінченної множини  $I = \{i_1, i_2, \dots, i_n, \dots\}$  аналогічно тому, як людина розпізнає гештальт. ІС також може сприймати асоціативні зв'язки між парами образів як елементи множини  $\omega \in \Omega$ , де  $\Omega \subseteq I \times I$  – довільна множина упорядкованих пар. Образною конструкцією (ОК) будемо вважати будь-яку підмножину  $\gamma \subseteq \Omega$ , яка є елементом  $\mathbf{F}$  –  $\sigma$ -алгебри підмножин з  $\Omega$ . Отже, всі ОК мають властивість  $\gamma \in \mathbf{F}$ .

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  $x_i \in \mathbf{F}$ ,  $y_i \in \mathbf{F}$ . Введено онтогенетичний принцип побудови системи  $S$  – множина  $\Omega$ , що визначає сенс її функціонування, отримується виключно шляхом послідовного накопичення оператором  $\oplus$  параметрів чергових  $\omega$  з зовнішнього «чорного ящику» та подальшого самовдосконалення  $\Omega$ . Формально базу знань системи  $S$  будемо як  $B = \bigcup_{i=1}^{m'} x_i$ , де  $m'$  – загальна кількість сприйнятих системою на даний час вхідних ОК.

Введено поняття мовного образу (МО), що відрізняється головною символічною ознакою у вигляді кореня слова. Таке поняття більш загальне за синсет з відомого WordNet, а також словарну статтю чи лексему, у формі яких будуються складові тезаурусів та онтологій. Формально підмножину  $I' \subset I$  доступних для інфологічної системи МО представлено за допомогою четвірки ос-

новних змістовних концептів:  $I' = \langle N; O; M; Q' \rangle$ , де  $N$  – поняття,  $O$  – об'єкт,  $M$  – метод,  $Q'$  – якість. Взаємозв'язок основних понять образного аналізу ПМК представлено на рисунку 1.



Рисунок 1 – Взаємозв'язок формальних понять образного аналізу ПМК

З метою визначення оператора  $\oplus$  концептуальної моделі для ОК  $\gamma \in \mathbf{F}$  задано формальну теорію  $Th$  як прикладну теорію першого порядку, а саме:

1. Введено скінчений алфавіт з  $Al = \{A, B, \dots, Z, x_1, x_2, \dots, x_n, t_1, t_2, t_3\}$  – змінних;  $Con = \{\emptyset, 1, \dots, n\}$  – констант;  $\{\backslash, \oplus\}$  – символів бінарних операцій ( $\backslash$  – зв'язок між двома образами в асоціативній парі  $\omega \in \Omega$ ,  $\oplus$  – операція об'єднання образних конструкцій «АБО ОК»);  $\{=\}$  – предикатного символу «дорівнює» у значенні теорії множин;  $\{\neg, \rightarrow, \forall\}$  – логічних зв'язок та кванторів, де  $\neg$  – заперечення (ні),  $\rightarrow$  – логічне слідування (якщо ..., то ...),  $\forall$  – квантор загальності; дужок “(“, “)” та коми “,”.

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

$$\langle \text{Терм} \rangle ::= x_i j \mid x_i \in Al, j \in Con;$$

$$\langle \text{Терм} \rangle ::= \langle \text{Терм} \rangle \langle \text{Терм} \rangle.$$

Літерами  $t_1, t_2, t_3 \in Al$  позначаються терми, побудовані в асоціативній нормальній формі (АНФ), а саме

$$\langle \text{АНФ}\omega \rangle ::= x_i \backslash x_j \mid x_i, x_j \in Al;$$

$$\langle \text{АНФтерм} \rangle ::= \langle \text{АНФ}\omega \rangle;$$

$$\langle \text{АНФтерм} \rangle ::= \langle \text{АНФтерм} \rangle \oplus \langle \text{АНФтерм} \rangle,$$

де  $\langle \text{АНФ}\omega \rangle$  названо елементарним термом у АНФ.

3. Виділено множину формул, які є логічними аксіомами (3.1÷3.3 – числення висловлювань, 3.4÷3.5 – числення предикатів першого порядку); власними ак-

сіомами (3.6÷3.11 – аксіоми комутативної напівгрупи, 3.12÷3.14 – прикладні аксіоми (продукції) теорії):

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. Визначено правила виведення

$A, A \rightarrow B \vdash B$  «*Modus ponens*»;  
 $A \vdash (\forall t)A$  «правило узагальнення»,

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

В теорії  $Th$  доведено такі власні теореми:

Теорема 1.  $\langle Терм \rangle \rightarrow \langle АНФтерм \rangle$ .

Теорема 2.  $\langle АНФтерм \rangle \rightarrow \langle АНФq \rangle \oplus \langle АНФ? \rangle \oplus \langle АНФа \rangle$ ,

де  $\langle АНФ\omega \rangle = x_i \setminus x_j \mid x_i, x_j \in Al$  позначається як  $\langle АНФ? \rangle$ ;  $\langle АНФа \rangle$  – всі елементарні терми з  $\langle АНФтерм \rangle$ , в яких символ  $x_j$  є першим, потім рекурсивно підставляється наступний символ за принципом пошуку у глибину (якщо в рекурсії знаходиться  $\langle АНФ? \rangle = x_j \setminus x_i$ , то символ  $x_i$  та всі наступні за ним не враховуються);  $\langle АНФq \rangle$  – всі інші окрім  $\langle АНФ? \rangle \oplus \langle АНФа \rangle$  терми з  $\langle АНФтерм \rangle$ .

Розглянуто модель формальної теорії  $Th$  як комутативну напівгрупу ОК. Під термом розуміється образна конструкція простого речення, а під формулою теорії – образний аналог логічного природномовного виразу. Отримана комутативна напівгрупа ОК на основі моделі теорії  $Th$  забезпечує представлення ОК, що відповідає синтагмі у вигляді 3-х складових питальної конструкції мовних образів.

Отримано формальні характеристики онтогенетичного методу створення бінарного нечіткого відношення образного сенсу  $Q$  інфологічної системи  $S$ , що задане на одній базисній множині (універсумі) образів  $I$ :

$$Q = \{ \langle i_l, i_j \rangle, \mu_Q(\langle i_l, i_j \rangle) \}, \quad (1)$$

де  $\mu_Q(\langle i_l, i_j \rangle)$  – функція належності бінарного нечіткого відношення, що задається як відображення  $\mu_Q : I \times I \rightarrow [0,1]$ . Через  $\langle i_l, i_j \rangle$  позначено елементарний терм в АНФ  $\langle АНФ\omega \rangle$  як кортеж з двох елементів, причому  $i_l \in I, i_j \in I$ . Відношення (1) застосовано як онтогенетична характеристика множини  $\Omega$ , тоді функцію належності

$$\mu_Q(\langle i_l, i_j \rangle) = f(k_{lj}, t_L), \quad (2)$$

де  $k_{ij}$  – кількість зафіксованих ІС зв'язків між  $l$ -м та  $j$ -м образами на момент часу  $t_L$ , запропоновано вважати природною чисельною мірою сенсу 1 *Sav* (синтагматичної асоціації вага) або *Saw* (Syntagmatic association weight).

У відповідності до концептуального підходу деталізовано функцію належності, що породжує бінарне нечітке відношення сенсу (1) на чотирьох послідовних рівнях, побудованих на базовому (2):

1. Рівень імовірнісного прогнозування – з метою нормування функції належності у проміжку  $[0,1]$  передбачено розрахунок статистичної оцінки  $\lambda$  (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 застосовано відому сигмоїдальну функцію

$$\mu_Q(< i_l, i_j >) = f_1(k_{lj}, \lambda) = 1 / (1 + e^{-k_{lj} + \lambda}). \quad (3)$$

Внаслідок нормування з'являється характерна властивість функції належності, отриманої за онтогенетичним методом, – середнє значення  $\mu_Q = \frac{1}{m} \sum_{j=1}^m \mu_{Qj} = 0,5$ .

2. Рівень врахування емоційного стану – введено можливість врахування ступеня узагальненої емоції «негатив–позитив» ІС за рахунок показника  $\mu = \{ \dots, -2, -1, 1, 2, \dots \}$ , тоді

$$\mu_Q(< i_l, i_j >) = f_2(k_{lj}, \lambda, \mu) = 1 / (1 + e^{\frac{k_{lj} - \lambda}{|\mu|}}). \quad (4)$$

При  $\mu = -1 \vee 1$  емоції не впливають на сенс функціонування ІС, а функція належності (4) вироджується у функцію (3). Збільшення показника  $\mu$  симетрично згладжує сигмоїдальну функцію  $f_2$ , що продемонстровано на рисунку 2.

3. Рівень врахування мотиваційної компоненти на основі образів-центрів потреб – запропоновано моделлю мотиву ІС вважати досягнення образу-центру потреби  $j'$  на момент часу  $t_L$ .

4. Рівень врахування рефлексів та результатів зовнішнього навчання – запропоновано фіксувати в ІС статистику результатів образної реакції  $y'_i$  на вхідну ОК  $x'_i$  при мотиваційній меті  $i'$  як зміну емоційного параметра  $\Delta\mu$ .

Обґрунтовано аксіоматику нечіткої міри сенсу  $Se : \mathbb{F} \rightarrow [0, \infty]$ :

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  $\gamma_1, \gamma_2 \in \mathbb{F}$  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 ( $\gamma_i \cap \gamma_j = \emptyset, i \neq j$ ), то  $Se(\bigcup_{i=1}^{\infty} \gamma_i) = \sum_{i=1}^{\infty} Se(\gamma_i)$  ( $\sigma$ -адитивність).

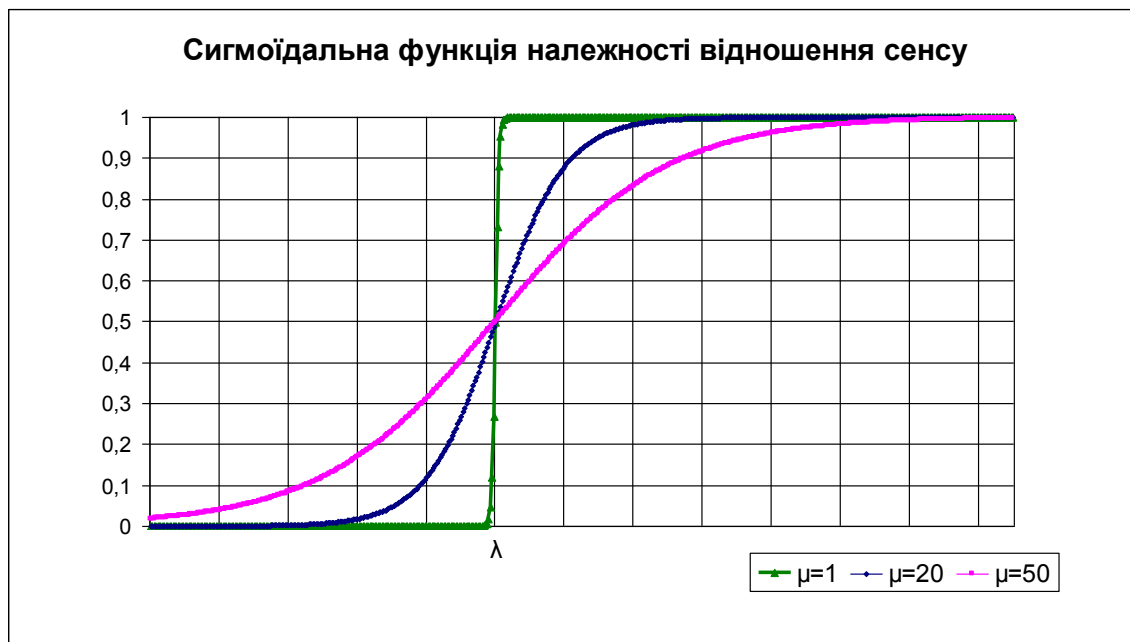


Рисунок 2 – Вплив показника  $\mu$  на функцію належності (4) відношення сенсу

На основі аксіоматики побудовано топологічний і квазіметричний простір упорядкованих пар образів  $(\Omega, \mathbf{F}, Se)$ . Надано пізнавальну інтерпретацію таким операціям з нечіткими відношеннями сенсу, як перетин, об'єднання, різниця та композиція. Отримано верхню оцінку співвідношення одиниць образного сенсу та інформації як логарифмічну згортку  $1 [Cav] = 2 \log_2 m [Bim]$ , що обумовлена системою обмежень у вигляді семантичної мережі АМО.

Конструктивну реалізацію абстрактного поняття ІС досягнуто шляхом створення структурно-функціональної моделі образної обробки ПМК. 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).

Надано характеристику основним функціональним блокам та інформаційним потокам образної природи, що вимагають обов'язкового врахування в ІС з метою формалізації функцій СОПМК та розв'язку семантико-залежних задач. Запропоновано структурну схему образної обробки ПМК, що 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.

На основі інтерпретації процесів образного мислення та формального визначення інформаційних потоків створено граф загальної структурно-функціональної моделі образної обробки ПМК. 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** присвячено обґрунтуванню методів моделювання процесів образної обробки природномовного контенту.

Покладено вважати вузлами АМО відомі вербальні ознаки  $n$  образів, які можуть поєднуватися між собою спрямованими асоціативними зв'язками від головного образу до підлеглого. В процесі функціонування ІС фіксує кількість  $k_{ij}$  виявлених зв'язків між  $i$ -м та  $j$ -м образами, 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 вважатимемо, що вага асоціативного зв'язку дорівнює

$$c_{ij} = \mu_Q(\langle i_i, i_j \rangle) = F_l(k_{ij}) = Se_{ij}, \quad l \in \{1, 2, 3, 4\}, \quad (5)$$

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

З метою формалізації концептуальної моделі процесів образної обробки ПМК побудовано граф  $G = \{V, E\}$ . Внаслідок обмеження кількості  $n$  відомих інфологічній системі  $S$  вершин нескінченного орієнтованого графа  $G$  на певний момент часу отримано скінчений граф  $G_z$ . Введено поняття ваги та шляху для двостороннього орієнтованого графа такого типу. Нехай  $(a_i, b_j) \in Y_z$ ,  $a_i \subset Y_+$ ,  $b_j \subset Y_-$ . Тоді вага  $c_{ij}$  для ребра  $(a_i, b_j)$  визначається згідно з (5). Особливість задач на графі  $G_z$  у відповідності до мовної практики полягає у тому, що вершина не може одночасно бути початковою та кінцевою в одному шляху.

На відміну від довільного орієнтованого графа запропонований метод моделювання 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 цих особливостей введено такі формальні визначення:

Мовний образ – вершина графа, що з'єднуються з іншими вершинами не менше, ніж двома ребрами. Кожна вершина має унікальну множину зв'язків з іншими суміжними вершинами, а між будь-якими двома вершинами графа

$G_z$  розмірністю  $n$  обов'язково існує скінчений шлях довжиною не більше за  $(n+1)/2$ .

Пара МО – ребро між двома суміжними вершинами графа з вагою, що визначається на основі (5). Пара МО як компонент події фіксує одиничний асоціативний зв'язок від головного образу до підлеглого, тому при внесенні нових синтагм до моделі образної пам'яті необхідно збільшувати на 1 лічильники  $k_{ij}$  для тих пар, з яких складається синтагма. 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 мовними образами пари – у цих випадках для перетворення орієнтованого графа  $G_z$  у неорієнтований граф  $G'_z$  запропоновано застосувати таке співвідношення для ваги ребра:

$$c'_{ij} = \mu_Q(\langle i_i, i_j \rangle) + \mu_Q(\langle i_j, i_i \rangle) = Se_{ij} + Se_{ji}. \quad (6)$$

Синтагма – підграф розмірністю  $m$  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 дорівнює  $m \ll n$ ;
- у кожній вершині дерева кількість підлеглих вершин не більша за  $m-1$ .

Розглянуто альтернативну можливість інтерпретації значимих концептів образного аналізу та синтезу шляхом ігнорування напрямку асоціативного зв'язку між парою образів та визначення його загальної сили згідно з (6). Будемо позначати через  $P(I)$  множину-ступінь (булеан) множини  $I$ . За визначенням гратка  $P(I)$  є скороченим позначенням для  $P(I) = \{Y \mid Y \subseteq I\}$ .

Вважатимемо графічне зображення множини-ступеня  $P(I)$ , що задовольняє попередньому опису, одиничним кубом  $n$ -вимірного простору (в  $n$ -вимірній системі координат). Приклад графа булеану для випадку  $n=4$  наведено на рисунку 3.

Багатошарова графічна проекція гратки булеану ( $n$ -вимірного багатогранника) на площину має такі ознаки:

- 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  $i$ ):

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

- загальне число зв'язків між вузлами графа:

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

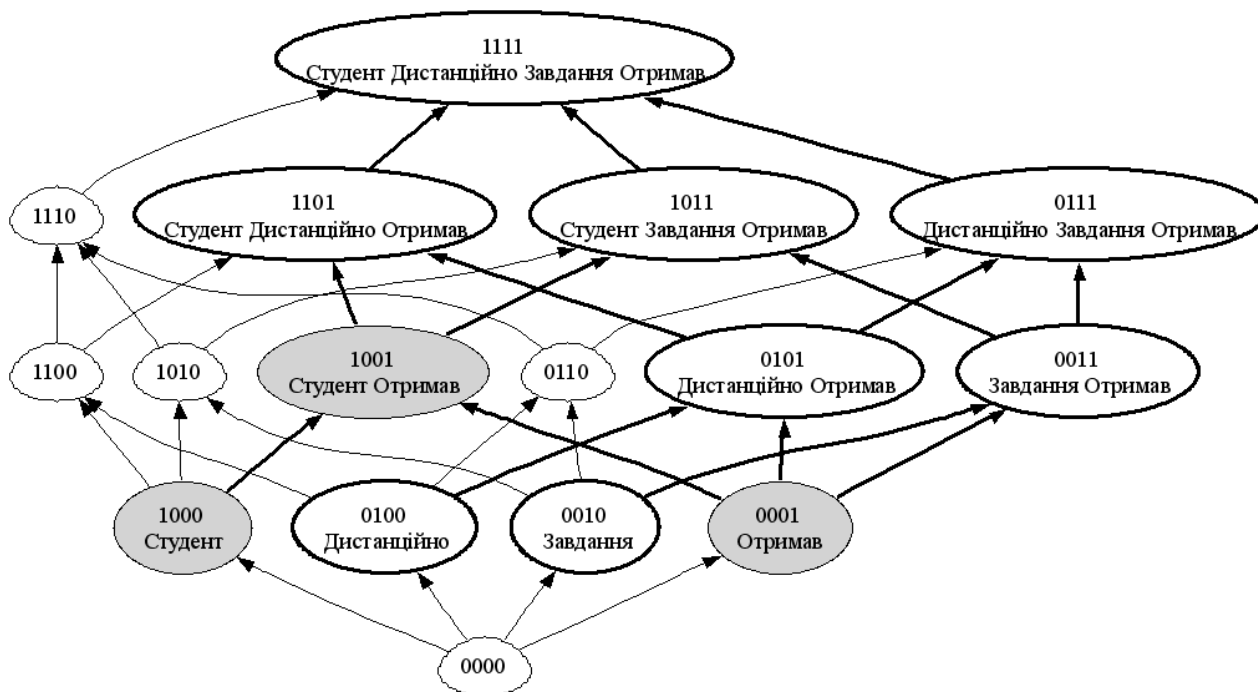


Рисунок 3 – Приклад графа булеану синтагми для випадку  $n=4$

Запропоновано систему кодування вузлів булеану на основі  $n$ -вимірних бінарних кодів з введенням глобальних нуля  $\Theta$  та одиниці  $\Delta$ . Забезпечено можливість використання відомих властивостей дистрибутивної обмеженої ґратки з доповненням до кожного елемента для розв'язання задач дослідження. Введено формальне поняття «піраміди сенсу», що дозволяє уніфікувати процеси образної обробки ПМК шляхом навігації між вузлами булеану на зразок синтагми з рисунку 3.

Запропоновано вважати булевою алгеброю сенсу (БАС) двохосновну алгебраїчну систему

$$BAS = \langle B; \Omega_b \rangle, \quad (7)$$

де  $B = \{Word, Number\}$  – основи, а  $\Omega_b = \{OP, RE, IF\}$  – сигнатура системи. В межах першої з основ  $Word$  варто розрізняти такі види символічних послідовностей (слів):

$$Word = \{Image-Word, Link-Type, Re-Word\},$$

де слова, що позначають мовні образи  $Image-Word$ , класифікуються в залежності від їх ролі в синтагмі, а саме:  $Object-Quality$  – якість об'єкта,  $Object$  – об'єкт,  $Notion$  – поняття,  $Method$  – метод,  $Method-Quality$  – якість методу:

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

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

$$\text{Link} - \text{Type} = \{r1, r2, r3, r4, r5, r6, r7\},$$

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* позначаються атрибути відношень, що мають бути внесеними до блоку пам'яті інфологічної системи. Як другу основу *Number* запропоновано розглядати числа таких видів:

$$\text{Number} = \{Bi, \text{Logic}, \text{Integer}, \text{Time}\},$$

де *Bi* – *n*-вимірний бінарний код вузлів булеану;

*Logic* – логічні значення 0 (False) та 1 (True):  $\text{Logic} = \{0,1\}$ ;

*Integer* – цілі невід'ємні числа:  $\text{Integer} = \{x \mid x \in \mathbb{Z}^+\}$ ;

*Time* – дійсні невід'ємні числа:  $\text{Time} = \{x \mid x \in \mathbb{R}^+\}$ .

Для *n*-вимірного бінарного коду прийнято позначення  $Bi_i$  – це код, в якому налічують *i* одиниць та, відповідно, *n* - *i* нулів. Очевидно, що вузол з кодом  $Bi_i$  належить до *i*-го шару графа булеану. На основі БАС отримано формальні теоретико-множинні визначення для таких феноменологічних понять моделі образної обробки ПМК, як синтагма, текст, фокус уваги, а також види пам'яті – довготермінова, асоціативна, оперативна та надоперативна

Сигнатура алгебраїчної системи  $\Omega_b$  складається з операцій *OP*, відношень *RE* та предикатів *IF*. До складу базових операцій БАС запропоновано відносити

$$\text{OP} = \left\{ \begin{array}{l} \text{Add} - \text{Image}, \text{Add} - \text{Twice}, \text{Evaluate} - \text{Level}, \text{Find} - \text{Image}, \\ \text{Level} - \text{Up}, \text{Level} - \text{Down}, \text{Change} - \text{Image}, \text{Sense} - \text{Pyramid} \end{array} \right\}.$$

Базовий склад предикатів БАС сформовано як

$$\text{IF} = \{\text{If} - \text{Image}, \text{If} - \text{A} - \text{Twice}, \text{If} - \text{Event}\}.$$

Відношеннями алгебраїчної системи, які дозволяють розглядати складові БАС у термінах реляційної моделі даних (рис. 4), вважаються

$$RE = \left\{ \begin{array}{l} \text{Image, Assoc - Twice, Construct, Event,} \\ \text{Interrogative - Pronoun, Link, Text, Words, Role} \end{array} \right\}. \quad (8)$$

Відношення  $RE$  увійшли до складу сховища ІС – декларативної складової бази знань ІС. Проведено порівняльний аналіз концептуальних понять dBase-машини та сховища ІС на рівнях органів відчуття (у вигляді ОК) та мовних засобів (як КМО та ПМК). Відсутність у сховищі ІС чітко визначених атрибутів даних приводить до відсутності сталої структури та розміру кортежів. Окрім цього відмінностями є:

- принципова неможливість модифікації даних у 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 природних інтелектуальних систем.

Показано, що обсяг сховища ІС має поліноміальну складність не більше 3-го порядку внаслідок включення до відношень БАС бінарних  $n$ -вимірних полів з кодами  $Vi_i$  на рівні образів, асоціативних пар, синтагм та текстів.

У розділі 4 розглянуто синтез функцій образного пошуку та генерації знань системи обробки природномовного контенту.

У відповідності до відомих ознак онтогенезу розвитку мовлення запропоновано формальну класифікацію образного пошуку, в якій враховано особливості парадигматичного мислення при безумовному пріоритеті синтагматичного. Класифікація формалізує 15 варіантів пошуку для можливих сполучень 3-х основних конструктивів моделі образного сенсу: 5 варіантів можна віднести до базового синтагматичного типу, а 10 – до похідного парадигматичного типу. Отриману класифікацію взято за основу побудови функціональної моделі СОПМК (таблиця 2) у відповідності до задач дослідження та можливості зручної інтерпретації 3-ма типами графів, визначених методами моделювання функцій образного аналізу та синтезу.

Задачу асоціативного пошуку представлено у побудові таких таблиць, де кожній з  $n$  вершин графа  $G'_z$  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  $G'_z(V, E')$  згідно з (6). Показано, що розмірність розроблених алгоритмів  $AList^{out}$  та  $AList^{in}$  залежить від обраного методу сортування, але не перевищує значення  $o(n \cdot \bar{m}^2)$ , де  $\bar{m}$  – середня кількість вихідних або вхідних ребер однієї вершини орієнтованого графа. Розмірність  $AList'$  зменшено від  $o(n^3 \cdot \bar{m}^2)$  до  $o(n \cdot \bar{m}^2)$  шляхом використання вже відомих алгоритмів  $AList^{out}$  та  $AList^{in}$ .

Таблиця 2 – Функціональна модель СОПМК

Складові СОПМК	Позначення	Базові алгоритми та прикладні задачі						
		Семантичний пошук			Інваріантний зміст та 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	$BAS$ $\langle B, \Omega_b \rangle$				$InsZX$ $InsZY$ $SynZZ$	$Find \cdot n'$	$Mech$ $Or \cdot Re$	$PartXY$ $PartXZ$ $PartYZ$

Задачу інсайтного пошуку інтерпретовано як пошук у множині вершин графа  $V_z = \{i \mid 1 \leq i \leq n\}$ . Розглянуто випадок, коли 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:

$$a) G_z(V, E) (a_{ik} + a_{kj}) \xrightarrow{Find \cdot 1} Max \text{ у відповідності до (3);}$$

$$b) G'_z(V, E') (a'_{ik} + a'_{kj}) \xrightarrow{Find \cdot 1'} Max \text{ у відповідності до (4).}$$

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  $m \ll n$  такий підхід вже може дати позитивний результат. Але, зважаючи на попередню відсортованість списків, обґрунтовано застосування методу обмежувачів з метою подальшого зменшення розмірності алгоритмів оптимізації типу 
$$\bigcap_{AList'_i, k = AList'_j, k} Max(AList'_i.a_{ik} + AList'_j.a_{kj}).$$

Знайдено мінімальний за довжиною  $n'$  шлях  $Find \cdot n'$  ( $Bi - I_1, Bi - I_2$ ) між вершинами  $Bi - I_1$  та  $Bi - I_2$  з  $G'_z$  з максимальною сумарною вагою ребер

$Se^{\Sigma} = \sum_{i=1}^{n'-1} c_{i,i+1}$ , якщо вага кожного ребра визначається згідно з (6). У загальному

випадку вирішення задачі вимагає прямого перебору розмірністю  $o((n-1)!)$  можливих варіантів маршруту, в який входять  $n$  вершин графа. Запропоновано врахувати визначені раніше особливості розрідженого графа АМО, що дозволило зменшити розмірність побудови маршруту до  $o(\bar{m}^L)$ , де  $\bar{m}$  – середня кількість суміжних з вершиною ребер неорієнтованого графа, а  $L$  – довжина маршруту у ребрах. Отримано оптимізаційний алгоритм, що формально узагальнює такі етапи:

- 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 виключенням вже задіяних у попередніх ітераціях.
- Для зберігання ланок маршруту використовується стек  $Route_i$ , причому внесення нової ланки маршруту (вершини) до стеку супроводжується перевіркою на зациклення, оскільки у стеку не повинно бути двох однакових вершин.
- Отриманий опорний маршрут оптимізується з метою зменшення  $L$ , оскільки другою складовою цільової функції виступає лаконічність висловлювання.

Розмірність розробленого алгоритму  $Find \cdot n'$  зменшується до  $o(L + \sum_{i=1}^{L-4} i)$ ,

де  $L \leq (n+1)/2$ , проте можна бути впевненим лише у існуванні одного прийняттого розв'язку задачі, а не у знаходженні найкращого з них. Для пошуку найкоротшого шляху  $s = v_1, v_2, \dots, v_{k-1}, v_k = t$  у  $G_z$  перевизначені елементи  $a_{ij} = \text{Max}_{(i,j) \in E_z} (c_{ij}) - c_{ij} + 1$  матриці суміжності  $A_Q$ . Задачу розв'язано на основі модифікації алгоритму вичерпної дії  $Find \cdot n$ , побудованого на основі принципу «пошук у глиб з поверненням».

Визначено поняття та сформульовано принципи, які дозволили побудувати метод моделювання механізму оперативної пам'яті СОПМК на основі формальної інтерпретації психофізіологічних феноменів ансамблю образів, фокусу уваги, орієнтувального рефлексу, вектору емоцій. Останній формалізовано множиною образів потужністю  $n_e \ll n$ , заданих як «центри потреб» згідно з рівнем 3 визначення функції належності (2) бінарного нечіткого відношення сенсу  $Q$ . Подальша деталізація запропонованого методу моделювання привела до алгебраїчної постановки задачі навігації між вузлами булеану від 5-го шару до 9-го, що відповідають ансамблю образів або бінарним кодам подій, з унеможливленням зациклення алгоритму.

Шляхом введення формальних обмежень орієнтувальний рефлекс представлено задачею оптимізації механізму оперативної пам'яті СОПМК (рисунок 4):

$$\begin{aligned} Bi - OM, \text{ Vector} - Set \rightarrow \text{Focus} - Bi | \\ \text{Focus} - Weight \geq \text{Weight} - Top, \end{aligned} \quad (9)$$



де *Bi-OM* – бінарний код оперативної пам'яті;  
*Vector-Set* – стек з множиною образів–складових вектора емоцій;  
*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*;  
*Weigh-Top* – достатнє (прийнятне) значення вектора емоцій.

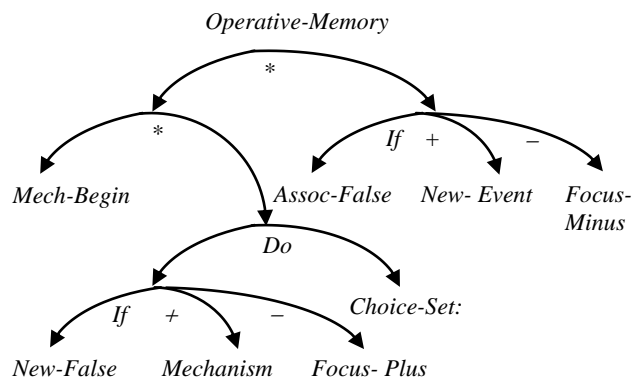


Рисунок 4 – Граф-схема оператора *Operative-Memory*

Оптимізаційну задачу (9) розв'язано оператором *Orient-Reflect* – оператор *Operative-Memory* виконується у циклі до тих пір, доки предикат *Stop-Find* не стане істинним:

$$\text{Orient - Reflect} ::= \{ [\text{Stop - Find}] \\ \text{Operative - Memory} \}.$$

Визначено правила та формально реалізовано можливість створювати парадигматичні класифікації образів через концепти об'єкта, метода та якості як складових МО. Запропоновано метод самовдосконалення бази знань інфологічної системи на основі моделювання таких складових парадигматичного устрою мови:

- 1)  $P_c$  – синонімічна асоціація або результат образного пошуку *SynXX*;
- 2)  $P_a$  – антонімічна асоціація (*AntXX*);
- 3)  $P_{\text{ч}}$  – асоціація типу часткове–загальне (гіпо–гіперонімія *PgXX*);
- 4)  $P_{\text{ц}}$  – асоціація типу частина–ціле (меронімія *CmXX*);
- 5)  $P_o$  – омонімічна асоціація як частковий випадок паронімії;
- 6)  $P_p$  – асоціація типу “у риму”.

Алгоритм *SynXX* визначення синонімічної асоціації  $P_{c_{ij}}$  для якості методу складається з таких кроків:

**[Крок 1]** Досліджується  $i$ -а 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]** Організується 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;

- b) визначається потужність  $\bar{x}$  множини  $X = \{l\} \cap \{m\}$ ;  
 c) 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  $i$ .

В алгоритмі *SynXX* використано 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  $\bar{x}_{\max} = 0$ .

У межах запропонованого методу отримано рекурсивні алгоритми для побудови 5-ти інших типів парадигматичних зв'язків на основі аналізу стану АМО. Формально введено загальну функцію сенсу ІС, яка залежить від кількості  $m$  ненульових елементів матриці  $A_Q$  та суми  $k_{\Sigma} = \sum_{i=1}^n \sum_{j=1}^n k_{ij}$  всіх елементів

матриці

$$Se^S = F(m, k_{\Sigma}). \quad (10)$$

Показано, що загальна кількість сенсу (10) в ІС наближається до  $Se^S \approx m/2$ . Збільшення сенсу досягається у випадках створення нових зв'язків між МО – синтагматичного (на 0.5 САВ), парадигматичного ( $\approx 3.5$ ), а також при внесенні у систему нового МО ( $m'$ ). Адекватність всіх запропонованих у розділі методів та алгоритмів образного пошуку і генерації знань СОПМК проілюстровано на експериментальних даних наскрізного тестового прикладу «WEB-технології: стандарти Semantic WEB».

**Розділ 5** присвячено реалізації інформаційної технології образної обробки природномовного контенту та розв'язанню семантико-залежних задач.

На основі отриманої у попередніх розділах роботи методології образного аналізу та синтезу структурно-функціональних моделей систем побудовано інформаційну технологію, що характеризується, на відміну від існуючих аналогів, сукупністю таких формальних обмежень:

- 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;
- 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;
- 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;
- на рівні синтезу структурно-функціональних моделей СОПМК – механізм онтогенезу ІС на основі кібернетичної інтерпретації нейропсихологічних даних; склад блоків, функцій та образних потоків моделі ІС; представлення нескінченних множин скінченними; вибір основ та сигнатури БАС; обмеження

функцій СОПМК класифікацією образного пошуку та функціональними вимогами до ІС;

- е) 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.

Розроблено схему організації інформаційних процесів та процедур технології образної обробки ПМК (рисунок 5), яка забезпечує внесення попередньо відібраного текстового матеріалу в напівавтоматичному або автоматичному режимах, побудову бази знань у вигляді семантичної мережі АМО та розв'язок ряду семантико-залежних задач на основі комплексу функцій СОПМК.

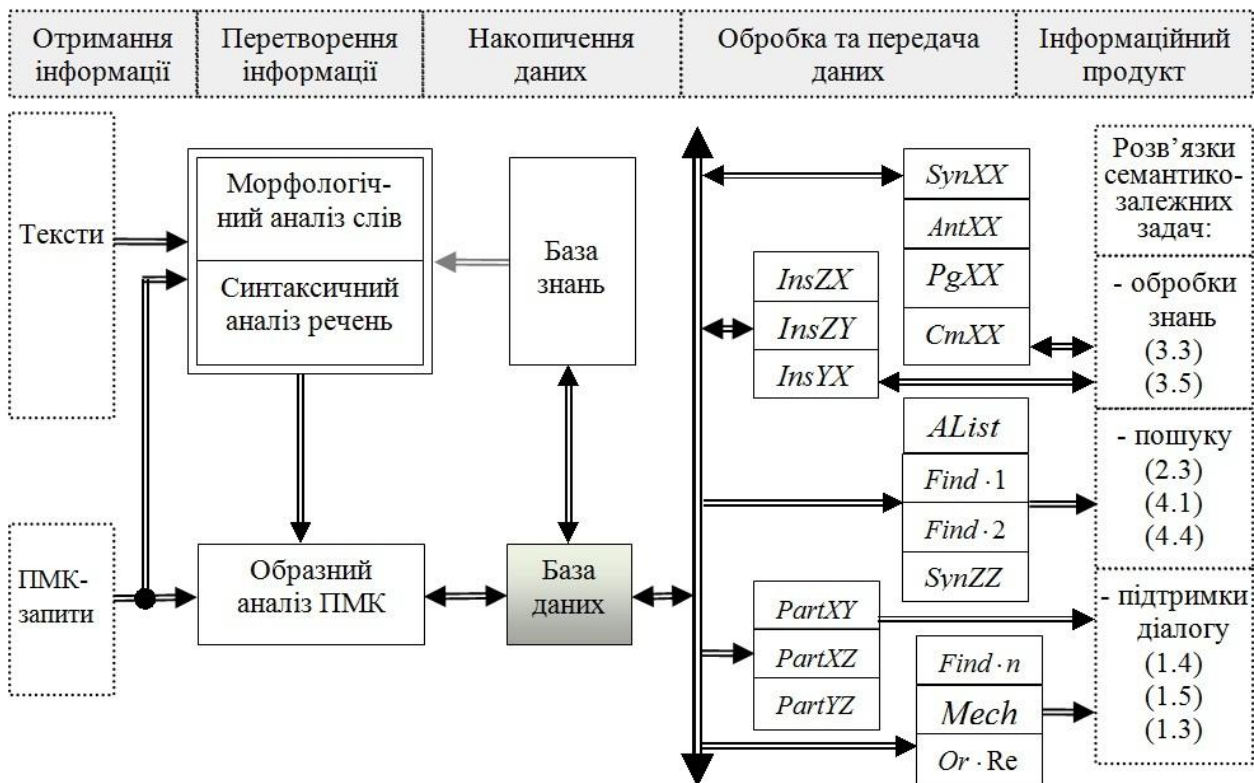


Рисунок 5 – Схема організації інформаційних процесів та процедур ІТ

Розглянуто змістовну, цільову, операційну та технологічну складові типового навчального процесу. Запропоновано в якості формальних ознак результатів пізнавальної діяльності вважати: для знань – топологічні характеристики фрагментів АМО, а для вмінь – чисельні оцінки образного сенсу окремих зв'язків, що забезпечує їх порівняння з елементами електронного контенту. Отримано та перевірено на даних наскрізного прикладу оцінки образного сенсу для головних конструктів ПМК – МО, сенсосполучення та синтагми, причому два останніх мають такі межі:

$$0 \leq Se_{ij}^p < n_i + n_j,$$

$$0 \leq Se_i^e < \sum_{j=1}^{n_i^e} Se_{(i)j}^p.$$

Уперше показано можливість конструювання образу-рішення проблемної ситуації засобами алгебраїчної системи *BAS* на основі формалізованих у розділі 4 операторів механізму оперативної пам'яті та орієнтувального рефлексу, постановкою оптимізаційної задачі конструювання образу-рішення вважатимемо

$$Check - Set, Vector - Set \rightarrow Focus - Bi |$$

$$Focus - Weight = \underset{i \in N}{Max}(Weight_i),$$

де *Check-Set* – стек для контролю за процесом вибору образу;

*Weight<sub>i</sub>* – вага асоціативних зв'язків *i*-го образу АО з образами *Vector-Set*.

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 ІС внаслідок зовнішньої верифікації відповідного парадигматичного зв'язку. Формально таке повідомлення для синонімії представлено за допомогою

$$\exists Pc_{ij} | Pc_{ij} \in P_c \rightarrow If - Synonim ,$$

$$Se^s + Se_{ij}^{Pc} \xrightarrow{Synonim} Se^s ,$$

де *If-Synonim* – предикат, що вимагає зовнішнього підтвердження;

*Synonim* – функція збільшення загального сенсу інфологічної системи внаслідок додавання синонімічного 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 часткове–загальне між образами нововиявленого кластера та збільшення загальної кількості  $Se^s$  сенсу системи. Загальну мету пізнавальної діяльності системи інтерпретовано у вигляді задачі оптимізації

$$Se^s \rightarrow \max | G_z .$$

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 мовні 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).

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 часткових 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.

Результати дисертаційних досліджень 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, Всеукраїнській громадській організації «Національна асоціація сільськогосподарських дорадчих служб України», 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), кафедр 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, автоматики та інформаційно-вимірювальної техніки Вінницького національного технічного університету. Об'єктами впровадження стали основні 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 основних результатів дослідження.

На підприємстві «AETHER. Inc» (США) впроваджено комплекс моделей, алгоритмів і програм для створення системи повнотекстового пошуку з удосконаленими функціями семантичного пошуку. Результати запропонованого методу пошуку порівнювалися з методами Гаусової, Жаккардової і косінусної подібності. Альтернативний пошук забезпечено програмою *search.py* з використанням бібліотеки *csv*. Результати програмного експерименту для 20 англійських запитів узагальнено та в графічному вигляді приведено на рисунку 6. Загальне зі всіма результатами поліпшення точності пошуку склало 11,3%, а збільшення повноти пошуку – 17,5 %.

У підприємстві «Інновінн» проведено програмний експеримент для технології пошуку зображень на основі формалізації фолксонометричного простору, що отримано на основі російськомовного опису зображень. Середнє за всіма результатами покращення точності пошуку склало 18,3%, а збільшення повноти пошуку – 38,8%.

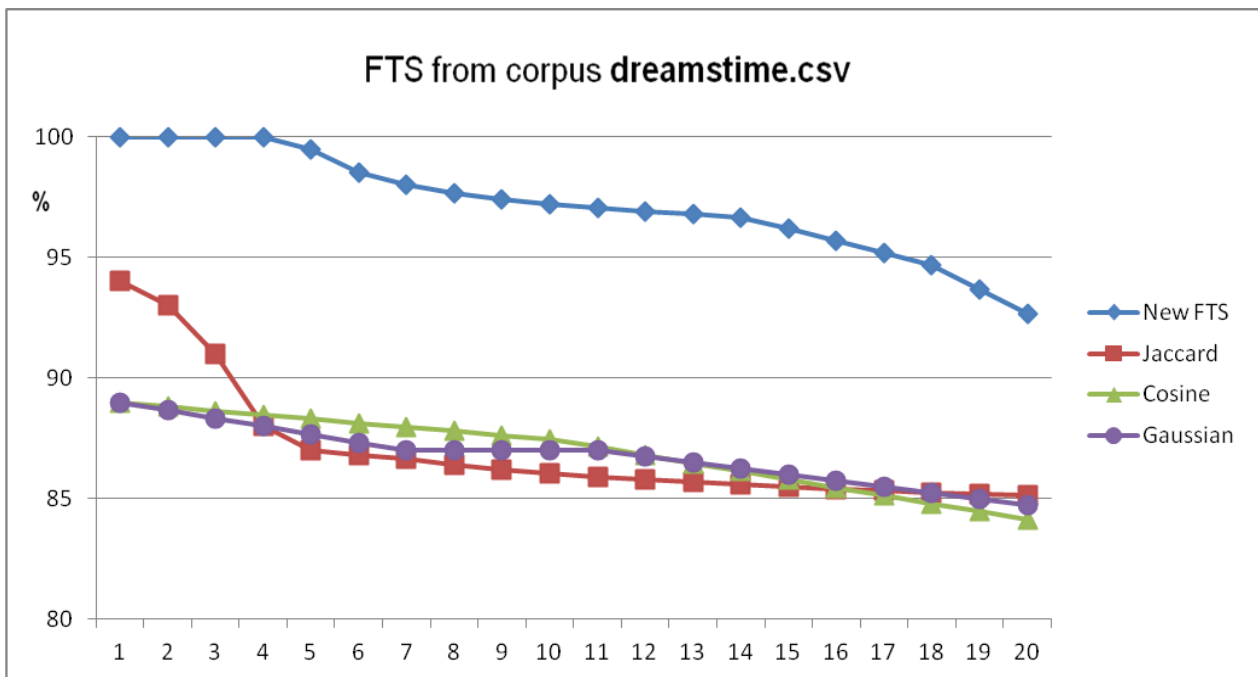


Рисунок 6 – Результати програмного експерименту в «AETHER. Inc»

Обсяг проведених експериментів для отриманої інформаційної технології пошуку склав 8660 англomовних назв та 47754 російськомовних. Первинна оцінка релевантності результатів пошуку здійснювалася незалежно 3-ма експертами за 4-рівневою шкалою (3 – відмінно, 2 – добре, 1 – задовільно, 0 – незадовільно).

У Центрі дистанційної освіти ВНТУ та Вінницькій ФМГ №17 проведено педагогічні експерименти для оцінки покращення успішності студентів та учнів, що досягнуто за рахунок нових засобів образної обробки ПМК. Отримані дані дозволяють стверджувати, що із надійністю 95% мінімальне покращення результатів навчання студентів становить 0,62 бали (4% від 16), а учнів – 0,2 бали (2% від 12). Максимальне покращення становить 1,6 бали (13% від 12) та 2,02 бали (13% від 16).

У додатках наведено приклади застосування формальної теорії  $Th$  для російськомовних речень, довідкову інформацію щодо наскрізного тестового прикладу даних з тематики «WEB-технології: стандарти Semantic WEB», тексти програм та результати тестування задач підтримки базових функцій пошуку та генерації знань СОПМК. Надано детальну характеристику діючого програмного прототипу інформаційної технології, а також представлено документи, що підтверджують впровадження наукових результатів дисертації.

## ВИСНОВКИ

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. У дисертаційній роботі були одержані такі результати:

1. 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 природномовного контенту з метою визначення образного сенсу як прагматичної першооснови змісту ПМК.

2. 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 множину обраних семантико-залежних задач було закладено до таких груп функцій системи, як підтримка діалогу, лінгвістичні, дидактичні та додаткові (пошук і переклад) можливості.

3. 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:

3.1. Уперше запропонований підхід до образного аналізу текстової інформації, що базується на обраних концептуальних поняттях, дозволив отримати чисельні ознаки загального образного сенсу природномовної конструкції.

3.2. 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.

3.3. Удосконалення методу створення нечіткого відношення образного сенсу завдяки приведення 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.

4. Синтез структурно-функціональних моделей систем на основі природномовної інформації досягнуто шляхом обґрунтування інтерпретацій відомих психофізіологічних феноменів інтелектуальної діяльності з визначенням архітектури бази знань та структури СОПМК. У тому числі:

4.1. 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 природномовного контенту.

4.2. Удосконалення методів моделювання процесів образної обробки ПМК на основі орієнтованих і неорієнтованих зважених графів та графа булеа-

ну через алгебраїчну систему БАС дозволило представити базові функції ІС у вигляді алгоритмів на графах і формалізувати архітектуру бази знань СОПМК.

4.3. 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.

5. У межах загальної методології побудовано інформаційну технологію образної обробки природномовного контенту на основі СОПМК, склад моделей для функцій якої відповідає функціональним вимогам до ІС. Зокрема:

5.1. Набули подальшого розвитку математичні моделі базових типів образного пошуку з включенням інсайту як часткового випадку, що дозволило зменшити розмірність пошуку на орієнтованих та неорієнтованих зважених графах в процесі розв'язання семантико-залежних задач.

5.2. Уперше розроблений метод моделювання образного механізму оперативної пам'яті СОПМК засобами двохосновної алгебраїчної системи БАС забезпечив отримання розв'язків семантико-залежних задач у вигляді алгоритмів навігації на графі булеану.

5.3. Уперше запропонований метод моделювання складових парадигматичного устрою мови на основі підходу до образного аналізу тексту дозволив отримати рекурсивні алгоритми самовдосконалення бази знань СОПМК.

6. Розроблена інформаційна технологія образної обробки природномовного 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.

7. Основні результати, отримані у дисертаційній роботі підтверджено 14-ма актами впровадження в організаціях, підприємствах, навчальних закладах та установах різної форми власності, в т.ч. за кордоном. Об'єктами впровадження стали основні складові СОПМК, автоматизоване робоче місце викладача, програмна оболонка для проектування електронних навчальних об'єктів та створені на її основі програмно-педагогічні засоби тощо, які на практиці чисельно підтвердили достовірність основних результатів дослідження.

8. Аналіз всіх результатів дослідження та впровадження дозволяє стверджувати, що на основі реалізованих у межах інформаційної технології функцій СОПМК досягнуто адекватні розв'язки визначених вимогами до ІС семантико-залежних задач. 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 природномовних складових, а практичні результати створюють нове експериментальне середовище для подальших досліджень у широкій предметній галузі інтелектуальних інформаційних технологій.



**СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ**

1. Бісікало О.В. Концептуальні основи моделювання образного мислення людини / Бісікало О.В. – Вінниця: ПП Балюк І.Б., ВДАУ, 2009. – 163 с. – ISBN 978-966-2959-62-3.
2. 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.
3. 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.
4. 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.
5. 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. – ISSN 1681-7893.
6. 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.
7. 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.
8. 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.
9. Бісікало О.В. Дослідження простору асоціативних пар в контексті бази знань електронного підручника / О.В. Бісікало // Вимірювальна та обчислювальна техніка в технологічних процесах. – 2006. – № 2 (28). – С. 109–113.
10. 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.
11. Бісікало О.В. Структура блоку пам’яті на основі моделі образного мислення людини / О.В. Бісікало // Искусственный интеллект. – 2007. – № 3. – С. 461–468. – ISSN 1561-5359.
12. 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. – ISSN 1562-9945.
13. Бісікало О.В. Методика побудови тезауруса навчальної системи на основі моделі образного мислення / О.В. Бісікало // Искусственный интеллект. – 2008. – № 4. – С. 730–735. – ISSN 1561-5359.

14. 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. – ISSN 0555-2656.
15. Бісікало О.В. Класифікація образного пошуку та моделювання інсайту / О.В. Бісікало // Вісник СумДУ (Серія “Технічні науки”). – 2008. – № 2. – С. 53-59.
16. 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](http://www.nbu.gov.ua/e-journals/VNTU/2009_2/2009-2.files/uk/09ovbapt_ua.pdf).
17. Бісікало О.В. Ассоциативный поиск для задач обучения на основе электронного тезауруса образов / О.В. Бісікало // Управляющие системы и машины. – 2009. – № 2. – С. 28–33. – ISSN 0130-5395.
18. 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](http://visnyk.sumdu.edu.ua/arhiv/2009/Tech_2_09/09bovoml.pdf).
19. Бісікало О.В. Формалізація образної пам'яті людини у вигляді графа / О.В. Бісікало // Бионика интеллекта. – 2009. – № 1 (70). – С. 127–131. – ISSN 0555-2656.
20. Бісікало О.В. Комплекс інфологічного моделювання образного мислення людини / О.В. Бісікало // Восточно-Европейский журнал передовых технологий. – 2009. – № 2/2 (38). – С. 14–17. – ISSN 1729-3774.
21. Бісікало О.В. Представлення пізнавальної діяльності на основі інфологічної моделі образного мислення людини [Електронний ресурс] / О.В. Бісікало // Оптико-електронні інформаційно-енергетичні технології. – 2009. – № 1 (17). – С. 90–97. – Режим доступу: [http://www.nbu.gov.ua/portal/natural/Oeiet/2009\\_1/17pdf/15.pdf](http://www.nbu.gov.ua/portal/natural/Oeiet/2009_1/17pdf/15.pdf).
22. Бісікало О.В. Представлення асоціативної мережі образів за допомогою графів / Олег Володимирович Бісікало, Роберт Геворкович Тадевосян // Вісник Держ. ун-ту ”Львівська політехніка”: Комп'ютерні науки та інформаційні технології. – 2009. – № 650. – С. 73–80. – ISSN 0321-0499.
23. Бісікало О.В. Моделювання складових парадигматичної будови мови / О.В. Бісікало // Искусственный интеллект. – 2009. – № 4. – С. 188–194. – ISSN 1561-5359.
24. 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.
25. 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. – ISSN 1999-9941.
26. Кветний Р.Н. Морфологічний аналіз слова на основі асоціативно-статистичного підходу / Р.Н. Кветний, О.В. Бісікало, І.А. Кравчук // Вісник Черкаського держ. технологічного університету. – 2010. – № 3. – С. 132-135.

27. 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. – ISSN 1562-9945.
28. 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. – ISSN 1812-5409.
29. 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. – ISSN 1681-7710.
30. 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. – ISSN 1997-9266.
31. 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. – ISSN 1028-9763.
32. Бісікало О.В. Онтогенетичний метод побудови нечіткого відношення сенсу / О.В. Бісікало // Штучний інтелект. – 2011. – № 1. – С. 134–140. – ISSN 1561-5359.
33. 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](http://www.nbu.gov.ua/portal/Natural/Vtot/2012_1/46kve.pdf).
34. 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. – ISSN 1028-9763.
35. Бісікало О.В. Образний аналіз текстової інформації з Wikipedia / О.В. Бісікало, І.О. Назаров // Вісник інженерної академії України. – 2012. – № 2. – С. 52–56. – ISBN 5-7763-8361-7.
36. 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 акредитації. Доповіді, виступи та повідомлення семінару, (Суми, 1–3 листопада 2000 р.). – Київ: Аграрна освіта, 2001. – С. 86–90. – ISBN 966-95661-8-5.
37. 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. – ISBN 966-641-063-X.
38. Бісікало О.В. Концепція проектування електронного навчального посібника / О.В. Бісікало // Наука і методика. – 2005. – № 3. – С. 73–78.

39. 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, 2005. – С. 41–43.
40. 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.
41. Бісікало О.В. Алгоритм пошуку інформації на основі моделі асоціативної пам’яті людини / О.В. Бісікало // ІНТЕРНЕТ–ОСВІТА–НАУКА–2006. П’ята міжнар. конф. ІОН–2006: зб. матеріалів конф., (Вінниця, 10–14 жовтня 2006 р.). – Вінниця: УНІВЕРСУМ-Вінниця, 2006. – Т. 2. – С. 560–565. – ISBN 966-641-193-8.
42. Bisikalo O.V. The Approach to Automation of Designing Knowledge Base in the Device-Making Industry / O. Bisikalo // Proc. of IEEE East-West Design & Test Workshop. EWDTW’06, (Sochi, 15–19 September 2006). – Sochi, Russia, 2006. – P. 440–442. – ISSN 966-659-124-3.
43. 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. – ISBN 978-966-641-210-5.
44. Бісікало О.В. Класифікація образного пошуку / О.В. Бісікало // Тези доповідей Першої міжнар. наук.-техн. конф. «Інтелектуальні системи в промисловості і освіті – 2007», (Суми, 7–9 листопада 2007 р.). – Суми, 2007. – С. 14–15.
45. Бісікало О.В. Проектування електронного підручника на основі формалізації пізнавальної діяльності людини / О.В. Бісікало // Перспективні технології навчання та освітні простори: зб. наук. праць. – Київ: МННЦ ІТiС, 2007. – Вип. 1. – С. 179–190. – ISBN 978-966-02-4307-1.
46. 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). – Kiev: Akadempriodika, 2007. – P. 336-344. – ISBN 978-966-02-4538-9.
47. Bisikalo O.V. The Determination of Results of Knowledge Test Control by Means of the Hierarchic Structures / Oleg Vladimirovich Bisikalo, Robert Gevorkovich Tadevosyan // Proc. of IEEE East-West Design & Test Symposium. EWDTTS’07, (Yerevan, 7–10 September 2007). – Yerevan, Armenia. – P. 585–588.
48. 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.

49. 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. – ISBN 978-966-02-4754-3.
50. 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. – ISBN 978-966-02-4874-8.
51. Bisikalo O.V. Model of image thinking based development of glossary / O.V. Bisikalo // Искусственный интеллект. Интеллектуальные системы: материалы IX междунар. научн.-техн. конф. – Донецьк: ІПШ «Наука і освіта», 2008. – Т. 2. – С. 87–91.
52. Бісікало О.В. Реалізація модульно-рейтингової системи у вигляді додатку до бази даних / О.В. Бісікало // Наука і методика. – 2008. – № 14. – С. 42–49.
53. Бисикало О.В. Субъективная единица смысла образных конструкций / О.В. Бисикало // Nauka: teoria i praktyka – 2009: materialy V miedzynar. naukowopraktycznej konf., (Przemysl, 7–15 sierpnia 2009). – Przemysl: Nauka i studia, 2009. – Vol. 6. – P. 9–12. – ISBN 978-966-8736-05-6.
54. 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: ІПШ «Наука і освіта», 2009. – С. 33–35.
55. 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.
56. Бісікало О.В. Онтогенетичний підхід до розуміння сенсу навчальних об'єктів / О.В. Бісікало // Теоретичні та прикладні аспекти побудови програмних систем ТАAPSD'2010: тези доповідей 7-ї міжнар. конф., (Київ, 4–8 жовтня 2010 р.). – Київ, 2010. – С. 85–94. – ISBN 978-966-171-299-6.
57. Бісікало О.В. Когнітивний простір образних конструкцій / О.В. Бісікало // Искусственный интеллект. Интеллектуальные системы. (ИИ–2010): материалы международной научно-техн. конф., (Кацивели, 20-24 сентября 2010 г.). – Донецьк: ІПШ «Наука і освіта». – 2010. – Т.1. – С. 17-21. – ISBN 978-966-7829-44-5.
58. 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.
59. Бісікало О.В. Розв'язання задач дистанційного навчання на основі формалізації психодидактичного середовища / Олег Володимирович Бісікало, Тетяна Василівна Вінійчук // Europejska nauka XXI powieka – 2011: materialy VII

- miedzynar. naukow-praktycznej konf., (Przemysl, 7–15 maja 2011). – Przemysl: Nauka i studia, 2011. – Vol. 19. – P. 6–9. – ISBN 978-966-8736-05-6.
60. 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>.
61. 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.
62. Ільїн В.В. Дидактичні та технологічні вимоги до програми-оболонки для підготовки та використання електронних навчальних посібників / Ільїн В.В., Теплюк В.М., Бісікало О.В. – Київ: Аграрна освіта, 2004. – 20 с.
63. Ільїн В.В. 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.: Наукометодцентр аграрної освіти, 2005. – 40 с. – ISBN 966-7906-14-0.

## АНОТАЦІЯ

Бісікало О.В. Інформаційна технологія образного аналізу тексту та синтезу структурно-функціональних моделей системи обробки природномовного контенту. – На правах рукопису.

Дисертація на здобуття наукового ступеня доктора технічних наук за спеціальністю 05.13.06 – інформаційні технології. – Вінницький національний технічний університет, Вінниця, 2012.

Дисертаційна робота присвячена вирішенню науково-прикладної проблеми розв'язання класу семантико-залежних задач на основі обробки і накопичування загальних знань образного сенсу. Розроблено основи теорії образного аналізу текстової інформації шляхом формалізації поняття інфологічної системи та онтогенетичного принципу її побудови, застосування моделі формальної теорії у вигляді комутативної напівгрупи образних конструкцій, створення методу генерації нечіткого відношення сенсу, що породжує простір з нечіткою мірою. Запропоновано методологію синтезу структурно-функціональних моделей систем на основі природномовної інформації та функціональну модель системи обробки природномовного контенту (СОПМК). Розроблено математичні моделі базових типів образного пошуку, методи моделювання образного механізму оперативної пам'яті СОПМК та складових парадигматичного устрою мови. Отримані складові інформаційної технології на основі СОПМК впроваджені та пройшли успішну апробацію в організаціях, підприємствах, установах та навчальних закладах в Україні та закордоном.

**Ключові слова:** аналіз, синтез, природномовний контент, образний сенс, структурно-функціональні моделі, граф, система, інформаційна технологія.

## АННОТАЦИЯ

Бисикало О.В. Информационная технология образного анализа текста и синтеза структурно-функциональных моделей системы обработки естественно-языкового контента. – На правах рукописи.

Диссертация на соискание ученой степени доктора технических наук по специальности 05.13.06 – информационные технологии. – Винницкий национальный технический университет, Винница, 2012.

Диссертационная работа посвящена решению актуальной научно-прикладной проблемы создания теоретико-методологических основ и технологических средств построения систем обработки естественно-языкового контента (СОЕЯК), способных решать широкий класс семантико-зависимых задач на основе накопления общих знаний образного смысла.

Ключевая идея работы заключается в формализации образного смысла естественно-языковой информации как результата функционирования инфологической системы (ИС) – абстрактной модели СОЕЯК. Обмен информацией между ИС и внешним миром происходит исключительно через образные конструкции – произвольные подмножества множества пар образов. Накопление таких конструкций приводит к появлению пространства образного смысла, в котором на уровне метаязыка разделяются понятия смысла-свойства и смысла-параметра. Предложено интерпретировать онтогенез ИС через построение ассоциативной сети образов (АСО) – базы знаний в форме семантической сети, для которой образный смысл-свойство интерпретируется как ее топология, а смысл-параметр – как числовая оценка любой подсети АСО с помощью впервые введенной в работе единицы смысла.

В работе формализовано понятие языкового образа на основе обобщения подмножеств однокоренных лексем. Языковой образ позволяет формально различать роли объекта, метода, свойства и понятия у элементов образной конструкции, при этом слова естественно-языковых конструкций являются вербальными признаками вершин ориентированного графа АСО. В результате достигается получение исходной информации ИС непосредственно из разделенных на предложения текстов.

Анализ известных методов моделирования процессов обработки естественно-языковой информации и построения баз знаний позволил выделить множество семантико-зависимых задач на основе 18 функциональных требований к ИС, объединенных в 4 группы: поддержка диалога, лингвистические и дидактические возможности, дополнительные возможности поиска и перевода.

Разработаны основы теории образного анализа текстовой информации, позволившие формализовать смысл-свойство и смысл-параметр. Предложена модель формальной теории 1-го порядка в виде коммутативной полугруппы образных конструкций, в которой бинарный оператор направленной ассоциативной связи обеспечивает представление вопрос-ответных конструкций из языковых образов. Усовершенствован метод построения нечеткого отношения образного смысла путем приведения аксиоматики пространства с нечеткой мерой к онтогенетическому принципу, формализации понятий мотивационной

цели и эмоционального состояния ИС, введения единицы смысла как численной оценки силы элементарной ассоциативной связи АСО. Исследовано соотношение единиц образного смысла и информации.

Обоснована методология синтеза структурно-функциональных моделей ИС на основе кибернетической интерпретации интеллектуальной деятельности человека. Формализация механизма обработки естественно-языкового контента достигнута в виде взаимосвязанной обработки данных 9-ю функциональными блоками-моделями головного мозга с помощью 19-ти информационных потоков образной природы.

Предложены методы моделирования процессов образной обработки естественно-языкового контента на основе ориентированных и неориентированных взвешенных графов, а также графа булеана через двухосновную алгебраическую систему, что позволило представить базовые функции инфологической системы в виде алгоритмов на графах и формализовать архитектуру базы знаний СОЕЯК. На основе новой классификации синтагматического и парадигматического типов образного поиска разработана функциональная модель СОЕЯК, обеспечивающая унификацию процессов обработки данных и знаний для решения базовых и прикладных семантико-зависимых задач.

Разработаны методы и алгоритмы для поддержки функций образного поиска и генерации знаний СОЕЯК. Дальнейшее развитие получили модели базовых типов образного поиска в АСО, в том числе инсайта на основе модификации известных алгоритмов поиска на разреженных графах, что привело к уменьшению вычислительной сложности алгоритмов. Предложенный метод моделирования образного механизма оперативной памяти СОЕЯК в виде алгоритмов навигации на графе булеана обеспечил решение задач исследования ориентировочного рефлекса, конструирования образа-решения, генерация сообщений и некоторых типов ответов на вопросы к системе. Разработанный метод моделирования компонентов парадигматического строя языка позволил получить рекурсивные алгоритмы самосовершенствования базы знаний СОЕЯК.

Разработана и программно реализована информационная технология образной обработки ЕЯК, объединившая в себе все полученные научные результаты в виде структуры и функций СОЕЯК. Экспериментальная проверка решения семантико-зависимых задач на основе технологии численно подтвердила улучшение релевантности поиска по точности и полноте, сокращение времени на создание программно-педагогических средств, повышение успеваемости учащихся. Основные результаты диссертации подтверждены 14 актами внедрения на предприятиях, организациях, в учреждениях и учебных заведениях в Украине и за рубежом. Объектами внедрения стали все основные модули СОЕЯК, программная оболочка для проектирования электронных обучающих объектов и созданные на ее основе программно-педагогические средства, модели и алгоритмы системы для борьбы с мошенничеством в сетях связи, подтвердившие на практике достоверность теоретических положений исследования.

**Ключевые слова:** анализ, синтез, естественно-языковой контент, образный смысл, структурно-функциональные модели, граф, система, информационная технология.



**ABSTRACT**

Bisikalo O.V. Information technology of image analysis of text and synthesis of structure-functional models of system of natural-language contents processing. – Manuscript.

Thesis for the Scientific Degree of the Doctor of Technical Sciences in the speciality 05.13.06 – information technologies. – Vinnytsia National Technical University, Vinnytsia, 2012.

The dissertation is dedicated to the solution of a problem of resolving class of semantic-dependent tasks based on processing and accumulation of general knowledge figurative sense. A basic theory of image analysis of textual information is developed. A basis of a formal theory for communicative subgroup of image structures, methods of generating a fuzzy sense relation, which produces a fuzzy measure space based on formalizing the notion of infologic system and ontogenesis principle is laid down. A methodology for the synthesis of structural-functional models of systems based on natural language information and functional model system for processing natural language content (SPNLC) is offered. Algorithms for implementing basic types of image search, methods for modeling the image mechanism of working memory of SPNLC and components of paradigmatic language structure are elaborated. These components of information technology based SPNLC been implemented and successfully tested in public and private organizations and academic institutions in Ukraine and abroad.

**Key words:** analysis, synthesis, natural-language content, imaging sense, structure-functional models, graph, system, information technology.

Підписано до друку 07.11.2012 р. Формат 29.7×42 ¼

Наклад 100 прим. Зам. № 2012-178.

Віддруковано в комп'ютерному інформаційно-видавничому центрі

Вінницького національного технічного університету

м. Вінниця, Хмельницьке шосе, 95. Тел.: 59-87-38

Свідоцтво суб'єкта видавничої справи

серія ДК №3516 від 01.07.2009 р.