

Підхід до побудови асоціативної пам'яті на основі моделі образного мислення

*Вінницький державний аграрний університет, Вінниця,
bisikalo@vsau.org*

Феномен асоціативної пам'яті людини вже давно знайшов своє відображення в математичних моделях та комп'ютерних технологіях, проте за своїми функціональними можливостями всі штучні конструкції суттєво відстають від свого природного аналога [1]. Одним з можливим підходів до наближення в розв'язанні даної проблеми є послідовна побудова моделі образного мислення [2], невід'ємною складовою якої виступає асоціативна пам'ять. Більше того, асоціативний пошук інформації в межах моделі образного мислення людини [3] вводиться в якості базової операції лінгвістичної системи [4].

Метою дослідження будемо вважати обґрунтування структурних особливостей блоку асоціативної пам'яті лінгвістичної системи. *Постановка задачі:* забезпечити основні функції асоціативної пам'яті на основі інтерпретації простору асоціативних пар у вигляді булеану [5].

Однією з головних задач моделі образного мислення [3] є накопичення зв'язків в асоціативній пам'яті на основі первинної інформації подій та/або синтагм (простих речень без внутрішніх підлеглих конструкцій) з природномовних текстів. В останньому випадку в модель простору асоціативних пар необхідно ввести додаткові дані щодо параметрів об'єднання підмножини пар в синтагму, в тому числі час запису синтагми, тип і напрямок зв'язку в кожній її парі, вид і правила представлення кожного образу в залежності від його ролі в певному асоціативному зв'язку. Крім того, модель має визначати і зберігати значення сили прямого і зворотного асоціативного зв'язку в кожній парі з метою подальшого використання цих важливих параметрів в лінгвістичній системі [4].

В якості прикладів природномовних конструкцій для ілюстрації підходу, що пропонується, розглянемо відомі віршовані рядки «Буря мглою небо кроет ...» та «Больно бьют по нашим душам голоса ...». Впорядковані асоціативні пари створюються за допомогою питальних займенників від головного до підлеглого образів цих конструкцій (наприклад – «Буря–чем?–мглою»).

Кожному образу конструкції присвоєно свій бінарний код з $n-1$ нулями та одиницею в потрібному розряді, який відповідає вузлу першого шару в графі булеану [5]. Дані приклади ілюструють підхід до кодування синтагм на основі графа булеану для випадків $n=4$ та $n=5$, де

n – кількість образів в синтагмі. Згідно з цим підходом на другому шарі має бути активізовано стільки вузлів (асоціативних пар), скільки питальних займенників було використано. Також треба врахувати, що в кожному реченні може бути суміщення питальних займенників в один «подвійний» вузол, який об'єднує підмет і присудок речення. На рис. 1 та рис.2, де приведено графи булеану для розглянутих прикладів, такі «подвійні вузли» на другому шарі разом з підлеглими вузлами першого шару позначено сірим кольором.

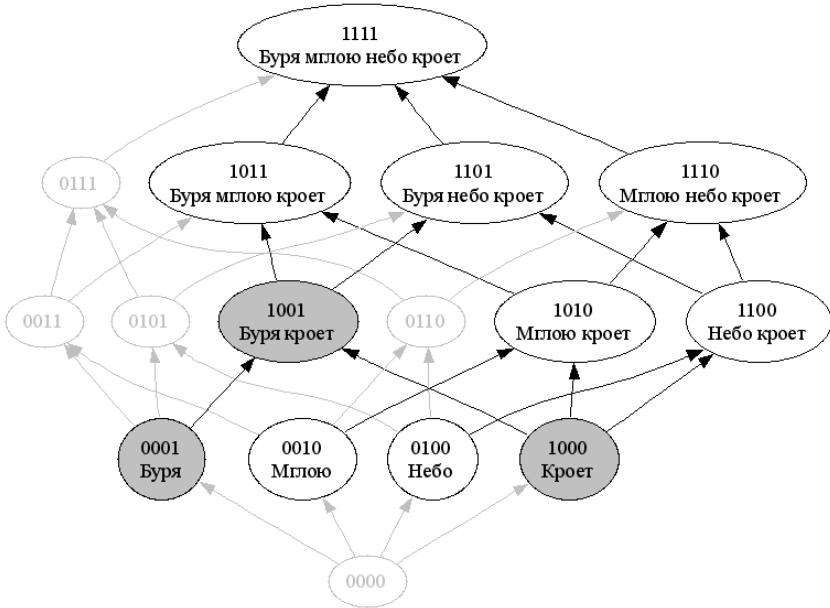


Рис.1. Приклад графа булеану синтагми для випадку $n=4$.

Наведені рисунки також ілюструють формальну можливість на основі інформації другого шару побудувати так звану «піраміду сенсу» окремої синтагми. Позначення приналежних до піраміди вузлів та зв'язків між ними (виділені на рисунках кольором) відбувається за наступним алгоритмом:

1. Активізуються всі вузли першого шару.
2. На другому шарі активізуються тільки ті вузли, які відповідають асоціативним парам саме цієї синтагми.
3. На всіх шарах булеану, вищих другого, активізуються тільки ті вузли, коди яких є результатом операції об'єднання \cup кодів всіх активованих на попередньому шарі вузлів.

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

Характерною для першого шару «піраміди сенсу» є така ознака підлеглості образу в даній синтагмі – вузол першого шару, що відповідає образу є підлеглим тоді і тільки тоді, коли на другий шар від нього активізовано лише один зв'язок, причому не в «сірій», а в «білий» вузол (вузли 10000 – «больно» та 00100 – «нашим» на рис.2).

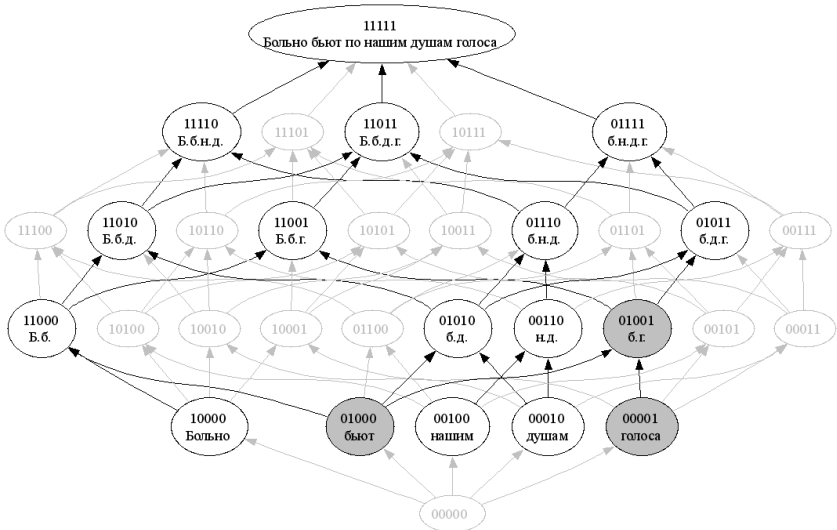


Рис.2. Приклад графа булеану синтагми для випадку $n=5$.

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

Таким чином, формалізація природномовних конструкцій за допомогою булеану дозволяє акумулювати у вигляді «пірамід сенсу» значимі параметри асоціативних зв'язків між образами кожної синтагми. Для n -вимірного бінарного коду будемо позначати як Bi_i – код з i одиницями та, відповідно, $n-i$ нулями. Очевидно, що вузол з кодом Bi_i належить до i -го шару графа булеану. Формально:

$$Bi = \{Bi_j, j = 1, n\}.$$

Введемо також наступні позначення:

- *Event-Id* – унікальний код синтагми (події);
- *Bi-Sy* – бінарний код синтагми;
- *Twice-Id* – унікальний код асоціативної пари;
- *Force⁺* та *Force⁻* – сила прямого та зворотного зв'язків асоціативної пари образів;
- *Construct-Id* – унікальний код конструкта асоціативної пари;
- *Pronoun-Id* – унікальний код питального займенника;
- *Logic* – напрямок асоціативного зв'язку: True – прямий та False – зворотний;
- *Text-Id* – унікальний код тексту, у складі якого знаходиться синтагма;
- *Text* – представлення синтагми у вигляді речення;
- *Bi-Te* – бінарний код тексту;
- *Time* – час створення тексту;
- *Name* – назва тексту;
- *Author* – автор тексту.

Тоді, на основі запропонованого підходу логічним виглядає наступна організація блока пам'яті у вигляді спискових структур:

1. булеан-модель блока пам'яті розмірністю n образів складається з n списків вузлів на кожному шарі булеану, причому початкові списки є пустими;
2. вхідною інформацією моделі будемо вважати послідовність (список) синтагм *Event* з ключовими елементами *Event-Id*, що відповідають непустим вузлам булеану 3-го шару та вище з бінарним кодом *Bi-Sy*;
3. булеан-модель блока пам'яті є динамічна конструкція, яка рекурсивно поповнюється (самовдосконалюється) в процесі внесення нових синтагм і збільшення кількості образів n та характеризується такими особливостями:
 - a. список першого шару *Image* (ключовий елемент – *Bi₁*) складається з n кодів булеану та інших ознак образів, що вже внесено до пам'яті;
 - b. список другого шару (власне асоціативна пам'ять *Assoc-Twice* з ключовим елементом *Twice-Id*) має найбільш складну структуру, оскільки для кожного непустого елемента з кодом *Bi₂* додаються значення лічильників *Force⁺* та *Force⁻*, а також відкритий підсписок *Construct* з елементів (*Construct-Id*, *Twice-Id*, *Event-Id*, *Pronoun-Id*, *Logic*), (... , ... , ... , ... , ...), ... – довжина такого підписку визначається кількістю синтагм, в яких була задіяна пара з цим кодом;

4. якщо потрібно фіксувати належність певної послідовності синтагм окремому тексту, то для шарів від 3-го та вище можна формувати ще один список *Text-RE* (ключовий елемент – *Text-Id*), де кожному непустому елементу з кодом *Bi-Te* може відповідати відкритий підсписок з елементів (*Event-Id, Bi-Sy, Text-Id, Text*), (... , ... , ... , ...), ... ; більш повна характеристика списку *Text-RE* буде досягнута за допомогою додаткових елементів *Time, Name* та *Author*.

В якості допоміжних списків блока пам'яті, необхідних для практичної реалізації лінгвістичної системи можна запропонувати:

- *Inter-Pronoun* – список всіх питальних займенників (ключовий елемент – *Pronoun-Id*);
- *Link* – список типів можливих асоціативних зв'язків у синтагмі (ключовий елемент – *Link-Id* – унікальний код типу зв'язку).

Таким чином, за допомогою структури з 7 взаємопов'язаних списків блоку пам'яті показана можливість врахувати значимі для функцій асоціативного мислення атрибути даних. Наукова новизна результатів дослідження пов'язана з побудовою структурних та функціональних особливостей блока пам'яті внаслідок інтерпретації простору асоціативних пар у вигляді булеану та введення поняття «піраміди сенсу». Практична доцільність результатів та їх прикладне значення полягає в обґрунтуванні спискової моделі даних в якості інформаційної основи лінгвістичної системи.

Список літератури

1. Анисимов А.В. Компьютерная лингвистика для всех: Мифы. Алгоритмы. Язык. – Киев: Наук. думка, 1991. – 208 с.: ил.
2. Бисикало О.В. Принципы построения концептуальной модели образного мышления. Первая международная конференция «Новые информационные технологии в образовании для всех», расширенные материалы конференции (29-31 мая 2006 г.). – Киев, 2006. – с. 25-34.
3. Бісікало О.В. Асоціативний пошук інформації в межах моделі образного мислення людини // Інформаційні технології та комп'ютерна інженерія – 2007. – № 1 (8). – с. 248-252.
4. Бисикало О.В. Принципы построения лингвистической системы на основе модели образного мышления. В сборнике «Труды восьмой международной научно-практической конференции «Современные информационные и электронные технологии», г. Одесса, 21-24 мая 2007 года». – Одесса, СИЭТ–2007, 2007. с.63.
5. Бісікало О.В. Дослідження простору асоціативних пар в контексті бази знань електронного підручника // Вимірювальна та обчислювальна техніка в технологічних процесах – 2006. – № 2 (28). – с. 109-113.