

## РОЗРОБКА МЕТОДІВ ФОРМУВАННЯ ФУНКЦІОНАЛЬНО ОРІЄНТОВАНИХ ІНФОРМАЦІЙНИХ ПОТОКІВ

*Аннотация.* Рассматриваются методы формирования информационных потоков, которые ориентированы на реализацию управляющих функций социальными объектами. Процесс управления путём передачи соответствующего информационного потока текстовую модель социального объекта. Исследуется возможность использования специальных семантических средств, которые усиливают действие информационного потока на объект управления.

*Ключевые слова:* модель, семантический параметр, информационный поток, управление.

Використання текстових моделей ( $TM_i$ ) та їх систем ( $STM$ ), для аналізу стану системи соціальних об'єктів ( $SO_i$ ) та для управління цими об'єктами, необхідно провести дослідження методів створення функціонально орієнтованих та ефективних інформаційних потоків ( $IP_i$ ), які є основним інструментом управління. Ефективність інформаційних потоків є інтегральним параметром ( $J^{IP}$ ), який залежить до цілого ряду факторів, до яких відносяться:

- структури соціального об'єкту  $SO_i$  і, відповідно, текстової моделі цього об'єкту  $TM_i$ ,
- зовнішніх семантичних параметрів  $TM_i$  і  $IP_i$ ,
- психосемантичних параметрів  $\mu$  моделі  $TM_i$  і  $IP_i$ ,
- внутрішніх семантичних параметрів  $\mu$  моделі  $TM_i$  і  $IP_i$ ,
- характеристики зв'язку між семантичними та психосемантичними параметрами.

В соціології відомі підходи до визначення загальних параметрів соціальної групи, коли останню представляють у вигляді удекорованого регулярного графу [1]. В цьому випадку, побудовані моделі дозволяють оцінити зміни стану  $SO_i$ , що відбуваються під впливом окремих фрагментів моделі, що відповідають вершинам відповідного графу. Такий вплив здійснюється на основі використання зв'язків між окремими елементами, що описуються ребрами. В залежності від структури такої моделі та енергії окремих елементів, що взаємно впливають на своїх сусідів, модель, в цілому, характеризується окремими станами, які мають свою інтерпретацію в соціальній предметній області. Модель окремого  $SO_i$  і, відповідно,  $TM_i$  розглядаються як динамічна структура, яка представляється у вигляді мережі, що відображає зв'язки між окремими елементами. Однією з таких моделей є

модель Ізінга, для формування мереж, яка описується наступним співвідношенням:

$$E(\{S_i\}) = \sum_i J K_i S_i < S^* >,$$

де  $S^* = (1 / \langle K \rangle N) \cdot \sum_{j \neq i} K_j S_j$ , де  $\langle K \rangle$  середнє зважене значення степені вершини. Такий підхід до моделювання загального стану та моделювання змін цього стану у випадку активізації локальної мережі, де зміна стану описується співвідношенням:

$$S(i + 1) = -\text{sign}[S_i(t); I_i(t) + h_i],$$

де  $S_i(t)$  і  $S(i + 1)$  два послідовні стани елемента  $i$ ,  $I_i(t)$  – описує соціальний вплив, який залежить від індивідуальних властивостей всіх елементів,  $h_i$  – параметр, що характеризує випадковість змін стану  $S_i$ .

Такий підхід, в загальній формі, описує зміни, які можуть відбуватися в  $TM_i$  і, відповідно, в  $SO_i$ , при дії на  $SO_i$  деякого інформаційного потоку  $IP_i$ . Приведений вище підхід не є достатньо конструктивним, оскільки оперує загальними параметрами, тому необхідно розглянути підхід, який ґрунтується на основі параметрів, використання яких дозволяє провести кількісну оцінку ефективності  $IP_i$ , яка в рамках наближень, що визначаються їх інтерпретацією дозволяє конструктивно оцінювати цей параметр. Для цього необхідно більш детально проаналізувати приведені вище параметри.

Структурний параметр, який будемо позначати  $p^S$ , визначається наступним чином.

*Визначення 1.* Структурний параметр  $p^S$  визначає середню зважену величину степені вершини графа, який відображає з заданим наближенням зв'язки між окремими компонентами  $TM_i$ , які ідентифікуються ребрами графа.

Формально, такий параметр визначається наступним співвідношенням:

$$\langle p^S \rangle \geq (\sum_{i=1}^m v_i^s) / m,$$

де  $m$  – число вершин графу  $G$ ,  $s$  – степені вершини  $i$  графа  $G$ ,  $v_i$  – вершина  $i$  графа  $G$ . Параметр  $p^S$  визначає середню величину міри зв'язності графа  $G_i$ . Міра зв'язності впливає на величину ефективності  $IP_i$ , яку будемо позначати  $\varepsilon(IP_i)$ . Зв'язок між величиною  $\varepsilon(IP_i)$  та  $p^S$  означає наступне. Чим більша величина параметру  $p^S$ , тим більша зв'язність між вершинами  $v_i$  в  $G_i$ , що відповідає більшій кількості комунікацій між елементами  $\xi_i$  з  $SO_i$ , або між  $tm_i \in TM_i$ , де  $tm_i$  описує відповідний  $\xi_i \in SO_i$ , або більш точно відповідає, у певному наближенні  $\xi_i$ , що описується у вигляді  $\xi_i \propto tm_i$ . Чим вищий рівень комунікації між  $\xi_i \in SO_i$ , тим більша інтенсивність комунікації між  $\xi_i$ , які використовуються для формування рішень про модифікацію відповідного  $SO_i$ , що пов'язана з  $IP_i$ , який діє на  $SO_i$ . Крім того, параметр  $p_i^S$ , де  $i$  номер вершини  $v_i$  визначає міру вагомості  $v_i$ , при прийнятті рішення стосовно активізації в  $SO_i$  дій, що пов'язані з  $IP_i$ .

Відомо, що  $tm_i \in TM_i$  відповідає опису окремого елемента  $\xi_i \in SO_i$ . Прийmemo наступне визначення.

*Визначення 2.* Вагомість реакції  $\omega_i(v_i)$ , або  $\omega_i(\xi_i)$ , якщо мова йде про

$SO_i$ , або  $\omega_i(tm_i)$ , якщо мова йде про  $TM_i$ , визначається розмірами відповідного  $tm_i$  з  $TM_i$ , який вимірюється кількістю фраз  $\varphi_j \in TM_i$  та відповідною кількістю слів  $x_{jk}$ .

Формально,  $\omega_i(tm_i)$ , чи  $\omega_i(\xi_i)$  визначається наступним співвідношенням:

$$\omega_i(tm_i) = \sum_{j=1}^m \sum_{k=1}^r sg(x_{jk}),$$

де  $x_{jk}$  слова в  $\varphi_j$ . Таким чином, важливість або вага  $\omega_i$  одного  $\xi_i \in SO_i$  визначається величиною тексту, який описує  $tm_i$  з точністю до кожної фрази. Сукупність фраз, що використовуються для опису окремого елемента  $\xi_i$ , який відповідає  $tm_i$ , визначає значимість  $\xi_i$ , або міру домінації  $\xi_i$  над  $\xi_j$ , якщо  $\omega_i(tm_i) > \omega_j(tm_j)$ . Інтерпретація такого параметра полягає у наступному. Чим більше тексту використовується в  $tm_i$  для опису  $\xi_i \in SO_i$ , тим більш вагомих є його рішення, яке стосується відповідного  $IP_i$ , що направлено на активізацію змін або модифікацій в  $SO_i$  і, відповідно, модифікацій в  $TM_i$ . Така інтерпретація співпадає з інтерпретацією семантичної значимості окремого слова  $x_{ji}$ , або з визначенням параметра  $\sigma^Z(x_{ji})$  [2]. На основі елементарної значимості окремого  $tm_i$  та семантичної значимості  $tm_j$ , кожен з яких є окремим абзацем, будемо визначати значення зовнішніх семантичних параметрів  $\mu(tm_i)$ . Прийmemo, що  $\mu^Z(tm_i) = \omega_i(tm_i)$ . Тоді, семантичні параметри, які визначаються між  $tm_i$  і  $tm_j$  є зовнішніми семантичними параметрами, які приймаються до уваги при побудові  $IP_i$ . З точки зору інтерпретації зв'язку між  $\varepsilon(IP_i)$  та  $\mu^Z(tm_i)$ , остання означає, що управляюча дія  $IP_i$  повинна в найбільшій мірі орієнтуватися на модифікацію  $tm_i$ , для яких  $\sigma^Z(tm_i) > \sigma^Z(tm_j)$ .

Що стосується внутрішніх семантичних параметрів  $\sigma^i(x_i)$ , чи  $\sigma^i(\varphi_i)$ , то їх величина також впливає на величину ефективності  $\varepsilon(IP_i)$ . Для цього прийmemo, що всі внутрішні  $\sigma^i$  повинні приймати оптимальні значення в рамках відповідної предметної області інтерпретації, але, при цьому, повинні не виходити за границю допустимих значень цих параметрів. Це означає, що чим більша величина  $\mu^Z(tm_i)$ , тим більш ефективним може бути  $IP_i$ .

На відміну від семантичних параметрів, значення яких визначаються на основі прийнятих в  $W_i$  методів оцінок значимості слів, чи окремих фраз для відповідної  $TM_i$ , чи  $W_i$ , в цілому, психосемантичні параметри безпосередньо пов'язані з ціллю, яка визначає зміни в  $SO_i$  і, відповідно, в  $TM_i$ . Це означає, що вибір параметрів  $\aleph_i$  та вибір їх значень здійснюється на основі визначення міри наближення до цілі відповідного  $TM_i$ , при дії  $IP_i$  на  $TM_i$ . Процедура вибору  $\aleph_i$  та визначення значень параметра вибраного  $\aleph_i$  реалізується на основі взаємодії  $IP_i$  з  $TM_i$ . Така взаємодія є окремою компонентою процесу формування  $IP_i$ .

Для підвищення ефективності дії  $IP_i$  на  $SO_i$ , для формування відповідних текстових потоків, використовуються, на рівні з семантичними параметрами, психосемантичні параметри  $\aleph_i$ . Ці параметри виходять за рамки вимог по

формуванню семантично приязного тексту, який оцінюється мірою семантичної узгодженості текстової моделі  $TM_i$  та текстового інформаційного потоку  $IP_i$ . Такі параметри відображають психологічні аспекти сприйняття тексту людиною, або групою людей, що складають  $SO_i$ . Очевидно, що психологічні аспекти досить важко оцінювати, а тим більше вимірювати, оскільки вони стосуються безпосередньо психології поведінки не тільки окремих осіб, а й колективу осіб. Тому, в рамках даної роботи, для кількісної оцінки психосемантичних параметрів прийємо наступні положення.

*Положення 1.* Синтаксичні елементи  $x_i$ , що описують психосемантичні ефекти, які можуть мати місце в текстовому описі інтерпретації  $j(x_i)$ , представляють собою підсилювачі семантичної узгодженості.

*Положення 2.* Психосемантичні підсилювачі  $U(\sigma^U)$  зв'язані з вибраними фразами  $\varphi_i$ , чи реченнями  $\psi_i$ , підсилення яких здійснюється по логарифмічному закону, кожний з яких має власний експотенціальний коефіцієнт.

Визначимо способи визначення величини значення підсилення, яке забезпечує різні психосемантичні параметри. Величина підсилення, яка забезпечується параметрами  $\aleph^R$  визначається кількістю повторень слова, яке вибрано для реалізації цього параметру. Нехай таким словом є  $x_i$  у відповідному місці речення  $\psi_i$ . Оскільки, ефективність дії  $\aleph^R$  описується функцією  $y = \ln(kx_i)$ , то для опису ефекту підсилення вибирається деякий інтервал  $[1, \alpha]$ , для її змінних. У відповідності з цією функцією, підсилення є ефективним тільки на вибраному інтервалі значень змінних  $x_i$ , який відповідає величині  $[1, \alpha]$ , де  $\alpha \ll \infty$ . Вибір такої функції обумовлюється тим, що механізми сприйняття зовнішніх подразників є захищеними від недопустимого збільшення їх значень. Це означає, що при збільшенні сили подразника, яким в даному випадку є кількість повторень деякого слова, відповідний механізм сприйняття переходить у стан насичення і, при переході значення подразника через цю границю, ефективність його дії зменшується. Коефіцієнт  $k$  вибирається експертами, що визначають їх значення для окремих випадків підсилення. У випадку  $\aleph^R$ , це означає, що кількість повторень слова  $x_i$  для здійснення ефекту підсилення є обмеженою. По даних, що ґрунтуються на психологічних експериментах, така кількість слів, в більшості випадків, не перевищує семи повторень.

Розглянемо механізми дії підсилювача  $\aleph^C$ , що розміщається в  $\varphi_i$ , чи реченні  $\psi_i$ , який полягає у прискоренні досягнення цілі. Для цього, більш детально розглянемо, що собою представляє ціль у даному випадку. Така ціль представляє собою деякий текстовий опис, який надає для  $SO_i$  в тій, чи іншій мірі наступну інформацію:

- опис нового стану  $SO_i$ , в який переходить об'єкт, що представляється з певною мірою повноти  $j[S_i(C_i)]$ ,
- опис обґрунтувань, що обумовлюють доцільність переходу  $SO_i$  в  $SO_i^*$ , або  $\Gamma_i[SO_i \rightarrow SO_i^*]$ .

Моделювання процесу сприйняття окремою особою та  $SO_i$  деякої інформації здійснюється наступним чином. Оскільки  $TM_i$  описує відповідну  $SO_i$ , то зміна опису в  $TM_i$  моделює процеси змін, які відбуваються в  $SO_i$ . Тому, розглянемо процес модифікації  $TM_i$ , при дії на  $TM_i$  відповідного  $IP_i$ . Цей процес реалізується в рамках наступних етапів.

1. Перетворення  $TM_i$  і  $IP_i$  в логічну систему формул.
2. Виводу з  $L(TM_i)$  і  $L(IP_i)$  логічної системи формул, або  $L(TM_i * IP_i) \rightarrow L(TM_i^*)$ .
3. Перетворення  $L(TM_i^*)$  в  $TM_i^*$  реалізується на основі використання  $TM_i$  та  $IP_i$  з врахуванням  $L(TM_i^*)$ .
4. Синтез текстових форм  $TM_i^*$  і  $IP_i$ , в результаті чого формується  $TM_i^{**}$ .
5. Перевірка семантичних параметрів  $TM_i^{**}$ .
6. Визначення відповідності  $TM_i^{**}$  поставленій цілі  $J[C_i(TM_i)]$ .
7. Визначення міри відповідності  $TM_i^{**}$  відповідній  $SO_i^* = f(TM_i^{**})$ .

Перетворення  $TM_i$  і  $IP_i$  в логічну систему формул реалізується у відповідності з розробленими методами таких перетворень. В результаті виконання таких перетворень отримаємо систему формул  $L(TM_i)$  і  $L(IP_i)$ . З приведеної системи формул, використовуючи систему правил виводу  $\Phi_i$ , виводимо формулу  $L_i^*(TM_i^*)$ , яка на логічному рівні описує текстову модифіковану модель. Завдяки такому виводу, ми отримуємо опис результату модифікації, який не вміщає логічних суперечностей, чи інших логічних аномалій [3]. Це забезпечується використанням несуперечної системи логічного виводу та за рахунок використання скінченної процедури виводу.

На наступному кроці реалізується вивід модифікованої моделі  $TM_i^*$ , який використовує  $TM_i$  і  $IP_i$  та враховує  $L_i^*(TM_i^*)$ . В рамках даної системи окрема логічна формула описує одне речення, а їх сукупність представляє собою підсистему логічних формул, які описують окремий абзац. Це описується наступним чином.

$$L(tm_i^a) = L_i(l_{i1} * \dots * l_{im}),$$

де  $tm_i^a$  – один абзац текстового опису  $TM_i$ ,  $l_{ij}$  – деяка логічна функція, що описує одне речення  $\psi_{ij} \in tm_i^a$ . В цьому випадку  $\mathcal{L}_i(TM_i) = \mathcal{L}_i[L_{i1} * \dots * L_{im}]$ . Функція  $\mathcal{L}_i$  описує логічні взаємозв'язки між абзацами  $tm_i^a$  в  $TM_i$ ,  $L_{ij}$  – описує взаємозв'язки між реченнями  $\psi_{ij} \in tm_i^a$ , а  $l_{ij}$  – описує логічні взаємозв'язки між окремими фразами  $\varphi_{ij} \in \psi_i$ . Описувати логічні зв'язки між окремими словами в рамках окремої фрази у випадку їх інтерпретації в текстових формулах не має сенсу, оскільки, в більшості випадків, елементарна інтерпретація представляє собою фразу, а фраза формується у відповідності з правилами граматики, яка використовується для формування відповідних текстів. Оскільки синтез текстових форм описує  $TM_i$  і  $IP_i$  повинен описуватися у вигляді конструктивних методик, то необхідно в рамках  $TM_i$  ввести структуру, яка орієнтована на відображення семантики

відповідної моделі. З точки зору граматики, як відомо,  $TM_i$  представляє собою послідовність деякого тексту, який складається з фраз і речень, які розділені абзацами та розділовими ідентифікаторами, що представляють собою числа [4]. Для формування семантичної структури будемо користуватися приведеними вище формальними засобами розподілу тексту, в якому семантично однорідною компонентою буде абзац. Кожний з абзаців ідентифікується номером, який визначається в рамках  $TM_i$  та мірою його семантичної значимості ( $\pi_i^Z$ ). Прийmemo, що абзац, який будемо позначати символом  $u_i$  і має найбільше значення  $\pi_i^Z$ , буде розміщатися в місцях  $TM_i$ , які позначаються найменшими числами, що ідентифікують ці абзаци. В загальному випадку, виконання зазначеної ситуації є не обов'язковим, оскільки міру семантичної значимості будемо ідентифікувати номером числа, яким позначається окремий абзац  $\pi_i$ . На основі приведеної розмітки структури  $TM_i$ , яку будемо позначати  $S(TM_i, r)$ , де  $r$  розмітка  $TM_i$  семантичними значимостями абзацу  $\pi_i^Z$ , або  $r_i(\pi_i^Z)$ , необхідно побудувати зв'язки між окремими абзацами, їх реченнями та фразами. Така структура будується на основі аналізу системи логічних формул, що описують формулу  $L(TM_i^*)$ , що була виведена з формул  $L(TM_i) \& L(IP_i)$ .

Формування зв'язків між елементами  $TM_i^*$  буде реалізовуватися наступним чином. В рамках  $L(TM_i^*)$  використовуються логічні змінні  $u_i$ , що ідентифікують окремі фрази. Ці змінні зв'язані між собою логічними елементарними функціями  $\{\&, \vee, \rightarrow\}$ , що використовуються в  $L(TM_i^*)$ . Всі зв'язки поділимо на наступні групи:

- зв'язки між фразами окремих абзаців,
- зв'язки між фразами в межах одного абзацу,
- зв'язки між абзацами в межах  $TM_i^*$ .

Оскільки  $SO_i$  представляє собою деяку соціальну групу, що об'єднується на основі спільних ознак  $H_i$ , то семантика окремих абзаців описує відповідні ознаки. Носями таких ознак є окремі члени соціальної групи, або деяка частина членів відповідної групи. Окрема ознака  $H_i$  характеризується сукупністю параметрів  $(P_{i1}, \dots, P_{in})$ , кожний з яких ідентифікується фразами  $(\varphi_{i1}, \dots, \varphi_{in})$ , а відповідні фрази на рівні логічних формул, що описують  $TM_i$ , ідентифікуються змінними  $(y_{i1}, \dots, y_{in})$ .

Розглянемо встановлення зв'язку між окремими фразами  $\varphi_i$  і  $\varphi_j$ , що знаходяться в різних абзацах  $\pi_k$  і  $\pi_g$ . Нехай  $k < g$ , що означає, що  $r(\pi_k) < r(\pi_g)$  і взаємне розміщення фраз описується співвідношенням  $\pi_k \rightarrow \pi_g$ . Припустимо, що в  $\pi_k$  використовується фраза  $\varphi_i$  еквівалентна змінній  $u_i$ ,  $\varphi_i \leftrightarrow u_i$ . В абзаці  $\pi_g$  використовуються фрази  $\varphi_i$  і  $\varphi_j$ , що зв'язані логічними функціями  $\&, \vee, \rightarrow$ . Тоді, встановлюється взаємозв'язок між  $\pi_k$  і  $\pi_g$  на рівні залежностей між фразами  $\varphi_i$  з  $\pi_k$  і  $\varphi_j$  з  $\pi_g$ , якщо в  $\pi_g$  існує фрагмент логічної формули, що записується у вигляді  $u_i * u_j$ . Формально, правило встановлення такого типу зв'язку описується наступним співвідношенням:

$$\{(\varphi_i \in \pi_k) \& (\pi_k < \pi_g) \& (\varphi_j \in \pi_g) \& [(\varphi_i * \varphi_j) \in \pi_g]\} \rightarrow \rightarrow \{[\varphi_i(\pi_k) \rightarrow \varphi_j(\pi_g)] \rightarrow (\pi_k \rightarrow \pi_g)\}. \quad (1)$$

Зв'язок між фразами в межах одного абзацу  $\pi_i$  наа логічному рівні виникає в тому випадку, коли  $\varphi_i \in \psi_k$ , а  $\varphi_j \in \psi_m$  та має місце  $(\varphi_i * \varphi_j) \in \psi_m$ . В цьому випадку, можна записати наступне правило встановлення зв'язку в  $\pi_i$ :

$$\{(\varphi_i \in \pi_k) \& (\varphi_j \in \pi_g) \& [(\varphi_i * \varphi_j) \in \pi_g]\} \rightarrow [\varphi_i(\pi_k) \rightarrow \varphi_j(\pi_g)] \quad (2)$$

Якщо зв'язків типу (1) між двома фразами  $\pi_k$  і  $\pi_g \in m$  і  $m > 2$ , то має місце співвідношення:

$$\{[(\varphi_{i1}, \dots, \varphi_{im} \in \pi_k) \& [(\varphi_{j1}, \dots, \varphi_{jm}) \in \pi_g] \& \& \{[(\varphi_{i1} \rightarrow \varphi_{i1}) \in \pi_g] \& \dots \& [(\varphi_{j1} \rightarrow \varphi_{jm}) \in \pi_g]\} \rightarrow (\pi_k \rightarrow \pi_g)\}. \quad (3)$$

Співвідношення (3) описує взаємозв'язок між абзацами  $\pi_k$  і  $\pi_g$ .

Приведена структура моделі  $TM_i$  відображає семантичні зв'язки між окремими фрагментами тексту. Логічна структура моделі описує такі взаємозв'язки між окремими елементами тексту, які не приводять до логічних суперечностей. Процес синтезу  $TM_i$  і  $IP_i$  ґрунтується на використанні семантичної сумісності між фрагментами з  $TM_i$  та фрагментами з  $IP_i$ . Такими фрагментами в рамках приведених структур є:

- абзаци тексту  $j(\pi_i)$ ,
- речення текстових описів  $j(\psi_i)$ ,
- текстові описи фраз  $j(\varphi_i)$ .

Між  $TM_i$  і  $IP_i$ , якщо останній орієнтований на те, щоб реалізувати управляючий вплив на  $SO_i$ , повинні існувати семантично узгоджені фрагменти. В рамках стратегії  $S_t(IP_i)$ , що описує процес управління  $SO_i$ , який проявляється у вигляді модифікації  $TM_i$ , реалізується принцип поступових модифікацій  $TM_i$ . Це означає, що окремий  $IP_i \rightarrow TM_i$  повинен діяти лише на один вибраний фрагмент, що найменше, на рівні  $\pi_i$ . Це обумовлене тим, що один фрагмент описує один тип особливостей, або характеристик, які описують семантичний зміст окремого  $SO_i$ , що відповідно відображається у текстовому описі  $j(TM_i)$ . Таким чином, в рамках стратегії поступових модифікацій, або  $S_t^{PM}(TM_i)$ , синтез реалізується в межах одного абзацу  $tm_i^a \in TM_i$ . Це означає, що окремий  $IP_i$  діє лише на окремий  $tm_i^a \in TM_i$ . Тому  $IP_i$  повинен вмщати наступні фрагменти текстових описів, до яких відносяться:

- фрагменти ідентифікації абзацу  $j[\pi_i(ip_i)]?$
- фрагмент, що описує модифікацію у відповідному абзаці,
- фрагмент, що описує ціль такої модифікації.

Формально, це можна описати співвідношенням:

$$j(IP_i) = F[j(ip_i^I), j(ip_i^M), j(ip_i^C)].$$

Фрагмент  $j(ip_i^I)$  представляє собою текстовий опис, що ідентифікує абзац. Такі ідентифікатори досить широко використовуються в традиційних

текстових описах і представляють собою назви окремих розділів тексту, які називають заголовками, короткі анотації змісту відповідного фрагменту тексту і, як різновидність більш формалізованого, з точки зору семантики фрагменту, можна до фрагментів типу  $ip_i^l$  відносити ключові слова. Очевидно, що  $j(ip_i^l)$  та  $j(tm_i^a)$ , або  $j(tm_i^l)$  не мусять бути абсолютно ідентичними. Для того, щоб можна було ввести міру ідентичності, на основі якої  $IP_i$  розпізнає відповідний  $tm_i^a$ , приймемо наступне визначення.

*Визначення 3.* Будь який абзац  $\pi_i(TM_i)$  представляє собою анотацію змісту абзацу  $tm_i^a(I)$ , яка повинна використовувати в скороченому описі цього змісту ключові слова абзацу, які будемо позначати змінною  $x_i^k$ .

*Визначення 4.* Фрагмент ідентифікації  $ip_i^l(TM_i)$ , що формується в  $IP_i$  і передається до  $TM_i$ , вміщає ключові слова  $(x_{i1}^k, \dots, x_{im}^k)$ , з допомогою яких вибирається необхідний абзац  $tm_i^a \in TM_i$ .

*Визначення 5.* Фрагмент ідентифікації  $ip_i^l(SO_i)$ , що формується в  $IP_i$  і передається до  $SO_i$ , вміщає анотацію модифікацій, що пропонуються для  $SO_i$ .

Компонента  $ip_i^M \in IP_i$  може представляти собою наступні семантичні сутності:

- опис процесу перетворень, які пропонуються в  $SO_i$ ,
- опис результату можливих перетворень, які пропонуються в  $SO_i$ , що представляють собою модифікований фрагмент з  $SO_i$  і, відповідно,  $TM_i$ ,
- синтезований опис, який вміщає фрагменти процесу перетворень та фрагменти опису модифікованого  $SO_i$  і  $TM_i$ .

Компонента  $ip_i^C \in IP_i$  представляє собою опис результатів перетворень, що реалізуються в  $SO_i$  під дією  $IP_i$ . Такий опис може представляти собою опис нових можливостей, що виникають в  $SO_i$ , для процесу його функціонування. Відповідні можливості можуть представляти собою опис позитивних змін, що виникають в результаті модифікації  $SO_i$  відповідним  $IP_i$  та опис негативних змін, до яких можуть привести відповідні модифікації. В рамках приведених семантичних інтерпретацій компонент  $IP_i$  приймемо, що будь який  $IP_i$  орієнтований на позитивні зміни в  $SO_i$ . При цьому, використання семантичних  $\sigma^i$  та психосемантичних параметрів  $\aleph^i$  орієнтовано на забезпечення ефективності впливу  $IP_i$  на  $SO_i$ . Для того, щоб виключити можливість використання такого явища, як дезінформацію у випадках реалізації управляючих дій  $IP_i$  на  $SO_i$  в рамках  $STM_i$  про узагальнені ознаки позитивних і несанкціонованих наслідків дії  $IP_i$  на  $SO_i$ . Такі уявлення детальніше розглядатися не будуть, оскільки вони стосуються соціальних досліджень [5].

1. Watts D.J., Strogatz S.Y., Collective dynamics of "small-world" networks/ Nature, 393: 440-442, 1998.
2. Афанасєва О.Ю. Дослідження семантичних особливостей відображення



графічного образу. / Моделирование та інформаційні технології, зб. наук. праць (ІПМЕ НАН України). – К. 2007, вип. 40.

3. Шенфілд Дж. Математическая логика. М.: Наука, 1976, -527 с.

4. Папина А.Ф. Текст: его единицы и глобальные категории. М.: Книжный дом «ЛИБРОКОМ», 2009. – 250 с.

5. Мухаев З.Т. Социология. М.: «Книга сервис», 2003. 320 с.

*Поступила 18.9.2013р.*

УДК 004.9

О.В.Тимченко<sup>1 2</sup>, І.О.Кульчицька<sup>2</sup>, О.О.Тимченко<sup>3</sup>

## **АЛГОРИТМИ ТА ФУНКЦІЇ ІНФОРМАЦІЙНОЇ СИСТЕМИ РОЗПІЗНАВАННЯ СИМВОЛІВ НА ОСНОВІ МЕТОДІВ ПОЛПШЕННЯ ЗОБРАЖЕНЬ**

*Анотація.* У даній статті розглянуто загальну структурну схему розпізнавання символів, описано основні функції кожного блоку даної схеми, визначено проблеми, що можуть виникати в кожному блоці.

*Annotation.* This paper discussed a general block diagram of character recognition, describes the main features of each unit of the scheme, identified problems that may arise in each block.

*Ключові слова.* Оптичне розпізнавання тексту, система розпізнавання, попередня обробка зображення.

*Keywords.* Optical character recognition (OCR), recognition system, image processing.

### **Вступ**

Розпізнавання образів є однією з дуже важливих і активно досліджуваних областей штучного інтелекту. Ця наука вивчає методи класифікації об'єктів. Таким об'єктом може бути цифрова фотографія (розпізнавання зображень), буква чи цифра (розпізнавання символів), запис мови (розпізнавання мови) тощо. Системи розпізнавання призначені для класифікації вхідних зображень або їх частин на категорії [3]. При цьому кожен фрагмент вхідного зображення можна розглядати як точку в деякому просторі всіх можливих фрагментів. На бінарному зображенні для фрагменту розміром  $p \times n$  елементів кількість всіх можливих фрагментів  $2^{n \times p}$  може бути дуже великою для відносно малого

---

<sup>1</sup> Uniwersytet Warmińsko-Mazurski w Olsztynie

<sup>2</sup> Українська академія друкарства

<sup>3</sup> Львівський Національний університет ім.Івана Франка