

МОДЕЛЮВАННЯ ОЦІННИХ СУДЖЕНЬ У ДОПИСАХ УЧАСНИКІВ ВЕБ-ФОРУМІВ

© Тимовчак-Максимець О.Ю., 2010

Розглянуто поняття і структуру оцінного висловлювання і запропоновано модель оцінного висловлювання як частини допису учасника форуму та схему виявлення й аналізу оцінних суджень у соціальних мережах WWW.

Ключові слова: оцінне судження, допис, смислово-емоційний аналіз, веб-форум.

This paper deals with the notion and structure of opinionated phrase. The model of an opinionated phrase as a part of forum member post as well as the scheme of identifying and analyzing opinionated phrases in WWW social environment are suggested.

Keywords: opinionated phrase, post, sentiment analysis, Web forum.

Вступ

В умовах розвитку ринкової економіки і посилення конкуренції між виробниками продукції відбувається боротьба за споживача. Кожен виробник намагається завоювати якомога ширше коло споживачів і вдається до різних заходів: розширення і частоті зміни асортименту продукції, проведення рекламних кампаній, розроблення різноманітних акційних пропозицій тощо. Сукупність цих заходів від різних виробників створює значний інформаційний шум, з якого кожен потенційний споживач повинен виділити інформацію про той продукт, який найкраще задовольнить його потреби та вимоги. Проте в інтенсивному потоці інформації споживачеві часто важко зорієнтуватися, чим, наприклад, відрізняються два однотипні товари, які переваги чи недоліки вони мають, якими характеристиками виділяється кожен з них. Щоб уникнути прикрої помилки, надаючи перевагу одному з варіантів, споживачі часто звертаються до досвіду та вподобань суспільства. Життєві історії, поради, схвальні або критичні відгуки про продукт часто стають ключовими факторами у прийнятті позитивного чи негативного рішення щодо його купівлі. Однак такий досвід не завжди є достатнім і/або наявним у найближчому соціальному оточенні споживача, тому виникає потреба пошуку інформації в інших суспільних джерелах.

Завдяки значному поширенню інформаційно-комунікаційних та інтернет-технологій, соціальне середовище WWW стало, фактично, рефлексором громадської думки у різних предметних областях. Знання, досвід, уподобання та поради учасників спільноти, акумульовані у вигляді дописів на форумах, особистих і колективних блогах, дошках повідомлень тощо, формують колективні знання. Дописи, які містять оцінні судження про продукт, що основані на особистому досвіді чи знаннях автора, становлять значну цінність, з одного боку, для потенційного покупця цього продукту, оскільки вказують на задовільне чи незадовільне ставлення автора допису до продукту. За результатами аналізу таких дописів, потенційний покупець підводить підсумки і визначає ставлення суспільства до продукту. З іншого боку, результати аналізу оцінних суджень користувачів певного продукту свідчать про рівень задоволення потреб та вимог користувачів цього продукту, що для виробника є надзвичайно важливою інформацією для забезпечення конкурентоспроможності його продукції.

Отже, оцінні судження користувачів допомагають потенційному покупцеві продукції зробити правильний вибір, а виробникові – отримати інформацію про попит на його продукцію.

Постановка проблеми

Щоб встановити думку учасників веб-спільноти про певний об'єкт, потрібно проаналізувати дописи про цей об'єкт на форумі від різних авторів за певний період. Дописи є продуктом людсь-

кого спілкування, і, хоча вони організовані тематично в різні гілки, учасники спільноти не є обмежені в обговоренні різних об'єктів в межах тематики, тому дописи не є структурованими і можуть містити інформацію, яка не становить цінності для встановлення уподобань соціуму. Отже, пошук та опрацювання великої кількості дописів вручну потребує значних затрат зусиль і часу. Виникає потреба у розробленні системи автоматизованого смислово-емоційного опрацювання оцінних суджень користувачів про об'єкт у соціальному веб-середовищі.

Оцінка, з одного боку, це універсальна категорія, яка виражає позитивне або негативне ставлення суб'єкта цієї оцінки до змісту висловлювання. Однак оцінка є соціально обумовленою, оскільки основана на системі цінностей носія, і в різних мовах реалізується, експліцитно або імпліцитно, в різних семантичних і синтаксичних структурах (словах, вигуках, словосполученнях, фразах тощо). Крім цього, реалізація одного і того самого значення оцінки в мовних конструкціях може варіюватись залежно від предметної області. Різноманіття мовних конструкцій для вираження певного оцінного значення ускладнює процес їх автоматизованого пошуку та аналізу.

Аналіз останніх досліджень та публікацій

Протягом останніх років, як свідчать публікації, смислово-емоційний аналіз текстів був сферою активних досліджень іноземних науковців. Напрями досліджень охоплюють смислову класифікацію текстів, тобто класифікацію текстів на такі, що мають позитивне чи негативне спрямування [15, 19], класифікацію суб'єктивності речень [10, 17, 21, 22, 23], смислово-емоційний аналіз тематики або властивостей (співвіднесення позитивного чи негативного значення з певними темами або властивостями об'єкта) [8, 9, 11, 12, 16, 18]. Формальне визначення різних аспектів проблематики смислово-емоційного аналізу та огляд основних досліджень та алгоритмів розглядає у своїх працях Liu [13,14]. Базою досліджень у цих працях є англійські тексти, тому, відповідно, ці напрацювання можуть бути лише частково застосовані для обробки україномовних текстів.

Вітчизняні дослідження категорії оцінки провадяться з теоретико-лінгвістичного погляду. Оцінні значення мовних виразів привертають увагу багатьох українських та зарубіжних мовознавців, їх розглядають у логіко-семантичному [1, 2, 4], семантико-прагматичному [5] та комунікативному [3] аспектах.

Психолінгвістична експертна система ВААЛ є однією з прикладних розробок, що здійснює емоційно-лексичний аналіз текстів та автоматичну категоризацію текстів для української, російської та англійської мов [7]. Однак ця система здійснює аналіз та категоризацію вибраного тексту, але не здійснює пошуку емоційно забарвлених текстів та оцінних висловлювань в масі слабкоструктурованої інформації.

Формулювання цілей статті

З огляду на складне мовне представлення (експліцитне або імпліцитне) оцінки об'єкта, у статті розглядається поняття оцінки і її класифікації, оцінне висловлювання з погляду його структури. На основі дописів на веб-форумі моделюється структура оцінного висловлювання як частини допису учасника форуму. Схема автоматизованого опрацювання дописів відображає послідовність дій, які виконує система для виявлення і аналізу оцінних суджень.

Виклад основного матеріалу

1. Поняття оцінки як категорії модальності

У мові відображається взаємодія людини і дійсності в різних аспектах, одним з яких є оцінний: мовець поділяє об'єктивну дійсність з погляду її ціннісного характеру, тобто “добра” і “зла”, “корисності” і “шкоди” тощо. Отже, оцінку можна розглядати як один з видів модальності, які накладаються на дескриптивний зміст мовного виразу. Висловлювання, яке містить оцінну чи іншу модальність, містить дескриптивну і недескриптивну, тобто модальну, компоненту, перша з яких описує реальну чи можливу ситуацію, а друга – ставлення мовця до цієї ситуації.

Оцінне висловлювання характеризується особливою структурою, яка містить обов'язкові та факультативні елементи. Цю структуру можна інтерпретувати як модальну рамку, яка накладається на висловлювання і не збігається ані з його логіко-семантичною будовою, ані з синтаксичною.

У книзі Е. М. Вольф “Функциональная семантика оценки” в основу оцінної модальності закладене відношення ArB , в якій А є суб’єктом оцінки, В – її об’єктом, а r – оцінним відношенням, яке має значення “добре/погано” [2, с. 12]. Отже, основними елементами оцінної модальної рамки є суб’єкт і об’єкт, які зв’язані оцінним предикатом. У природній мові відношення r може мати втілення як у різноманітних словах (*добрий, поганий, цінний, дешевий, поважати, зневажати* тощо), так і в семантиці висловлювання загалом. Оцінне значення, яке є частиною модальної рамки, поєднується з дескриптивним, яке виражене в пропозиційних структурах.

Предикат r у виразі ArB (суб’єкт – оцінка – об’єкт) характеризується таким набором ознак, які відображають специфіку оцінного ставлення суб’єкта до об’єкта:

- емотивність або оцінність;
- емоційність/раціональність;
- афективність.

Суб’єкт оцінки, експліцитний чи імпліцитний, – це особа або соціум, з погляду якої/якого дається оцінка. Об’єкт оцінки – це особа, предмет, подія чи стан речей, яких стосується оцінка. Крім цього, до модальної рамки входять, зазвичай імпліцитно, шкала оцінок і стереотипи, на які орієнтована оцінка в соціальних уявленнях мовців.

Ці елементи оцінної структури відповідають компонентам оцінки з погляду логіки. Проте в природній мові оцінна структура побудована значно складніше і містить, крім цього, ще ряд компонентів. Так, суб’єкт і об’єкт оцінки часто поєднуються аксіологічними предикатами, такими як предикати думки, відчуття, сприйняття (*здаватися, вважати, бути переконаним тощо*).

Семантичний зв’язок оцінних слів та позначень об’єкта оцінки здійснюється на базі аспекту оцінки (основної змінної), який визначає властивості об’єкта, за якими його оцінюють. Оцінне висловлювання може також містити і факультативні елементи – мотиватори, класифікатори, різноманітні засоби інтенсифікації та деінтенсифікації. У випадку порівняльної оцінки до модальної рамки входять додаткові елементи – те, з чим порівнюється, властивість чи ознака, за якою дається порівняння, мотиватори порівняння тощо [2, с. 13].

Модальна рамка оцінки стосується прагматичного аспекту висловлювання, однак вона тісно пов’язана також з семантикою оцінної структури, які разом утворюють амальгамовані конструкції, де прагматичний і семантичний фактори не завжди легко розділити. Тобто, якщо суб’єкт оцінки і предикат судження можна з впевненістю віднести до модальної рамки, то об’єкт оцінки завжди входить також до семантичної (дескриптивної) структури висловлювання.

Як бачимо, оцінна структура складається з багатьох елементів, що ускладнює її виокремлення та обробку порівняно з описовою структурою.

2. Суб’єктивний і об’єктивний фактори в оцінці

У тексті оцінка матеріалізується у вигляді певних оцінних суджень, які містять як власне вираження оцінки (суб’єктивне), так і деякі якісні характеристики об’єкта, який оцінюється. В оцінці завжди є суб’єктивний фактор, який взаємодіє з об’єктивним, і в оцінному висловлюванні, навіть якщо в ньому прямо не виражений суб’єкт оцінки, мається на увазі ціннісне відношення між суб’єктом і об’єктом. Висловлення чи приписування цінності (оцінювання) є встановленням певного відношення між суб’єктом чи суб’єктами оцінки і її об’єктом [4].

Наявність оцінного суб’єкта передбачає деякі особливі властивості оцінного судження і, передусім, можливість суперечки про оцінку, в якій зіштовхуються різні думки. Таке зіткнення оцінних суджень виникло на “Форумі Рідного Міста” щодо популярної кінострічки “Аватар” [6]:

(А) *Як на мене, не дуже цікаво, не розумію, чому фільм викликав такий ажіотаж у світі, можливо, хіба завдяки технологіям...*

(В) *Шикарна річ! На мій смак – трішки затягнуті динамічні сцени, особливо бойові, та все ж – фільм дуже сподобався!*

(С) *Я "Аватаром" не розчарована, бо побачила саме те, що й очікувала – красиво зроблену фантастику.*

Суб'єктивний аспект визначає також склад аксіологічних предикатів як одного із елементів оцінної структури. Такі предикати, як-от: *як на мене, на мій смак* тощо – відображають суб'єктивний характер оцінки.

Існують різні класифікації типів оцінки, але вони спираються на розуміння суб'єктивного і об'єктивного аспекту значення оцінних слів і висловлювань. Загалом оцінки поділяються на такі, що мають загальнооцінне та частково-оцінне значення. Загальна оцінка висловлюється поняттями “добре”- “погано” (“краще”- “гірше”) і відповідними до них синонімами і містить в собі оцінку об'єкта загалом. Частково-оцінні значення стосуються окремих рис, сторін чи ознак об'єкта, який при цьому розглядається з одному певному ракурсі.

Загальна оцінка *добре*”- “погано” здійснюється відносно соціокультурної норми і має зміст лише при порівнянні з нею. Зауважимо, що значення “*добре*” відповідає нормі або навіть перевищує її, тоді як “*погано*” – це завжди відхилення від норми або її зрушення. Вважається, що норма відображає дещо ідеалізовану картину дійсності.

Першу класифікацію оцінок за типами об'єктів запропонував Г.Х. фон Врігт у книзі “Різновиди добра” [20].

Фон Врігт запропонував класифікацію типів оцінки, що основана на видах об'єктів і семантиці сполучуваності зі словом “хороший” (good). Ця класифікація налічує шість різновидів оцінок:

- 1) інструментальна оцінка (instrumental) – оцінка знарядь, інструментів: *хороший молоток, годинник*;
- 2) технічна оцінка (technical) – оцінка здібності, майстерності: *хороший працівник, музикант*;
- 3) медична оцінка (medical) – оцінка функцій органів тіла та здібностей розуму: *хороший зір, хороша пам'ять*;
- 4) гедоністична (hedonic) – оцінка задоволення: *хороший запах, хороша сигара*;
- 5) оцінка сприятливості (beneficial) – сприяння досягненню позитивного результату: *хороший намір, хороший вчинок*;
- 6) утилітарна (utilitarian) – оцінка корисності, придатності: *хороша нагода, хороша новина*.

Н.Д. Арутюнова запропонувала класифікацію частково-оцінних значень, основану на взаємодії суб'єкта оцінки з її об'єктом [1]. Дослідниця виділяє три групи частково-оцінних значень:

- сенсорні;
- сублімовані;
- раціоналістичні.

Ці групи поділяються на категорії, які в мові найчастіше виражені прикметниковими формами.

До групи сенсорних оцінок входять сенсорно-смакові та психологічні оцінки. Сенсорно-смакові, або гедоністичні, оцінки – найіндивідуалізованіший різновид, який пов'язаний із фізичними відчуттями та психічним досвідом (наприклад: *смачний – несмачний, привабливий–відразливий*). Категорія психологічних оцінок поділяється на інтелектуальні (*цікавий, захоплюючий, багальний*) та емоційні (*радісний, бажаний*).

До другої групи належать сублімовані або абсолютні оцінки. Тут виділяють естетичні оцінки, основані на синтезі сенсорних і психологічних (*гарний, чудовий*), та етичні, що стосуються норм (*добрий, чесний*).

Раціоналістичні оцінки пов'язані з практичною діяльністю людини, повсякденним досвідом людини, з фізичною чи психічною користю, з відповідністю стандартам. До них належать утилітарні оцінки (*корисний*), нормативні (*правильний, здоровий*) і телеологічні оцінки (*ефективний, вдалий*).

Очевидно, що аналіз оцінок в лінгвістичному плані спирається на розуміння суб'єктивного і об'єктивного аспекту значення оцінних слів і висловлювань. В оцінних висловлюваннях суб'єктивне нерозривно пов'язане з об'єктивним, а їх відношення є обернено пропорційним.

3. Модель оцінного судження у веб-дописі

Загальна позитивна характеристика, що висловлюється в тексті щодо певного об'єкта, не доводить, що автор позитивно ставиться до окремих його аспектів чи властивостей. І навпаки. У типових емоційно-забарвлених текстах автори висловлюють як позитивні, так і негативні оцінки,

хоча їх ставлення до об'єкта та узагальнювальне судження може бути або позитивним, або негативним. Окрім цього, тести часто містять факти, а також інформацію, яка може виражати позитивну/негативну оцінку, але не стосуватися об'єкта.

Нехай (t) – це текст, тобто допис на форумі, у блозі чи огляд певного продукту, який містить оцінні судження про цей продукт. Оскільки переважна більшість текстів складаються з більш ніж одного речення (s) , то $t=(s_1, s_2, \dots, s_n)$

Розглянемо коментар дописувача з “Форуму Рідного Міста” щодо електронної книжки [6]:

“(1) Враховуючи ціну на їхньому сайті, то ще зекономив. (2) Виправдання собі вже придумав – 'воно себе окупить'. (3) Задоволений. (4) По функціональності – так, було б класно з неї в Інтернет ходити чи текст по-людськи набирати. (5) Як саме електронна книжка дивайс своє ім'я виправдовує. (6) Легенький, приємно взяти до рук, шрифти і їх розміри – без обмежень. (7) Тішить око, одним словом. (8) Бокові кнопки, правда, хитаються, але що вже з тайванських китайців візьмеш 🙄.”

Щодо ставлення автора – речення s_1, s_2, s_3, s_5, s_6 і s_7 передають позитивні емоції, а s_4 і s_8 – негативні, хоча вони мають різне спрямування, тобто стосуються різних характеристик цього об'єкта. Так, речення s_1 стосується ціни, s_3 – функціональності пристрою, s_4 – пристрою загалом, тобто самого об'єкта, s_6 – фізичних властивостей і функціонального забезпечення, а s_8 – функціональних кнопок. Речення s_2 , хоч і має позитивне емоційне забарвлення, але не стосується пристрою. Отже, для реалізації ефективного автоматизованого аналізу оцінних суджень слід враховувати їх емоційну спрямованість.

Для визначення спрямованості оцінного судження слід визначити, що є його об'єктом. Об'єктом є певна сутність, що може бути продуктом, людиною, організацією, подією або темою, яку оцінюють. Об'єкт може мати набір компонентів (або частин) та набір атрибутів (або якостей), які повинні мати власний набір компонентів та атрибутів, які, своєю чергою, також можуть поділятися. Отже, об'єкт ієрархічно можна розкласти на елементи, використовуючи відношення “частина від цілого”.

Формально ми можемо описати об'єкт (o) як сутність, для якої існує множина властивостей $F=(f_1, f_2, \dots, f_n)$, кожна з яких характеризується парою (C, A) , де C – ієрархія компонентів, а A – набір атрибутів для f . Якщо цю структуру зобразити у вигляді графа, то об'єкт буде вершиною графа, а його властивості $f_n(C_n, A_n)$ – вузлами цього графа.

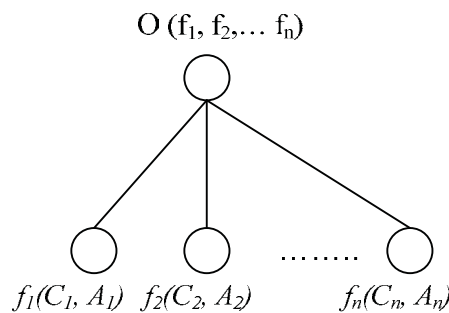


Рис. 1. Структура об'єкта у вигляді графа

Одне речення може містити оцінні висловлювання про одну і більше властивість (1), з іншого боку оцінні судження про одну і ту саму властивість можуть бути частиною кількох речень емоційно забарвленого тексту (2):

(1) Джойстиком гортати сторінки дуже зручно, але, оскільки корпус у нього зручний і легкий, його легко можна тримати лише одною рукою.

(2) В налаштуваннях пропонують немало дуже корисних можливостей. Особливо корисним є пункт “Призначення клавіш”.

Отже, оцінним уривком тексту (t) про властивість (f) об'єкта (o) є послідовність речень тексту (t) , в яких міститься позитивна або негативна оцінка властивості (f) .

Окрім явно виражених (експліцитних) властивостей: “шрифти і розміри без обмежень”, де компоненти (шрифти і розміри) та їх атрибути (без обмежень) чітко виражені, трапляються також

неявно виражені властивості. Наприклад, послідовність слів “*Легенький, присмно взяти до рук*” виражає імпліцитну властивість *ваги*.

Якщо властивість (*f*) або хоча б одне синонімічне значення цієї властивості (якщо *екран* є властивістю, то *дисплей* або *монітор* є його синонімічними представленнями) безпосередньо наявні в реченні (*s*), то така властивість є експліцитною. Якщо ні властивість (*f*), ні жодне з її синонімічних представлень не виявлені в реченні, але є певний вказівник на цю властивість, то така властивість є імпліцитною. Такий вказівник називається індикатором властивості. У ролі індикаторів властивостей можуть виступати прикметники та прислівники. Деякі індикатори є загальними і використовуються для характеристики будь-якої властивості, наприклад, *хороший, поганий*, але більшість застосовується для модифікації конкретних властивостей: вказівник *привабливий* модифікує властивість “зовнішній вигляд”, *великий* модифікує “розмір” тощо.

Отже, кожному властивості $f_i \in F$ можна виразити скінченним набором слів або словосполучень $W_i = \{w_{i1}, w_{i2} \dots w_{im}\}$, які є синонімами цієї властивості, або описати певним скінченним набором індикаторів властивості $I_i = \{i_{i1}, i_{i2} \dots i_{iq}\}$.

Групування синонімічних представлень властивостей, а також нагромадження індикаторів властивостей є одним з ключових завдань для здійснення автоматизованого аналізу оцінних суджень. Для створення базового набору властивостей та індикаторів використовуються тезауруси, але оскільки більшість індикаторів є залежними від предметної області, то подальше розширення цього набору здійснюють, відбираючи їх з текстів певної предметної області.

Оцінне судження завжди має суб'єкт оцінки. Однак автор допису не завжди є суб'єктом оцінного судження.

(1) *На мій погляд, пристрій більш ніж достойний.*

(2) *Кілька знайомих мають вже кілька років, жодних скарг і маса задоволення.*

Дописи (1) і (2) мають позитивне емоційне забарвлення, але в першому автором допису і суб'єктом судження є “я”, тобто автор допису, а в другому автор допису передає думку суб'єкта, тобто “кілька знайомих”. З прикладу видно, що автор допису сам може бути суб'єктом оцінного судження або передавати судження іншого індивіда, соціуму тощо. Отже, суб'єктом (*h*) оцінного судження (*E*) стосовно об'єкта (*o*) або його властивості (*f*) є особа, яка висловила це судження. Для кожного судження справджуються такі умови:

1) один і той самий суб'єкт може висловлювати оцінні судження стосовно різних об'єктів і їх властивостей;

2) одне і те саме судження не може мати двох різних суб'єктів.

Для автоматизованого опрацювання тексту та визначення спрямованості чи полярності висловлювання особа автора не відіграє ключової ролі, оскільки визначається семантика тексту, а не його достовірність. Особа автора, а точніше, його авторитетність, є одним з ключових факторів на завершальній стадії опрацювання дописів на форумах, оскільки впливає на довіру до цього допису, а отже, і на ранг цього допису. Огляди продуктів чи послуг у блогах здійснює, зазвичай, сам автор тексту, тому автор ототожнюється з суб'єктом думки. Допис на форумі може містити оцінні судження самого автора допису, а також судження інших суб'єктів. Відповідно, довіра до суджень інших суб'єктів у цьому дописі буде рівнозначною довірі до автора допису.

Як вже згадувалося, найважливішою особливістю оцінки є те, що в ній завжди є суб'єктивний фактор, що взаємодіє з об'єктивним. Суб'єктивний компонент передбачає позитивне чи негативне ставлення суб'єкта оцінки до її об'єкта. Це ставлення відображається оцінним предикатом, який може набувати двох основних значень – “за” і “проти”, “добре” і “погано”, “+” і “-”. Отже, оцінним предикатом (*ep*) судження (*E*) стосовно об'єкта (*o*) або його властивості (*f*) є ставлення суб'єкта оцінки до її об'єкта.

На перший погляд здається, що признаки “добре” і “погано” завжди є семантичним компонентом оцінного слова: *привабливий, якісний, недорогий* (“+”), *ненадійний, громіздкий*, (“-”). Однак у багатьох слів, які мають компонент оцінки, цей знак не є детермінованим. Знак (“+”) чи (“-”) часто не можна приписати не лише словам, але і словосполученням, якщо їх розглядати поза контекстом. Так, наприклад, слова *дешевий, простий, важкий*, словосполучення *може бути, більш*

мени, хотілось би кращого є складними для автоматизованого визначення полярності, оскільки не мають чітко детермінованого знака, а тому їх полярність визначається за висловлюванням загалом, тобто місцем відповідної ситуації у ціннісній картині світу (рис. 2).

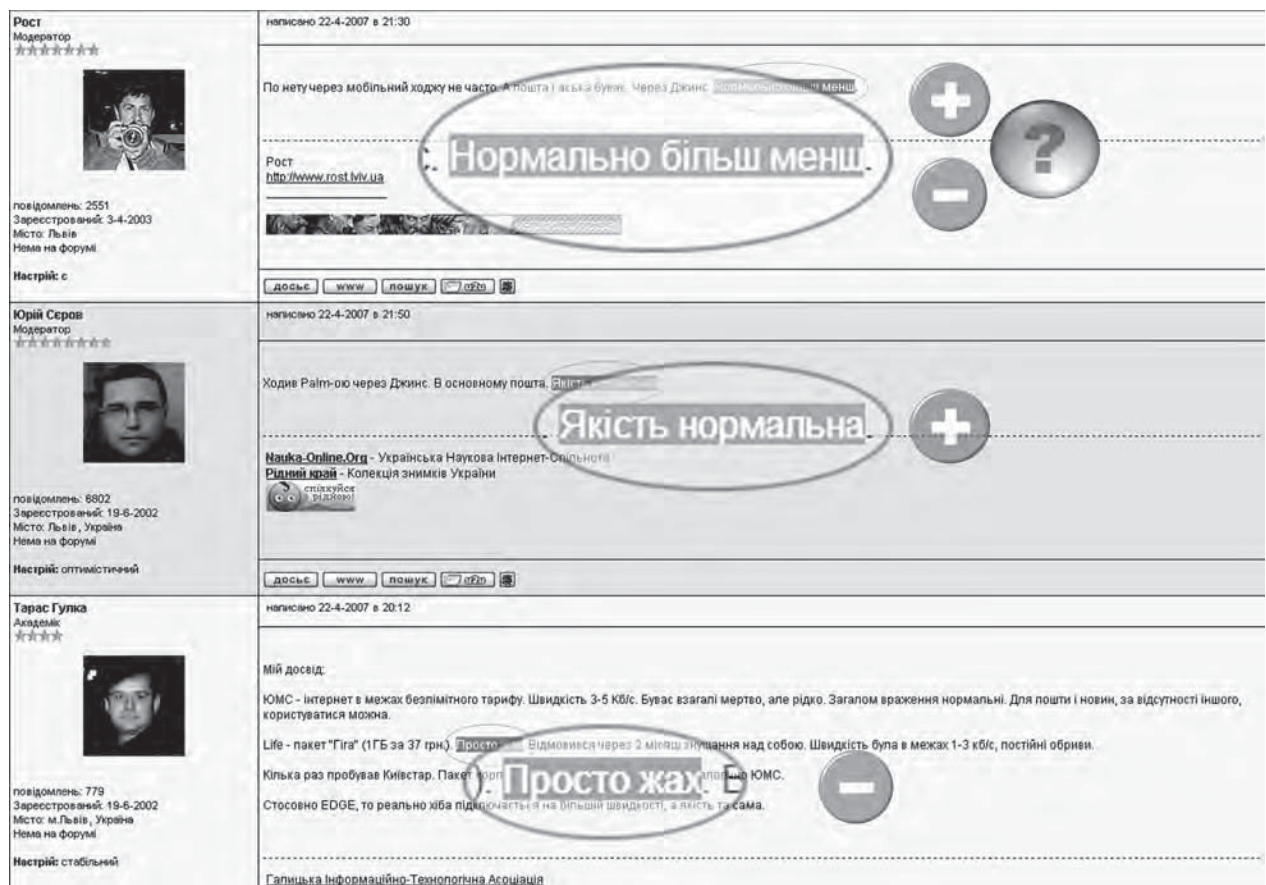


Рис. 2. Приклади детермінованого і недетермінованого знака оцінного предикату

Кожен допис у блозі чи на форумі, а відтак і оцінне судження, характеризується часом його опублікування (tt). Ця характеристика є ключовою для встановлення актуальності оцінного судження. Наприклад, для встановлення динаміки оцінних суджень про певний об'єкт необхідно відфільтрувати оцінні судження за певний період часу.

Для того щоб виділити з тексту аксіологічне висловлювання про об'єкт, треба виокремити такі елементи:

$$E = (o_j, f_{jk}, ep_{ijkl}, h_i, tt_i), \quad (1)$$

де E – оцінне судження; o_j – об'єкт оцінного висловлювання; f_{jk} – властивість об'єкта оцінного висловлювання; ep_{ijkl} – оцінний предикат (полярність) оцінного судження щодо властивості f_{jk} ; h_i – суб'єкт цього судження; tt_i – час, коли автор опублікував допис.

Ця модель використовується в системі обробки оцінних суджень, які розміщені у дописах учасників веб-спільнот, принцип роботи якої описаний в наступному розділі.

4. Система обробки оцінних суджень у веб-спільнотах

Завданням системи обробки оцінних суджень є отримання консолідованих результатів інтелектуального аналізу оцінних суджень учасників соціальних спільнот. Є три основні етапи опрацювання оцінних суджень про об'єкт:

- 1) пошук і відбір речень, які містять оцінні судження;
- 2) інтелектуальний аналіз відібраних речень;
- 3) консолідація та представлення результатів аналізу.

Схема системи обробки оцінних суджень у соціальних спільнотах зображена на рис. 3.

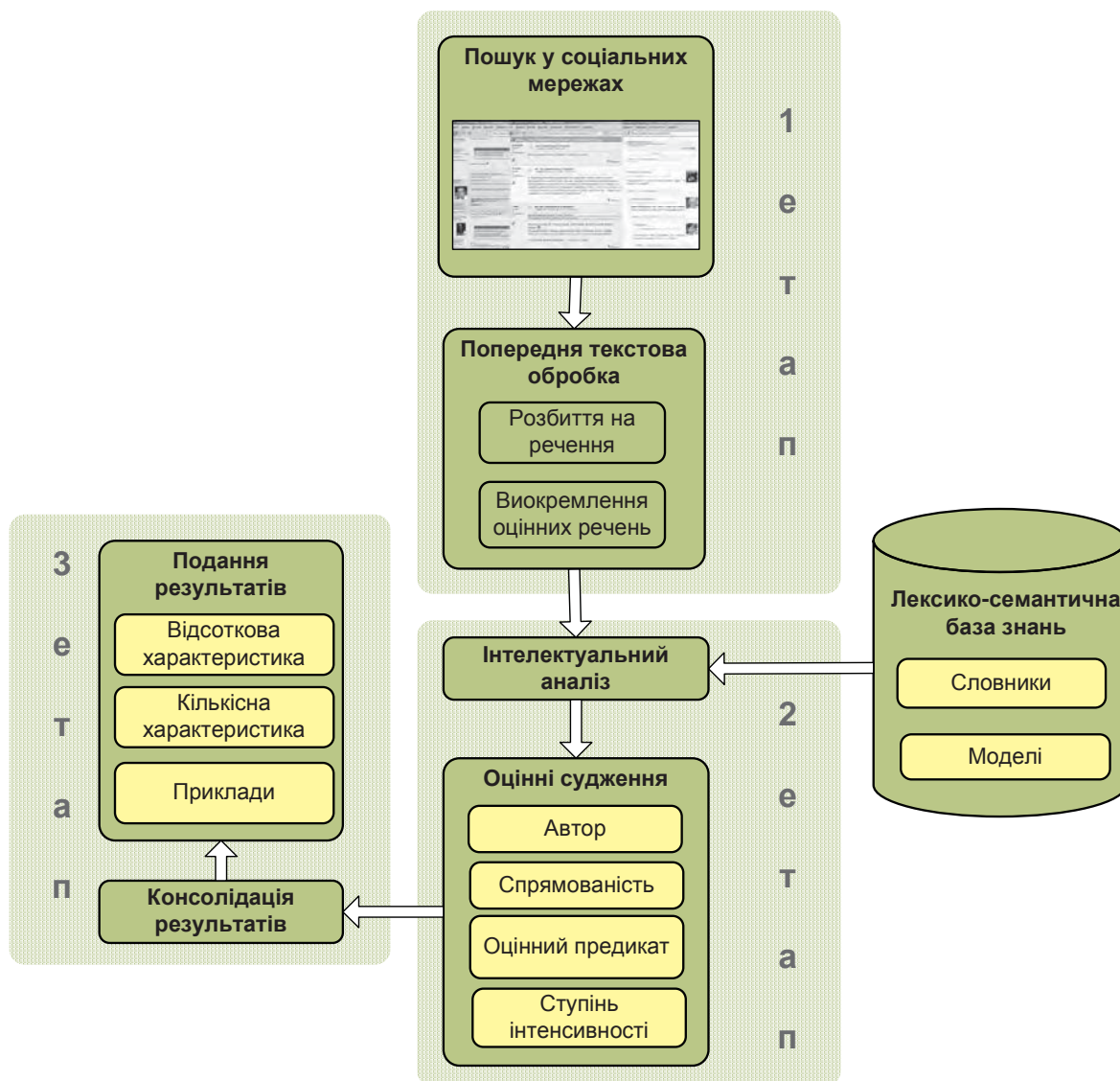


Рис. 3. Схема роботи системи обробки оцінних суджень у соціальних мережах

На першому етапі здійснюється пошук текстів, які містять оцінні висловлювання щодо вибраного об'єкта і попередня обробка текстів, тобто поділ тексту на речення та виділення речень, які містять безпосередні оцінні судження, а також речень, які є оточенням оцінних суджень.

Автоматизований пошук текстів, які містять оцінні висловлювання, у масивах слабо-структурованої інформації, таких як форуми, блоги, сайти оглядів продуктів тощо, відрізняється, наприклад, від пошуку тематичної інформації. Пошук тематичної інформації здійснюється за ключовими словами, до яких входить слово-дескриптор або синонімічні конструкції, які описують об'єкт, а також слова, які уточнюють пошук (наприклад, *фотокамера* або *фотоапарат* – слово-дескриптор, *Canon* – уточнювальне слово). На відміну від фактичної інформації, яка висловлюється прямо, оцінна, тобто суб'єктивна інформація може висловлюватися нечітко чи навіть ухильно, що ускладнює її опис певним ключовим словом чи навіть фразою. Для пошуку суджень до слова-дескриптора, крім уточнювального слова, слід додавати слова, які передбачають отримання суб'єктивної інформації, такі як *поради*, *задоволений*, *не варто*.

Для відбору речень, які містять оцінні судження, кожне речення з тексту перевіряють на наявність лексичних одиниць та синтаксичних структур, які характерні для оцінних суджень, що містяться у базі знань у вигляді словників та наборів моделей, сформованих на основі тренувальних корпусів текстів та ручної вибірки з масиву дописів у соціальних мережах.

Наступний етап передбачає інтелектуальний аналіз відібраних речень з метою виявлення об'єкта та його властивостей, щодо яких висловлюється оцінне судження, полярність висловлювання щодо кожної з властивостей, тобто “*позитивна/негативна*”, суб'єкта висловлювання або автора, якщо це одна і та сама особа, а також час публікації допису, який містить це оцінне висловлювання.

На останньому етапі здійснюється аналіз отриманих результатів, тобто визначення кількісних та відсоткових характеристик позитивно та негативно спрямованих суджень. Такі характеристики визначають як для об'єкта загалом, так і для окремих його властивостей, стосовно яких були ідентифіковані оцінні судження.

Така система дає змогу оцінити рівень задоволеності споживацьких потреб певним продуктом, виявити його переваги та недоліки, а також порівняти однотипні продукти.

Висновки

Поняття оцінки, з одного боку, є універсальною категорією з погляду логіки, оскільки виражає позитивне або негативне ставлення суб'єкта цієї оцінки до змісту висловлювання. Стосовно реалізації в мові, оцінка є соціально обумовленою і контекстно-залежною. Мовне представлення логічної оцінки може бути імпліцитним, що значно ускладнює її обробку. Оцінне висловлювання характеризується особливою структурою, яка містить обов'язкові та додаткові елементи. Запропонована модель оцінного судження з веб-допису містить обов'язкові елементи оцінного висловлювання, такі як об'єкт, суб'єкт та оцінне відношення, а також елементи, зумовлені середовищем оцінного судження (веб-форум), наприклад, час опублікування допису. Система обробки оцінних суджень на веб-форумі дає змогу оцінити об'єкт, як загалом, так і його окремі властивості, його відповідність потребам споживачів, показує рівень зацікавленості споживачів у цьому продукті.

1. Арутюнова, Н.Д. *Типы языковых значений : Оценка. Событие. Факт [Текст]* / Н.Д. Арутюнова. – М. : Наука, 1988. – 336 с.
2. Вольф, Е.М. *Функциональная семантика оценки [Текст]* / Е.М. Вольф. – М. : Наука, 1985. – 232 с.
3. Гуйванюк, Н.В. *Формально-семантические співвідношення в системі синтаксичних одиниць [Текст]* / Н.В. Гуйванюк. – Чернівці : Рута, 1999. – 336 с.
4. Ивин А.А. *Основание логики оценок [Текст]* / А.А. Ивин. – М. : Изд-во Моск. ун-та, 1970. – 230 с.
5. Космеда Т. *Категорія оцінки в українському мовознавстві [Текст]* / Т. Космеда // *Українська філологія: школи, постаті, проблеми: Зб. наук. пр.* – Л., 1999. – Ч. 2. – с. 138–146.
6. *Матеріали веб-форму “Форум Рідного Міста” [Електронний ресурс]* – Режим доступу: <http://misto.ridne.net/index.php>.
7. *Реализация и области применения системы ВААЛ-2000 [Електронний ресурс]* – Режим доступу: <http://www.vaal.ru/proekt/vaal2000.php>.
8. Carenini, G. *Interactive Multimedia Summaries of Evaluative Text [Електронний ресурс]* / G. Carenini, R.Ng, A. Pauls // *Proceedings of the 11th international conference on Intelligent user interfaces.* – New York, NY, USA: ACM, 2006. – P. 124 – 131. – Режим доступу : <http://www.eecs.berkeley.edu/~adpauls/PAPERS/iui06.pdf>.
9. Hu M. *Mining and summarizing customer reviews [Електронний ресурс]* / M. Hu, B. Liu // *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining.* – New York, NY, USA : ACM, 2004. – pp. 168 – 177. – Режим доступу: <http://www.cs.uic.edu/~liub/publications/kdd04-revSummary.pdf>.
10. Kim, S. *Determining the Sentiment of Opinions [Електронний ресурс]* / S. Kim, E. Hovy // *Proceedings of the COLING conference.* – Geneva, Switzerland, 2004. – pp. 1367-1373. – Режим доступу: <http://www.isi.edu/natural-language/people/hovy/papers/04Coling-opinion-valences.pdf>.
11. Kobayashi, N. *Extracting Aspect-Evaluation and Aspect-of Relations in Opinion Mining [Електронний ресурс]* / N. Kobayashi, K. Inui, Y. Matsumoto // *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning.* – Prague, 2007. – pp. 1065–1074. – Режим доступу : <http://www.aclweb.org/anthology/D/D07/D07-1114.pdf>.
12. Ku, L.-W. *Opinion Extraction, Summarization and Tracking in News and Blog Corpora [Електронний ресурс]* / L.-W. Ku, Y.-T. Liang, H.-H. Chen // *Proceedings of AAAI-2006 Spring Symposium on Computational Approaches to Analyzing Weblogs.* AAAI Technical Report. – California: AAAI Press, 2006. – pp. 100-107. – Режим доступу: <http://www.aaai.org/Papers/Symposia/Spring/2006/SS-06-03/SS06-03-020.pdf>.
13. Liu, B. *Sentiment Analysis and Subjectivity [Електронний ресурс]* /

[editors: N. Indurkha and F. J. Damerau] // *Handbook of Natural Language Processing, Second Edition*. – NW: Chapman & Hall/CRC, 2010. – pp. 627-660. – Режим доступа : <http://www.cs.uic.edu/~liub/FBS/NLP-handbook-sentiment-analysis.pdf>. 14. Liu, B. *Web Data Mining: Exploring Hyperlinks, Content and Usage Data* [Текст] / B. Liu. – Berlin: Springer Berlin Heidelberg. – 2009. – 552p. 15. Pang, B. *Thumbs up? Sentiment Classification Using Machine Learning Techniques* [Электронный ресурс] / B. Pang, L. Lee, S. Vaithyanathan // *Proceedings of the ACL-02 conference on Empirical methods in natural language processing*. – Morristown, NJ, USA : Association for Computational Linguistics, 2002. – P.79–86. – Режим доступа: <http://acl.ldc.upenn.edu/acl2002/EMNLP/pdfs/EMNLP219.pdf>. 16. Popescu, A-M. *Extracting Product Features and Opinions from Reviews* [Электронный ресурс] / A-M. Popescu, O. Etzioni // *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*. – Morristown, NJ, USA : Association for Computational Linguistics, 2005. – P.339–346. – Режим доступа : http://turing.cs.washington.edu/papers/emnlp05_opine.pdf. 17. Riloff, E. *Learning extraction patterns for subjective expressions* [Электронный ресурс] / E. Riloff, J. Wiebe // *Theoretical Issues In Natural Language Processing: Proceedings of the 2003 conference on Empirical methods in natural language processing*. – Morristown, NJ, USA: Association for Computational Linguistics, 2003. – P. 105–112. – Режим доступа: <http://www.cs.utah.edu/~riloff/pdfs/emnlp03.pdf>. 18. Titov, I. *A Joint Model of Text and Aspect Ratings for Sentiment Summarization* [Текст] / I. Titov, R. McDonald // *Proceedings of 46th Meeting of Association for Computational Linguistics (ACL-08)*. – Columbus, OH, USA, 2008. – P. 308–316. 19. Turney, P. *Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews* [Электронный ресурс] / P. Turney // *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)*. – Philadelphia, 2002. – P. 417–424. – Режим доступа: <http://acl.ldc.upenn.edu/acl2002/MAIN/pdfs/Main425.pdf>. 20. Von Wright, G.H. *The Varieties of Goodness* [Текст] / G. H. Von Wrigh. – London : Routledge & K. Paul, 1963. – 222 p. 21. Wiebe, J. *Learning to Disambiguate Potentially Subjective Expressions* [Текст] / J. Wiebe, T. Wilson // *International Conference On Computational Linguistics: Proceedings of the 6th conference on Natural language learning*. – Morristown, NJ, USA : Association for Computational Linguistics, 2002. – P. 1–7. 22. Wilson, T. *Just how mad are you? Finding strong and weak opinion clauses* [Текст] / T. Wilson, J. Wiebe, R. Hwa // *AAAI Conference On Artificial Intelligence: Proceedings of the 19th national conference on Artificial intelligence*. – California: AAAI Press / The MIT Press. – 2004. – P. 761–767. 23. Yu, H. *Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences* [Электронный ресурс] / H. Yu, Y. Hatzivassiloglou // *Proceedings of 2003 Conference on Empirical Methods in Natural Language Processing*. – 2003. – P. 129–136. – Режим доступа: <https://pantherfile.uwm.edu/hongyu/www/files/articles/p129-yu.pdf>.