Dorogyy Y.Yu., Nykolyn O.I.
National Technical University of Ukraine
«Kyiv Polytechnic Institute»

COMPLEX ACCESS LEVEL SYSTEM FOR OUT-OF-BAND MANAGEMENT NETWORK

Summary
Project describes creation of system of allocation of access rights for out-of-band management network. It contains analysis of existing solutions and selection of system components. Special attention was paid to the validation of automation functional tests, designing a web interface and preparing documentation.
Keywords: Access level, management network, web server, database, system design, security, IT Security Procedure.

УДК 81'322.4

СЕМАНТИЧНИЙ РОЗБІР РЕЧЕННЯ, ЯК НЕОБХІДНА СКЛАДОВА АЛГОРИТМУ ТРАНСФЕРНОГО ПЕРЕКЛАДУ

Мацаєва Т.В.
Національний технічний університет України
«Київський політехнічний інститут»

Лінгвістична обробка природно-мовних текстів. Основні етапи синтаксичного аналізу природно мовного тексту. Побудова семантичного дерева відношень.
Ключові слова: морфологічний аналіз, синтаксичний аналіз, семантичний розбір, трансферний переклад, семантичне дерево відношень.

Для трансферного перекладу необхідно використовувати семантичний розбір речення, тому це перший крок для розробки алгоритму трансферного перекладу. Лінгвістична обробка природномовних текстів є однією з центра-льних проблем інтелектуалізації інформаційних технологій.

Велика кількість науково–дослідних програм у світі спрямовані на розвиток лінгвістичних інформаційних систем. Значні зусилля науковців спрямовані на розробку математичних алгоритмів та комп'ютерних програм обробки текстів природною мовою. Але більшість досліджень та розробок присвячені розумінню англійської мови, водночас як головною проблемою для нашого географічного регіону є обробка текстів українською мовою. Тому дана робота спрямована на дослідження лінгвістичної обробки саме україномовних текстів.

Речення, в якому проведений семантичний розбір, легко подати у вигляді простої семантичної мережі, яка повністю описуватиме його структуру й за якою легко здійснювати пошук зв'язків з іншими реченнями. Вказана семантична мережа матиме зірчасту структуру, в основі якої лежатиме підмет та присудок речення, а додаткові члени речення (додатки та обставини) являтимуть собою гілки графа.

При обробці текстової інформації та подальшому її зберіганні можна відкидати сполучники, прийменники, частки, вигуки та деякі відмінки займенників, бо ці слова є службовими частинами мови або просто несуть відтінок почуттів і волевиявлення і не несуть інформаційної цінності, а потрібні для зв'язку слів у реченні.

Поєднання двох або більше повнозначних слів на основі підрядного зв'язку утворює синтаксичну одиницю – словосполучення. Будова словосполучень наведена на рис. 1. Види словосполучень наведені в таблиці 1.

Згідно статистики пошукових запитів, наведену статистами Google, 89,9% запитів на пошук в текстових документах складається з словосполучень, що містять у собі іменник, який виступає підметом, потім іде його опис – представлений прикметниками або дієприкметниками, та дієслово, яке виступає ознакою дії цього іменника і являється присудком. Згідно такого висновку статистів можна побудувати чітку схему аналізу тексту прогнозуючи запити.

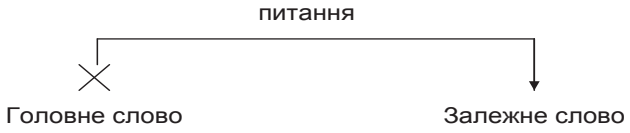

Рис.1. Будова словосполучень

Таблиця 1

Види словосполучень

За будовою	Прості – складаються з двох повнозначних слів, між якими один вид синтаксичних відношень
	Складні – два чи більше компоненти, між якими два або більше видів синтаксичних відношень
За поєднанням компонентів	Вільні – кожний компонент – окремий член речення.
	Невільні – словосполучення – один член речення.
	Синтаксично нерозкладні.
	Лексичні.
За головним словом	Іменні – головне слово – іменник, прикметник, числівник, займенник.
	Дієслівні – головне слово – форми дієслова.
	Прислівникові – головне слово – прислівник.
За способом підрядного зв'язку	Узгодження – залежне слово вживається у тому самому роді, числі й відмінку, що й головне
	Керування – головне слово вимагає від залежного певного відмінка, при зміні головного слова залежне не змінюється
	Прилягання – залежне слово має незмінну форму і приєднується до головного тільки за змістом.

© Мацаєва Т.В., 2014

Іншими словами зв'язування слів у реченні доцільно починати зі словосполучення, що визначає головне відношення (відношення між підметом і присудком) у цьому реченні.

Як відомо речення складається з підмета, присудка, додатка, означення та обставини. Підмет – головний член речення, що називає предмет, про який говориться в реченні; присудок – головний член речення, який означає те, що говориться про підмет; означення – це другорядний член речення, який вказує на ознаку предмета, його якість, властивість, приналежність, а також на кількість і порядок предметів, у реченні означення залежить від іменника або інших слів, ужитих в значенні іменника; обставина – це другорядний член речення, який указує на місце, час, спосіб, мету, умову, причину дії; додаток – це другорядний член речення, що означає предмет, на який спрямована дія.

В українській мові за членами речення не закріплене певне (незмінне) місце. Але зміна порядку слів у реченні іноді призводить до зміни граматичного або стилістичного значення. Тому при побудові речень треба дотримуватися таких правил:

1. Керований іменник від керуючого слова відривати не треба;

2. Не можна розривати однорідні члени речення іншими словами;

3. Не слід міняти місцями слова у словосполученні, якщо це впливає на зміст вислову;

4. У словосполученнях, у яких поєднуються іменники однакого роду, треба уникати поширення їх підрядними частинами

Згідно цих правил можна зробити висновок, що речення доцільно поділяти на словосполучення різних типів, для визначення головного словосполучення (відношення між підметом і присудком) та другорядних словосполучень (відношення між предметом та його ознакою, дією й об'єктом, на який спрямована дія та дією та її ознакою).

Розглянемо основні етапи роботи блоку синтаксичного аналізу природно мовного тексту. Нехай на вхід аналізатора поступає текстовий документ українською мовою, тоді алгоритм обробки документа буде мати вигляд:

Крок 1. Аналізується весь текст та розділяється на окремі речення. Це відбувається по такому принципу: шукаються усі крапки – якщо крапка знайдена, то це означає кінець речення і початок нового.

Крок 2. Відбувається морфологічний аналіз: речення розбиваються на окремі незалежні слова.

Крок 3. В кожному реченні відкидаються вигуки, сполучники, частки та прийменники – це стоп-слова, які не несуть смислового навантаження і тому не будуть заноситись в базу.

Крок 4. Виділяються усі займенники, що знаходяться у реченні. Це робиться для вирішення задачі анафоричних зв'язків, або заміни займенників у тексті на відповідні поняття (іменники). Заміщенню підлягають деякі особові, відносні, вказівні, присвійні та зворотні займенники.

Крок 5. Визначаються усі відмінки відповідних слів речення, для правильного виділення зв'язку між ними.

Крок 6. Виконується синтаксичний аналіз речення. Згідно відмінків слів речення поділяються на словосполучення типу: підмет-присудок, іменник-прикметник(и) (дієприкметник), дієслово-прислівник(и) (дієприслівник). З цих словосполучень у майбутньому будуть побудовані дерева семантичних відношень. Для визначення залежності між іменниками і прикметниками (дієприкметниками), що їх характеризують, використовується положення, що вони повинні бути узгодженні по відмінку та знаходитись поруч. Теж саме робиться з дієсловами і прислівниками (дієприслівниками), які характеризують опис дієслова. Іменник у називному відмінку пов'язується з дієсловом, це зв'язок підмета і присудка. Числівники, що стоять поруч з іменниками вважаються їх характеристикою, тому приписуються у дерево відносин так само як прикметники. Якщо в реченні немає іменника у називному відмінку, а є займенник, то відбувається звернення до пункту 3 цього алгоритму, для визначення іменника. Якщо ж немає іменника та займенника, шукаються прикметники (дієприкметники) або числівники, після цього можна скористатись алгоритмом пункту 3, але врахувати що у нас замість займенника є прикметник чи числівник. В такому випадку будуть записуватись у словосполучення та у дерево не прикметник чи числівник, а знайдений іменник, а також його характеристики у іншому словосполученні. Якщо речення, що аналізується – складне, то згідно алгоритму будуть знайдені усі підмети і присудки, та для кожного такого ядра будуть побудовані синтаксичні дерева відношень. В реченні в якому немає підмета, а тільки присудок, словосполучення буде складатись лише з одного слова, а дерево семантичних відносин у корені буде мати лише дієслово.

Крок 7. Проводиться морфологічний аналіз слова, згідно якого слово ділиться по будові на основу та закінчення.

Крок 8. Відбувається першочергове збереження лексем речень без зв'язків, яке можна зобразити за допомогою діаграми класів, представленої на рис. 2. Як бачимо, на діаграмі визначено 5 класів, які характеризують слова з україномовного тексту:

– Клас «Основа слова» буде містити усі основи слів.

– Клас «Тип слова» містить частину мови слів.

– Клас «Відмінок слова» міститиме усі можливі відмінки слів.

– Клас «Закінчення слова» містить відрізані закінчення слів.

– Клас «З'єднання слова» фактично містить відомості про слова, які колись оброблялись.

Таким чином будується словник семантичного забезпечення системи обробки тексту.

Крок 9. Відбувається першочергове семантичне дроблення. Для цього будується синтаксичне дерево відношень онтологічної мережі, яке показує онтологію зв'язків основ речення. В якості концеп-

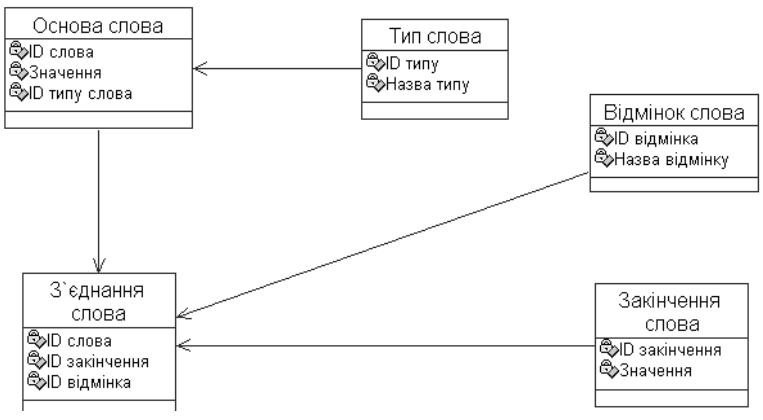

Рис. 2. Діаграма класів першочергового збереження інформації

тів онтології виступають такі поняття: об'єкт, процес та характеристика. В якості об'єкту можуть виступати лише іменники, в якості процесу – дієслова, а в якості характеристики – прикметник, дієприкметник, прислівник та дієприслівник Зв'язок між такими концептами визначається як опис головного концепту – об'єкту. На рис. 3 графічно зображено першочергове семантичне дерево відношень концептів, в якому є повна характеристика головного концепту згідно кроку 5 цього алгоритму. Якщо при синтаксичному аналізі не визначено головним концептом об'єкт, або він відсутній, то автоматично буде визначатись інша лексична одиниця як головний концепт та записуватись окремим деревом відношень.

Крок 10. Для визначення кінцевої смислової структури тексту необхідно визначити семантичний зв'язок об'єктів кожного речення. Для цього відношення об'єктів кожного речення представляється алгебраїчною системою виду:

$$M=<O, A, L, H, R_1, R_2, R_H>,$$

де O – множина головних концептів виділених в реченні; A – множина ребер, що зв'язують концепти із O; L – множина семантичних відношень, виявлених в реченні, що використовуються в якості міток ребер із A, H – множина класів, що зв'язують концепти із O, R_1 – відношення інцидентності на , де N – підмножина ідентифікаторів учасників відношення моделі M, R_2 – відношення інцидентності на AxL, R_H – відношення класової приналежності на OxH.

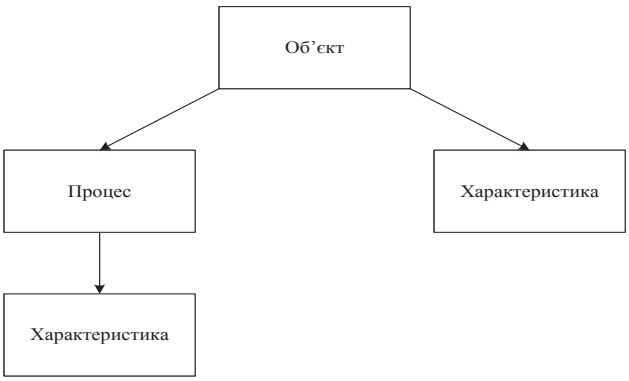

Рис. 3. Першочергове семантичне дерево відношень

Таку структура описує граф, у якого вершинами являються об'єкти, а будь-яке ребро помічене семантичним відношенням і з'єднує ті вершини-об'єкти, які знаходяться один з одним в деякому відношенні. Ідентифікатори учасників записуються в вигляді міток при кожній вершині. Приблизний вигляд графу семантичних відношень об'єктів представлений на рис. 4.

Відношення класової приналежності (ребра графа) визначається згідно шаблонів побудови словосполучень в числовому еквіваленті. Кожному шаблону призначається відповідна вага словосполучення у реченні. Вагу словосполучення визначає відношення об'єктів між собою, а саме синтаксична роль об'єктів у реченні. Зокрема більшу вагу отримують словосполучення в яких участує іменник, який виступає у реченні підметом, а потім в вже – в ролі прямого чи непрямого доповнення і т.п. Вага різник видів відношень об'єктів у словосполученні зазначена в таблиці 2.

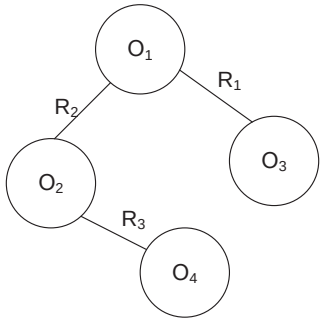

Рис. 4. Приклад графу семантичних відношень для чотирьох концептів

Таблиця 2

Вага типових шаблонів словосполучень

Правило поєднання об'єктів	Вага
Об'єкт підмет + об'єкт місце	10
Об'єкт підмет + об'єкт причина	9
Об'єкт підмет + об'єкт часова приналежність	8
Об'єкт підмет + об'єкт умови	8
Об'єкт підмет + об'єкт допустимості	7
Об'єкт підмет + об'єкт міри і ступеня	7
Об'єкт підмет + об'єкт приналежність	10
Об'єкт + об'єкт місце	6
Об'єкт + об'єкт причина	5
Об'єкт + об'єкт часова приналежність	4
Об'єкт + об'єкт умови	4
Об'єкт + об'єкт допустимості	3
Об'єкт + об'єкт міри і ступеня	1
Об'єкт + об'єкт приналежність	6

Вага у таблиці 2 вказана у відносних одиницях. При подальшому аналізі відповідності тексту пошуковим запитам кінцева вага для словосполучень, в яких присутнє більше двох об'єктів може бути отримана простим сумуванням ваги тих простих конструкцій зазначених в таблиці 2, на які можна буде розбити загальний запит.

Крок 11. Формується база даних таблиці відношень слів. В неї записується тільки основи лексем, без закінчень. Відкидані закінчення будуть також зберігатись згідно пункту 8 цього алгоритму, бо коли у майбутньому буде вводитись новий пошуковий запит, то спершу в базі слів будуть знаходитись усі слова, визначатися їх основа та аналізуватися усі варіанти словоформ пошукового запиту. Після чого будується дерево синтаксичного відношення запиту для подальшого порівняння його з існуючими відношеннями в онтологічній мережі бази знань.

Для трансферного перекладу, необхідно знати відповідність між словами з українського словника та німецького(англійського) словника. Для цього і потрібно робити семантичний аналіз. За допомогою нього можна сформувати правила, для трансферного перекладу.

Список літератури:

1. Analysis, Statistical Transfer, and Synthesis in Machine Translation / Brown P. – К.: Proceedings of TMI-92, 1992.
2. Statistical Machine Translation by Parsing / Melamed I.D. – К.: Proceedings of the ACL. Barcelona, 2004.
3. Ярмолюк А.Э. Трансферная модель машинного перевода основанная на пу-тях. 2013 – 3-rd Conference «European Applied Sciences: modern approaches in scientific researches».

Мацаева Т.В.
Национальный технический университет Украины
«Киевский политехнический институт»

СЕМАНТИЧЕСКИЙ РАЗБОР ПРЕДЛОЖЕНИЯ,КАК НЕОБХОДИМАЯ СОСТАВЛЯЮЩАЯ АЛГОРИТМА ТРАНСФЕРНОГО ПЕРЕВОДА

Аннотация
Лингвистическая обработка естественно-языковых текстов. Основные этапы синтаксического анализа. Построение семантического дерева отношений.
Ключевые слова: морфологический анализ, синтаксический анализ, семантический разбор, трансферный перевод, семантическое дерево отношений.

Matsaeva T.V.
National Technical University of Ukraine
«Kyiv Polytechnic Institute»

SEMANTIC ANALYSIS AS A NECESSARY COMPONENT OF TRANSFER TRANSLATION ALGORITHM

Summary
Linguistic processing of natural language texts. Main stages of the syntactic analysis. Building a semantic tree relations.
Keywords: morphological analysis, syntactic analysis, semantic analysis, transfer translation, semantic relations tree.

УДК 534.84

БИНАУРАЛЬНЫЙ СЛУХ И АНАЛИЗ АКУСТИЧЕСКИХ СВОЙСТВ ПОМЕЩЕНИЙ

Москалец М.П., Богданов А.В.
Национальный технический университет Украины
«Киевский политехнический институт»

Исследованы вопросы бинаурального слуха человека. Выполнен анализ свойств бинауральной слуховой системы человека. Описана работа модели бинаурального модуля, имитирующего слуховую систему человека и приведен пример анализа бинауральных трехмерных карт. Создана модель слуховой системы человека.
Ключевые слова: локализационная модель, кросс-корреляция, временная задержка прибытия звука, интенсивностная разница, трехмерный спектр.

Проектирование помещений с хорошей акустикой является в настоящее время скорее искусством на научной основе, чем наукой, причем число специалистов, владеющих этим искусством, в мире очень ограничено. Искусство создания залов с хорошей естественной акустикой в значительной степени утрачено, достаточно сравнить качество звучания музыки в залах XVIII-XIX веков с современными концертными комплексами.

Начиная с конца XIX века предпринимаются значительные усилия по поиску объективных параметров, позволяющих адекватно оценивать акустические свойства залов. Однако, несмотря на значительный прогресс в исследованиях, однозначной связи между перечисленными выше параметрами и оценкой качества звучания музыки и речи в различных залах окончательно установить не удается.

Наличие двух приемников слуха обеспечивают человеку возможность воспринимать пространственный звуковой мир и оценивать перемещение звуковых сигналов в пространстве. Возникает вопрос, почему не использовать данное свойство человека для анализа акустических свойств помещений?

Для построения модели бинауральной слуховой системы человека необходимо изучить ее свойства.

К числу основных свойств бинаурального слуха можно отнести: пространственную локализацию, эффект предшествования, бинауральное суммирование громкости, бинауральные биения и слияние звуков.

Пространственная разнесенность ушных раковин, и экранирующее влияние головы и торса за счет дифракционных эффектов приводит к значительным различиям между сигналами, поступающими в правое и левое ухо, что позволяет произвести локализацию звукового источника в пространстве, обусловленную тремя физическими факторами:

а) временным (ITD);
б) интенсивностным (ILD);
в) спектральным.

Несмотря на то, что в обычных условиях в оба уха звуки поступают с определенным различием во времени, по интенсивности и спектру, мы воспринимаем один слуховой образ. Механизм бинаурального слияния звуков описан в виде математической модели, которая основывается на поиске центральной слуховой нервной системой перекрестных корреляций между звуковыми сигналами в обоих ушах. Другими словами, звуки, поступающие в уши, рассматриваются как статистические события, а механизм бинаурального слияния использует поиск общности между ними.

В общем виде эффект предшествования заключается в том, что в пределах определенного отрезка времени ранее поступивший звуковой сигнал (фронт звуковой волны) доминирует в слуховом восприятии над звуками, поступившими позднее (эхо).

© Москалец М.П., Богданов А.В., 2014