

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
"КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ім. ІГОРЯ СІКОРСЬКОГО"

Природномовна комунікація з прикладною програмою

Виконав: студент 6-го курсу ТЕФ, гр. ТІ-61м,
Бугайов Сергій Володимирович

Науковий керівник: к.т.н. доц. Стативка Ю. І.

Київ – 2018

Актуальність

Сучасні природномовні діалогові системи є обмеженими в розумінні природної мови, зокрема української мови. Тому задача створення системи підтримки природномовного діалогу українською мовою з прикладною програмою є актуальною.

Мета дослідження

Мета дослідження полягає в створенні алгоритмічного та програмного забезпечення для автоматизованої обробки природної писемної мови з ціллю підтримки природномовного діалогу з прикладною програмою.

Об'єкт: Програмне забезпечення систем автоматичної обробки природної мови.

Предмет: Програмне забезпечення систем підтримки природномовного діалогу.

Задачі дослідження

- проаналізувати існуючі програмні засоби, алгоритми та методи реалізації діалогових систем;
- розробити модель діалогу та удосконалити алгоритм класифікації реплік користувача;
- створити програмне забезпечення для підтримки природномовного діалогу з прикладною програмою.

Наукова новизна

Удосконалено алгоритм класифікації реплік користувача за рахунок використання:

- граматичних тегів;
- тегів іменованих сутностей.

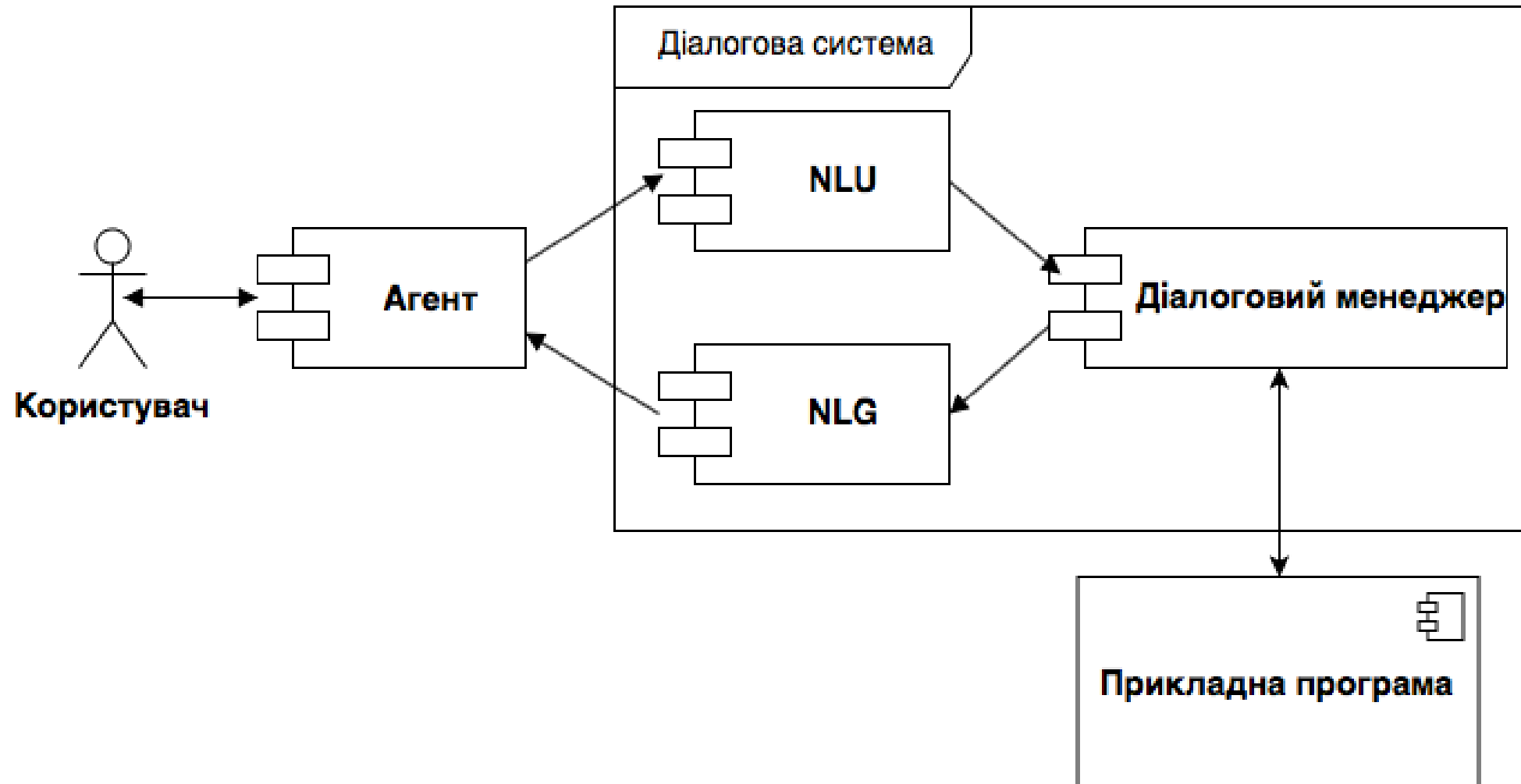
Результат:

- більш висока точність класифікації реплік користувача;
- покриття більшої кількості перифраз репліки.

Діалогова система

- обробляє запит/репліку природною мовою;
- генерує осмислену відповідь;
- намагається імітувати мовну поведінку людини.

Загальна архітектура діалогових систем



NLU (Natural Language Understanding) – модуль розуміння природної мови

NLG (Natural Language Generator) – модуль генерації природної мови

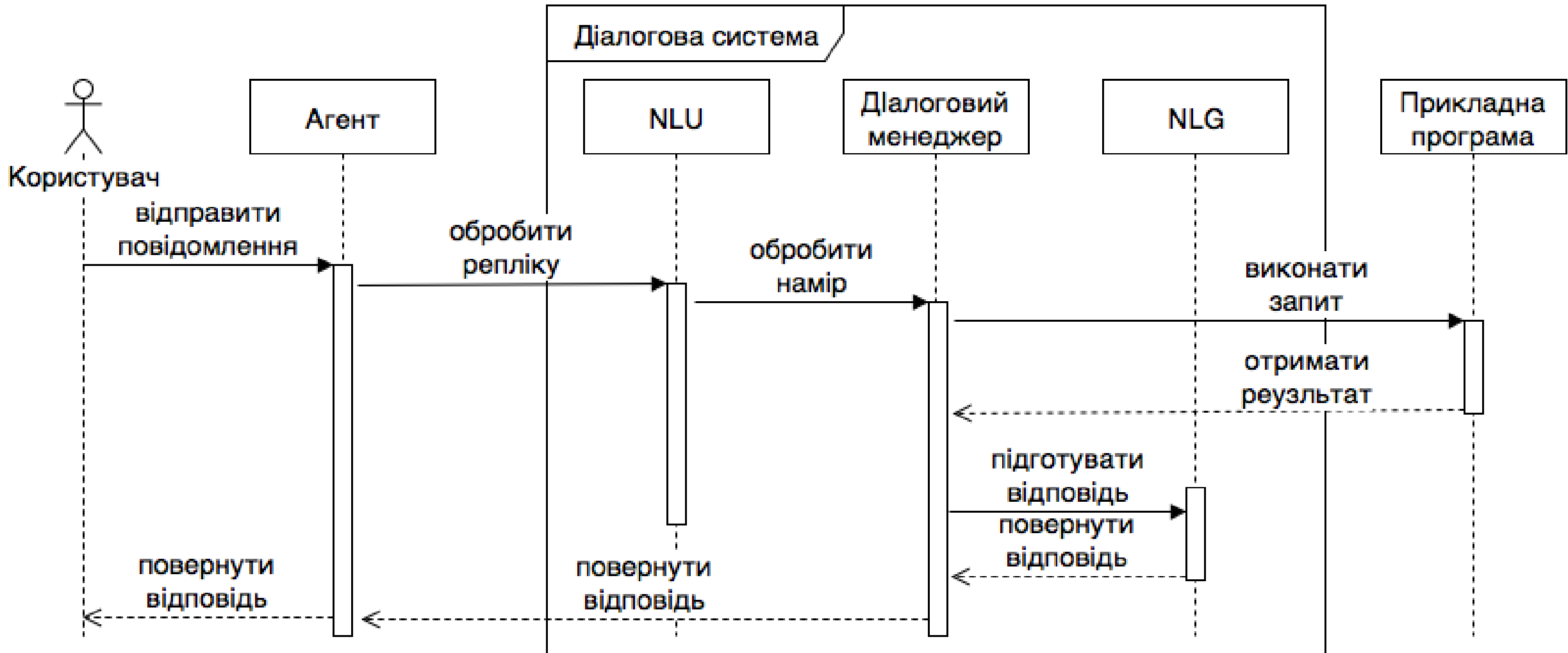
Модель діалогу

Намір користувача		
Шаблони реплік користувача		
Параметри	Питання	Шаблони відповідей
Запит до прикладної програми		
Відповіді системи		

Приклад наміру користувача

Дізнатися погоду		
Яка погода в LOCATION ?		
LOCATION	Де саме Вас цікавить погода?	В LOCATION
Зробити запит до програми прогнозу погоди в LOCATION		
В LOCATION максимальна температура X градусів.		

Процес обробки репліки користувача

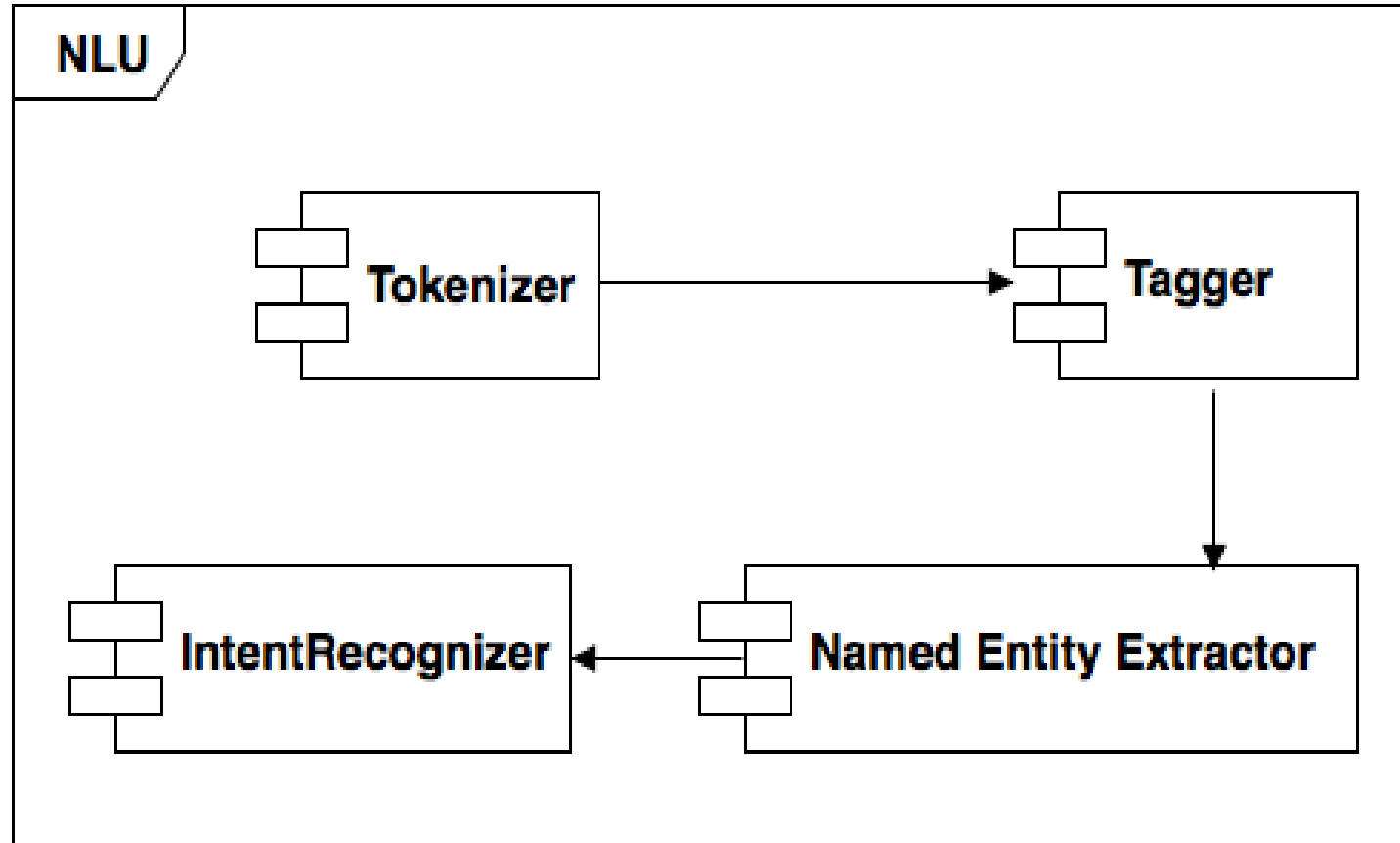


Модуль розуміння природної мови

Задачі:

- токенізація;
- морфологічний аналіз: частиновомне тегування;
- розпізнавання іменованих сутностей;
- класифікація реплік користувача.

Модуль розуміння природної мови



Tokenizer – токенизація;

Tagger – частининомвне тегування;

Named Entity Extractor – розпізнавання іменованих сутностей;

Intent Recognizer – класифікація репліки коритувача.

Модуль розуміння природної мови

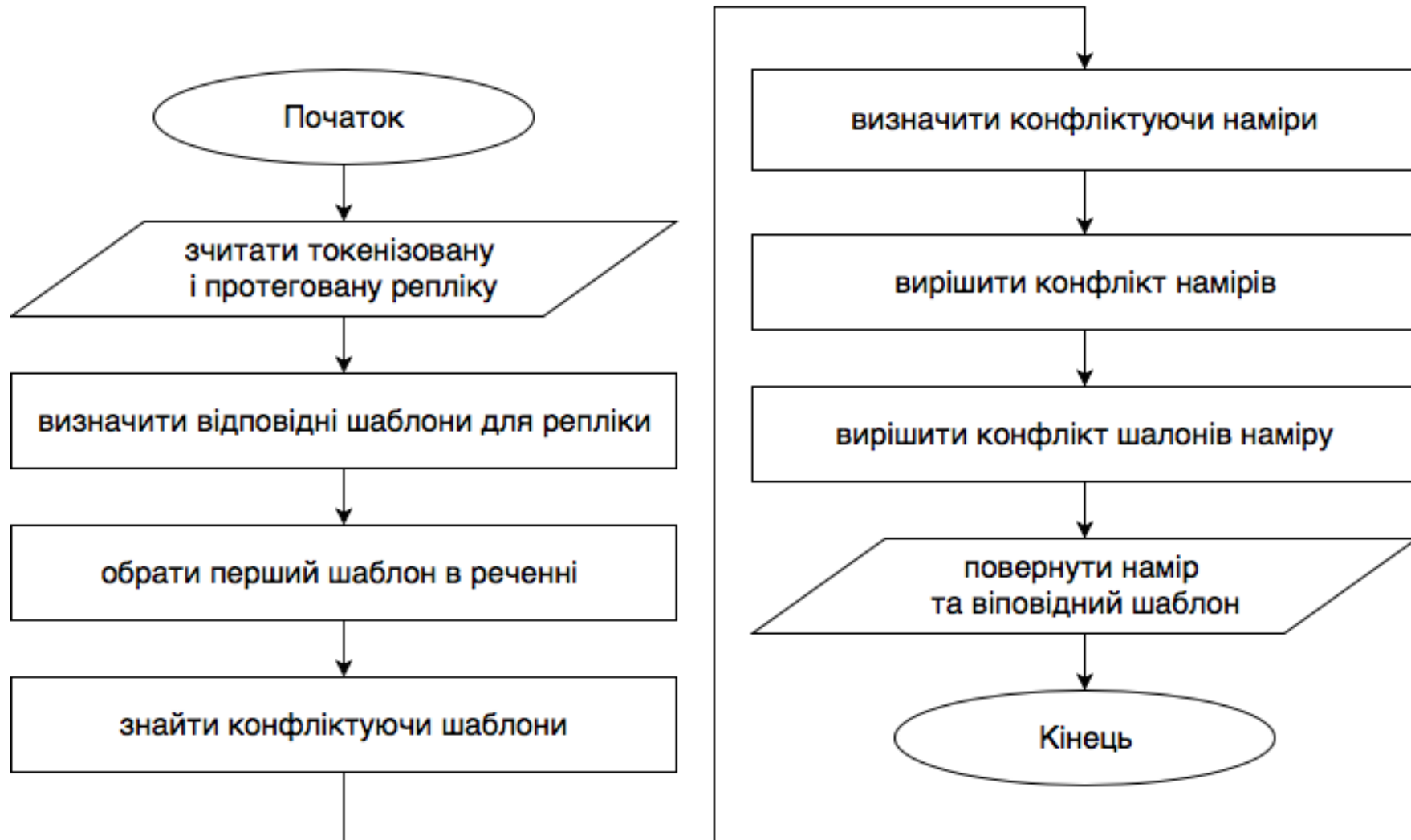
Приклад обробки репліки користувача

Репліка: Яка погода в Києві?

Токени	Теги частин мови	Іменовані сутності
який	&pron:int:rel:def (питальний займенник)	
погода	noun:inanim:f:v_naz (іменник неістота)	
в	prep (прийменник)	
Київ	noun:inanim:m:v_mis:pro р (іменник неістота, власна назва)	LOCATION (локація)
?		

Модуль розуміння природної мови

Модифікований алгоритм класифікації реплік користувача



Модуль розуміння природної мови

Приклад морфологічного шаблону

Лема	Тег частини мови	Тег іменованої сутності
погода	noun:inanim	
в	prep	
	noun:inanim	LOCATION

```
("погода",  
 ( "в",  
 ( null,  
   asList("noun:inanim"),  
   asList("prep"),  
   asList("noun:inanim"),  
   emptyList()),  
   emptyList()),  
   asList(LOCATION))
```

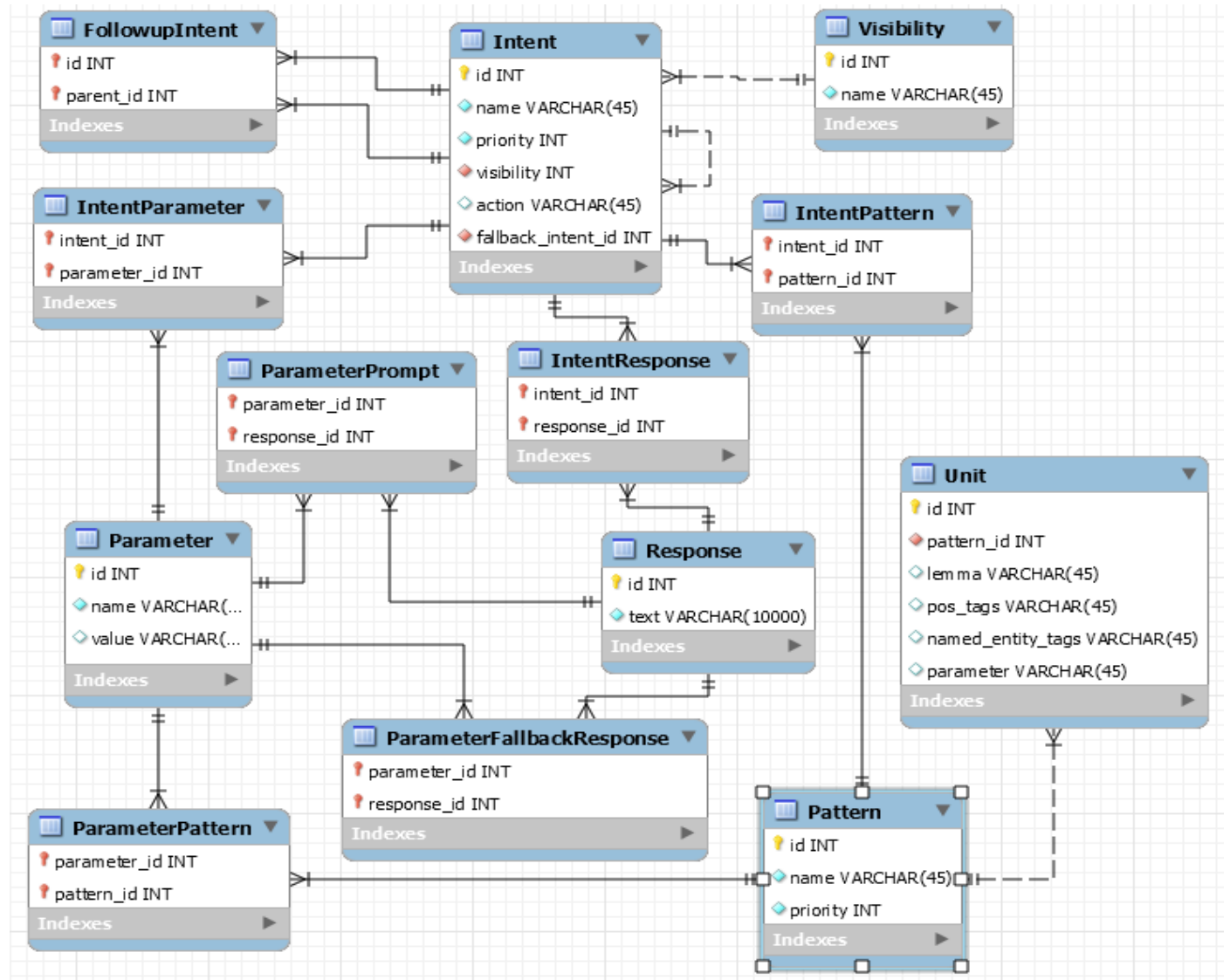
Діалоговий менеджер

- змінює стан діалогу відповідно до наміру користувача;
- реалізований за фреймовим підходом;
- підтримує вкладені діалоги;
- підтримує змішану ініціативу.

Стан діалогу складається з репліки користувача, наміру, параметру та відповідей системи.

Діалоговий менеджер

Схема бази даних



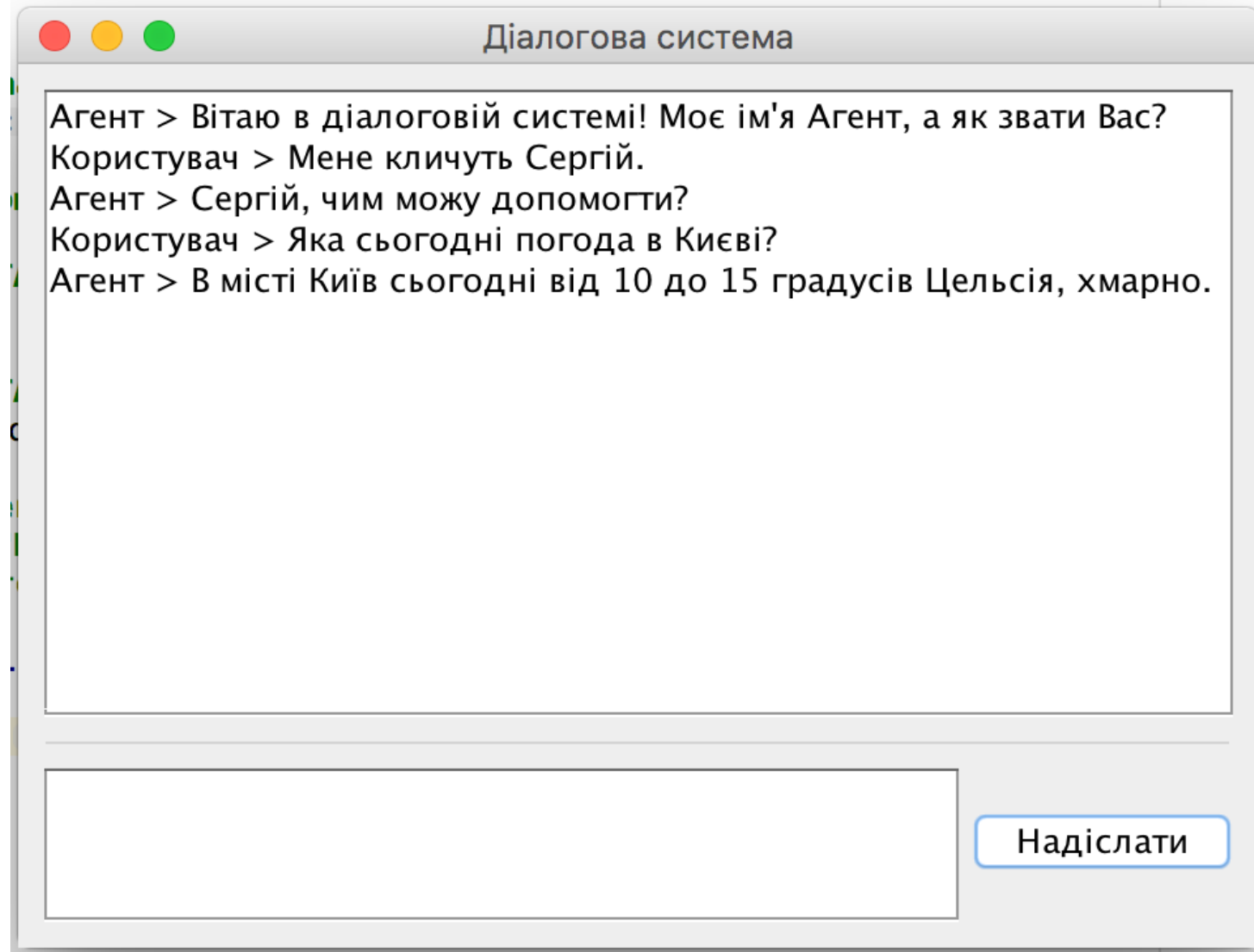
Модуль генерації природної мови

Приклад роботи

В місті {@Location} сьогодні від {@Min} до {@Max} градусів Цельсія.

В місті Київ сьогодні від 15 до 20 градусів Цельсія.

Приклад роботи програми



Висновки

- проаналізовано існуючі системи, методи створення діалогових систем, алгоритми обробки природної писемної мови;
- удосконалено алгоритм класифікації реплік користувача за рахунок використання морфологічних шаблонів та розроблено модель діалогу;
- розроблено програмне забезпечення для підтримки природномовного діалогу з прикладною програмою;
- результати дослідження прийняті до впровадження підприємством ТОВ "Netcracker".

Дякую за увагу!