

Вінницький національний технічний університет

Факультет комп'ютерних систем та автоматики

Кафедра комп'ютерних систем управління

Дипломний проект

На тему: **Синтез граматики голосового інтерфейсу
багатомовного програмного продукту**

Виконала: ст.гр. ЗКСУА-15сп

Зарванська А.І.

Керівник: к.т.н., доц. кафедри КСУ

Грищук Т.В.

Вінниця, ВНТУ – 2016 року

Мета: підвищення ефективності процесу розробки голосового інтерфейсу.

Об'єкт дослідження: голосовий інтерфейс.

Предмет дослідження: методи синтезу граматики голосового інтерфейсу.

Задачі ДП:

- ✓ провести аналіз стандартів голосового інтерфейсу;
- ✓ розробити новий метод синтезу граматики голосового інтерфейсу;
- ✓ розробити алгоритм формування глосарію системи;
- ✓ розробити структуру системи;
- ✓ розробити програмне забезпечення, що реалізує роботу даного методу;
- ✓ розрахувати економічну ефективність від впровадження нової системи автоматизації розробки граматики голосового інтерфейсу.

Постановка задачі методу генерування правил КВ-граматики

Вхідні дані: для лінгвістичного опису голосової команди формується набір фраз природною мовою.

*{open a document,
open file,
open document,
open a file}*

Обмеження: максимальна кількість слів голосової команди, ≤ 10 ; слова голосової команди не повторюються в межах одного варіанту; кожне слово голосової команди міститься в глосарії нетермінальних символів граматики; якщо слово відсутнє в глосарії, то потрібно його додати в автоматизованому режимі.

Глосарій (фрагмент):

OpenV -> "open";

A -> "a";

DocN -> "document";

FileN -> "file";

Вихідні дані: на основі набору фраз команди та глосарію системи синтезувати лінгвістичне правило виведення фраз команди, що складається тільки з нетермінальних символів та символів операцій.

OpenV [A] (DocN | FileN).

Синтез правила грамматики на основі матриці суміжності

Крок 1. Формуємо матрицю суміжності на основі фраз голосової команди.

Крок 2. Сортиємо рядки за кількістю одиниць.

	<u>OpenV</u>	A	<u>DocN</u>	<u>FileN</u>
<u>OpenV</u>	0	1	1	1
A	0	0	1	1
<u>DocN</u>	0	0	0	0
<u>FileN</u>	0	0	0	0

Крок 3. Знаходимо однакові стовбці матриці, знаходячи паралельні вершини.

Крок 4. Об'єднуємо паралельні вершини в одну за допомогою операції АБО.

	<u>OpenV</u>	A	<u>(DocN FileN)</u>
<u>OpenV</u>	0	1	1
A	0	0	1
<u>(DocN FileN)</u>	0	0	0

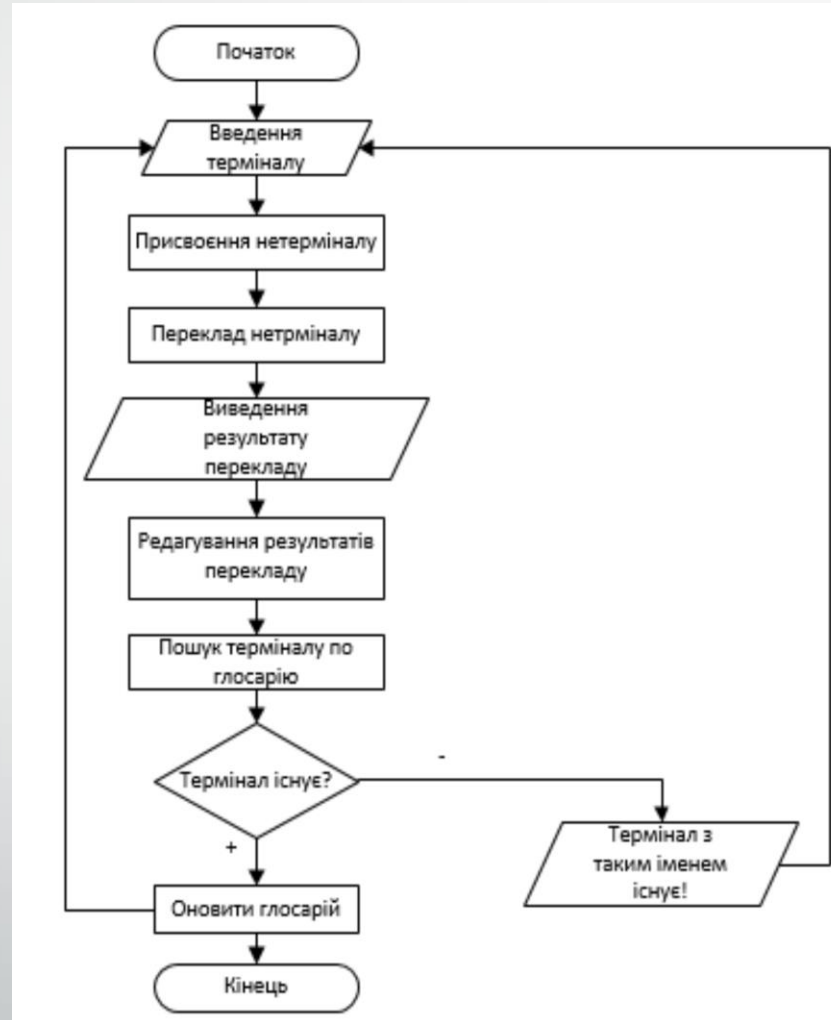
Крок 5. Знаходимо рядки, між якими є лише одна відмінність. Позначаємо цю вершину як опціональну.

	<u>OpenV</u>	<u>[A]</u>	<u>(DocN FileN)</u>
<u>OpenV</u>	0	1	1
<u>[A]</u>	0	0	1
<u>(DocN FileN)</u>	0	0	0

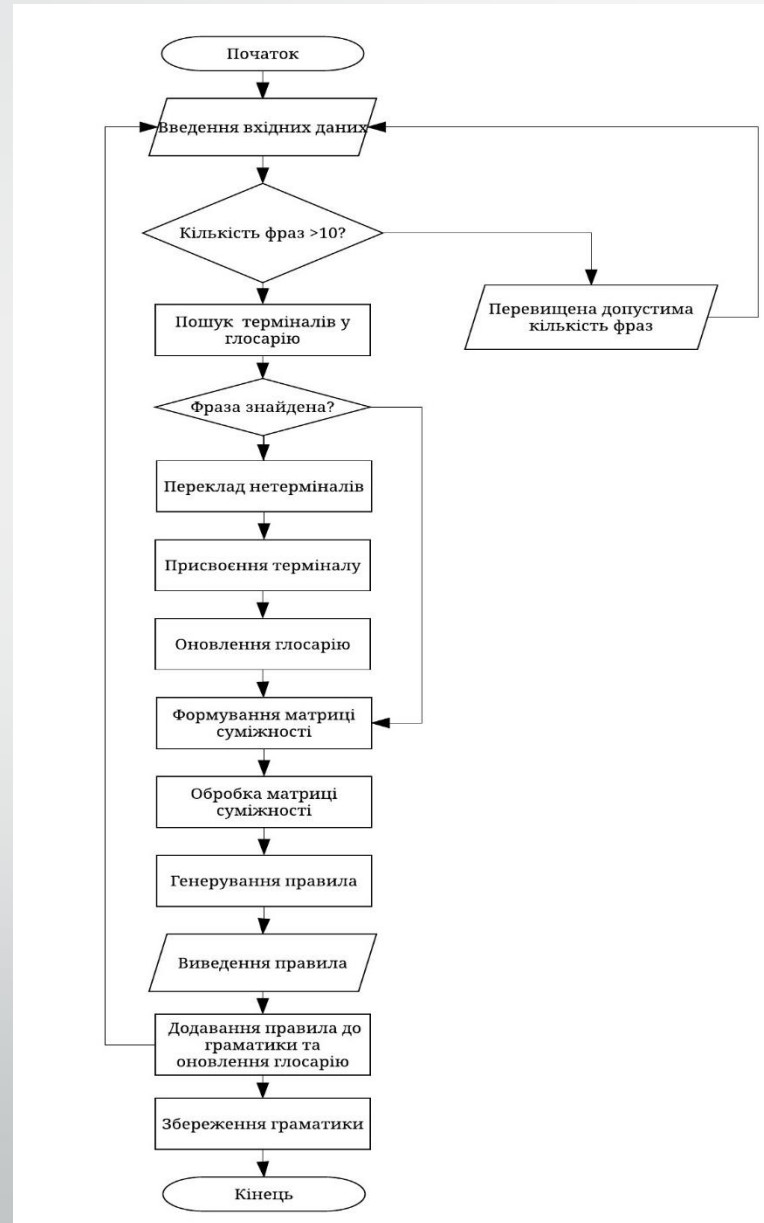
Таким чином, в результаті обробки матриці суміжності отримано правило грамматики, що співпадає з виведеним у прикладі:

OpenV [A] (DocN | FileN).

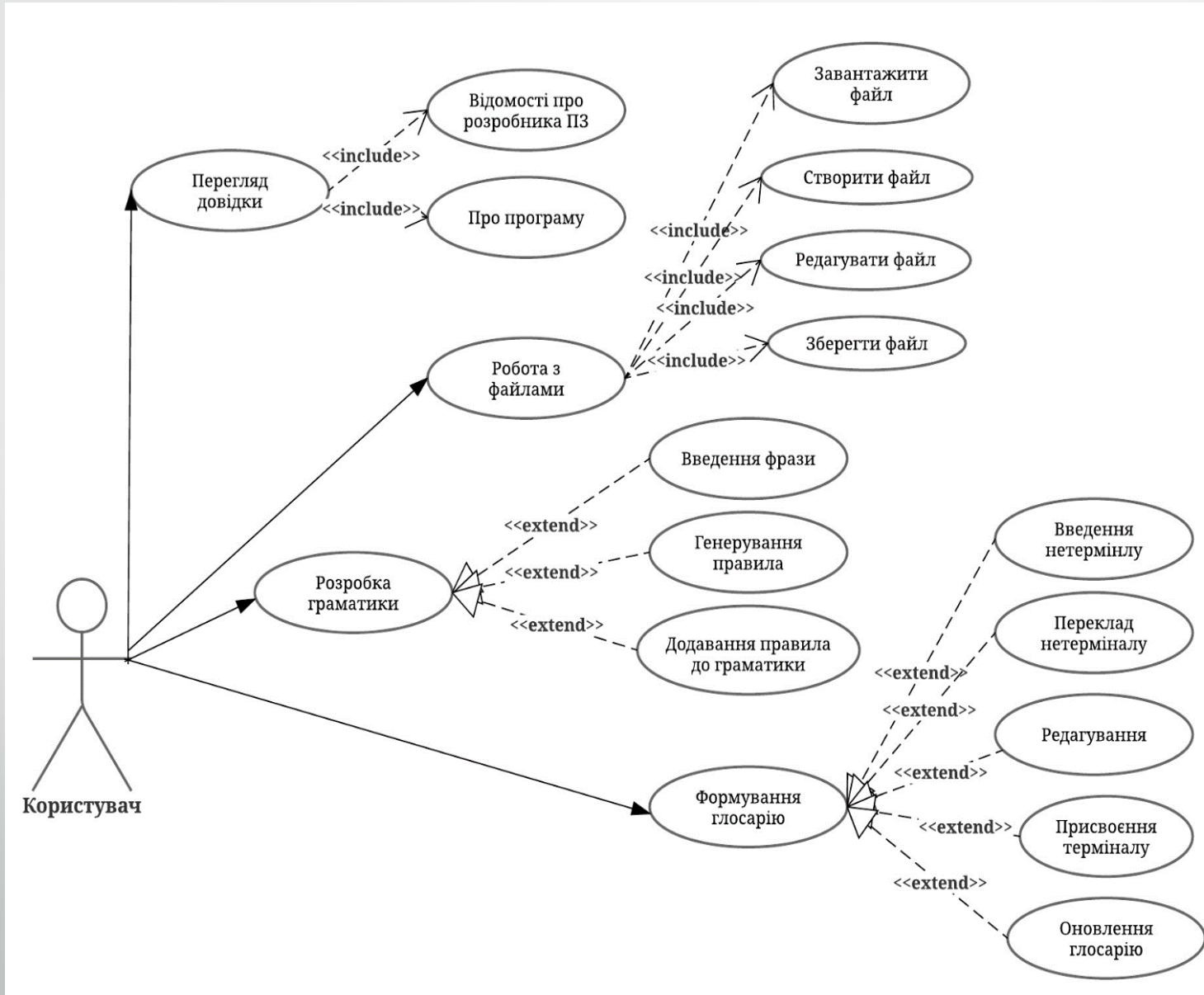
Блок-схема алгоритму формування глосарію



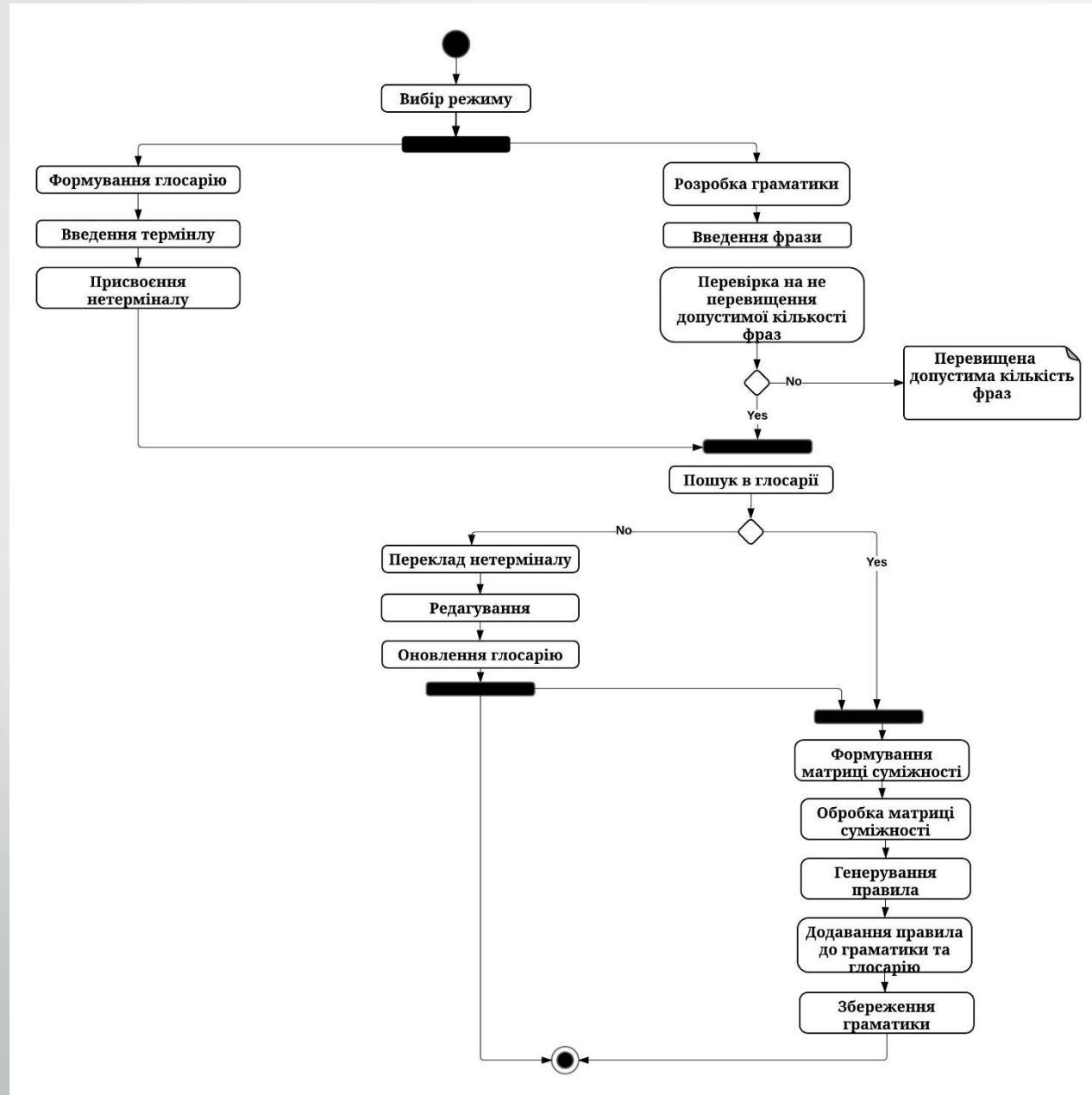
Блок-схема алгоритму синтезу правил КВ-граматики



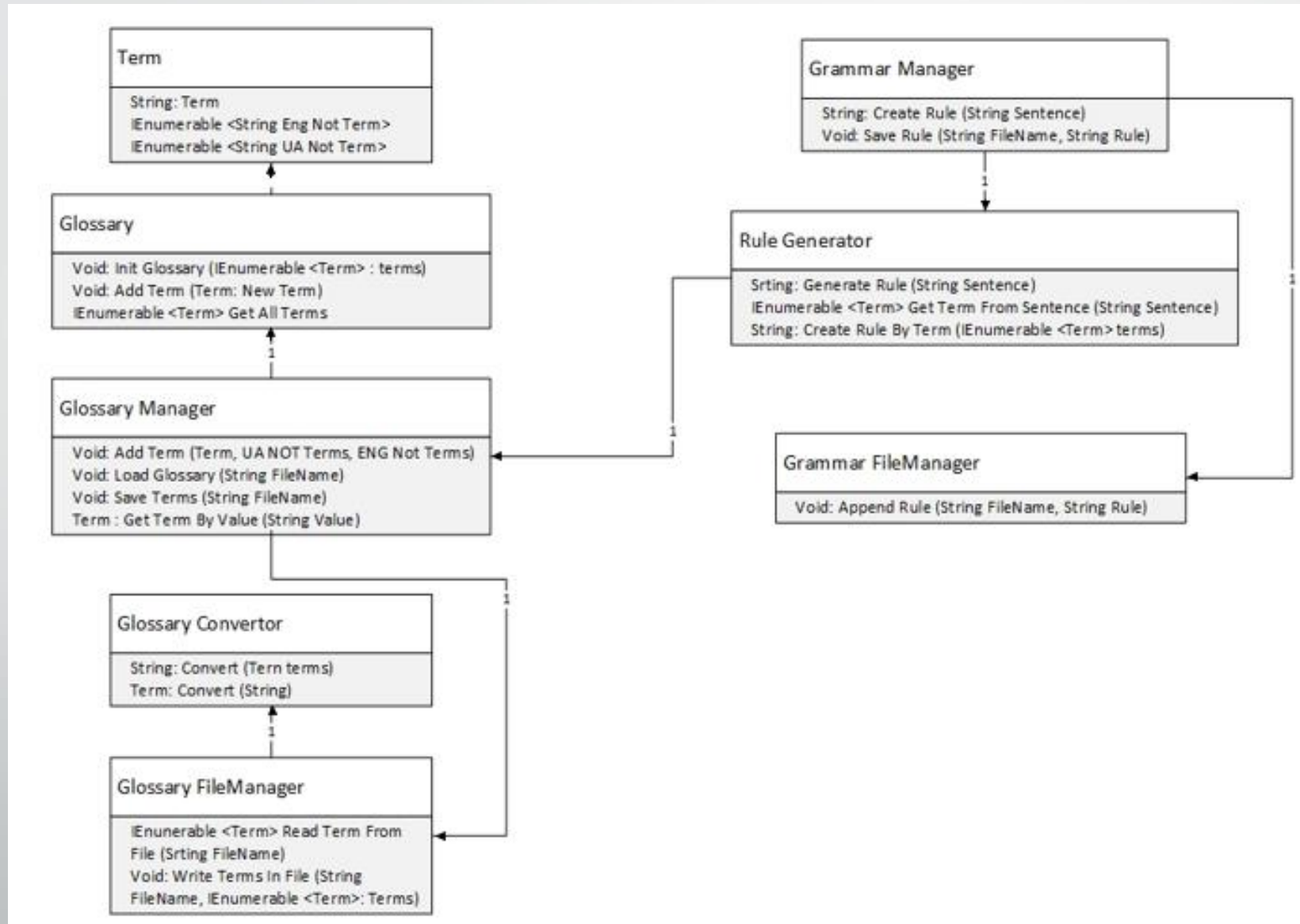
UML-діаграма варіантів використання



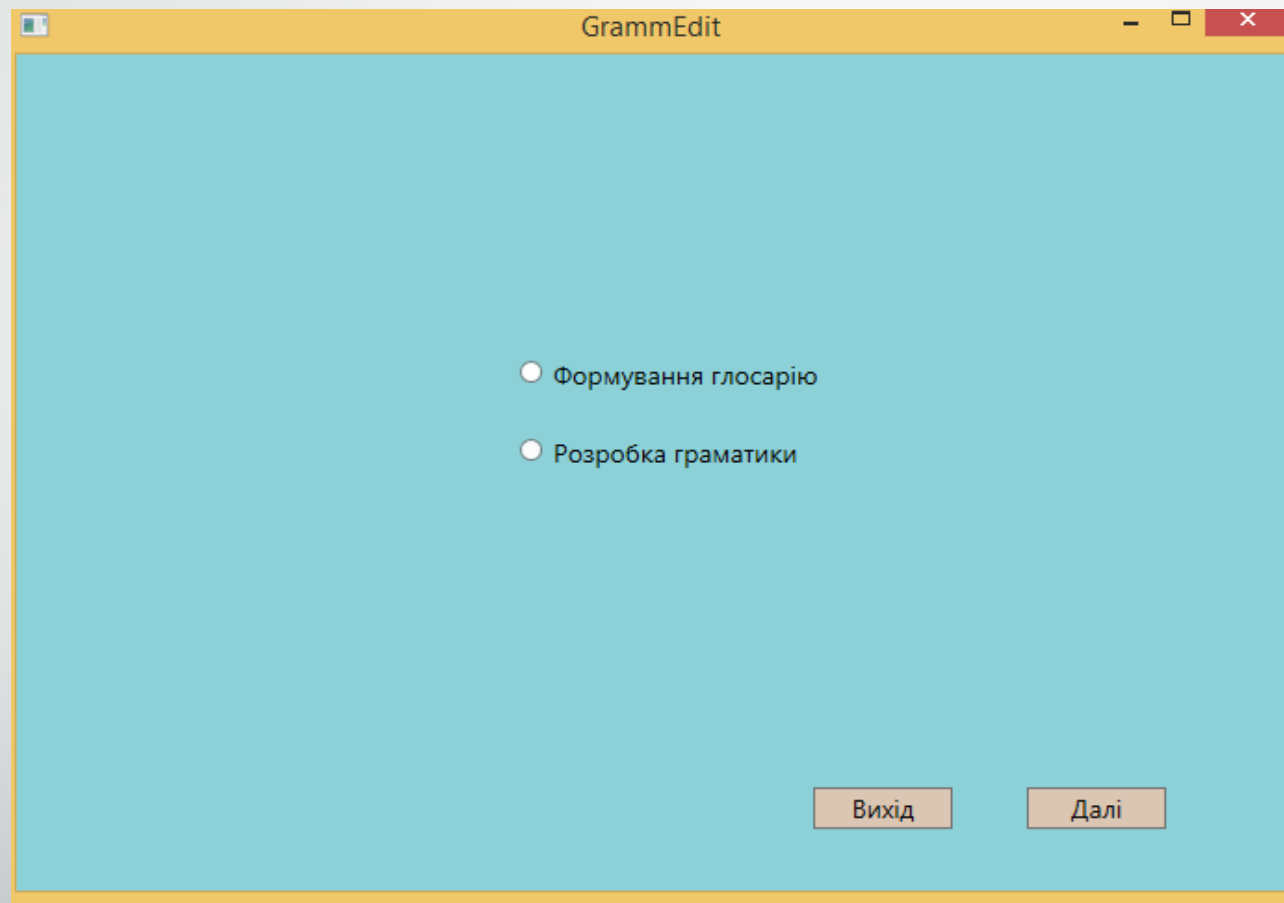
UML-діаграма діяльності



UML-діаграма класів



Головне вікно програми



Фрагмент сформованого глосарію

```
глосарій англ. - Notepad
File Edit Format View Help
A --> "a";
AlignObjectNP --> Pronoun | SelectionNP;
AlignV --> "align";
Aligned --> "aligned";
AlignedRLC --> CenterA | (Right | Left) Aligned;
And --> "and";
AppColors --> Black | Maroon | Green | Olive | Navy | Purple | Teal | Gray | Silver | Red | Lime | Yellow | Blue |
Fuchsia | Aqua | White;
AppName --> "WordPad";
Aqua --> "aqua";
Bigger --> "bigger" | "larger";
Black --> "black";
Blue --> "blue";
Bold --> "bold" | "bolded" | "boldface";
BoldItalic --> "bold" ("italic" | "italics");
BoldV --> "bold" | "boldface";
BulletStyle --> Bulleted | "bullet" "style";
BulletV --> "bullet";
Bulleted --> "bulleted";
Caps --> "cap" | "caps" | "capitals";
Center --> "center" | "centre";
CenterA --> "centered" | "center" "aligned" | "centred" | "centre" "aligned";
CenterV --> "center" | "centre";
CloseExitV --> "close" | "exit" | "quit";
ColorN --> "color" | "colour";
```

Фрагмент розробленої граматики

```
Cmd - Notepad
File Edit Format View Help
UndoThatCmd --> UndoV Pronoun;

UndoMultCmd --> UndoV [The] (LastPrev Num_2_9 CommandsN |
[LastPrev] CommandN) |
UndoV Pronoun (Num_2_9 TimesN | Twice);

UnFontStyleCmd --> UnFontStyleV ObjectNP |
SetV (ObjectNP |
FontStyleNP) To UnFontStyles |
MakeV Pronoun UnFontStyles;

SystemCmds --> CloseAppCmd |
UndoThatCmd |
UndoMultCmd;

SetFontNameAndSizeCmd --> SetChangeV FontNP To FontCombo |
SetChangeV ObjectNP To FontCombo |
MakeV Pronoun FontCombo;

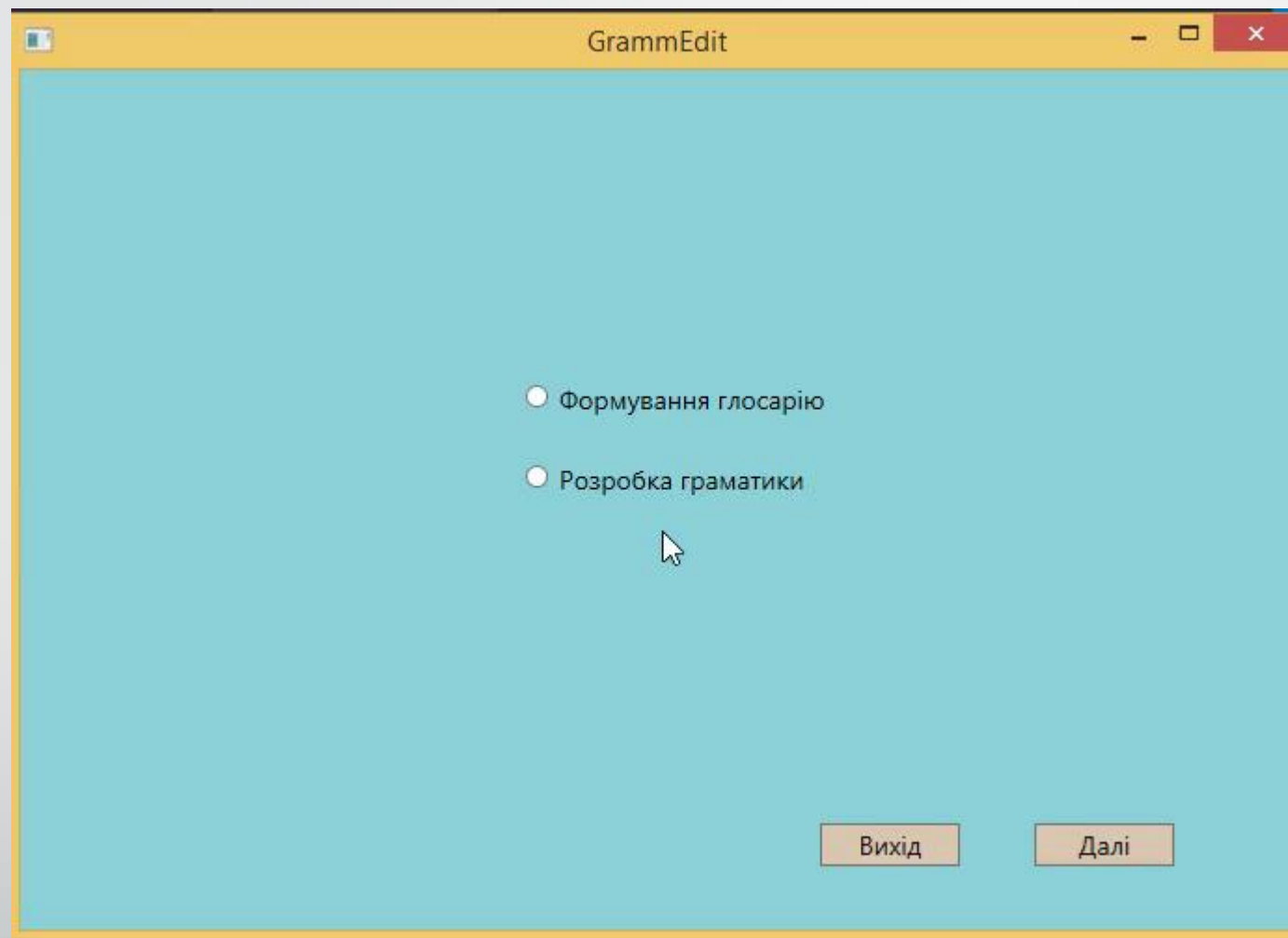
SetFontNameCmd --> SetChangeV FontNP To FontName |
SetChangeV ObjectNP To FontName |
MakeV Pronoun FontName;

SetFontSizeCmd --> SetChangeV FontSizeNP To NumPointsNP |
SetChangeV ObjectNP To NumPointsNP |
MakeV Pronoun NumPointsNP;

IncreaseFontSizeCmd --> EnlargeV FontSizeNP |
MakeV Pronoun Bigger;

FormattingStyleCmd --> FormattingStyleV ObjectNP |
SetV ObjectNP To FormattingStyles |
MakeV Pronoun FormattingStyles;

FontStyleCmd --> FontStyleV ObjectNP |
SetV (ObjectNP | FontStyleNP) To FontStyles |
```



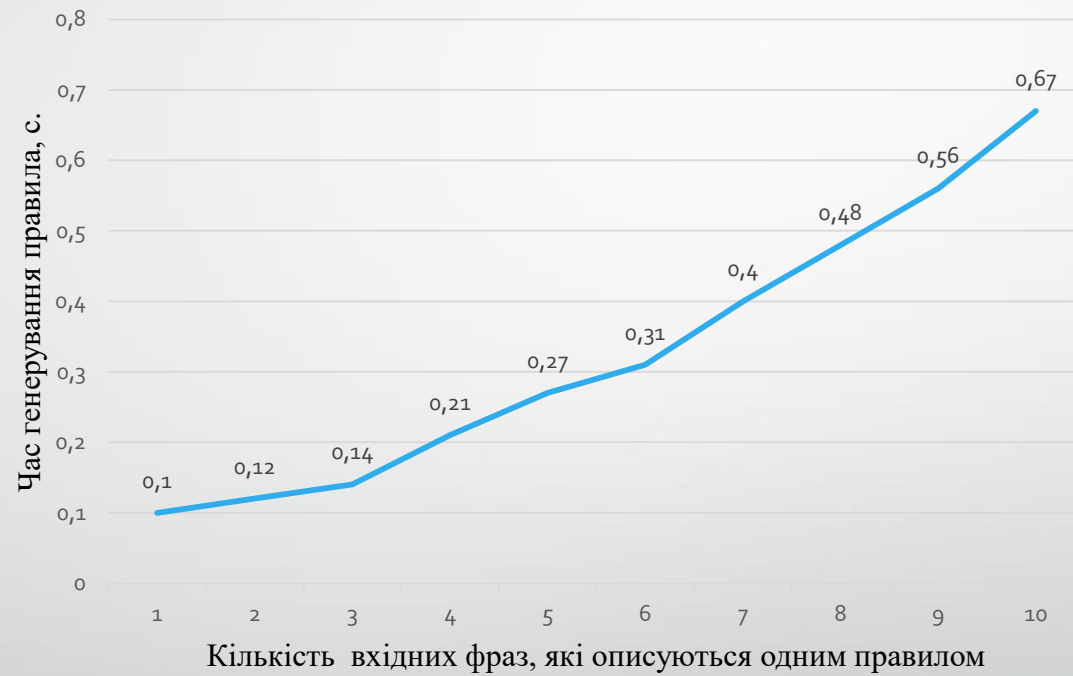
GrammEdit

- Формування глосарію
- Розробка граматики

Вихід

Далі

Аналіз швидкодії ПЗ від кількості фраз, які описуються одним правилом



Показники економічної ефективності

Показники	Задані у ТЗ	Досягнуті у дипломному проекті
1. Витрати на розробку	Не більше 20 тис. грн	15 тис. грн.
2. Коефіцієнт використання результатів роботи	2	2
3. Коефіцієнт результативності роботи	3	3
4. Комплексний показник рівня виконаної роботи	не менше 1,5	1,8



Дякую за увагу!