

# Система для розв'язування задач з геометрії

Доповідач: Смиш О. Р.  
Керівник: Жежерун О. П.

## Мета дослідження

Аналіз наявних рішень обробки текстів українською мовою та огляд можливостей використання такого рішення для створення системи з розв'язування задач.

## Завдання роботи

Створення системи, що дає змогу розв'язувати задачі, які подано природною українською мовою.

# Natural Language Processing

- ◆ Сегментація тексту
- ◆ Розмічування частин мови
- ◆ Лематизація
- ◆ Видобування інформації

## Задачі відкритого типу з короткою відповіддю

У трикутнику  $ABC$  відомо, що  $\angle A = 30^\circ$ ,  $\angle B = 45^\circ$ ,  $CM$  — висота,  $AC = 10$  см. Знайдіть відрізок  $BM$ .

# Аналіз задач з геометрії

№	Слово	Повторюваність
1	дорівнювати	137
2	см	119
3	трикутник	95
4	кут	82
5	знайти	80
6	градус	54
7	i	45
8	сторона	44
9	висота	32
10	y	29
11	abc	27
12	основа	26
13	що	22
14	прямокутний	20

у трикутник ABC відомо, що кут A дорівнювати 30 градус, кут B дорівнювати 45 градус, CM — висота, AC дорівнювати 10 см. знайти відрізок BM.

## Аналіз задач з геометрії

У раўнабедраным трохвугольніку адна старана роўна 5 см, другая — 10 см. Знайдзіце перыметр трохвугольніка.

В равнобедренном треугольнике ABC с основанием AC проведена биссектриса AD. Найдите углы этого треугольника, если  $\angle ADB = 110^\circ$ .

Triangle ABC has side lengths of  $AB = 10$ ,  $BC = 24$ , and  $AC = 26$ . Find the three angles of the triangle.



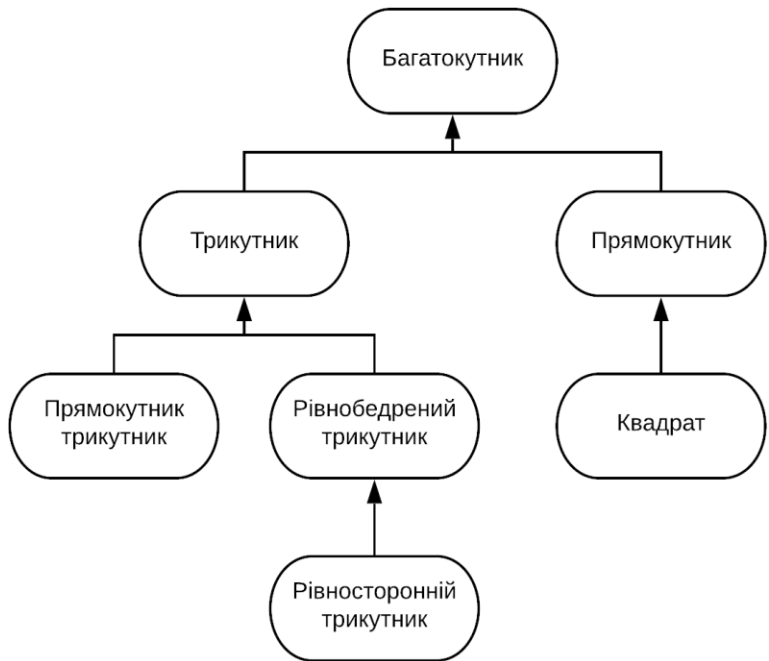
# Аналіз готових рішень

Великий  
електронний  
словник  
української мови  
(**ВЕСУМ**)

Морфологічний  
аналізатор та  
генератор для  
української та  
російської мов  
**Руморphy2**

**UDPipe**, навчена  
на золотому  
стандарті

# Побудова класів та методів для опису планіметрії



# Морфосинтаксовий аналізатор в задачах планіметрії

- ◆ Попереднє опрацювання тексту
- ◆ Числа з комою
- ◆ Задачі зі словом “дорівнює”
- ◆ Задачі зі словом “проведений”
- ◆ Опрацювання шуканого в тексті

# Вивід задачі аналізатором тексту

[1, 'У', 'у', 'ADP', '\_', 'Case=Loc', '2', 'case', '\_', '\_']  
[2, 'трикутнику', 'трикутник', 'NOUN', '\_', 'Animacy=Inan|Case=Loc|Gender=Masc|Number=Sing', '4', 'obl', '\_', '\_']  
[3, 'ABC', 'ABC', 'X', '\_', 'Foreign=Yes', '4', 'nsubj', '\_', '\_']  
[4, 'відомо', 'відомо', 'ADV', '\_', 'Degree=Pos', '0', 'root', '\_', 'SpaceAfter=No']  
[5, '...', 'PUNCT', '\_', '\_', '9', 'punct', '\_', '\_']  
[6, 'що', 'що', 'SCONJ', '\_', '\_', '9', 'mark', '\_', '\_']  
[7, 'кут', 'кут', 'NOUN', '\_', 'Animacy=Inan|Case=Nom|Gender=Masc|Number=Sing', '9', 'nsubj', '\_', '\_']  
[8, 'А', 'А', 'X', '\_', 'Foreign=Yes', '7', 'nmod', '\_', '\_']  
[9, 'дорівнює', 'дорівнювати', 'VERB', '\_', 'Aspect=Imp|Mood=Ind|Number=Sing|Person=3|Tense=Pres|VerbForm=Fin', '4', 'ccomp', '\_', '\_']  
[10, '30', '30', 'NUM', '\_', 'Case=Acc|NumType=Card|Uninflect=Yes', '11', 'nummod:gov', '\_', '\_']  
[11, 'градусів', 'градус', 'NOUN', '\_', 'Animacy=Inan|Case=Gen|Gender=Masc|Number=Plur', '9', 'obj', '\_', 'SpaceAfter=No']  
[12, '...', 'PUNCT', '\_', '\_', '15', 'punct', '\_', '\_']  
[13, 'кут', 'кут', 'NOUN', '\_', 'Animacy=Inan|Case=Nom|Gender=Masc|Number=Sing', '15', 'nsubj', '\_', '\_']  
[14, 'В', 'В', 'X', '\_', 'Foreign=Yes', '13', 'nmod', '\_', '\_']  
[15, 'дорівнює', 'дорівнювати', 'VERB', '\_', 'Aspect=Imp|Mood=Ind|Number=Sing|Person=3|Tense=Pres|VerbForm=Fin', '9', 'parataxis', '\_', '\_']  
[16, '45', '45', 'NUM', '\_', 'Case=Acc|NumType=Card|Uninflect=Yes', '17', 'nummod:gov', '\_', '\_']  
[17, 'градусів', 'градус', 'NOUN', '\_', 'Animacy=Inan|Case=Gen|Gender=Masc|Number=Plur', '15', 'obj', '\_', 'SpaceAfter=No']  
[18, '...', 'PUNCT', '\_', '\_', '21', 'punct', '\_', '\_']  
[19, 'СМ', 'СМ', 'X', '\_', 'Foreign=Yes', '21', 'nsubj', '\_', '\_']  
[20, '—', '—', 'PUNCT', '\_', 'PunctType=Dash', '21', 'punct', '\_', '\_']  
[21, 'висота', 'висота', 'NOUN', '\_', 'Animacy=Inan|Case=Nom|Gender=Fem|Number=Sing', '15', 'parataxis', '\_', 'SpaceAfter=No']  
[22, '...', 'PUNCT', '\_', '\_', '24', 'punct', '\_', '\_']  
[23, 'АС', 'АС', 'X', '\_', 'Foreign=Yes', '24', 'nsubj', '\_', '\_']  
[24, 'дорівнює', 'дорівнювати', 'VERB', '\_', 'Aspect=Imp|Mood=Ind|Number=Sing|Person=3|Tense=Pres|VerbForm=Fin', '21', 'parataxis', '\_', '\_']  
[25, '10', '10', 'NUM', '\_', 'Case=Acc|NumType=Card|Uninflect=Yes', '26', 'nummod:gov', '\_', '\_']  
[26, 'см', 'см', 'NOUN', '\_', 'Abbr=Yes|Animacy=Inan|Case=Gen|Gender=Masc|Number=Plur|Uninflect=Yes', '24', 'obj', '\_', 'SpaceAfter=No']  
[27, '...', 'PUNCT', '\_', '\_', '24', 'punct', '\_', '\_']

# Зразок розв'язку задачі

У трикутнику ABC відомо, що  $\angle A = 30^\circ$ ,  $\angle B = 45^\circ$ , CM — висота, AC = 10 см. Знайдіть відрізок BM.

## Трикутник

Назва - ABC

## Сторони:

AB = 13.66 - ----

BC = 7.07 - ----

CA = 10.0 - ----

## Кути:

ABC (B) = 45.0 - , сторони: AB, BC; ----

BCA (C) = 105.0 - , сторони: BC, CA; ----

CAB (A) = 30.0 - , сторони: CA, AB; ----

## Висота:

CM = 5.0

## Трикутники, утворені від висоти:

Назва - BCM

## Сторони:

BC = 7.07 - гіпотенуза ----

CM = 5.0 - катет ----

MB = 5.0 - катет ----

## Кути:

BCM (C) = 45.0 - гострий, сторони: BC, CM; ----

CMB (M) = 90 - прямий, сторони: CM, MB; ----

MBC (B) = 45.0 - гострий, сторони: MB, BC; ---- ABC

Назва - CAM

## Сторони:

CA = 10.0 - гіпотенуза ----

AM = 8.66 - катет ----

MC = 5.0 - катет ----

## Кути:

CAM (A) = 30.0 - гострий, сторони: CA, AM; ---- CAB

AMC (M) = 90 - прямий, сторони: AM, MC; ----

MCA (C) = 60.0 - гострий, сторони: MC, CA; ----

['side', 'name']

['відрізок', 'BM']

## Відповідь:

MB = 5.0

# Висновки

- ◆ З'ясовані сучасні можливості та тенденції в галузях обробки природної мови.
- ◆ Виконано аналіз тексту задач з геометрії українською і не тільки та їхньої структури.
- ◆ Зроблено критичний огляд наявних на сьогодні рішень з обробки природної української мови.
- ◆ Створено систему для розв'язування задач з геометрії на основі створеної логіки, що описує предмет з використанням морфосинтаксового аналізатора.

**Дякую за увагу!**