

# Дипломная работа



АЛГОРИТМЫ ПРЕДСЕМАНТИЧЕСКИХ ТЕКСТОВЫХ  
ПРОЦЕССОРОВ СИСТЕМЫ АВТОМАТИЗИРОВАННОГО  
ПОСТРОЕНИЯ ОНТОЛОГИЙ

Середа А.А.  
НТУУ «КПИ», ИПСА  
гр. ДА-62

# Существующие работы



- Каталог лингвистических ресурсов в сети:  
([www.rvb.ru](http://www.rvb.ru))
- Работы Старостина С.А. – проект StarLing  
([starling.rinet.ru](http://starling.rinet.ru))
- Работа группы «МикроСистемы»  
([www.analyst.ru](http://www.analyst.ru))
- Проект «ДИАЛИНГ» группы АОТ ([www.aot.ru](http://www.aot.ru))

# Задачи графематического анализа

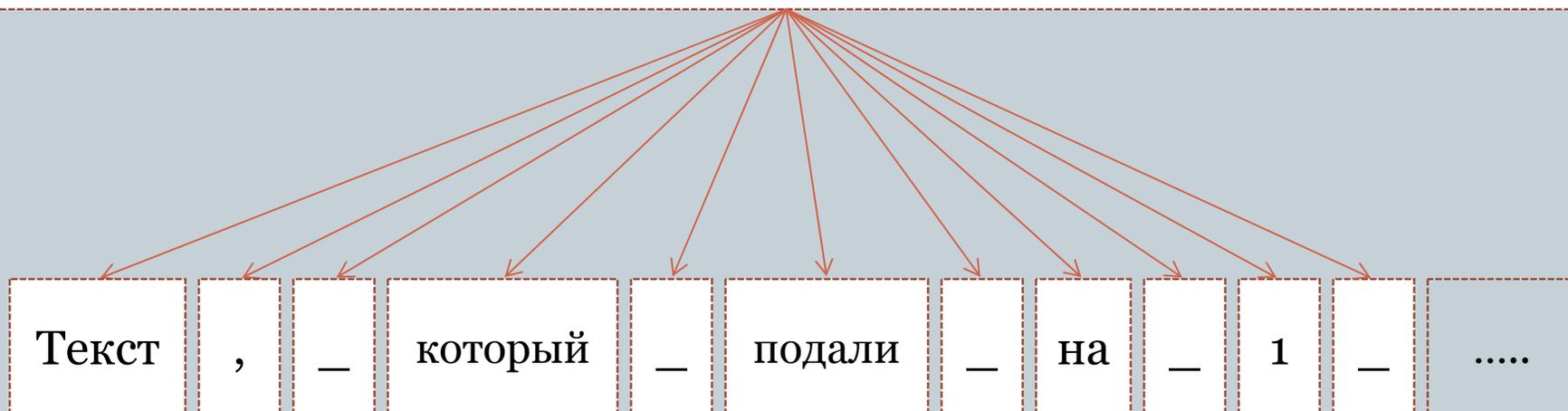


- Разделение входного текста на слова, разделители и т.д.
- Сборка слов, написанных в разрядку
- Выделение:
  - Устойчивых оборотов
  - ФИО
  - Электронных адресов и имен файлов
  - Предложений

# Результат работы ГА



Текст, который подали на 1 этап анализа текста – ГА.



# Результат работы ГА



Первое простое предложение.

## Graphemathical Alalysis

word	descriptors
	0 0 BEG DOC
Первое	0 6 RLE Aa NAM?
простое	7 7 RLE aa
предложение	15 11 RLE aa
.	26 1 PUN CS? SENT_END

# Морфологический анализ (МА)



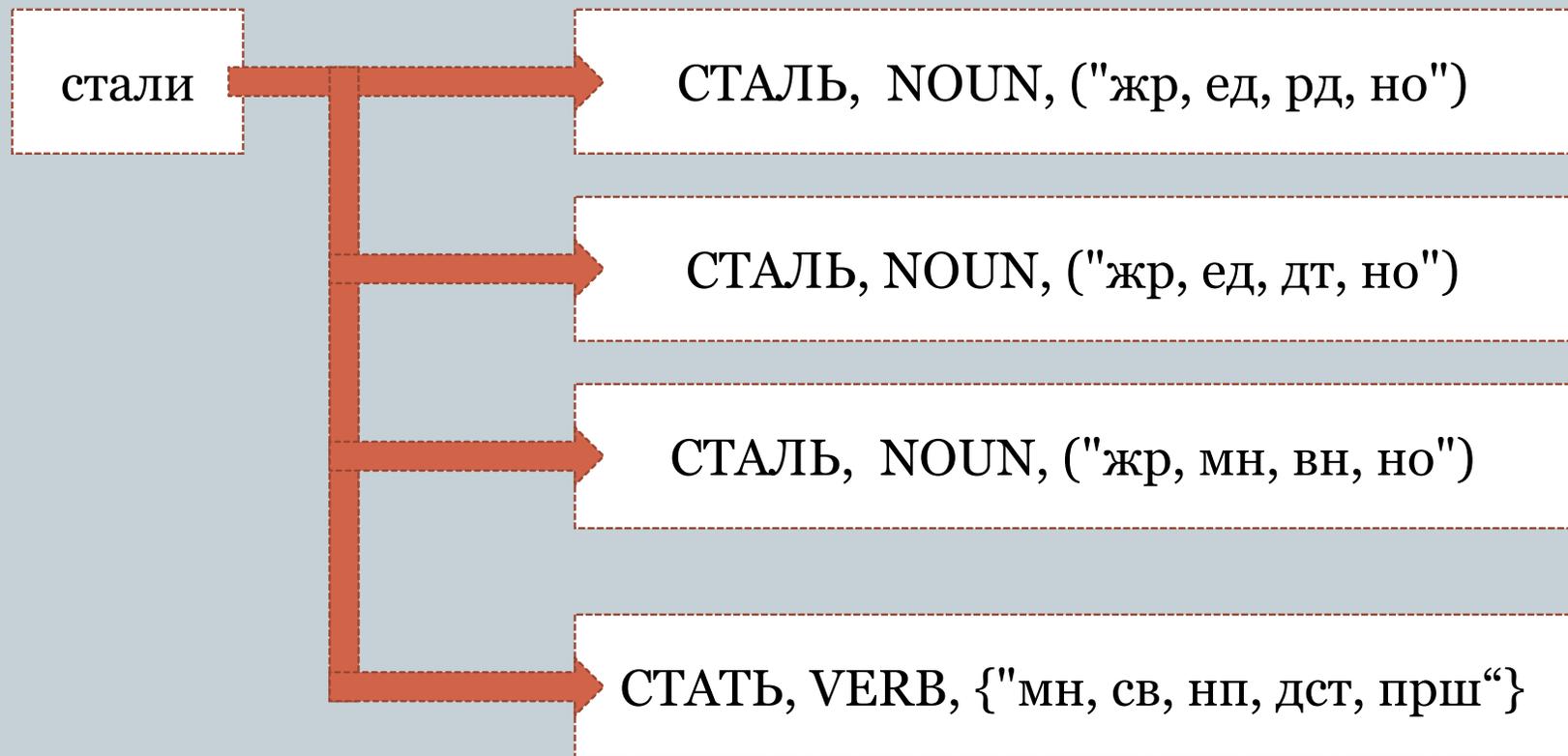
- Работает на уровне отдельных слов
- Определяет морфологические характеристики слова и его лемму

# Подходы к реализации МА



1. На основе словаря («четкая морфология»)
2. На основании некоторых правил («нечеткая морфология»)

# Результат работы МА



# Результат работы МА



Первое простое предложение.

## Morphological Analysis

word	ColumnHeader
ПЕРВОЕ	С ср,ед,им
ПЕРВОЕ	С ср,ед,рд
ПЕРВЫЙ	ЧИСЛ-П мр,ед,им,од,но
ПЕРВЫЙ	ЧИСЛ-П мр,ед,рд,од,но
ПРОСТОЙ	П мр,ед,им,од,но
ПРОСТОЙ	П мр,ед,рд,од,но
ПРОСТОЙ	С мр,ед,им
ПРЕДЛОЖЕНИЕ	С ср,ед,им
ПРЕДЛОЖЕНИЕ	С ср,ед,им

# Синтаксический анализ (СА)

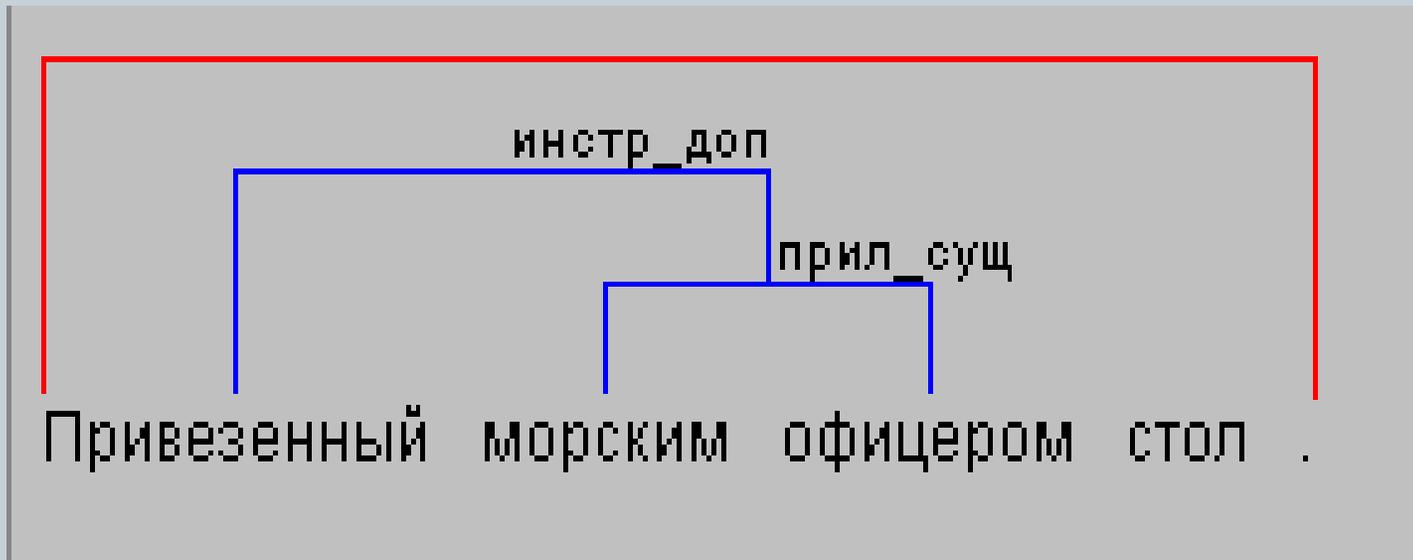


- Определение всех синтаксических признаков и синтаксических связей слов
- Необходим для семантического анализа

# Результаты работы СА



Привезенный морским офицером стол.



# Результаты работы СА



Привезенный морским офицером стол.

привезенный стол



ПРИЛ-СУЩ

морской офицер



ПРИЛ-СУЩ

привезенный офицером



ПРИЛ-СУЩ

# Результат работы СА



## Первое простое предложение.

Syntactical Analysis	Syntactical Analysis	Syntactical Analysis
pair	pair	pair
type	type	type
С ср.едим <> П мр.ед...	ЧИСЛ-П мр.едим,одн...	Первое + простое
С ср.едим <> С ср.едим	Первое + простое	Ч - П
Первое + предложение	ЧИСЛ-П мр.едим,одн...	ЧИСЛ-П мр.едрд,одно...
С - С	П мр.едим,одно <> С ...	П мр.едрд,одно <> С с...
П мр.едим,одно <> С ...	ЧИСЛ-П мр.едим,одн...	ЧИСЛ-П мр.едрд,одно...
С ср.едим <> П мр.ед...	Первое + простое	Первое + простое
С ср.едим <> С ср.едим	Ч - П	Ч - П
Первое + предложение	ЧИСЛ-П мр.едим,одн...	ЧИСЛ-П мр.едрд,одно...
С - С	П мр.едим,одно <> С ...	П мр.едрд,одно <> С с...
П мр.едим,одно <> С ...	ЧИСЛ-П мр.едим,одн...	ЧИСЛ-П мр.едрд,одно...
С ср.едим <> П мр.ед...	Первое + простое	Первое + простое
С ср.едим <> С ср.едим	Ч - П	Ч - С
Первое + предложение	ЧИСЛ-П мр.едим,одн...	ЧИСЛ-П мр.едрд,одно...
С - С	П мр.едрд,одно <> С с...	С мр.едим <> С ср.едим
П мр.едрд,одно <> С с...	ЧИСЛ-П мр.едим,одн...	ЧИСЛ-П мр.едрд,одно...
С ср.едим <> П мр.ед...	Первое + простое	Первое + простое
С ср.едим <> С ср.едим	Ч - П	Ч - С
Первое + предложение	ЧИСЛ-П мр.едим,одн...	ЧИСЛ-П мр.едрд,одно...
С - С	П мр.едрд,одно <> С с...	С мр.едим <> С ср.едим
П мр.едрд,одно <> С с...		

# Заключение



- Проведен обзор разработок и алгоритмов
- Проведен анализ результатов работы МА
- Выделены слабые места разработки, требующие улучшения
- Предложены идеи улучшения работы алгоритмов

# ВОПРОСЫ ?



**СПАСИБО ЗА ВНИМАНИЕ**

Середа А.А.  
НТУУ «КПИ», ИПСА  
гр. ДА-62