

Интерфейс на основе нейросетевых языковых моделей для взаимодействия человека и робота на русском языке

А. Г. Сбоев ^{1,2}, А.В. Грязнов ¹, Р.Б. Рыбка ¹, М.С. Скороходов ^{1,2}, И.А. Молошников ¹

¹ Национальный исследовательский центр «Курчатовский институт», Москва, Россия

² Национальный исследовательский ядерный университет «МИФИ», Москва, Россия

Актуальность

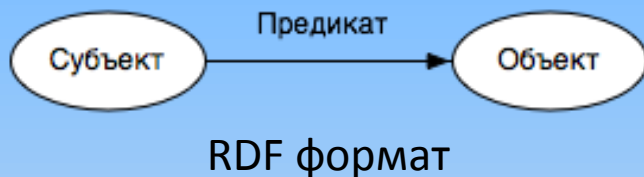
Разработка эффективных интерфейсных систем «человек-машина» для управления робототехническими устройствами при помощи естественного языка является востребованной задачей, в особенности, когда такое управление осуществляется оператором без специальной подготовки.

Проблема

Управление робототехническими устройствами в полевых условиях, в том числе операторами без специальной подготовки, требует создания гибкой и точной системы преобразования команд на естественном языке в формализованный формат команд, распознаваемый робототехническим устройством. Главным отличием формализованного формата от естественного языка является отсутствие неоднозначностей толкования команды. Данное отсутствие ставит актуальную задачу построения системы обработки русскоязычных команд в формализованный формат, которая включает обработку свободного порядка слов, синонимов, омонимии и т.д.

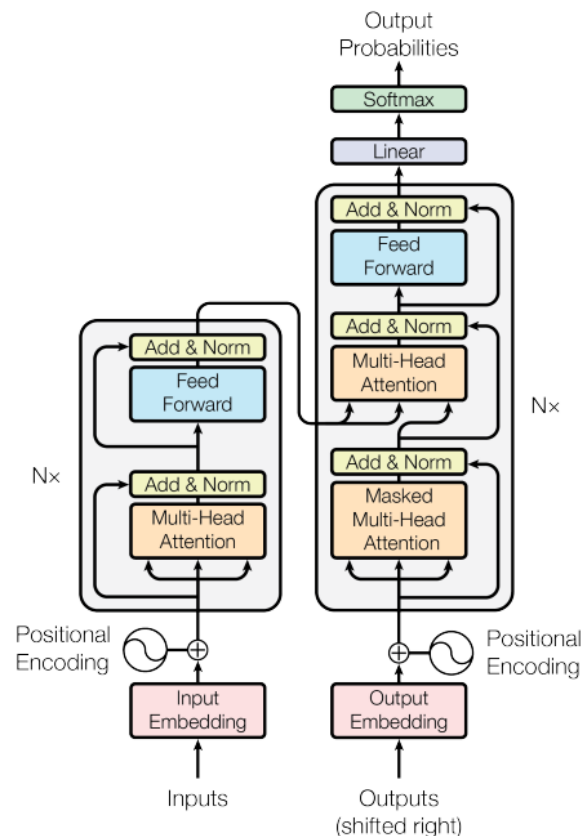
Цель

Использование методов глубокого обучения на основе языковых моделей с архитектурой трансформер для обработки русскоязычных команд в формализованный графовый формат RDF для управления робототехническим устройством.



Задачи

1. Имплементация нейросетевой языковой модели поиска и замены местоимений и нейросетевого интерфейса обработки сложных команд [1]
 2. Анализ влияния модели поиска и замены местоимений на время обработки и точность формализации команд естественного языка
-
1. Нейросетевой интерфейс конвертации сложных русскоязычных текстовых команд в формализованный графовый вид для управления робототехническими устройствами / Сбоев А.Г., Грязнов А.В., Рыбка Р.Б., Скороходов М.С., Молошников И.А. // Вестник Национального исследовательского ядерного университета МИФИ. – 2022. – Т.11 № 2. – С. 153-163.



Обзор работ

Управление робототехническими устройствами на естественном языке реализуется при помощи нейросетевых моделей с топологией трансформер с последующей формализацией в управляющие команды.

Авторы [2] представили систему на базе языковой модели BERT для интерпретации естественного языка в функции языка программирования Python для управления манипулятором Dobot Magician.

В другой работе [3] авторы используют BERT модели для классификации естественного языка по атрибутам и составления последовательности подзадач на примере управления ассистентом в виртуальной среде.

В [4] представлено использование большой языковой модели PaLM, которая справляется с двусмысленными командами на естественном языке для управления кухонным помощником.



2. Translating Natural Language Instructions to Computer Programs for Robot Manipulation / S.V. Gubbi, R. Upadrashta, B. Amrutur // arXiv:2012.13695. – 2020
3. FILM: Following Instructions in Language with Modular Methods / S. Y. Min [и др.] // arXiv:2110.07342. – 2021.
4. Do As I Can and Not As I Say: Grounding Language in Robotic Affordances / M. Ahn [и др.] // arXiv:2204.01691. – 2022.

Наборы данных

1. Синтетический набор данных - генератор по заданным шаблонам, использующий словари с синонимами

№	Описание команды	Примеры генератора
1	Взаимодействие с ближайшим объектом	Осмотри ближайшее дерево
2	Последовательные 3 команды	Осмотри этот камень, далее подъезжай к дому, а потом найди человека
3	Направление взгляда	Поезжай к этому человеку
4	Команды, связанные местоимением	Поворачивай к ближайшему дому и осматривай его

2. Набор данных, полученный в процессе краудсорсинга в данной работе - целевой группе участников были выданы инструкции, кратко описывающие возможные команды на естественном языке

Команда	Движение на кол-во метров в направлении
Возможные атрибуты команды	Действие: двигаться Направление: восток, направо, запад, налево Количество метров: 5 метров, 5 м
Шаблон	[действие]+[направление]+[количество метров]
Примеры генератора	Пойди на 2м; Поезжай на 24 м вперед;

3. Открытые данные [5] - набор новостных, художественных и технических текстов, а также тексты из социальных сетей.

Предложения	Обучение	Тестирование	Валидация
С пропусками	5542	680	1382
Всего	16406	2045	4142

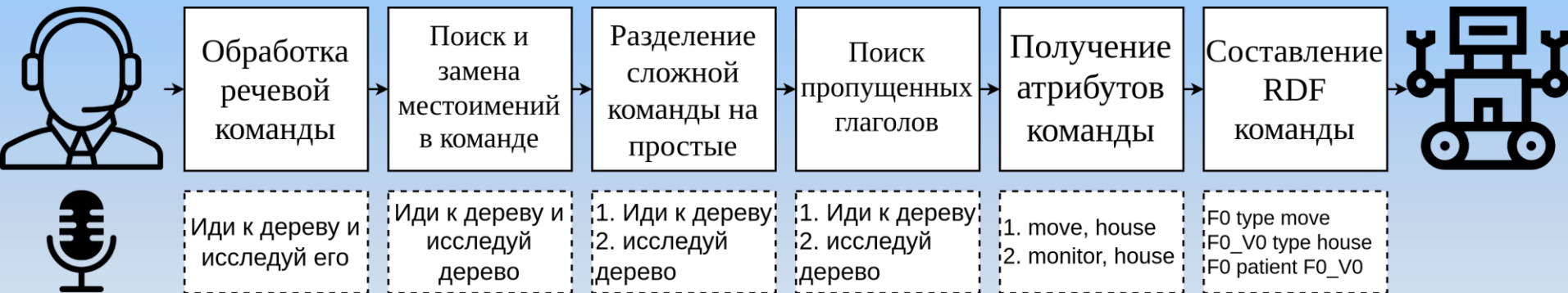
Методы

Модель RuT5 для генерации текстовых последовательностей:

- Поиск и замена местоимений в команде

Модели MultilingualBERT, RuBERT-tiny [1]:

- Восстановление пропущенных глаголов
- Разделение сложной команды на простые
- Классификация атрибутов простой команды



1. Нейросетевой интерфейс конвертации сложных русскоязычных текстовых команд в формализованный графовый вид для управления робототехническими устройствами / Сбоев А.Г., Грязнов А.В., Рыбка Р.Б., Скороходов М.С., Молошников И.А. // Вестник Национального исследовательского ядерного университета МИФИ. – 2022. – Т.11 № 2. – С. 153-163.

Эксперимент

Для анализа эффективности представленных методов была проведена оценка точности выполнения команд как с использованием поиска и замены местоимений в сложных командах (модель RuT5, две версии base и small отличающиеся количеством параметров), так и без данного этапа. При оценке точности проверялись атрибуты команды в сформированном RDF формате, в основном класс объекта

Обработка	Входная последовательность	Выходная последовательность
Генерация текста команды	Подойди к дому и осмотри его	Команда: подойди к дому и осмотри дом
Генерация пары объект-местоимение	Поворачивай к этому человеку, осматривай его	Цель: человека его

Процесс обработки	Успешно обработанных команд, %	Среднее время обработки, с
Ru-T5-small, текст команды	32.6	0.232
RuT5-base, текст команды	69.8	0.264
RuT5-small, объект-местоимение	95.4	0.132
RuT5-base, объект-местоимение	98.7	0.117
Без замены местоимений	5	0.044

Заключение

Полученная система на основе нейросетевых языковых моделей позволяет эффективно обрабатывать сложные русскоязычные команды в формализованный формат RDF для управления робототехнической платформой.

Данная система работает наиболее точно с моделью для поиска и замены местоимений ruT5-base, которая заменяет местоимение в исходной команде, вместо генерации полного текста команды.

Проведенные исследования показывают, что модель поиска и замены местоимений существенно увеличивает точность интерпретации нейросетевого интерфейса на основе восстановления пропущенных глаголов, декомпозиции сложных команд на простые и классификации атрибутов простых команд