

Министерство образования Республики Беларусь  
Учреждение образования  
«Брестский государственный технический университет»  
Кафедра ИИТ

Лабораторная работа №3

По дисциплине: «Естественно-языковой интерфейс ИС»

Тема: «Разработка системы машинного перевода документов»

Выполнил:

Студент 4 курса

Группы ИИ-21

Романко Н. А.

Проверил:

Булей Е. В.

Брест 2024

**Цель:** освоить на практике основные принципы машинного перевода документов.

### Ход работы:

№	Направление перевода	Язык
4	Англо-немецкий	Научные статьи по computer science, Сочинения по литературе

- на входе – естественно-языковой текст на входном языке, подлежащий процедуре машинного перевода;
- подсчитать количество слов во входном тексте, количество переведенных слов, определить грамматическую информацию (теги частей речи и их расшифровка, для этого следует использовать функциональность лр. No3 весенний семестр).
- на выходе: – перевод входного текста на выходной язык; – упорядоченный по частоте встречаемости в тексте список слов и их переводов на выходной язык с грамматической информацией (вкладка 1) (можно использовать функциональность лр. No1 весеннего семестра); – построенное дерево синтаксического разбора выбранного предложения (вкладка 2).
- обеспечить наличие утилиты автоматического пополнения/корректировки полученного словаря (таблицы БД).
- обеспечить сохранение и распечатку результатов перевода и упорядоченных по частоте встречаемости в тексте списков слов и их переводов на выходной язык с грамматической информацией в файл формата txt кодировки Unicode.
- интерфейс системы должен быть простым и доступным для пользователей любого уровня, содержать понятный набор инструментов.

### Код программы:

Чтение PDF файла:

```
def open_pdf(self):
    file_path = filedialog.askopenfilename(filetypes=[("PDF Files", "*.pdf")])
    if file_path:
        try:
            doc = fitz.open(file_path)
            self.temp_file = tempfile.NamedTemporaryFile(delete=False, mode='w+', encoding='utf-
8')

            for page_num in range(doc.page_count):
                page = doc.load_page(page_num)
                page_text = page.get_text("text")
                self.temp_file.write(page_text + '\n\n')

            self.temp_file.flush()
            self.temp_file.seek(0)

            self.text_area.delete(1.0, tk.END)
            self.text_area.insert(tk.INSERT, self.temp_file.read())

            self.translated_text_area.delete(1.0, tk.END)
            self.translate_button.config(state=tk.NORMAL)
        except Exception as e:
            messagebox.showerror("Ошибка", f"Не удалось открыть PDF: {e}")
```

Перевод текста:

```
def translate_text(self):
    if self.temp_file:
```

```

try:
    self.temp_file.seek(0)
    self.translated_text_area.delete(1.0, tk.END)

    self.translate_button.config(text="Переводим...")

    for line in self.temp_file:
        if self.stop_translation_flag:
            break

        if line.strip():
            translated = self.translator.translate(line.strip(), src='en', dest='de').text
            # перевод с английского на немецкий
            self.translated_text_area.insert(tk.END, translated + '\n\n')
            self.translated_text_area.update_idletasks()

    self.translate_button.config(state=tk.NORMAL, text="Перевести")
    self.open_button.config(state=tk.NORMAL)
    self.stop_button.config(state=tk.DISABLED)

except Exception as e:
    messagebox.showerror("Ошибка", f"Не удалось перевести текст: {e}")
    self.translate_button.config(state=tk.NORMAL)
    self.stop_button.config(state=tk.DISABLED)
    self.open_button.config(state=tk.NORMAL)
else:
    messagebox.showwarning("Внимание", "Сначала откройте PDF файл.")

```

### Результат:

Открытие PDF-файла:

[Открыть PDF](#)

## Automating Thought of Search: A Journey Towards Soundness and Completeness

Daniel Caol, Michael Katz<sup>2</sup>, Harsha Kokel<sup>2</sup>, Kavitha Srinivas<sup>2</sup>, Shirin Sohrabi<sup>2</sup>

<sup>1</sup> Cornell University

<sup>2</sup> IBM Research

### Abstract

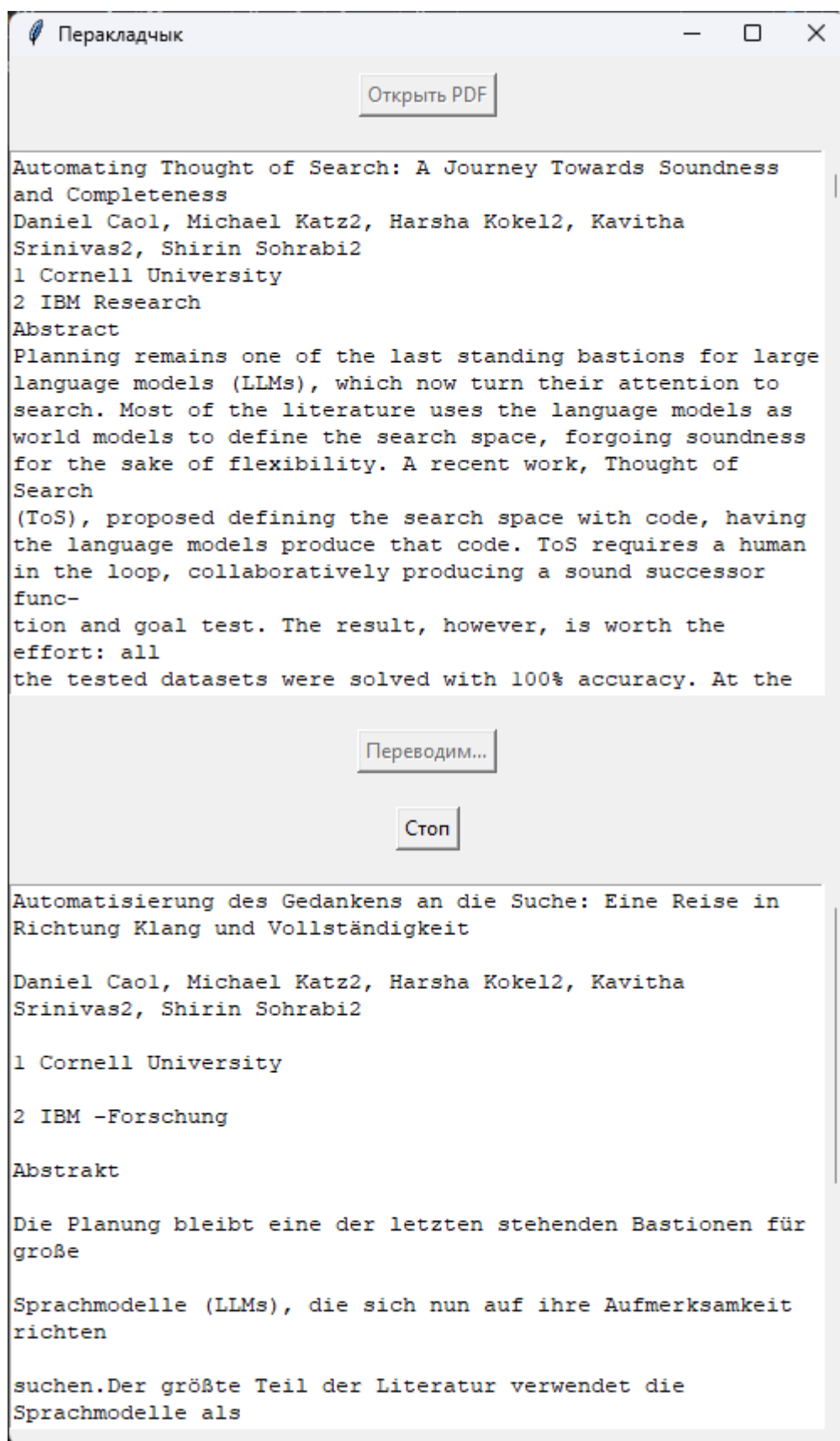
Planning remains one of the last standing bastions for large language models (LLMs), which now turn their attention to search. Most of the literature uses the language models as world models to define the search space, forgoing soundness for the sake of flexibility. A recent work, Thought of Search

(ToS), proposed defining the search space with code, having the language models produce that code. ToS requires a human in the loop, collaboratively producing a sound successor function and goal test. The result, however, is worth the effort: all

the tested datasets were solved with 100% accuracy. At the

[Перевести](#)[Стоп](#)

Перевод текста:



**Вывод:** в ходе выполнения лабораторной работы освоил принципы машинного перевода документов.