

ИНТЕГРАЦИЯ МЕТОДОВ СИНТАКСИЧЕСКОГО И СЕМАНТИЧЕСКОГО АНАЛИЗА ТЕКСТОВ

Халалов Закир Абдужалалович
доцент кафедры узбекского
языка и иностранных языков
Академии МВД РУз

Аннотация. В статье рассматривается проблема интеграции синтаксического и семантического анализа при автоматической и лингвистической обработке текстов. Анализируются основные подходы отечественной лингвистики к синтаксису и семантике, а также современные вычислительные методы их объединения. Приводятся примеры взаимодействия синтаксических структур и смысловых ролей. Показано, что интеграция двух уровней анализа повышает качество интерпретации текста и эффективность прикладных лингвистических систем.

Ключевые слова: синтаксический анализ, семантический анализ, интеграция, зависимый парсинг, валентность, смысловые отношения, корпусная лингвистика, русский язык.

Проблема интеграции синтаксического и семантического анализа занимает центральное место в современной лингвистике и компьютерной обработке текстов. Синтаксический анализ определяет структурные отношения между словами (кто является подлежащим, дополнением, обстоятельством и т. д.), а семантический — выявляет смысловые связи (кто совершает действие, над чем совершается действие, в каком контексте и с какой целью).

Отдельное применение этих подходов часто приводит к неполному пониманию текста. Например, грамматически правильно разобранный предложение может оставаться семантически неясным без анализа смысловых

ролей. Следовательно, их интеграция является необходимым шагом к созданию интеллектуальных систем обработки языка (Валгина, 2003; Захаров, 2005).

- Теоретические основы интеграции

Интеграция синтаксического и семантического анализа предполагает объединение двух уровней описания языка — структурного (формального) и содержательного (смыслового).

Как отмечает Шведова Н. Ю. (1980), синтаксис неразрывно связан с семантикой, поскольку грамматические отношения выражают смысловые зависимости.

А Золотова Г. А. (1973) указывает, что каждое предложение имеет смысловую структуру, в которой каждому элементу соответствует определённая семантическая роль: агенс (действующий), пациенс (объект действия), инструмент и т. д.

Таким образом, интеграция двух видов анализа позволяет установить не только «кто с кем согласуется», но и «кто что делает» и «почему».

- Методы интеграции анализа

- Лингвистический (правилевой) подход

В традиционной отечественной лингвистике синтаксико-семантическая интеграция строится на описании валентности глагола и смысловых ролей.

Например, в предложении:

(1) Студент читает книгу в библиотеке.

Синтаксический анализ определяет:

Студент — подлежащее,

читает — сказуемое,

книгу — прямое дополнение,

в библиотеке — обстоятельство места.

Семантический анализ уточняет:

Студент — агенс (действующее лицо),

книга — пациенс (объект действия),
в библиотеке — локатив (место действия).

Интеграция этих двух уровней позволяет сформировать полную синтаксико-семантическую схему:

[АГЕНС: Студент] → [ДЕЙСТВИЕ: читать] → [ПАЦИЕНС: книга] →
[ЛОКАТИВ: библиотека]

Такая модель используется при построении фреймовых структур в семантическом анализе (Селиверстова, 2005).

-ьКорпусно-статистические методы

Корпусная лингвистика позволяет выявлять статистические закономерности между синтаксическими зависимостями и смысловыми ролями. Например, в корпусе можно наблюдать, что глагол читать в 90% случаев управляет объектом в винительном падеже, который выполняет роль пациенса.

В системах автоматического анализа (например, в проекте «Русский синтаксический корпус») создаются аннотации, где синтаксические теги дополнительно сопровождаются семантическими метками. Это позволяет обучать нейросетевые модели, которые предсказывают не только грамматическую структуру, но и значение отношений.

- Нейросетевые и гибридные методы

Современные вычислительные модели (например, RuBERT, DeepPavlov, ChatGPT) интегрируют синтаксическую и семантическую информацию в контекстных векторных представлениях слов.

Пример автоматического синтаксико-семантического анализа:

(2) Мальчик открыл дверь ключом.

Синтаксический разбор:

Мальчик — подлежащее (NOUN),

открыл — глагол (VERB),

дверь — дополнение (NOUN, Acc),

ключом — обстоятельство инструмента (NOUN, Instr).

Семантическая интерпретация:

Мальчик — агенс,

дверь — пациенс,

ключом — инструмент.

Интеграция двух уровней позволяет автоматически построить семантический граф:

АГЕНС(мальчик) —ДЕЙСТВИЕ(открыть)—> ПАЦИЕНС(дверь) →
ИНСТРУМЕНТ(ключ)

Такой граф может быть интерпретирован машиной и использован, например, для систем вопросов-ответов.

- Применение интегрированных методов

Машинный перевод – сохранение синтаксической структуры при передаче семантического смысла.

Анализ тональности – выделение субъектов и объектов оценки.

Автоматическое реферирование – выявление смысловых центров текста.

Поиск по смыслу – определение связей между действиями и участниками, а не только совпадений слов.

Например, запрос «кто изобрёл лампу?» может быть обработан системой, распознавшей семантическую структуру агенс + действие + объект, даже если в тексте встречается фраза «Лампа была создана Эдисоном».

Проблемы и перспективы

Отсутствие единых русскоязычных стандартов семантической разметки.

Сложность автоматического выделения скрытых смыслов и эллипсиса.

Необходимость объединения грамматических правил с нейросетевыми методами для объяснимости результатов.

В перспективе — создание гибридных систем, где нейронные сети опираются на лингвистические принципы русской грамматики. Это направление активно развивается и исследуется.

Интеграция синтаксического и семантического анализа является ключом к глубокому пониманию текста. Для русского языка, обладающего богатой морфологией и свободным порядком слов, такие методы особенно важны. Объединение формальных структур синтаксиса и смысловых моделей семантики обеспечивает создание интеллектуальных систем, способных интерпретировать текст не только по форме, но и по содержанию.

Библиография

1. Шведова Н. Ю. (ред.) Русская грамматика. Том II. Синтаксис. — М.: Наука, 1980.
2. Валгина Н. С. Современный русский язык: Синтаксис. — М.: Высшая школа, 2003.
3. Золотова Г. А. О смысловой структуре предложения. — М.: Наука, 1973.
4. Селиверстова О. Н. Типы синтаксико-семантических отношений в русском языке. — М.: Языки славянской культуры, 2005.
5. Захаров В. П. Корпусная лингвистика. — СПб.: СПбГУ, 2005.
6. Перфильев А. А. Методы синтаксического анализа и сопоставления конструкций естественного языка, ориентированные на применение в информационно-поисковых системах. — М.: ИПС РАН, 2011.