

ОМОНИМЫ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

Г.С. Самандарова

Старший преподаватель Alfraganus University, Ташкент

Аннотация: В статье рассматривается место омонимии в искусственном интеллекте в эпоху научно-технической революции. Омонимы стали особенно значимыми в контексте развития информационных технологий, автоматизации и искусственного интеллекта.

Выделяются основные проблемы интеграции искусственного интеллекта, и намечается возможные способы их преодоления и нивелирования.

В статье рассматриваются трудности работы с омонимами, при машинном переводе в связи с проблемой многозначности слов в коммуникации между человеком и машиной. Также рассмотрены методы решения проблем омонимии в ИИ. Автоматическая обработка текстов, машинный перевод, создание интеллектуальных систем, связанных с использованием естественного языка, – снова выдвинули проблему омонимии на первый план.

Ключевые слова: омонимы, омонимия, искусственный интеллект, научно-техническая революция, многозначность, контекст

НТР конца XX и начала XXI веков, конечно, отразилась и на лингвистике мира. «Сущность НТР — прочная и организованная связь науки, техники и производства...»¹ «Значительное воздействие на русский язык оказала и продолжает оказывать научно-техническая революция. Ее влияние распространяется на разные стороны языка и имеет различные проявления - от стирания диалектных различий до бурного роста терминологии. Но говоря о новом в языке, не следует забывать, что язык - устойчивая система, его развитие происходит постепенно, с обязательной опорой на традицию. Самые значительные, революционные преобразования общества не ведут к "революции языка".²

В эпоху научно-технической революции (НТР) омонимы стали особенно значимыми в контексте развития информационных технологий, автоматизации и искусственного интеллекта. В этот период начали активно развиваться вычислительные системы, способные обрабатывать естественный язык, что подчеркнуло проблему многозначности слов в коммуникации между человеком и машиной.

С появлением новых технологий и понятий увеличилось количество слов с техническими и общеупотребительными значениями, что привело к возникновению новых омонимов. Например, слово «мышь» в контексте компьютерной техники стало означать устройство управления, что отличается от его биологического значения.

В условиях глобализации и интернационализации резко возросла потребность в автоматическом переводе и анализе текстов, что обострило проблему омонимии. Машины

¹ И.И. Лейман Очерки методологии исследования. «Общественное развитие и НТР» Ленинград . Наука, 1982. - 268 с.

² С. И. Виноградов. [Русский язык эпохи НТР: споры, мнения, оценки](#) Русская речь. 1986. № 1, С. 11 - 21

часто сталкиваются с трудностями при определении правильного значения слова, особенно если контекст недостаточно ясен.

Во время НТР стали создаваться обширные текстовые корпуса, которые используются для обучения систем обработки естественного языка. Это помогло лучше понять распространенность омонимов и разработать методы их разрешения.

Из-за роста технологического контента и необходимости быстрого освоения новых терминов, пользователи сталкиваются с двойными значениями слов, что может приводить к недоразумениям в технической документации и коммуникации.

Приведём примеры омонимов, обострившихся в эпоху НТР.

Файл: может означать документ на компьютере или ряд чего-либо (например, «файл солдат»).

Блок: в технике – это модуль или часть системы, а в обыденной речи – крупный кусок чего-то (например, «блок льда»).

Код: программный код или шифр для защиты информации, но также и закон или свод правил. Эти примеры подчеркивают, что НТР не только ускорила технологический прогресс, но и вызвала новые лингвистические вызовы, требующие эффективных решений для точной передачи информации в условиях многозначности языка.

Омонимы в контексте искусственного интеллекта (ИИ) представляют собой интересную задачу, поскольку омонимы могут вызывать неоднозначность при обработке естественного языка (NLP). ИИ должен правильно интерпретировать омонимы в зависимости от контекста, что требует применения сложных алгоритмов анализа текста.

Омонимы – это слова, которые звучат или пишутся одинаково, но имеют разные значения (например, "ключ" как инструмент и "ключ" как источник воды). Для человека контекст помогает различить значение, но ИИ может столкнуться с проблемой интерпретации.

Пример: В предложении "Он нашёл ключ от дома возле ключа", ИИ должен различить оба значения слова "ключ" (инструмент и источник воды) на основе окружающих слов.

Для решения проблем омонимии в ИИ используются следующие методы.

Контекстуальный анализ. Современные системы ИИ, такие как модели глубокого обучения (например BERT, GPT), проводятся на огромных объёмах текстов и используют контекст для распознавания значений омонимов.

Семантические сети. Одним из подходов является использование семантических сетей, где слова и их значения связаны между собой. Это позволяет ИИ находить более вероятное значение омонима в зависимости от других слов в предложении.

Использование словарей и баз данных. В дополнение к обучению на текстах, системы ИИ могут использовать существующие базы данных, такие как WordNet, которые содержат информацию о значениях слов и их связях с другими словами.

Машинный перевод также сталкивается с проблемой омонимов, особенно при переводе с одного языка на другой. Неправильная интерпретация омонимов может привести к ошибкам в переводе.

При переводе фразы "Он заказал стол для переговоров" на английский, ИИ должен правильно интерпретировать "стол" как "table" в значении предмета мебели, а не "desk" (рабочее место).

Современные генеративные языковые модели, такие как GPT, обучены на огромных массивах данных и способны эффективно различать значения омонимов благодаря обучению на

контекстуальных примерах. Такие модели могут анализировать не только отдельные предложения, но и абзацы, чтобы понимать, о каком значении слова идёт речь.

Модели ИИ постоянно совершенствуются, и в будущем ожидается ещё более точное распознавание омонимов благодаря улучшенным методам контекстного анализа.

В специализированных областях (медицина, юриспруденция и т.д.) омонимы могут иметь особые значения, поэтому ИИ может адаптироваться к таким условиям, обучаясь на специфичных текстах.

Развитие специализированных наборов данных, содержащих примеры омонимов, поможет улучшить обучение моделей, что повысит их точность при обработке естественного языка.

Омонимы представляют собой сложную задачу для искусственного интеллекта, но современные методы NLP, такие как глубокое обучение и контекстный анализ, позволяют успешно решать эту проблему. Текущие и будущие исследования в этой области направлены на улучшение качества распознавания омонимов, что особенно важно для приложений, связанных с машинным переводом и генерацией текста.

Омонимия в контексте искусственного интеллекта представляет собой проблему, связанную с многозначностью слов, которые пишутся и произносятся одинаково, но имеют разные значения. В области обработки естественного языка (NLP), омонимия может затруднять правильное понимание и интерпретацию текста ИИ-системами, так как алгоритмам сложно определить, какое именно значение слова уместно в текущем контексте.

Существует ряд проблем омонимии в искусственном интеллекте.

Неправильное определение значения омонима может привести к неверному анализу текста или созданию некорректного ответа.

Омонимы могут быть переведены неправильно, если ИИ не понимает контекста.

Чат-боты могут давать неуместные ответы, если не различают значения омонимов.

Проблемы омонимии решаются следующими методами.

1. Контекстный анализ. Использование нейронных сетей, таких как трансформеры (например, BERT), для учета контекста при интерпретации значения слов.

2. Обучение на больших корпусах данных. Чем больше данных использовано для обучения ИИ, тем лучше модель может понимать контекст и различать омонимы.

3. Многозадачное обучение. Использование моделей, обученных на решении нескольких связанных задач (например, анализ тональности и классификация текста), может улучшить понимание контекста.

Преодоление омонимии важно для повышения точности ИИ в обработке языка, что делает эту задачу ключевой в развитии передовых NLP-систем.

Таким образом «проблема омонимии стара как лингвистика. В то же самое время она переживает вторую молодость. Сначала автоматическая обработка текстов, далее машинный перевод, а совсем недавно – создание интеллектуальных систем, связанных с использованием естественного языка, – снова выдвинули ее на первый план».³

References

³ А.И. Головня «ОМОНИМИЯ КАК СИСТЕМНАЯ КАТЕГОРИЯ ЯЗЫКА» Минск 2007, стр.3.

- .И. Лейман Очерки методологии исследования. «Общественное развитие и НТР» Ленинград . Наука, 1982. - 268 с.
- . И. Виноградов. Русский язык эпохи НТР: споры, мнения, оценки Русская речь. 1986. № 1, С. 11 – 21
- .И. Головня «ОМОНИМИЯ КАК СИСТЕМНАЯ КАТЕГОРИЯ ЯЗЫКА» Минск 2007, стр.3.
4. Стивен Клейн, Иван Бёрд и Эдвард Лопер "Natural Language Processing with Python" Published by O'Reilly Media, June 2009 p. 60
5. Daniel Jurafsky and James H. Martin "Speech and Language Processing", Upper Saddle River, New Jersey 07458 2007, p. 705-747