

## К ПРОБЛЕМЕ МОДЕЛИРОВАНИЯ ЯЗЫКОВОГО ПЕРЕВОДА СПЕЦИАЛИЗИРОВАННОГО НАУЧНО-ТЕХНИЧЕСКОГО ТЕКСТА.

В настоящей работе рассматривается задача, которая относится к проблеме моделирования языкового перевода специализированного научно-технического текста. Задача заключается в определении интенций адресанта в речевом акте, оформленном в виде научной статьи. Интенционный анализ может применяться в автоматизированном переводе научно-технических текстов, в силу высокой контекстной обусловленности последних и, как следствие, наличие в нем «форматных» языковых конструкций, клише.

Согласно модели «интенция-текст» каждая условная часть научно-технического текста содержит некоторый замысел автора, ответ на вопрос «С какой целью данная часть воспроизведена в тексте?». Так автор научной статьи по механике сплошной среды намеревается (с некоторой вероятностью) в вводной части познакомить читателя с рассматриваемой задачей, показать актуальность исследуемой проблемы, в частности, посвященной описанию математической модели, ввести понятие об исследуемом физическом объекте или задать множество, а в обосновании используемого им численного метода обозначить используемую в доказательстве вспомогательной теоремы математическую функцию. Условная часть текста может состоять из одного или последовательностью предложений, например: *«Одним из вариантов расчетной схемы теплозащитной конструкции является стержень из стеклопластика при одностороннем высокотемпературном нагреве, который рассматривается в предлагаемой работе.»* и *«При этом в режиме спуска на наружной «горячей» поверхности теплозащитной оболочки КСА температура повышается до 1000 и более градусов Цельсия, а температура внутренней «холодной» поверхности на несколько сотен градусов ниже. Таким образом, имеет место односторонний высокотемпературный нагрев.»*. В первом примере автор называет исследуемую проблему, а во втором обозначает суть проблемы, используя описание физического явления в ее основе.

В силу избыточности научно-технических текстов интенции выражаются шаблонными языковыми конструкциями, одинаковыми целиком или по своей структуре для различных текстов различных авторов. Так предложение *«Для изучения точности рассматриваемого метода использована модельная задача, имеющая точное аналитическое решение.»* является более «универсальным», чем любое из приведенных в примере ранее.

В модели интенционного анализа в качестве исходных данных выступает научно-технический текст в заданной предметной области, и для каждой условной части исследуемого текста требуется определить упорядоченное подмножество в некотором

множестве интенций. Последнее представляет собой набор интенций, например: «*доказать теорему*», «*привести описание физического явления*», «*ввести обозначение*» и т.д. Каждый элемент подмножества обладает весом, который обозначает, насколько точно анализируемая часть текста отражает интенцию. Вес задается некоторой числовой функцией, учитывающей лексикографические особенности рассматриваемой части и статистические данные, полученные в результате ручной обработки тестового набора научно-технических текстов.

К лексикографическим особенностям предложения относится набор синтагм, являющихся типичным для выражения той или иной интенции. Например, наборы {*B; настоящей; работе; рассматривается; ранее представлен; мной; в*} и {*B работе; мы рассматриваем; ранее представлен; нами; в*} выражают замысел автора обозначить в начале статьи саму исследуемую проблему и тот факт, что она уже изучалась, ссылаясь на свою предыдущую работу. При этом в наборе учитывается также порядок следования синтагм, расстояние между ними и позиция относительно начала предложения. Замечу, что в данном случае синтагма необязательно является сочетанием слов удовлетворяющем грамматическим нормам языка, а последовательность символов, часто совпадающая с такими сочетаниями.

Очевидно, что информации из поверхностного синтаксического анализа исследуемой части текста недостаточно, чтобы в каждом конкретном случае определить адекватное подмножество множества интенций. Поэтому для образования подмножества и вычисления весов его элементов требуются статистические данные, характеризующие тенденции на уровне структуры текста. Учитывается позиция анализируемой части относительно текста целиком и соседних частей, управление в синтаксических конструкциях, составляющих данную часть, возможность замены языковой единицы (предложения, словосочетания, слова) на иную, передающую то же семантическое значение, возможность сокращения (отбрасывания языковых единиц) с сохранением смысла.

Примером, где лексикографический анализ даст пустое подмножество интенций, служит часть, состоящая из двух предложений (см. пример в начале работы). Причина, по которой автор приводит описание физического явления в первом предложении, становится ясной только из второго предложения. В первом предложении нет сколько-нибудь значимой синтагмы, отражающей замысел автора. Такое предложение можно употребить и в части, посвященной описанию экспериментального опыта, и в введение, как это и было сделано в реальной научной статье [1, Стр. 19-22]. А синтагма «*таким образом*» во втором предложении указывает на наличие некой посылки, обозначенной ранее, для вывода, сделанного в данном предложении. Окружение каждого из двух предложений, позволило

определить замысел автора. При этом здесь можно говорить об интенции части и об интенциях каждого из предложений: для первого – автор знакомит читателя с физическим явлением, лежащем в основе исследуемой проблемы, для второго – автор обозначает характеристику физической модели, ссылаясь на явление, имеющее место в реальности.

Таким образом, интенции можно различать по уровням. Так «*доказать теорему*» и «*доказать теорему, используя метод математической индукции*» - разные интенции, хотя последняя является частным случаем первой. Более точные интенции позволяют находить наиболее приемлемое выражение в речи, сохранять стиль изложения. Особенно данные результаты приобретают актуальность в случае, когда интенционный анализ используется при переводе текстов на языках разных типов и присущих носителям различным социально-культурных сред. Например, с русского языка – фузионно-флективного – на английский язык – агглютинативный. Множество интенций заданного дискурса выступает посредником между речью на разных языках, позволяет переводить языковые конструкции с одного естественного языка на другой без акцентирования на том, какие именно переводимый и переводящий языки.

Сам процесс интенционного анализа для автоматизированного перевода представляет собой сравнение части исходного текста (например, предложения) с различными шаблонами – наборами синтагм и характеристик – выделяется упорядоченное множество интенций, которые в речи реализуются данной частью текста. Полученное множество интенций используется на этапе синтеза эквивалентных частей на переводящем языке согласно концепции, изложенной выше.

Автоматизация интенционного анализа обуславливает задачу создания словаря шаблонов и выработки четких принципов вычисления характеристик каждого шаблона, требуется сбор и обработка статистических данных, определение весовых функций. В итоге необходимо получить необходимый и достаточный набор информации для синтеза эквивалентных конструкций, для перевода научных статей не из тестового набора, т.е. применимый в реальных условиях.

1. Современные проблемы механики сплошной среды: В 2-ух т./Под ред. А.В.Наседкина - Ростов-на-Дону: Изд-во ООО «ЦВВР», 2006. Т.2.–333с.