

## УРОВНИ РЕАЛИЗАЦИИ МАШИННОГО ПЕРЕВОДА

К. Фабриц

1. Машинный перевод представляет собой раздел вычислительной лингвистики. Вычислительная лингвистика, в свою очередь, есть часть языкознания /частного и общего/, изучающая системы для понимания и порождения естественного языка, осуществляющихся и осуществляемых при помощи ЭВМ.

Данная область лингвистики, называемая также "инженерная лингвистика", преследует как практические, так и теоретические цели. К практическим целям вычислительной лингвистики относятся машинный перевод, автоматизация обработки информации, составление автоматических словарей, а также создание интерактивных систем с ориентацией на все большее использование естественного языка для работы с ЭВМ.

Что касается теоретических стремлений в вычислительной лингвистике, то они прежде всего сводятся к трем направлениям. Во-первых, теоретики вычислительной лингвистики считают ЭВМ удобным /или, по крайней мере, более удобным, чем традиционные/ средством проверки состоятельности различных грамматик: при помощи ЭВМ сравнительно легко и, что не менее важно, объективно можно провести тесты в связи с различными правилами, например, порождения предложения<sup>1</sup>. Во-вторых, с начала семидесятых годов стали вырабатываться основные положения в области "представления знаний". Представление знаний -- это конструирование такой системы, которая способна использовать различные моменты общих для людей процессов мышления. Система представления знаний опирается на операции, имеющие место в решении самых разнообразных задач. Комплекс операций, ведущих к верному и наиболее оптимальному решению задач, лежит в основе исследований по искусственному интеллекту<sup>2</sup>, одной из интереснейших областей сочетания человеческого мышления, проявляющегося в естественном языке, электронно-вычислительной машины, созданной в результате изысканий человека, и

"повседневной жизни", ставящей в своей комплексности ряд задач, решение которых в силу сложности явлений представляется реально возможным лишь с помощью указанного сочетания. В-третьих, теоретическая вычислительная лингвистика изучает вопросы создания моделей, применяемых для описания тех психологических процессов, которые лежат в основе языкового понимания<sup>3</sup>. Как представление знаний выходит за пределы "собственного" языкознания /которое, между прочим, и до сих пор характеризовалось постоянным расширением круга исследуемых им вопросов/ и соприкасается с областью искусственного интеллекта, так и моделирование языкового понимания имеет отношение к новой дисциплине. Моделирование понимания, разработка которого вовлекает, кроме лингвистов, также и психологов, философов и других ученых, представляет собой предмет когнитивной науки. /См. схему 1./

Разумеется, практические исследования проходят в неразрывном соотношении с теоретическими. Те грамматические теории, которые оказываются более пригодными для имплементации в программе в силу их правильной формализованности и надежной в ходе верификации работы, оцениваются как более состоятельные, чем теории, формальное доказание которых затрудняется из-за недостаточной их разработанности, или же непригодности в целом.

Это заключение особенно действительно по отношению к машинному переводу, отличающемуся своей многоплановостью. В ходе построения системы машинного перевода должны быть учтены концепции различных языковых теорий. Многие лингвисты придерживаются мнения, что машинный перевод -- это не что иное, как задача сугубо практическая, решение которой не должно выходить за рамки программирования практических сведений о фактах той или иной пары языков. Думается, что сущность известной неполноты систем машинного перевода, отсутствие у них полного автоматизма заключается не столько в объективных причинах несовершенности технического оборудования /трудности такого характера все быстрее снимаются за счет появления более мощных ЭВМ и более развитых методов хранения огромных массивов информации/, сколько в "субъективных", сугубо языковедческих моментах, таких, как сложная взаимосвязанность явлений

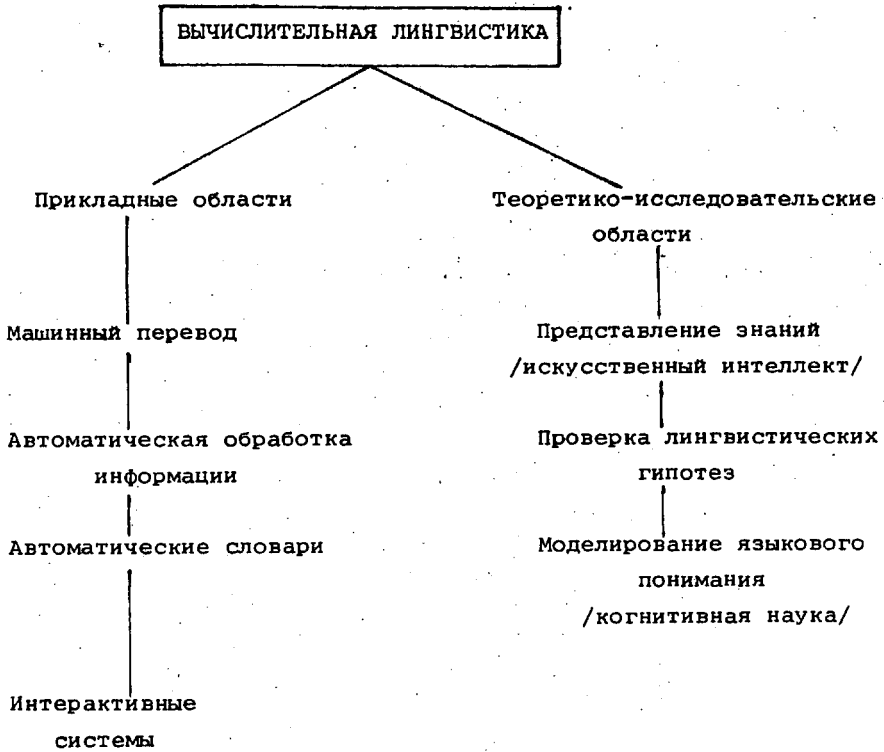


Схема 1. Предметные области исследований по вычислительной лингвистике.

разнообразных языковых ярусов, отсутствие полной структурированности языковых средств передачи информации, а также отсутствие "точных" и всеобъемлющих теорий, допускающих максимальную формализацию данных и сведений о языке.

Подобная формализация может быть достигнута лишь в результате такого, весьма сложного и трудоемкого представления информации о языке, которое предполагает иерархизацию и параллельную структуризацию всего массива языковых сведений. Цель настоящей работы и состоит в том, чтобы коротко описать архитектуру вырабатываемой в Сегедском университете системы машинного перевода.

2. Машинный перевод начался в США в 1947-ом году. За сорок лет своего развития машинный перевод прошел нелегкий путь. На смену энтузиазму, вызванному мыслью о возможности быстрого достижения высококачественного автоматизированного перевода, в середине шестидесятых годов пришло огорчение и разочарование исследователей, постепенно осознавших чрезмерную сложность решаемой задачи<sup>4</sup>. Несмотря на кризис машинного перевода в США, в других странах данная проблематика продолжала разрабатываться. Особенно высокого уровня достигли проекты по машинному переводу в Западной Европе, Советском Союзе и Японии<sup>5</sup>. Вплоть до конца семидесятых годов машинный перевод ценился в кругу лингвистов прежде всего потому, что он послужил первым толчком к возникновению области вычислительной лингвистики в целом. Ныне интерес к машинному переводу вновь возрос благодаря тому, что результат, который можно ожидать от работающих систем автоматизированного перевода, для специалистов информационной промышленности является небезынтересным в финансовом отношении.

В настоящее время машинный перевод как область вычислительной лингвистики характеризуется комплексом собственно лингвистических, алгоритмизационных и программных задач. Установление взаимосвязи этих задач, распределение представления различных сведений в процессе перевода играют решающую роль с точки зрения успешности работы системы перевода. Из названных кругов задач здесь

коснемся первого и второго: программный элемент в архитектуре системы перевода носит сугубо практический характер. Иерархизацию задач, связанных с реализацией машинного перевода, схематически можно изобразить так:

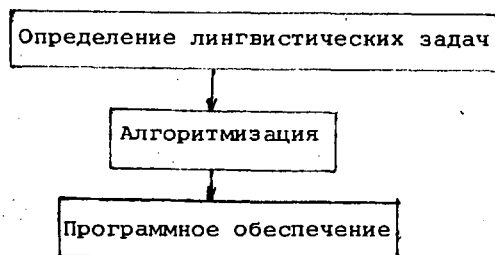


Схема 2. Иерархия задач для реализации машинного перевода

Определение лингвистических задач состоит в основном в группировании сведений, связанных с языком и процессом перевода, на различные уровни. Это обозначает детализацию всей информации, с использованием которой осуществляется перенос содержания текста с одного языка на другой /или другие/. Вслед за представлением всевозможных лингвистических сведений идет работа алгоритмизации, т.е. организации полученных сведений с целью их "упорядочения" в такой форме, согласно которой и становится возможным запрограммирование алгоритмов и алгоритмизированных операций для последующей обработки машиной переводимого текста.

3.1. В кругу лингвистических задач при подготовке языкового материала для машинного перевода выделяются, с одной стороны, задачи лексикографические /тезаурусные/, морфологические, синтаксические и семантические, а с другой стороны, -- задачи переводческие.

Что касается лексикографических задач, то здесь наблюдается тенденция стремления к наиболее полному представлению информации

всевозможных языковых сведений, связанных с появлением той или иной лексемы<sup>6</sup>. Автоматический словарь, таким образом, представляет собой базу, на которой строится вся система машинного перевода. Такой словарь содержит, кроме самих лексем, различные указания на "поведение" определенного слова в парадигматическом и синтагматическом отношении к другим словам. Структура автоматического словаря разбивается в отдельных системах машинного перевода<sup>7</sup>, характерным в настоящее время является появление, кроме базисного словаря исходного языка, также и специального словаря, задача которого состоит в снятии большой доли двусмысленности слов или словоупотреблений. Такие специальные словари работают на основе информации, получаемой о своеобразии сочетания слова с другими словами. Подобный словарь, называемый "контекстологическим", имеется в системе Ю.Н. Марчука "АМПАР", по утверждению которого данный словарь гарантирует 95 %-е решение двусмысленностей<sup>8</sup>.

Основное отличие машинного словаря от словарей двуязычных заключается в своеобразном распределении слов по различным категориям. Соответствующая требованиям машинного перевода классификация слов во входном языке -- решающий шаг по пути к надежной системе. Входной словарь представляет лексику входного языка в таком виде, который позволяет работать с этим фондом на всех уровнях реализации машинного перевода. Для этого требуется снабжение отдельных лексем такими элементами лингвистических сведений в форме цифровых кодов, которые обеспечивают нужный и достаточный запас информации для последующей работы.

На основе словарной информации входного машинного словаря строятся этапы как морфологического, синтаксического, семантического анализа, так и нахождения переводных эквивалентов. Это означает, что лексема, представленная в "полном" виде во входном словаре, должна быть дополнена таким набором кодов, который окажется пригодным для манипуляции на последующих этапах. При этом следует иметь в виду, что те "пучки" информации, которые приписываются к

отдельным лексемам, используются на различных этапах целиком. Так, информация о принадлежности одной лексемы к разряду существительных, будет рекурсивно использована в ходе морфологического, синтаксического и т.д. анализа. Если данное существительное /скажем, scale/ в анализируемом тексте выступает во множественном числе, то информация о правилах употребления данного слова в этой форме числа будет использоваться не только на уровне морфологии и синтаксиса, но также и на уровне семантики, который в архитектуре машинного перевода выступает как высший уровень анализа текста на входном языке.

Итак, составление информационных пучков преследует комплекс целей, главной из которых является обеспечение нужных для обработки текста лингвистических сведений, позволяющих механизацию разноуровневых анализов.

Значение составления информационных пучков, а также полнота и общая приемлемость прилагаемых к лексемам пучков проявляются на различных уровнях по-разному.

3.1.1. На уровне лексики пучок как придаток к словам имеет опосредованное значение. Роль автоматического словаря сводится к выполнению следующих задач: содержать ту лексику входного /и выходного/ языка, при помощи которой осуществляется основная масса передачи информации /см. в этом отношении проблематику подязыка<sup>9</sup>/; обеспечить распознавание лексем входного текста; осуществить трансфер /переход/ первого уровня с входного языка на выходной. Последняя задача, т.е. предварительный /частичный/ перевод лексем входного языка на выходной язык, выполняется как раз при использовании информационных пучков.

Как известно, непосредственный перевод /дословный перевод/ одной лексемы, ее замещение лексемой выходного языка представляет собой сущность первых систем машинного перевода. Ошибочность такого упрощенного подхода к проблемам межъязыковых соответствий вскоре осозналась, а впоследствии системы машинного перевода стали применять принцип помещения в процесс перевода некоей промежуточной фазы, в рамках которой происходит "избавление" вход-

ного текста от трудностей, вызванных спецификой входного языка.

Однако в настоящее время наблюдается тенденция к возвращению включения непосредственного перевода в процесс автоматизированного перевода. Это вызвано, с одной стороны, тем, что степень сложности систем автоматических словарей настолько высока, что работа с ними без их упрощения становится весьма медленной. С другой стороны, в процесс перевода желательно ввести по возможности больше контрольных средств для определенной "самопроверки" приемлемости перевода<sup>10</sup>. Одним средством такой проверки может служить проверка по информационным пучкам. Таким образом достигается некоторое убыстрение процесса перевода вместе с большей степенью его адекватности.

Перевод лексический выполняет в системе машинного перевода две функции: во-первых, после распознавания лексем входного языка обеспечивается предварительный перевод; во-вторых, благодаря данному представлению лексики в процесс перевода включается проверочное средство, работающее на основе принципа предпочтения /статистической вероятности/.

3.1.2. На уровне морфологии и информационный пучок выступает как ориентир для морфологического анализа. Анализ ведется с целью определения машиной актуальной структуры состава слова. Так, на основе сведений, полученных от словарных форм лексем, производится сегментация словоупотребления на составные элементы, отделяются корни и аффиксы, определяются типы чередований и т.п. Часть результатов морфологического анализа переходит в трансфер для морфологического порождения выходных словоформ, а другая часть информации морфологического анализа встраивается в информационный пучок и передается на синтаксический анализ. Роль морфологического анализа в случае английского языка не настолько значительна, насколько она важна для генерации венгерских словоформ. Этим обстоятельством объясняется определенное отсутствие баланса между английским и венгерским морфологическими уровнями: малоразвитый в отношении морфологической структуры английский язык на этом уровне дает мало информации по сравнению с той долей информации, которая может быть



извлечена в ходе синтаксического анализа.

По сути дела, морфологический анализ английского языка в целом занимает сравнительно мало места в комплексе анализов на различных уровнях. Объем английской морфологии не превышает размера той части венгерской морфологии, в рамках которой решается проблематика объектного и субъектного спряжения глаголов. Венгерская морфология, наоборот, представляет собой систему с высокой степенью сложности. Она обусловлена не только спецификой агглютинативности венгерского языка, но также и тем, что морфологическая ступень генерации есть последняя фаза перевода, где получается конечная форма выходного текста.

Этим объясняется упомянутое отсутствие баланса, которое налагает большие трудности на конструкцию системы машинного перевода /см. в связи с этим раздел об алгоритмизации. "3.2"/.

3.1.3. У р о в е н ь с и н т а к с и ч е с к и й в случае английского языка предполагает анализ самых разнообразных явлений. На этом уровне вскрывается структура синтагм и предложений. Здесь решаются проблемы распознавания границ синтагм на основе грамматического анализа, являющегося некоторым амалгамом систем Винограда, Маркуса и Н. Сайгер<sup>11</sup>.

Хотя систему синтаксического анализа /и генерации венгерских предложений/ нельзя считать завершенной, относительно основных характерных черт машинного синтаксиса мы уже в состоянии высказаться.

Благодаря большой степени информативности лексем, снабженных в машинном словаре указанным информационным пучком, синтаксический анализ проходит слева направо на основе заложенных в памяти ЭВМ правил синтаксического построения с учетом актуально релевантных пучков. В настоящее время синтаксический анализ распространяется на три слова /или эквивалента слова/, то есть анализ охватывает синтаксические связи не более, чем между тремя словами<sup>12</sup>. Те конструкции, синтаксическая принадлежность которых на основе трехсловных единиц не выявляется, попадают на семантический анализ. Трехсловный анализ, однако, в подавляющем боль-

шинстве случаев оказывается достаточным для определения синтаксических связей внутри синтагм и предложений. Такой подход синтаксического анализа называется "детерминистским" на том основании, что в данном случае нет повторных возвращений к уже проанализированным единицам. Для этого требуется довольно обширное представление возможных комбинаций компонентов синтаксических конструкций. Работа над актуальным рядом слов поэтому базируется, с одной стороны, на детальное описание правил возможных синтаксических построений и, с другой стороны, на информационные пучки.

О приемлемости детерминистского подхода к синтаксическому анализу ведутся споры<sup>13</sup>. Здесь мы ограничимся всего лишь замечанием о том, что детерминистский подход с точки зрения машинного перевода представляется предпочтительным прежде всего потому, что при его применении снимаются затраты времени на различные недетерминистские операции /нет рекурсивных "проскоков", повторных анализов в случае заблуждений из-за омоничности конструкции и т.д./.

Замысел детерминистского подхода в тривиальном случае состоит в том, что при обеспечении полного набора правил, применяемых к описанию синтаксических отношений, пословно проходя по предложению, чем дальше мы отходим от левой границы предложения, тем меньше остается потенциально возможных конструкций, применительно к пройденной и оставшейся частям данного предложения.

При порождении структуры получаемого венгерского предложения используется не только сведение о синтаксисе актуального английского предложения, но также и информация, полученная от последней ступени анализа английских входных предложений об их семантических особенностях.

3.1.4. На уровне семантического анализа извлекается информация о таких релевантных свойствах данного предложения, как прономинализация, восполнение различных типов эллипсисов и т.д. Данные типы анализа можно назвать "синтаксико-семантическими". Их задача заключается в особом рода "постсинтаксическом" анализе, направленном на обеспе-

чении приписания синтаксической структуры таким конструкциям, которые на предыдущем уровне анализа оказались неопределимыми. Так, напр.; предложение

/1/ What did the man and the woman come all this way here for?

/Зачем мужчина и женщина прошли всю эту дорогу сюда?/

содержит две пары слов /обозначенные "——" и "——"/, из которых одна /"——"/ ускользает из "рук" синтаксического анализа. Неспособность синтаксического анализа правильно разобрать данное /"——"/ сочетание объясняется тем, что использованный в ходе анализа структуры предложения "буффер" работает в указанном трехсловном режиме. Из этого следует, что первая пара /"——"/ "распознается" в ходе анализа, тогда как вторая "избегает" отождествления с заданной той или иной структурой /см. схему 2/.

Спецификой семантики данной системы англо-венгерского машинного перевода, однако, является не синтаксико-семантический анализ, а модульная семантическая репрезентация содержания входного предложения. В основе такой репрезентации лежит определенная семантическая запись, при помощи которой как бы переписывается содержание входного предложения в виде модуля или модулей, связываемого или связываемых с другими модулями репрезентации для представления "общего" абстрактного смысла текста.

Стремлений в этом плане повышения качества машинного перевода немало. Имеются попытки перепи́си содержания входного текста в определенный семантический метаязык<sup>14</sup>. Преимущество такого подхода сказывается в том, что на абстрактном уровне осуществляется определенное "понимание" содержания входного текста. Это значит, что в ходе генерации выходного текста в основу данного процесса можно поставить общий абстрактный смысл, который впоследствии выступает средством, управляющим процессом генерации. Далее, из такого подхода следует, что возникает возможность пе-

ревода на любой язык, а эта положительная черта высоко оценивается при конструировании таких систем, которые служат для перевода с нескольких языков на несколько языков<sup>15</sup>.

С другой стороны, моделирование понимания текста, основанное на лексико-грамматическом своеобразии конкретного входного текста, предполагает весьма сложную конструкцию такой системы, требует огромных затрат времени и сил, не говоря о необходимости более глубокого понимания когнитивных процессов, заключенных в лексико-грамматической системе языка.

В свете вышесказанного машинный перевод при его автоматическом осуществлении предстает как комплекс знаний о фактах языков, психологических процессов понимания текстов и представления их в виде переводов, о возможностях и границах использования экспертных систем и применении искусственного интеллекта, а также о связанных с этими задачами вопросах алгоритмизации и программирования.

3.2. Что касается вопросов алгоритмизации процесса машинного перевода, то на данном уровне осуществляется организация работы, предполагающей автоматическое выполнение всех процессов с момента ввода английского текста до получения его венгерского перевода.

Алгоритмизация при машинном переводе проводится в двух планах. Во-первых, создаются алгоритмы тех языковых данных, которые в процессе перевода составляют различные банки данных, содержащие описания языковых фактов в зафиксированном виде. Сюда относятся различные словари с лексическим или лексико-грамматическим материалом, а также самые разнообразные банки правил для анализа и синтеза на различных уровнях реализации машинного перевода. Эти алгоритмы вместе должны снабжать систему перевода полным материалом для реализации всевозможных операций, направленных на "обработку" входного текста с целью получения выходного. Эти операции облекаются в форму таких алгоритмов, которые обеспечивают динамику обработки<sup>16</sup>.

Алгоритмизация в построении систем машинного перевода осно-

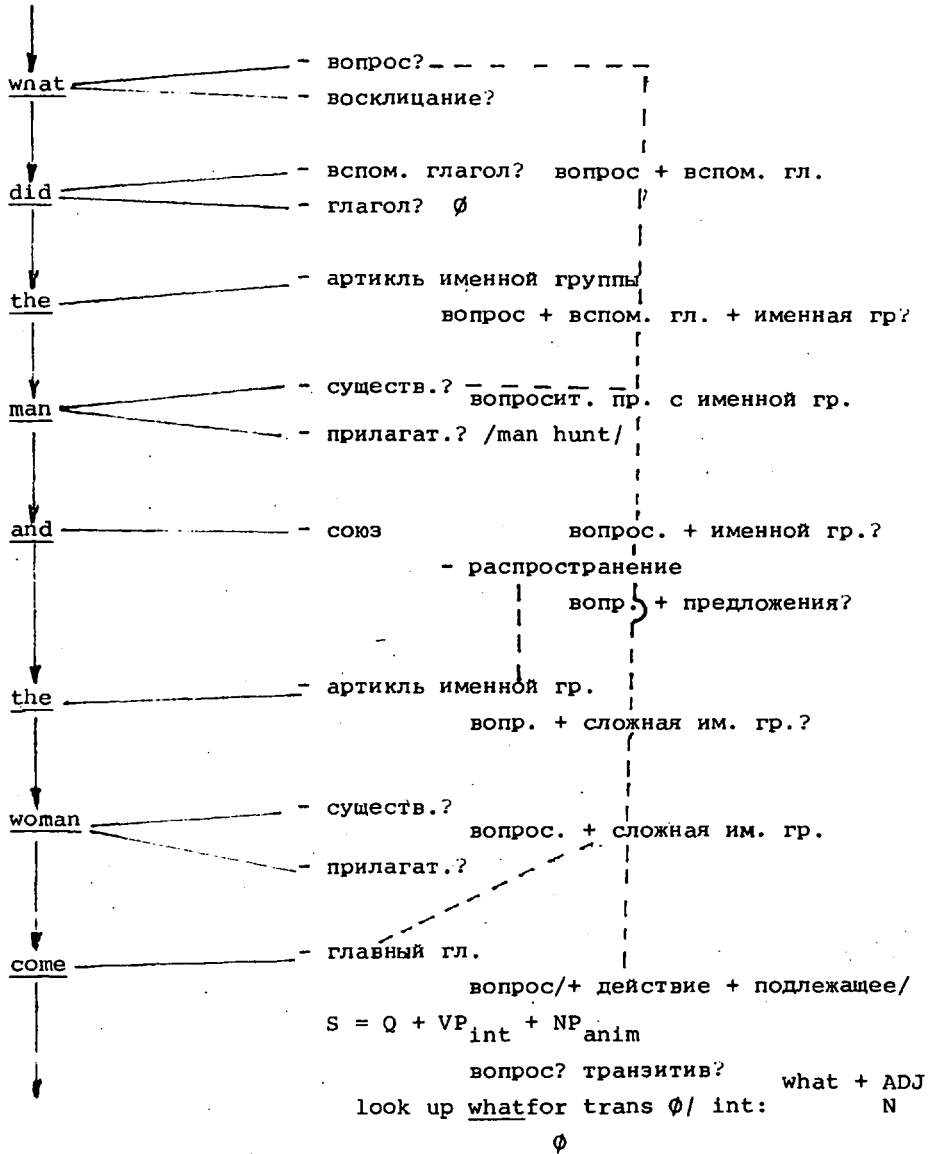
вызывается на стратегии организации комплекса задач, которая /стратегия/ совмещает в себе как элементы эвристики, так и взгляды алгоритмизатора на те или иные лингвистические, психологические, в частности, когнитивные, теории.

Стратегия алгоритмизации предопределяет надежность вырабатываемой системы. Если, например, после предварительных экспериментов на маленьком корпусе система оказывается трудно дополняемой или мало обозримой, а расширение корпуса приводит к неразрешимым противоречиям, то реорганизация системы обусловлена недостатками в выборе подходящей стратегии. Решающим, как кажется, фактором при определении подходящей стратегии является не только /а возможно, и не столько/ правильное надстраивание различных языковых информационных, а также и выбор оптимальной фильтрации этих информационных.

Дело в том, что одной из наиболее существенных трудностей машинного перевода является разделение сведений, полученных о входном тексте, на релевантные и нерелевантные с точки зрения выходного текста. Хотя такое разделение в принципе вполне возможно, на практике оказывается, что релевантность и нерелевантность есть признаки, определяемые не столько априори, сколько "на ходу", т.е. в зависимости от конкретного контекста.

Например, в одних случаях английский порядок слов с точки зрения венгерского перевода является коммуникативно релевантным, а в других случаях он не играет важной с точки зрения перевода роли. Или же набор глагольных форм в одних случаях не имеет значения с точки зрения адекватной передачи содержания, а в других различие, скажем, двух форм настоящего времени /he sees и he is seeing/ необходимо в силу имеющейся разницы в значении двух форм. Перечень таких нечетко определенных границ между релевантными и нерелевантными сведениями можно было бы продолжить. С такими трудностями, думается, можно справиться лишь тогда, если указанное выше "понимание" содержания охватывает не только отдельно взятые предложения, но также и целые куски текста, в рамках которых могут решиться трудности, выходящие за пределы предложения.

Synt. analysis



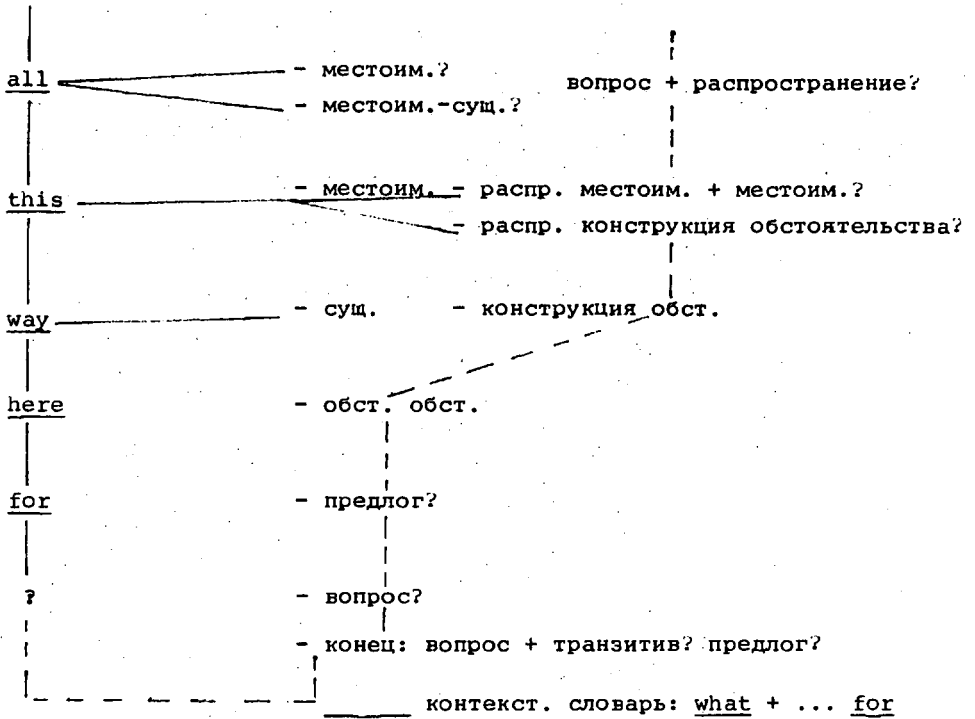


Схема 3. Работа буфера в коде синтаксического анализа

4. Как известно, стопроцентных систем машинного перевода еще не существует. Работающие системы достигают в общем 80--90 % адекватности перевода. Но, естественно, остальные 10--20 % ошибочных переводов нельзя оценить как незначительное количество, ибо даже при нескольких процентах непонятных переводов текста в целом остаются бессодержательными. Когда высказываются мнения по поводу распределения информационной насыщенности текста, согласно которым 60--70--80 % релевантной информации текста отходит на лексический состав предложения, то, по-видимому, не учитывается в полном объеме та роль, которую играют грамматические единицы различных уровней в оформлении смысла текста.

Успешность работы системы машинного перевода в шестидесятих годах рассматривали как зависящую от практического использования лингвистических теорий. Оказалось, однако, что решение проблемы машинного перевода -- это не только практическое применение накопленных наукой о языке знаний. Машинный перевод как часть вычислительной лингвистики представляет собой ту область исследований, в которой интегрируются достижения теоретической лингвистики /и, в частности, переводоведения/, когнитивной науки и искусственного интеллекта. Практическим в своей ориентации машинный перевод является в том отношении, что связанные с ним исследования претендуют на непосредственное применение результатов в производстве.

Практическая установка далее проявляется в том, что эти системы /по крайней мере в настоящее время/ направлены на переводы специальных текстов, прежде всего, научно-технической литературы. Однако, усилия, заложенные в построение такой системы, носят скорее теоретический, нежели практический /прикладной/ характер.



Примечания

1. См. об этом: R. Grishman. Computational Linguistics. An introduction. "Cambridge University Press". 1986. pp. 5--7.
2. См. об этом хотя бы: P.H. Winston. Artificial Intelligence. /Second Edition/. "Addison-Wesley", Reading, Massachusetts, 1984.
3. Примером подхода к синтаксическому анализу со стороны моделирования психологических процессов языкового понимания может служить синтаксический анализатор Маркуса, см.: M. Marcus. A Theory of Syntactic Recognition for Natural Language. "MIT Press", Cambridge, Massachusetts, 1980.
4. После 1966 г., вслед за опубликованием Национальной Академией Наук США доклада "АЛПАК" о нецелесообразности поддержки проектов в области машинного перевода, исследования по указанным темам практически прекратились. Подробнее о влиянии этого доклада см., напр.: Jonathan Slocum. A Survey of Machine Translation: its History, Current Status and Future Prospects. In: Computational Linguistics. Vol. 11, No. 1, 1985. pp. 1--18.
5. См. в этом отношении такие системы, как ЭУРОТРА в Западной Европе, АМПАР, НЕРПА и ФРАП в Москве /Всесоюзный центр переводов/, а также исследования в Японии, связанные с "пятым поколением" электронно-вычислительных машин.
6. См. об этом, напр., систему машинного словаря в проекте автоматического перевода с английского языка на чешский /АПАЧ/ в работе Е. Хайчовой: E. Hajičová. Machine Translation in Czechoslovakia. /A manuscript/. 1986.
7. См. об одном решении проблематики конструкции автоматического словаря: Ю.Д. Апресян. Толково-комбинаторный словарь в формальной модели языка и разрешение многозначности. -- В кн.: Актуальные вопросы практической реализации систем автоматического перевода. Ч. 2. /Отв. ред. В.М. Андрищенко/. "Изд. Московского университета". 1982, с. 108--128.
8. Ю.Н. Марчук /составитель/. Контекстологический словарь для машинного перевода многозначных слов с английского языка на русский. М., 1976. Ч. 1--2.
9. Проблематика подязыка рассматривается в сборнике: Sublanguage: Studies of Language in Restricted Semantic Domains. /Ed. by R. Kittredge, J.L. Lehrberger. "de Gruyter", Berlin, New York, 1982.
10. См. Хайчова, ук. соч.
11. См. упомянутый в Примечании 3 синтаксический анализатор М. Маркуса, а также: N. Sager. Natural Language Information Processing. "Addison-Wesley", Reading, Massachusetts, 1981; T. Winograd. Language as a Cognitive Process. Volume I. Syntax. "Addison-Wesley", Reading, Massachusetts, 1983.

12. Подобная стратегия разработана у Маркуса, см. ук. соч.
13. Теория М. Маркуса неоднократно подвергалась критике. Пожалуй, наиболее обоснованными можно считать критические замечания Сампсона в кн.: Parsing Natural language. /Ed. by M. King/. "Academic Press", London, New York etc. 1983. G.R. Sampson. Deterministic Parsing. pp. 91--116.
14. В специальной литературе на русском языке сравнительно полное описание метаязыка читаем в работе: Н.Н. Леонтьева, Е.В. Урысон. Алгоритм построения информационной записи текста. 1 этап. -- НТИ, сер. 2, 1973, № 12.
15. Такая многоязыковая система машинного перевода строится в Западной Европе под названием "ЭУРОТРА", которая включает все 9 официальных языков Европейского экономического сообщества. См., напр.: R. Johnson, M. King, L. des Tombe. EUROTRA: a Multilingual System under Development. -- In: Computational Linguistics, Volume 11, No 2--3, 1985, pp. 155--169.
16. Динамика в архитектуре системы машинного перевода АМПАР описана в книгах Ю.Н. Марчука: Ю.Н. Марчук. Проблемы машинного перевода. "Наука", М., 1983; его же. Методы моделирования перевода. "Наука", М., 1985.