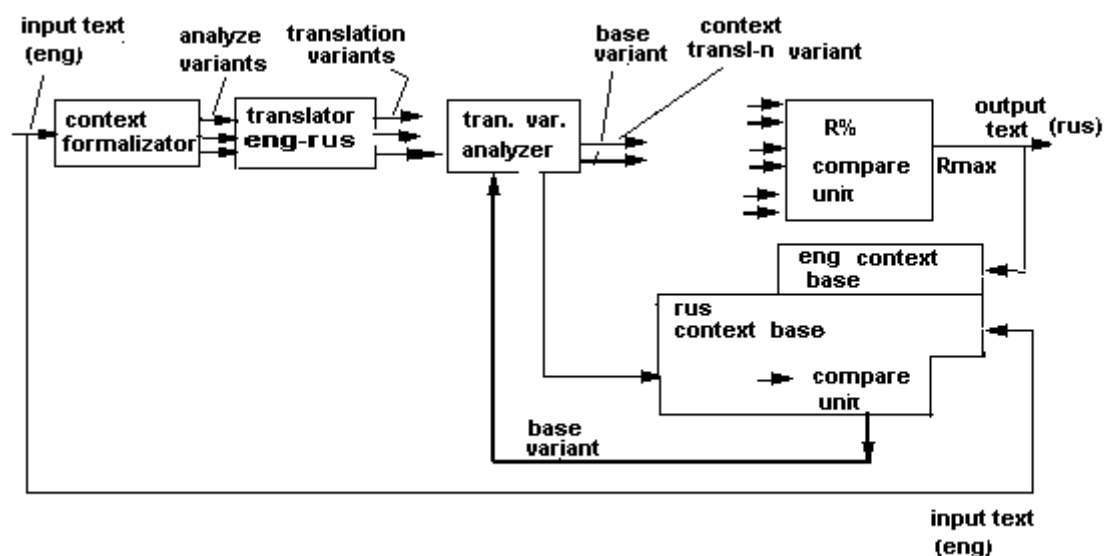


# Варианты разрешения неопределенностей при трансляции текста

Неопределенности обуславливаются:

- а) различной схемой разбивки контекста на группы
- б) различными вариантами трансляции какой-либо группы
- в) различными вариантами трансляции какого-либо элемента группы (например, слова).



Входной текст поступает в Формализатор контекста.

Этот блок:

разбивает контекст на части (например, на группы подлежащего, сказуемого и т.д.) по различным вариантам.

A few of samples from real contexts

(несколько примеров из реального контекста):

Sample 1

input: **Maximum permitted rated mains voltages.**

Translation variants

Output (var1) : Максимум разрешил расчетные напряжения сети

Output (var2): Максимальные разрешенные расчетные напряжения сети

Output (var3): Максимум разрешенный рассчитал напряжения сети

Формализатор блока дробит input на куски и выдает в транслятор варианты анализа.

В данном случае формализатор оперирует следующими лингво-признаками: элемент стоящий на первом месте может быть подлежащим (1), определяющим словом (2), определяемым словом (3).

В (1) ФК присваивает этому элементу категорию существительного и т.д.

Sample 2) **Manufacturers of machinery, boilers and auxiliary equipment must supply a sufficient number of operating and maintenance notices and manuals together with the**

**equipment.**

Sample 3) **The flash point of liquid fuels for the operation of boilers and diesel engines may not be lower than 60 °C.**

Sample 4) **In exceptional cases, for ships intended for operation in limited geographical areas or where special precautions subject to the Society's approval are taken, fuels with flash points between 43 °C and 60 °C may also be used.**

Из ФК формализованный вариант  $i$  (у нас 1,2,3) поступает в Транслятор (такой как, например, наш старина Сократ), а из него вариант трансляции залезает в Анализатор вариантов трансляции (АВТ). АВТ передает анализ варианта = русский вариант трансляции в Блок сравнения, который «накладывает» данный вариант на «клише» из Базы Контекстов.

При их совпадении на более, чем  $R\%$  (к примеру 70%) выходной сигнал - вариант трансляции (rus) выдается «на гора» (экран, принтер, винчестер), а также в БД контекстов. При вводе в БД (записи в память) вых сигнал записывается в Rus блок, а входной сигнал (input text) в Eng блок. БД может пополняться также в интерактивном режиме с участием эксперта контекстов. Ее предварительная «накачка» производится на представительном массиве контекстов по экспертно определенной тематике.

В нашем случае

(var2): Максимальные разрешенные расчетные напряжения сети

при трансляции в БД может оказаться клише:

Максимальное разрешенное расчетное напряжение сети

Блок сравнения должен присвоить варианту (2) 100% сходство с клише.

При этом он «знает» о синтаксической структуре - данного варианта от (АВК) и клише от (БД).

Поэтому при сравнении разница в окончаниях во внимание не принимается. Структура «клише» и варианта анализа совпадают на 100%.  $Rstr=100$ .

(var3): Максимум разрешенный рассчитал напряжения сети «клише» : Максимальное разрешенное расчетное напряжение сети

Структура клише и варианта не совпадают:

разные элементы являются подлежащим. Местное совпадение: на контексте - напряжения сети - ( $Rloc=100$ , параметр совпадения на локальном контексте).

(var1) : Максимум разрешил расчетные напряжения сети

Местное совпадение: на контексте - расчетное напряжение сети

$Rloc=100$ , но протяженность контекста  $L1>L3$  поэтому (var1) предпочтительнее чем (var3).

По параметру совпадения структуры наиболее предпочтителен (2)

По параметру совпадения контекста наиболее предпочтителен (2)

Sample2

**Manufacturers of machinery, boilers and auxiliary equipment must supply a sufficient number of \ambiguity1\ operating and maintenance notices and manuals together with the equipment \ambiguity2\ .**

(1) Производители машинного оборудования, котлы и дополнительное оборудование должны поставлять достаточное количество ... \неопределенность 1 \

(2) Производители машинного оборудования, котлов и дополнительного оборудования должны поставлять достаточное количество ... \неопределенность 1 \

В БД могут оказаться клише :

Производители машинного оборудования

Производители котлов

Производители дополнительного оборудования

К БД в качестве опции, выбираемой экспертом перед автоматической трансляцией, можно прицепить «аппендикс» - Анализатор Смысла с включением таких процедур как: (,) = + ; and = +.

После прохождения через АС контекст (1) должен получить приоритет относительно контекста (2). Алгоритм следует разрабатывать.

Но пока нам не до аппендиксов. Все строится на формализации контекстов и сравнении их с набором базовых.

Итак «к нашим баранам». В (1) и (2) разнятся группы подлежащего.

В(1) группа подлежащего делится на несколько самостоятельных элементов. В (2) в группе подлежащего один основной элемент с его определениями.

**Sample 3) The flash point of liquid fuels for the operation \ambiguity 1 \ of boilers and diesel engines may not be lower than 60 °C.**

Var1) В контексте 1 \ fuels = сказуемое

Var2) В контексте 1 \ fuels = входит в группу подлежащего после прохождения через Блок Сравнения контекста 1 вероятно, что трансляция по var2 получит приоритет из-за наличия в БК образцов : eng- liquid fuel и rus - жидкое топливо. Для flash point - температура воспламенения, также найдутся образцы в БД.

**Sample 4) In exceptional cases, \1 for ships intended for operation in limited geographical areas or where special precautions subject to the Society's approval are taken, fuels with flash points between 43 °C and 60 °C may also be used.**

**Sample 5) The formulae for plate panels subjected to lateral pressure \1 as given in the following Sections are based on the assumption of an un-curved plate panel having an aspect ratio b/a Æ 2,24.**

Context1 : The formulae for plate panels subjected to lateral pressure

var1) subjected = сказуемое

var2) subjected = определение

Translate var1) Формулы для плоских панелей подвергли боковому давлению...

Translate var2) Формулы для плоских панелей подвергаемых боковому давлению...

*Sample of ambiguity when an infinitive constructions translating:*

(неопределенности при трансляции инфинитивных конструкций)

**If the reason for overhaul is to renew the column bearings, insert the ignition key and turn it to release the steering column lock.**

1) to renew = инфинитив

2) is to renew = оборот долженствования

Варианты трансляции

1) Если причина ремонта - восстановить подшипники колонки,...

2) Если причина ремонта должна восстановить подшипники колонки,...

После прохода через механизм трансляции, изображенный на схеме, вар2) получит приоритет, если БД содержит клише:

Если причина ремонта - ...

*Sample of ambiguity when a gerundive constructions translating:*

(неопределенности при трансляции конструкций на основе герундия)

### **Fit the bearing securing clips.**

MT Translating: Монтируйте подшипник фиксируя скобы.

Надо : Монтируйте (установите) стопорные зажимы (кольца) подшипника

Bravo the Program, but what about the sense?

Поможет ли здесь База Знаний а la CYCORP? В обоих вариантах нет претензий по поводу смысла. Только некая БД (в голове или на винчестере) содержит «клише», накапливаемые статистически по мере чтения текстов.

Неопределенности последнего контекста не разрешаются с помощью Базы Знаний типа

Сусорг, [Cycorp Knowledge Base](#),

а также по предлагаемому методу, включающему многовариантную трансляцию контекста с последующим сравнением транслированных вариантов и «клише» из Базы Контекстов.

*Sample of ambiguity when translating of sense*

(неопределенности при трансляции смысла)

### **First check that the insulation is sound.**

Сначала контроль, что изоляция является звуком. (вариант перевода программой Socrat)

Если БД содержит клише : eng - insulation is sound и rus - изоляция в исправности, неопределенность разрешается.

**This newly defined technology embraces/1 these developing substrate technologies for their ability to enable standardization at competitive cost using packaged rather than bare die.**

Это по-иному определило технологические объятия...

Эта недавно определенная технология обнимает...