

УДК 81'322  
ББК 81.184  
Х82

Рецензенты:

заслуж. деятель науки РФ, проф., докт. филол. наук  
Р. С. Гиляревский  
проф., докт. техн. наук В. А. Цветкова  
Научный редактор – заслуженный деятель науки РФ  
доктор технических наук, профессор К.К. Колин

**Хорошилов, Ал-др А.**

Х82 Фразеологический машинный перевод текстов.  
Теоретические основы и технологические решения /  
Ал-др А. Хорошилов, А. В. Кан, Ал-ей А. Хорошилов ; под  
научной ред. К. К. Колина. – Москва : Берлин : Директ-Медиа,  
2019. – 466 с. DOI: 10.23681/563869

ISBN 978-5-4499-0089-0

В монографии рассматриваются теоретические основы и технологические решения в области разработки систем машинного перевода текстов нового поколения – систем фразеологического машинного перевода. Концепция создания таких систем была предложена проф. Г. Г. Белоноговым в 1975 г. в его предисловии к книге Д. А. Жукова «Мы - переводчики». В рамках этой концепции впервые было сформулировано утверждение, что в качестве основных единиц смысла в словарях систем машинного перевода должны использоваться наименования понятий - фразеологические словосочетания. В монографии описаны основные принципы и технологии создания и функционирования таких систем, позволяющих получать высокое качество перевода текстов. Наряду с техническими решениями, в ней также рассмотрены процессы управления функционированием системы, ее настройки на конкретные предметные области и взаимодействия пользователей с системой в процессе диалогового общения.

Монография рассчитана на широкий круг научных сотрудников и специалистов в области научно-технической информации, компьютерной лингвистики и информационных технологий, а также преподавателей, студентов и аспирантов системы высшего образования.

УДК 81'322  
ББК 81.184

ISBN 978-5-4499-0089-0 © Хорошилов Ал-др А., Кан А. В., Хорошилов Ал-ей А.,  
2019  
© Издательство «Директ-Медиа», оформление, 2019

# Содержание

Высокое качество перевода текстов – необходимое условие развития современного общества и эффективного международного научно-технического сотрудничества .....	9
Предисловие .....	14
Введение .....	16
Глава 1. Единицы языка и речи в системах автоматической обработки текстовой информации.....	20
Глава 2. Концепция фразеологического машинного перевода текстов с одних естественных языков на другие .....	41
2.1. Предварительные замечания.....	41
2.2. Концепция фразеологического машинного перевода текстов .....	49
2.3. Пути реализации концепции фразеологического машинного перевода .....	51
Выводы.....	53
Глава 3. Общие принципы создания и функционирования систем ФМП .....	56
3.1 Основные подходы к решению проблемы машинного перевода .....	56
3.2 Общий порядок функционирования систем ФМП .....	58
3.3 Первоначальный процесс создания системы ФМП .....	63
3.4 Общая базовая архитектура системы ФМП .....	65
Выводы.....	72

Глава 4. Машинные грамматики естественных языков .....	73
4.1 Общие принципы создания машинных грамматик .....	73
4.2. Машинная грамматика русского языка .....	77
4.3 Машинная грамматика языков стран Западной Европы и Юго-Восточной Азии .....	111
Выводы.....	138
Глава 5. Семантико-синтаксический анализ текстов.....	140
5.1 Синтаксическая модель русского языка на основе обобщенных синтагм .....	140
5.2 Разрешение грамматической неоднозначности английских слов с помощью метода аналогии .....	147
5.3 Разрешение грамматической омонимии русских слов .....	154
5.4 Общие принципы построения процедур семантико-синтаксического анализа текстов .....	157
5.5 Принципы синтаксического анализа текстов на основе использования синтаксических правил.....	161
5.6 Процедура синтаксического анализа русских текстов на основе синтаксических правил.....	165
5.7 Процедура синтаксического анализа английских текстов, основанного на синтаксических правилах.....	169
5.8 Автоматическое установление структурного сходства предложений .....	171
5.9 Построение процедуры синтаксического анализа английских текстов на основе обобщенных синтагм .....	175

Выводы.....	187
Глава 6. Концептуальный анализ текстов .....	189
6.1 Общие принципы реализации процедуры концептуального анализа текстов .....	189
6.2 Концептуальный анализ с контролем по словарю ЭКС.....	191
6.3 Концептуальный анализ текстов на основе «логической шкалы» эталонного словаря .....	196
6.4 Концептуальный анализ текстов на основе синтаксических структур эталонного словаря .....	200
6.5. Концептуальный анализ на основе обобщенных синтагм .....	209
6.6 Сравнительный анализ частотных словарей, полученных различными методами .....	215
Выводы.....	219
Глава 7. Трансфер.....	221
7.1 Основные принципы реализации процедуры трансфера .....	221
7.2 Структура и содержание двуязычных словарей системы ФМП .....	223
7.3 Преобразование текстового представления текста в совокупность поисковых представлений фрагментов текста.....	233
7.4 Поиск переводных соответствий фрагментов исходного текста в комплексе словарей системы ФМП ....	234

7.5 Выбор приоритетных переводных соответствий для фрагментов исходного текста.....	237
7.6 Соотнесение исходных текстовых фрагментов с их приоритетными переводными соответствиями .....	239
Выводы.....	243
Глава 8. Семантико-синтаксический синтез текстов .....	244
8.1 Принципы реализации процедуры синтаксического синтеза текстов .....	244
8.2 Синтаксический синтез глагольных словосочетаний.....	247
8.3 Локальный синтаксический синтез словосочетаний.....	251
8.4 Семантико-синтаксический синтез переводного предложения .....	256
8.5 Перевод текстов на основе модели обобщенных синтагм .....	259
8.6 Перевод текстов на основе установления смыслового сходства синтаксических конструкций предложений.....	274
Выводы.....	277
Глава 9. Технологии создания декларативных средств для системы ФМП.....	279
9.1 Предварительные замечания.....	279
9.2 Исследования тематических реферативных баз данных ВИНТИ .....	280

9.3. Автоматическое составление словарей наименований понятий без контроля по тезаурусу.....	293
9.4 Технологии составление словарей по отраслевому корпусу текстов .....	299
9.5 Автоматизация составления словарей по параллельным двуязычным текстам .....	333
9.6 Общий порядок автоматизированного составления и ведения фразеологических машинных словарей.....	325
Выводы.....	330
Глава 10. Технологии функционирование систем ФМП .....	331
10.1 Основные виды переводческой деятельности.....	331
10.2. Модификации системы ФМП.....	332
10.3. Лингвистическое обеспечение систем ФМП .....	345
10.4. Порядок работы лингвиста-переводчика в системе ФМП .....	347
Глава 11. Технологии создания новых направлений перевода для системы ФМП.....	388
11.1. Возможности создания универсального мультязычного переводчика .....	388
11.2. Разработка технологий создания новых направлений перевода системы ФМП.....	394
Выводы.....	410
Глава 12. Опыт эксплуатации модификаций систем ФМП .....	411
12.1. Начальный этап разработки системы .....	411
12.2. Дальнейшее развитие системы RETRANS.....	419

12.3. Отраслевые модификации системы ФМП .....	421
12.4. Программно-лингвистическая платформа	
MetaFraz .....	427
Выводы.....	443
Заключение .....	445
Литература .....	450