

Частное образовательное учреждение
высшего образования
«ИНСТИТУТ ИНОСТРАННЫХ ЯЗЫКОВ»

Специфика современного машинного перевода

Антипова Вероника Алексеевна

Выпускная квалификационная работа бакалавра

**Научный руководитель
Кузьмич И.В.**

Санкт-Петербург

2021

Содержание

	Введение	3
	Глава 1. Теоретические основы машинного перевода	6
1.1	История машинного перевода как научно-прикладного направления переводоведения	6
1.2	Сущность цели и задачи машинного перевода в переводоведении	12
1.3	Классификация систем машинного перевода, его схемы и общий алгоритм действия	14
	Выводы по главе 1	23
	Глава 2. Особенности научно-технических текстов и их перевода	25
2.1	Сравнительный анализ перевода научно-технического текста, выполненного с помощью двух систем машинного перевода	25
2.2	Анализ адекватности перевода научно-технического текста	32
2.3	Анализ адекватности перевода научно-технического текста	43
	Заключение	48
	Список использованных источников	52
	Приложение №1	57
	Приложение №2	63

Введение

Общеизвестно, что наше время характеризуется глобальной интеграцией всех стран мира. Это ведёт к увеличению необходимости перевода языковых потоков в режиме реального времени. Более того, в современном мире перевод всё больше уходит от исключительно творческого характера и дополнительно приобретает характер ремесла и даже массовой индустрии. В первую очередь это относится к переводу технических и юридических (и частично научных) текстов. Назовём этот тип перевода «промышленным». Как любой индустрии, промышленному переводу необходимы адекватные инструменты.

Объектом исследования выступает процесс машинного (автоматизированного) перевода.

Предметом исследования были выбраны особенности передачи научно-технических текстов с ИЯ на ПЯ с использованием систем машинного перевода PROMT и Google, на примере статьи О. Г. Василенко «Рекуррентная Каскад-Корреляция обучающей архитектуры» [51] («The Recurrent Cascade-Correlation Architecture» [53]). Данный текст был выбран из-за его актуальности для специалистов исследуемой в нем отрасли, а также для проверки тезиса «Соответствует ли уровень адекватности перевода МП переводу человека». Наука не стоит на месте и развивается, потому было принято решение провести анализ между современными машинными переводчиками и переводом О.Г. Василенко от 1991 года, чтобы была возможность прийти к выводу о том, достаточно ли были улучшены программы сервисов МП, что достичь того уровня перевода, который был у человека в 1991 году.

Актуальность темы исследования определяется возрастающей сложностью современных коммуникативных процессов и необходимостью взаимного понимания участников общения на уровне межличностной, внутрикультурной и межкультурной коммуникации.

Также актуальность обусловлена тем, что машинный перевод - это эффективное средство для просмотра и поиска информации на иностранном языке, а именно эта функция является главной при работе в Интернете. Интернет технологии дали новое развитие машинному переводу, помогли вывести его на новую стадию развития. Современное состояние машинного перевода позволяет получать относительно корректный текст перевода с большинства языков. И хотя полностью автоматический высококачественный перевод невозможен, однако, имеется программное обеспечение, которое облегчает сам процесс перевода. В результате настройки на предметную область и интеграции с другими программами обработки документов средство машинного перевода позволяет автоматизировать получение перевода.

Сейчас, когда для перевода текстов с одного языка на другой компьютер используется самыми разными способами – от двуязычных и многоязычных электронных словарей до систем типа «память», или «архив» переводов translation memory, этот вопрос оказывается не таким уж простым. Мы будем понимать машинный перевод как процесс, при котором компьютер по заданному тексту на одном языке производит новый текст на другом языке, которого раньше в этом компьютере не было: понятно, что ни словари, ни архивы переводов таким свойством не обладают.

Цель данной работы – провести комплексный анализ особенностей передачи научно-технического текста с ИЯ на ПЯ с использованием систем машинного перевода. В процессе исследования нами были поставлены следующие **задачи**:

1. Выявить сущность, определить цели и задачи машинного перевода в переводоведении;
2. Привести и проанализировать классификацию систем машинного перевода, его схемы и общий алгоритм действия;
3. Проанализировать особенности научно-технических текстов и их перевода;

4. Провести сравнительный анализ и анализ адекватности перевода научно-технического текста, выполненного с помощью двух систем машинного перевода (PROMT и Google).

5. Сравнить образцы перевода с переводом, сделанным О. Г. Василенко. Анализ причин расхождения в итоговом тексте на ПЯ.

Характеристика источников. Для данного дипломного исследования использовались основополагающие труды отечественных и зарубежных лингвистов и переводчиков, рассматривающие как общие вопросы переводоведения, так и проблемы машинного (автоматизированного) перевода, а также статьи действующих переводчиков и преподавателей данного направления.

Методология исследования. В выпускной квалификационной работе использованы следующие методы исследования: теоретические (анализ) и эмпирические (сравнение, изучение литературы и результатов сравнения).

Научная новизна исследования заключается в анализе сравнения перевода одной статьи, выполненного в разные года человеком и машиной, для получения актуальной информации об уровне МП на данный момент.

Теоретическая значимость работы состоит в дополнении опыта предыдущих исследователей темы машинного перевода.

Практическая значимость состоит в том, что результаты исследовательской части могут быть использованы при программировании в настройке МП для повышения уровня перевода.

Структура. Дипломная работа состоит из введения, двух глав, заключения, списка литературы и двух приложений. В первой главе будут рассмотрены теоретические вопросы машинного перевода как научно-прикладного направления переводоведения. Вторая (практическая) глава посвящена сравнительному анализу перевода научно-технического текста, выполненного с помощью двух систем машинного перевода. В приложениях представлена рассматриваемая в работе статья на русском и английском языках.

Глава 1. Теоретические основы машинного перевода

1.1. История машинного перевода как научно-прикладного направления переводоведения

Термин «машинный перевод» может иметь два значения. В узком смысле - это процесс перевода текста с одного естественного языка на другой, который полностью или почти полностью реализуется компьютером. В ходе этого процесса на вход машины подается текст, который не сопровождается какими-либо дополнительными инструкциями, а на выходе получается текст на другом языке, который является переводом ввода и преобразованием. Вывод происходит без вмешательства человека (иногда допускается постредактирование).

Машинный перевод в самом широком смысле - это область исследований на стыке таких наук, как лингвистика, математика и кибернетика, которая направлена на создание систем, реализующих машинный перевод в узком смысле [52, с. 3].

Впервые идея расширения границ перевода привычных черт появилась у английского изобретателя, знаменитого математика Чарльза Бэббиджа в конце 1830-х годов. Это был проект первого в истории компьютера - «аналитической машины». Суть работы будущей машины заключалась в использовании потенциала машинной памяти для применения словарей. Однако идея ученого так и не была реализована.

Развитие машинного перевода и его современную картину можно представить как взаимодействие, борьбу двух направлений, двух основных подходов к проблеме [19, с. 35 - 36].

Особенностью первого подхода является ориентация на использование наиболее мощного универсального смыслового языка. Первая система реализовывала «прямой» подход к переводу, при котором все операции, выполняемые при переводе, рассматривались как операция межъязыкового

перехода - прямое преобразование текста на исходном языке в текстовый перевод.

В машинном переводе, в собственном смысле этого слова, такой подход приводит к так называемым «тотальным стратегиям» [21, с. 96]. Общей характеристикой этих стратегий является стремление получить полностью автоматизированный высококачественный машинный перевод за счет максимального использования семантического уровня языка. Этот подход внес значительный вклад как в теорию, так и в практику машинного перевода.

Хронологически второй подход появился раньше первого. Он основан на идее промежуточного языка и тесно связан с идеей переводческих соответствий на чисто лингвистическом уровне. Методологическая сущность этого подхода заключается: в изучении поведения языковых единиц, особенно в лингвистическом контексте; в моделировании уровня владения человеческим языком, особенно в процессе межъязыкового перевода; при переходе от простого к более сложному.

В конце 40-х - начале 50-х годов, после появления первых компьютеров (ЭВМ), начались серьезные работы в области машинного перевода. Исследователи считают, что датой зарождения машинного перевода (как области исследований) считается 1947 год, когда в послевоенное время задачу перевода в первую очередь сравнивали с задачей расшифровки текстов. Последнее в то время уже начали проводить на компьютере [6, с. 17]. Директор естественных наук «Фонда Рокфеллера» (американский благотворительный фонд) Уоррен Уивер написал статью, в которой проблема перевода текста с одного языка на другой определена как еще одна область прикладной технологии дешифрования. Идея машинного перевода вызвала бурную дискуссию, сразу за которой последовала теоретическая разработка первых технологий. Ходили слухи о полной замене переводчика-человека электронными системами, что вызвало обеспокоенность профессиональных переводчиков.

В 1954 г. состоялся так называемый «джорджтаунский эксперимент». Общественности был представлен первый электронный переводчик - русско-английская система под названием IBM Mark II, содержащая словарь на 250 единиц и 6 грамматических правил [1, с. 112]. В течение следующих 10 лет технологии машинного перевода (здесь и далее МП) продолжали быстро развиваться, но вскоре стало ясно, что машина не сможет переводить тексты так же правильно, как человек. В 1967 г. члены-корреспонденты Национальной академии наук США выступили с докладом, в которой представили данные о нерентабельности машинного труда, исходя из реальной ситуации в США и показателей стоимости различных методов перевода. По результатам этого выступления было решено прекратить дальнейшее финансирование исследований в области автоматического перевода, что на долгие годы затормозило развитие МП. Тем не менее, несмотря на официальный отказ от разработки новых технологий перевода, первые системы перевода продолжали пользоваться спросом в научных и военных учреждениях США и СССР.

Принципы машинного перевода, существовавшие в то время, с современных позиций можно было охарактеризовать как обработку текста с помощью словаря [23, с. 18], которая осуществляется с помощью определенных специфических процедур, полностью подчинена задаче перехода с входного языка на выходной (т. е. не предусматривает каких-либо дополнительных этапов, функции которых были бы независимыми и могли бы определяться отдельно от этого межъязыкового перехода).

Задача МП тогда ставилась как чисто практическая [27, с. 31]. Предложения на английском языке преобразовывались в предложение на другом языке с помощью двух основных операций. Первая операция - отдельные слова заменяются их собственными переводами, вторая - переведенные слова переставляются и модифицируются в соответствии с правилами переводимого языка.

Предполагалось, что для научно-технических текстов быстрее всего будет спроектировать машинные системы, обеспечивающие их перевод в промышленных масштабах. Однако данные ожидания не могли быть оправданы, поскольку основывались на двух фундаментальных заблуждениях.

- Была переоценена степень схожести тех операций, которые могли быть выполнены на компьютере, к элементарным функциям человеческого интеллекта.

- Недооценивалась изначальная фундаментальная сложность естественного языка как инструмента интеллектуальной деятельности человека и перевода как одного из самых сложных видов этой деятельности [47, с. 105].

На ранних этапах исследованиях МП авторы в первую очередь сосредоточили внимание на технических проблемах размещения большого словаря в памяти компьютера и возможности обеспечения в нем эффективного поиска. Большое количество исследователей, например, такие как Ю. Н. Марчук, Л. Л. Нелюбин, И. И. Ревзин, считали, что программное обеспечение для автоматизированного перевода с грамматикой основано на существовавших на тот момент теориях строения языка в сочетании с изобретенными в спешке правилами [25 - 28, 32, 41].

Перевод, произведенный разработанными программами, был настолько плохим, что его было совершенно невозможно понять. Проблема была в том, что естественное языковое значение текста зависит не только от самого предложения, но и от общего контекста.

Скорее всего, эти первые проекты действительно не дали реальных результатов. Однако было выявлено множество основных проблем перевода текстов на естественный язык: «многозначность слов и синтаксических конструкций, практическая невозможность глобального описания семантической структуры мира даже в ограниченной предметной области,

отсутствие эффективных формальных методов». для описания языковых образов» и др. [41, с. 43].

В дополнение к системам машинного перевода, которые уже упоминались выше, были также созданы более частные «станки» для автоматизации отдельных операций перевода, которые могли помочь переводчику и редактору: автоматические словари и терминологические базы данных, компьютерные тезаурусы, инструменты редактирования экрана, системы проверки орфографии, терминологическая и грамматическая коррекция текстов.

Особое внимание было уделено новому разделу лексикографии для построения электронных словарей, и уже сформировались представления об алгоритмах морфологического, синтаксического и лексического анализа для МП [19].

Наиболее плодотворным следует считать второй период в истории развития машинного перевода - это период с 1960 по 1967 годы. Именно в этот период началось активное развитие в области машинного перевода, которое впоследствии привело к тому, что машинный перевод стал представлять собой конгломерат всех видов наук, не только лингвистического порядка, но в первую очередь связанных с компьютерами. Следует обратить особое внимание на появление идеи промежуточного языка [47, с. 10].

Также стоит отметить, что именно в этот период возникла идея создания машинного словаря в помощь переводчику-человеку. Машинный словарь нашел большое практическое применение. В конце 60-х годов А. К. Жолковский и И. А. Мельчук подняли тему центральной роли словаря в описании языка [11, с. 42].

Следующий этап развития технологий МП приходится на конец 70-х - начало 80-х годов. XX и связан с появлением новых вычислительных технологий. На этом этапе удалось успешно совместить новые достижения технической кибернетики с новыми результатами теории и практики

лингвистических исследований в области автоматической обработки текстов. Этот союз возник благодаря осознанию социальных потребностей в области компьютерного перевода - преодоления языковых барьеров [19]. В рамках этих исследований человек отошел на второй план. Программисты отказались от идеи и создания «идеальной» машины-переводчика: все последующие новые системы разрабатывались с целью многократного увеличения скорости трансляции информации, но, конечно, с участием человека для получения идеальной качества окончательного текста.

Новый этап исследований приходится на 90-е гг. Прошлый век ассоциируется с невероятным прогрессом персональных компьютеров и появлением сети Интернет.

Текущее состояние машинного перевода характеризуется некоторым слиянием результатов двух подходов. Однако это не простая механическая связь результатов, а объединение их на основе новых моделей и т.д., причем основное внимание уделяется трансляционному аспекту владения естественным языком. В современном МП фокус исследований - моделирование действий переводчика-человека, особенно с точки зрения использования им двусмысленных и многозначных переводческих соответствий при переводе с исходного языка на переведенный. Процесс перевода очень сложен, а правильное использование преимуществ программного обеспечения не в последнюю очередь определяет качество перевода.

Возможности перевода в режиме онлайн позволяют преодолевать языковой барьер и осуществлять бесплатную навигацию по иностранным сайтам и общение с иностранцами.

На данный момент есть две технологии машинного перевода информации: Machine Translation и Translation Memory. Технология ТМ подробно описана ниже в одноименном разделе.

Программы, разработанные на основе технологии машинного перевода (Machine Translation), выполняют последовательный перевод текста с

использованием определенных лингвистических алгоритмов. Прежде всего, система анализирует структурные элементы входного предложения, затем преобразует его в соответствии со структурой целевого языка и синтезирует окончательный вариант. Программа также должна уметь распознавать фиксированные выражения и фразеологизмы, а также иметь большой словарный запас для повышения качества перевода. При переводе тематических текстов обычно требуется подключение специализированных словарей. С помощью систем с технологией машинного перевода вы можете мгновенно получить черновик перевода, что весьма полезно, когда вам нужно быстро понять общий смысл текста.

1.2. Сущность, цели и задачи машинного перевода в переводоведении

Необходимость систем машинной обработки текстов напрямую связана с большим ростом объема электронной текстовой информации, данные из которой необходимо извлекать в любой момент и по возможности полностью автономными процедурами. Однако развитие МП – активно развивающейся области научных исследований и разработок СМП (систем машинного перевода), где в процессе перевода с одного естественного языка на другой используется ЭВМ – определяется как потребностями делового мира в качественных машинных переводчиках, так и чисто научными стимулами.

На базе экспериментальных стабильных МП проводятся проверки различных аспектов общей теории понимания, речевого общения, преобразования информации, а благодаря им возможно создание новых, более эффективных моделей самого МП.

Разработка технологии автоматизированного перевода с иностранного языка на русский и обратно является логичным продолжением исследований, проводимых в области искусственного интеллекта и лингвистического анализа текстов. Российскими учеными проводятся активные работы в данной научной сфере. Разработанные ими в лексикографии уникальные

методы описания словарей, лингвистической структуры языка, схем семантико-синтаксических конструкций, а также мощный инструментарий обработки и преобразования словарей и текстовых данных позволяют строить системы перевода для различных иностранных языков разного уровня сложности [32, с. 50].

В тоже время, в прикладных задачах, связанных с созданием разнообразных информационных систем и/или систем для работы с документами разной структуры часто возникает необходимость применения методов автоматизированного перевода текстов для обработки потоков информации, поступающей и/или исходящей на иностранном языке. Например, применение инновационных технологий МП и многоязыкового семантического анализа информационных потоков поможет сильно повысить информационное обеспечение управления крупными и средними предприятиями.

Задача любого переводчика состоит в том, чтобы передать смысл текста (письменного или устного) на исходном языке средствами переводного языка. Задача машинного перевода заключается в том же.

Сам по себе процесс перевода состоит из нескольких частей. Когда переводчик (или компьютер) получает подлежащий переводу текст, он, прежде всего, должен этот текст понять, т.е. постичь его смысл. Этот смысл надо впоследствии облечь в слова, т.е. создать новый текст. Таким образом, перевод складывается из двух последовательных задач: задачи понимания текста и задачи производства текста [12, с. 59].

В жизни каждый человек при общении с другими людьми попеременно оказывается то слушающим, то говорящим, ему и приходится решать как задачу понимания, так и задачу производства текста.

Именно эти две задачи составляют основу лингвистики – науки, предметом которой является человеческий язык. Знаменитый лингвист И.А. Мельчук говорил, что язык есть универсальный преобразователь смысла в текст и обратно, а его лингвистическая теория так и называется: теория

«Смысл – Текст» [30, с. 72]. Достижения современной лингвистической науки в изучении таких процессов, как анализ и синтез естественно-языкового текста очень велики. Именно поэтому эти достижения стремятся как можно больше использовать разработчики СМП. В процессе анализа входного языка системы автоматизированного перевода используют средства, которые лингвистика предлагает для моделирования механизмов понимания, а в процессе синтеза выходного языка – средства, наработанные лингвистикой для моделирования механизмов производства текста.

1.3. Классификация систем машинного перевода, его схемы и общий алгоритм действия

Перевод текста на компьютере может быть осуществлен как просто самой машиной, так и человеком. В первом случае это называется машинным переводом (МП), во втором – автоматизированным переводом (АП) (Computer-Aided Translation). При автоматизированном переводе всю работу производит человек, а компьютер позволяет сделать это быстрее и качественнее.

Изначально после создания компьютеров активно разрабатывалась концепция МП, и исследования в области компьютерной лингвистики преобладали именно в этом направлении. Однако сами переводчики не считали эту концепцию верной и продвигали идею использования компьютерных технологий, в первую очередь, в помощь себе. Предшественник нынешнего Евросоюза – Европейское объединение угля и стали – в 60-е годы XX века начало создание терминологических баз данных Eurodicautom. Аналогом таких баз данных в СССР послужил ВИНТИ – Всероссийский институт научной и технической информации РАН.

В 1980г. Мартин Кей, один из пионеров компьютерной лингвистики, в статье сформулировал идею АП в современном виде, обозначив следующее: «компьютер берет на себя рутинные операции и освобождает человека для

операций, требующих человеческого мышления» («by taking over what is mechanical and routine, it (computer) frees human beings for what is essentially human») [26, с. 49].

Сегодня для осуществления письменного перевода компьютеры используют различными способами. Таковыми являются, например, так называемые корпуса – объемные коллекции текстов на одном либо на нескольких языках. Они содержат краткое описание того, каким образом разнообразные слова и выражения фактически используются в том или ином языке в целом, а также в конкретном тематическом сегменте. Сюда же относится «память переводов», которая хранит образцы прежде переведенных текстов, глоссарии и словари.

Машинный перевод, который также называется электронным, автоматическим или компьютерным, представляет собой процесс изменения текста на одном языке в равный по содержанию текст на другом языке и продукт этого процесса. Система машинного перевода (СМП) имеет программный комплекс, в состав которого входят средства грамматического анализа и двуязычные словари, включающие разнообразные виды грамматической информации (синтаксической, семантической, морфологической), требуемой для соблюдения переводных соответствий.

- Системы прямого перевода (СПП) – это программно-аппаратные комплексы, которые анализировали текст буквально, «слово за словом», при этом смысл переводимого фактически не соблюдался. В данных СПП их возможности напрямую зависели от величины памяти компьютера, которая обуславливала доступные объемы словарей. Примером такой СПП являлась IBM Mark II, благодаря которой «джорджтаунский эксперимент» стал возможен. Таковыми были «электронные переводчики» первого поколения.
- Следующим шагом в развитии СПП стали Т-системы (англ. Transfer – «преобразование»). В этом случае при переводе учитывались синтаксические структуры текста, знания о коих получают в средней

школе. Т-системы позволяли определить синтаксическую структуру переводимого текста согласно грамматическим правилам входного языка благодаря его анализу и впоследствии перестроить его в синтаксическую структуру выходного языка, и создать новый текст с вставкой необходимых слов этого языка. Однако на сегодняшний день эта система признана бесперспективной, поскольку установлено, что фактическая система соответствий более сложная, и для ее корректного перевода требуется принципиально отличающийся алгоритм операций.

- Далее системы машинного перевода, количество которых увеличилось, в соответствии с принципами их действия начали делить на ТМ-комплексы и МТ-программы.

Пакет Translation's Workbench фирмы TRADOS – самый распространенный в мире профессиональный ТМ-комплекс. Он используется, в первую очередь, переводчиками-профессионалами, которые понимают преимущества от определенной доли автоматизации своего труда компьютерными технологиями в тех случаях, когда повторяющиеся тексты имеют близкие структуру и тематику.

Главным замыслом «памяти переводов» является исключение перевода повторяющегося текста повторно. Данная технология осуществляет сопоставление текста, который требуется перевести, с уже имеющейся в заранее созданной «входной» базе информацией. Если найденный фрагмент подходит под предварительно заданные критерии, перевод принимается из «выходной» базы. Результат этой машинной операции впоследствии подлежит постобработке человеком.

Известная канадская система МЕТЕО, которая осуществляет перевод прогнозов погоды с английского на французский язык и наоборот, представляет собой действительно успешное применение МТ-программы, поскольку ей уже около 30 лет, а она до сих пор активно используется. Главный акцент при разработке МЕТЕО делался на том, что для автоматизированного машинного перевода первоочередным условием

является искусственное ограничение как грамматических правил, так и запаса слов языка, что и было воплощено в жизнь.

Системы машинного перевода в 1990 г. стали одним из основных направлений развития компьютерных технологий, достигнув более высокого уровня качества после прохождения сложной стадии переосмысления и взаимной интеграции. По сей день актуальна классификация этих систем, предложенная Зари Чайлдсом: «электронные переводчики» делятся на 3 группы [43, с. 32]:

- ФАМТ (Fully-Automated Machine Translation) – приложения целиком автоматизированного машинного перевода.
- НАМТ (Human-Assisted Machine Translation) – инструменты для автоматизированного машинного перевода текстов, который осуществляется с участием человека; при этом в данную группу включены как программы, ведущие происхождение от МТ- и ТМ-систем, так и основанные на отличных принципах – фразеологическом, статистическом и других, а также комплексные решения.
- МАНТ (Machine-Assisted Human Translation) – вспомогательные инструменты осуществления перевода людьми с применением компьютерных технологий.

Системы ФАМТ в данный момент пребывают в стадии разработки. Причиной тому является сложность понимания, перевода и синтеза «живых» текстов машиной, которую на данном этапе и в самые ближайшие годы вряд ли удастся преодолеть. Программы, позволяющие осуществлять перевод устной речи «на лету», представляют собой один из видов ФАМТ-программ. Однако их функции довольно ограничены.

НАМТ-системы немногочисленны и пока далеко не универсальны, поскольку их реализацию сдерживают все те же объективные факторы. В этой области добилась больших успехов российская фирма PROMT; интересные разработки есть и у других компаний. Мы подробно расскажем о пакете PROMT Professional 8.0.

Реализацию систем НАМТ сдерживают все те же объективные обстоятельства. Самой успешной компанией в использовании этих систем является российская PROMT (речь о пакете PROMT Professional 8.0. пойдет ниже). Также некоторые успехи в этой сфере имеют и другие фирмы.

Категория МАНТ представлена сегодня наибольшим числом компьютерных словарей (онлайновых и программных) и электронных переводчиков.

Системы машинного перевода FAMT и НАМТ уже практически подошли к тому, чтобы называться системами искусственного интеллекта. Эти системы в определенной степени выполняют функции мозга человека. Так, они осуществляют построение текста на выходном языке на основе входного, используя для этого набор заданных правил, которые определены алгоритмами и структурами данных.

Существует также иная классификация, связанная со спецификой реализации программ и применением конкретных инструментов [12, с. 56]:

Самой объемной и пользующейся спросом группой является машинный перевод. Она представлена главным образом независимыми приложениями, предназначенными прежде всего для целиком автоматизированного перевода. От пользователя требуется только ввести направление перевода и изредка его тематику. Скорость перевода, его качество, некоторые другие потребительские характеристики (развитые инструменты пополнения словарной базы, взаимодействие с другими инструментами обработки данных (браузерами, текстовыми процессорами, почтовыми клиентами), удобство интерфейса) представляют собой основное требование к продуктам данной группы;

Продукты группы «памяти перевода» близки к системам перевода МАНТ и прежде всего их относят к вспомогательным средствам для переводчика. Принцип работы этих систем заключается в том, что накапливаются и повторно применяются наиболее используемые грамматические и лексические структуры и эталонные переводы. Такого

типа инструменты служат обычно для решения специальных технических задач вроде производства терминологических баз данных, локализации программных продуктов, либо перевода типовых промышленных документов (договоров, инструкций). Для работы же со связными текстами такие продукты мало пригодны;

Следующая группа – контролируемый язык и машинный перевод на основе базы знаний. В данном случае вместо свободного входного языка применяются контролируемый входной язык. Это требует конкретного ограничения грамматической, лексической и семантической составляющих. Благодаря чему входной текст становится более простым и структурированным, что положительно сказывается на качестве и точности перевода;

Далее следует группа «Инструменты для перевода» в Интернете. Сюда относятся службы, которые дают возможность работы как с отдельными словами (словари), так и с фрагментами текста (непосредственно переводчики) в режиме онлайн. Их главная выгода заключается в том, что не требуется установки никаких дополнительных программ, и пользователю без задержек предоставляется перевод напрямую в окне браузера, а сама услуга эта бесплатная. В ряде фирм предлагаются услуги машинного перевода, где в зависимости от степени участия человека стоимость может быть выше или ниже [20].

Сегодня web-сервисы перевода практически ничем не уступают программам перевода, которые устанавливаются на ПК. Более того, благодаря доступу к мощным серверам у web-сервисов даже есть некоторые преимущества. Однако основным недостатком остается то, что web-систему невозможно настроить под себя путем создания словарей пользователя, а также нет специализированных словарей.

Все интернет словари можно разделить на три полезные группы:

- толковые словари;
- двуязычные словари;

- переводчики [15, с. 54].

Преимущество толковых словарей – их специализация. В этих словарях содержатся термины, имеющие отношение к различным научным дисциплинам, в частности, медицинские термины, а также кратко объясняется их значение. Следовательно, при работе с такими словарями у пользователя должен быть определенный уровень владения английским языком. Словари, в которых нет ни одного русского слова, также необходимы. Поскольку существуют отдельные термины, значение которых можно найти только в англоязычном толковом словаре, где оно раскрывается с помощью понятных и распространенных слов.

Последовательность строго определенных операций с текстом, которые требуются для установления оптимальных соответствий перевода при заданном направлении перевода для данной пары языков называется алгоритмом перевода.

Принимая во внимание, что разные системы перевода, будь то FАМТ-, НАМТ- или даже МАНТ-системы, основаны на отличных подходах к переводу, можно заключить, что алгоритмов полностью универсальных нет. Но все же нечто общее можно отразить некой упрощенной схемой. И при этом надо иметь ввиду, что и при машинном переводе, и при переводе непосредственно человеком, осуществляется большой диапазон разного рода операций. Рассмотрим важнейшие составляющие.

1. При переводе текста человек вводит его в компьютер, а затем осуществляет действия по настройке СМП: определяет базовую тематику, терминологические словари для специальной лексики, задает параметры для входного и выходного языков, устанавливает ограничения при переводе имеющихся в тексте собственных имен, и прочее.

2. Программа производит морфологический анализ входной информации, в словаре входного языка осуществляет поиск словоформ. При анализе (по МТ-технологии, фразеологической или статистической) возможно появление и других данных.

3. Система перевода ищет соответствия. Сперва она осуществляет перевод фразеологизмов, идиом, штампов данной предметной сферы. Следующий шаг – установление важнейших грамматических параметров всех прочих составляющих входного текста, таких как времена глаголов, числа существительных и прочее. На данной стадии характерно отделение однозначных слов от многозначных, а затем для перевода многозначных слов применяются специализированные словари. Перевод же однозначных слов осуществляется по спискам эквивалентов. В конце осуществляется лексический анализ и непосредственно перевод. На данной стадии человек может принять участие в процессе, так как необходимо скорректировать в тексте некоторые языковые особенности, которые машина не способна воспринять. Как пример, необходимость выбора правильного значения слова из нескольких возможных.

4. На следующем этапе система осуществляет грамматический анализ произведенного перевода, при котором учитываются данные выходного языка. Таким образом устанавливаются отсутствующие грамматические характеристики, в частности, какое время глагола из трех возможных в русском языке соответствует времени Present Progressive английского глагола.

5. Идет создание выходных словоформ и текста в целом.

6. На завершающем этапе снова в работу машины вмешивается человек, чтобы устранить недочеты и исправить ошибки автоматического перевода машины.

СМП первого поколения или прямого перевода (П-системы) осуществляли перевод с одного языка на другой без учета смысловых связей. Они производили только морфологический анализ (сокращенно - Ман) текста, беря во внимание лишь то, от какого слова произошла и какой его формой является данная словоформа. Пример: словоформа «стекло» может быть и производным от глагола «стекать», и собственно существительным «стекло». Результат морфологического анализа называется морфологическим

представлением предложения. В СМП первого поколения морфологическое представление предложения входного напрямую преобразовывалось в то же представление выходного.

Системы второго поколения (Т-системы) являются более сложными. Они параллельно с Ман также производят синтаксический анализ (Сан). При этом происходит построение синтаксической структуры текста в виде дерева с разметкой. Ветви дерева представляют синтаксические связи, а узлы – либо только словоформы (если это дерево зависимостей), либо как словоформы, так и группы словоформ (если это дерево составляющих).

При осуществлении перевода в Т-системах можно выделить три главные фазы: анализ (Ман + Сан), трансфер (реализация перехода от одного языка к другому путем преобразования дерева входного предложения в дерево выходного), синтез, подразделяющийся на морфологический синтез Мсинт (синтез слов из аффиксов и основ) и синтаксический синтез Ссинт (переход от синтаксического дерева предложения к последовательности морфологических представлений словоформ).

Существует еще один тип систем – «интерлингвовые» (И-системы). Они направлены на приобретение в результате анализа представления, которое не зависит от языка (интерлингвовое представление). В этом случае система пытается достичь того же результата перевода, который осуществляется человеком.

СМП больше подходят для работы с материалами научного, технического и образовательного характера, где присутствует более высокая строгость изложения. Тексты такого рода по большей части однозначны и не допускают двоякого толкования. Таким образом, применение машинного перевода на современном уровне развития СМП довольно узко. Оно пригодно для перевода публицистических, разговорных текстов, но при условии, что в них будет минимум всевозможных оборотов. В этом случае перевод текста для ознакомления вполне подойдет, а вот для его полного грамотного изложения требуется участие человека. Что касается

художественной литературы, поэзии, то машинный перевод здесь не подходит. Смысл переводимого текста будет полностью искажен. Следует учитывать, что СМП не может выйти за те логические параметры, которые в них заложили.

Многозначность машиной не может быть воспринята правильно, в результате чего текст будет ошибочно интерпретирован, а смысл утерян. Поэтому, как уже говорилось прежде, переводимый текст нужно проверять как до начала его перевода машиной, так и обрабатывать по окончании перевода ручным способом. Только так возможно устранить все несоответствия.

Понимание тематики оригинального текста, лексики и грамматики, а также корректное пользование программами словарей предопределяет качество перевода текста.

Один из главных залогов успешного перевода – правильное использование словарного запаса, всевозможных оборотов, словоформ. При своевременном внесении в специальные словари новых терминов перевод текста будет наиболее приближен к оригиналу, с максимальным сохранением его стилистики.

И также следует отметить, что программная реализация на высоком уровне не менее важна, чем хорошо реализованная лингвистика, и в настоящее время разработчики СМП это понимают.

Выводы по первой главе

Работа переводчика заключается в доступном донесении смысла иностранного повествования до целевой аудитории.

В настоящее время культура напечатанного листа постепенно вытесняется культурой экрана (screen culture). В связи с этим изменяются функции и сущность перевода. Информация порождается, передается и воспринимается напрямую через компьютер. Поэтому и основным

инструментом современного переводчика сейчас является именно он. Часто можно встретить фразы вроде «появление компьютера явилось прорывом в переводческой деятельности» [8, с. 77].

На данный момент существуют две технологии для автоматизированного перевода текстов - машинный перевод и технология «памяти переводов». В первой технологии подразумевается связный перевод текста, с использованием определенных лингвистических алгоритмов. Прежде всего, система анализирует структурные элементы входного предложения, затем преобразует его в соответствии со структурой целевого языка и синтезирует окончательный вариант. Программа также должна уметь распознавать фиксированные выражения и фразеологизмы, а также иметь большой словарный запас для повышения качества перевода. При переводе тематических текстов обычно требуется подключение специализированных словарей. С помощью систем с технологией МП вы можете мгновенно получить черновик перевода, что весьма полезно, когда вам нужно быстро понять общий смысл текста.

Технология Translation Memory основывается на сравнении документа, который нуждается в переводе, с данными, которые хранятся в предварительно созданной пользователем базе переводов. Когда система перевода находит сегмент, подходящий под установленные критерии, его перевод заимствуется из базы переводов Translation Memory. Данный принцип работы наиболее эффективен при переводе повторяющихся текстов, сходных по своей тематике и структуре. Системы с технологией ТМ полезны для профессиональных переводчиков, они могут пригодиться в том случае, если вы собираетесь постоянно переводить большое количество литературы.

Глава 2. Автоматизированный перевод английских научно-технических текстов

2.1. Особенности научно-технических текстов и их перевода

Говоря о таком широко применяемом понятии, как «научно-техническая литература» мы подразумеваем совокупность разной по своей составляющей части литературы, например: монография, учебные материалы (пособия), выдержки из журналов, описания технических параметров изделий, справочники, инструкции, патенты и прочие разновидности. Перечисленные нами типы научно-технической литературы различаются между собой языком. К примеру, язык монографий, научных трудов и статей из журналов, как правило, намного ярче и богаче, чем, скажем, у технических справочников и руководств к действию, инструкций.

Научно-техническая литература и ее лексические аспекты. Общеизвестно, что лексика научно-технической литературы включает в себя слова, которые чаще других можно встретить в социуме, а также широкий спектр специфических терминологий. Первая часть общеупотребительных слов, к примеру, *to work* - работать, *to know* - знать, *place* - место, *new* - новинка, хорошо знакома, соответственно, перевод речевых оборотов с одного языка на другой выполняется легко и просто. Вторая часть слов, как правило, изначально незнакома субъектам, а потому и формирует собой тот самый фундаментальный лексический резерв, требующий усвоения в ходе обучения. Как утверждают эксперты, словарный запас, который предстоит заучить состоит из порядка 2000-2500 слов [14, с. 186].

Так, к широко популярным лексико-грамматическим специфичным особенностям, которые свойственны для научно-технических текстов, напечатанных на английском языке, можно отнести — степень информативности, передовые терминологии и специальная лексика,

достижение номинативности, а еще всевозможные атрибутивные группы и повсеместное задействование эллиптических конструкций.

Примечательно, что в современном мире англоязычные научно-технические информационные работы ориентированы на заимствование тех аспектов, которые в прошлом можно было лицезреть исключительно в газетно-публицистических или художественных материалах. Впрочем, это совершенно не значит, что хорошо знакомый всем нам научный стиль изложения претерпит какие-либо видимые или невидимые изменения. Единственный момент — появление определенных сложностей у абитуриентов техникумов в момент осуществления перевода профессионально-ориентированных текстов.

К характерным лексическим особенностям, присущим современным научно-техническим изданиям относят: применение стилистически разноплановой лексики, избыточность неологизмов, характерных по большей степени для компьютерной тематики и, конечно, метонимии и метафоры, наиболее часто используемые для художественных рукописей.

Процесс перевода стилистически разноплановой лексики в научно-технический контент должен осуществляться на базе широко применяемого принципа следования цели коммуникации и сохранности научного стиля. Кроме того, следует учитывать и установленные нормы русского языка, поскольку в тех же англоязычных работах использование разговорной речи вполне приемлемо, а вот в русскоязычном материале уместнее будут нейтральные.

Рассмотрим несколько примеров, которые можно встретить в статьях научно-технических журналов, e.g. *The scientists also drive home problems with making predictions from models.* (Science: 2005) - Ученые умы всячески способствуют осознанию текущих проблем через прогнозирование, базирующееся на моделировании (дословный перевод).

Одновременно с интерпретацией специфической терминологии из научно-технических изданий, имеет смысл раскрывать глубокое содержимое

языка оригинала, а также выполнять его передачу на языке перевода за счет последнего. Ввиду того, что языковые системы русского и английского языков относительно схожи, в некоторых случаях, когда предложения на данных языках не имеют больших структурных различий, а текстуальные значения лексики не вызывают трудностей, перевод осуществляется прямым путем, т.е. в тексте на переводящем языке используются соответствия, которые «напрашиваются» сами собой: e.g. *An amplifier is a device designed for increasing power.* - Усилитель является устройством, предназначенным для повышения мощности.

Однако в подавляющем большинстве случаев из-за лексико-синтаксических трудностей и структурных расхождений текста оригинала, перед переводчиком появляется ряд задач: лексико-грамматический анализ «определенных мест» текста-источника, определение текстуальных значений языковых единиц и поиск адекватных им соответствий в языке перевода, т.е. в этих случаях «прямой перевод» в такой ситуации невыполним, вследствие чего переводчик обращается к различного рода преобразованиям. Лексические трансформации обусловлены [16, с. 85]:

а) семантической структурой слова одного языка, отличающегося от структуры слова другого;

б) одни и те же термины в звучании на одном языке могут не обладать средствами выражения, а вот во втором, напротив, выражаться очень ярко;

в) словоформы обоих языков, интерпретирующих одно и то же понятие, могут характеризоваться отличительным употреблением и такой же несхожей сочетаемостью в речи;

г) одна и та же трактовка в нескольких языках может сопровождаться различным количеством средств выражения.

Достоверный перевод неологизмов, содержащихся в научно-технических материалах, как правило, выполняется с учетом ключевых тенденций, ориентированных на заимствование новоявленных понятий или, как вариант, их калькирования. За последние десятилетия лексический запас

русского языка дополнился несметным количеством технических, инженерных и компьютерных терминологий, занявших твердые позиции в узусах, таких как Интернет-ресурсы, чипы и т. п. И хотя сверхмерное заимствование понятий все так же находится на второй позиции после лексических замен, в обиходе нередко можно наткнуться на не всегда обоснованное заимствование с введенных в русский язык понятий или интерпретаций. Собственно, предвидя такого плана ситуаций, среднестатистический преподаватель, работающий в том или ином учебном заведении должен приложить усилия для того, чтобы его подопечные (обучающиеся) использовали фактор заимствования как можно реже, в идеале — никогда, а вместо этого обращали внимание на вариации перевода. Например, *to upgrade* - развить, расширить, вместо: совершить (исполнить) апгрейд.

Осуществляя перевод научно-технического материала, специалист, как правило, повсеместно задействует общепринятые метафоры и, само собой, метонимии, приближая их тем самым тем самым к художественной стилистике. Вместе с тем при таком подходе к переводу контента образ, в большинстве случаев, растворяется (размывается) по причине очевидных различий в картинах мира обоих языков, e.g. ... *demonstrated masterful use of its lung power*. (Motor Trend: 2002) - ... наглядным примером показать возможности двигателя.

Современные материалы, изложенные на английском языке и содержащие в себе научно-технические термины пестрят причинно-следственными и логическими союзами, плавно переключавшимися из середины сложного речевого оборота в начало абзаца. Очевидно, что такого плана словоформы делают текст более связным и логичным, выполняя роль союза, соединяющего между собой определенные части, а в определенных моментах и вовсе несут эмоциональную нагрузку (к примеру, *firstly, secondly, still, also, consequently, moreover, nevertheless, however, yet, likewise, etc*), e.g. *Moreover, Java's ability to work with different platforms is important in the*

fragmented cellular-phone market. (Computer: 2002) - Возможности программирования Ява, позволяющие функционировать на самых разных площадках, крайне необходимы для успешного фрагментирования на рынке моделей телефонов.

Нынешний научно-технический текст теперь включает в себя грамматические особенности, которые когда-то можно было видеть исключительно в газетных выдержках. В частности: акцентирование внимания на текущем времени с целью обозначения тех или иных мероприятий, осуществленных в недавнем времени (как например, речь может идти об экспериментах в прошлом [35, с. 191]), e.g. *To make atoms heavier than uranium, physicists bombard it with neutrons, or squash smaller atoms together.* (The Economist: 2001) - С целью создания атомов, которые бы были массивнее урана, ученые осуществили определенные действия: "атаковали" его нейронами, а еще стиснули атомы, состоящие из более легких элементов.

The vessel to be completed in summer 2006 is equipped with... (Diesel & Gas Turbine Worldwide: 2005) - Посудину, строительство коей планируется завершить в летний период 2006 года вдобавок оснастят...

Английские и русские языковые конструкции, распространенные в научно-технических текстах напрочь исключают любые совпадения. Это становится понятно в тот самый момент, когда анализируется сравнительная частота использования отдельных речевых частиц [49, с. 88], e.g. *The engine is the source of power that makes the wheels go round and the car move.* - Мотор является полноценным источником энергии, запускающей автоматическое передвижение машины по дорожному полотну; e.g. *The fuel system is designed to store liquid gasoline and to deliver it to the engine cylinders in the form of vapor mixed with air.* - Встроенная система питания способствует заправке бака топливом и подачей его в цилиндры; e.g.. *A similar approach has marked the EU's efforts to expand the current club of 15 countries to embrace former communist countries further east.* - Аналогичный подход используется для

усиления ЕС с целью расширения клуба, включающего в себя 15 государств через присоединение к нему тех республик, что входили в состав СССР; e.g. *The exhibited new components are not the only harbingers of spring. In electronics, the beginning of spring coincides with the beginning of preparations for attending two important components show* (небольшое вступление статьи, размещенной в журнале под названием «Electronics», рассказывающей об основах производства полупроводниковых приборов).

Буквальная интерпретация данного предложения звучит так (... «все эти новоявленные экспонаты не могут выступать в качестве достоверного источника информации, предвещающего скорейшее наступление весенней поры. Так, в той самой электронике начало весны отождествляется с активной подготовкой к принятию участия в обеих выставках оборудования, представляющего немалый интерес...») Ввиду невероятно строжайших критериев, предъявляемых в отношении русской научно-технической подачи информации, такие лексические обороты по типу «предвестник наступления весенней поры» (что-то «абсолютно новенькое») или «наступление весенней поры» (разработка оборудования, влияющего на начало чего-то новоявленного в этой сфере техники) в итоговых переводах является недопустимой.

Вместе с тем переводчик так же не имеет права на подобного рода ошибки, то есть для квалифицированного специалиста исключены такие «вольности». Кроме того, абсолютно неуместными и лишними в переводе будут считаться и такие эпитеты, как, например, *dramatic, successful, excellent* и прочие. В свою очередь, обязательной лексико-грамматической трансформации будут подлежать и такие, к примеру, словоформы, как «*Whither Electronics?*» (оглавление статьи из журнала; *whither* (adv) куда?, (союз) куда?; (сущ.) место назначения) - Перспективы развития электроники. Взявшись за перевод речевого оборота, состоящего из словоформы «*The startling thing we notice is...*» не стоит применять такое оценочное суждение,

как «поразительный» или «потрясающий», в идеале прозвучит: «Отметим крайне интересный факт».

Practical electronic device (дословный перевод: практический электронный прибор) необходимо преподнести, как «электронный прибор, предназначенный для практического использования»).

А вот такого плана обороты речи, как «*Let us write (rewrite) the equation as follows*», «*Equations 2 and 3 may be combined*» и тому подобное трактуются обиходными русскими изречениями: «Пишем наше уравнение в следующем формате», «Соединив между собой уравнения под номером 2 и 3, на выходе получим...» [16, с. 63].

Собственно, становится очевидным, что те критерии, которые выдвигаются в отношении научно-технического стиля, в русско-английском переводе обозначаются как самостоятельная теоретическая и практическая проблема. Под понятием "перевод" - как правило, подразумевается дословная передача сути содержательной части текста посредством использования наиболее эффективных методик и языковых средств, свойственных для научно-технического стиля. Впрочем, достаточное количество аспектов, присущих английскому научно-техническому материалу (к примеру, информативность, наполненность специфической терминологией и тому подобное), допускаются и для русского научно-технического стиля.

Примечательно, что как русскому, так и английскому языкам свойственны конкретизация и требовательность формулировок и четкое воспроизведение специфической терминологии. Впрочем, следует заметить, что та же требовательность соблюдается больше в русском научно-техническом стиле, нежели английском. Так, в английских научно-технических текстах очень часто можно наткнуться на выражения, окрашенные в эмоциональный спектр, которые, к слову, категорически неприемлемы для идентичных работ, написанных уже на русском. А ввиду того, что данное расхождение носит больше стилистический характер, то в

ходе осуществления перевода специалист вынужден адаптировать стиль и прочие нюансы английского контента к нормам и правилам русского языка.

Правильность перевода в том и заключается, чтобы преподнести полученную информацию с максимальной достоверностью, не имеющей фактических расхождений с текстом первоначального источника.

Стоит учитывать, что, порой без адаптации под ту или иную стилистику в процессе перевода не обойтись совершенно. Например, в случае, если мы имеем дело с определенными стилистическими признаками, проявившимися в научно-технических материалах как в оригинале, так и в переведенном материале [7].

2.2. Сравнительный анализ перевода научно-технического текста, выполненного с помощью двух систем машинного перевода

В машинной обработке научно-технических текстов существует ряд прикладных задач, которые должны быть решены в процессе:

- литературно-научное редактирование,
- извлечение текстовых знаний,
- реферирование,
- аннотирование

Для этого помимо определенных алгоритмов терминологического анализа необходимы также процедуры выявления и учета композиционно-логической, дискурсивной организации текстов.

С учетом разобранных особенностей перевода научно-технических текстов постараемся проанализировать процедуру распознавания в заданном тексте приемов и соответствующих текстовых сегментов. Процедура основывается на поверхностном синтаксическом анализе фраз текста и словаре общенаучной лексики, который охватывает слова и выражения общенаучной речи.

Ниже приведены 4 образца текстов, используемых в дальнейшем анализе:

1. The Architecture (оригинал)

Cascade-Correlation [Fahlman, 1990] is supervised learning architecture that builds a near-minimal *multi-layer network topology* in the course of training. Initially the network contains only inputs, output units and the connection between them. This single layer of connections is trained (using the Quickprop algorithm [Fahlman, 1988]) to minimize the error. When no further improvement is seen in the level of error, the network's performance is evaluated. If the error is small enough, we stop. Otherwise we add a new hidden unit to the network in an attempt to reduce the residual error.

To create a new hidden unit, we begin with a pool of candidate units, each of which receives weighted connections from the network's inputs and from any hidden units already present in the net. The outputs of these candidate units are not yet connected into the active network. Multiple passes through the training set are run, and each candidate unit adjusts its incoming weights to maximize the correlation between its output and the residual error in the active net. When the correlation scores stop improving, we choose the best candidate, *freeze its incoming weights*, and add it to the network. This process is called «tenure». After tenure, a unit becomes a permanent new feature detector in the net. We then re-train all the weights going to the output units, including those from the new hidden unit. This process of adding a new hidden unit and re-training the output layer is repeated until the error is negligible or we give up. Since the new hidden unit receives connection from the old ones, each hidden unit effectively adds a new layer to the net. (See figure 1)

Cascade-correlation eliminates the need for the user to guess in advance the network's size, depth, and topology. A reasonably small (though not minimal) network is built automatically. Because a hidden-unit feature detector, once built,

is never altered or cannibalized, the network can be trained incrementally. A large data set can be broken up into smaller “lessons,” and feature-building will be cumulative.

Cascade-Correlation learns much faster than backprop for several reasons: First only a single layer of weights is being trained at any given time. There is never any need to propagate error information backwards through the connections, and we avoid the dramatic slowdown that is typical when training backprop nets with many layers.

Second, this is a “greedy” algorithm: each new unit grabs as much of the remaining error as it can. In a standard backprop net, the all the hidden units are changing at once, competing for the various jobs that must be done - a slow and sometimes unreliable process.

Cascade-correlation, like back-propagation and other feed-forward architectures, has no short-term memory in the network. The outputs at any given time are a function only of the current inputs and the network’s weights. Of course, many real-world tasks require the recognition of a sequence of inputs and, in some cases, the corresponding production of a sequence of outputs.

2. Архитектура (перевод О. Г. Василенко)

Каскад-корреляция [Fahlman, 1990] является обучаемой системой с учителем, которая строит почти минимальную *многослойную сеть* во время курса обучения. Изначально сеть содержит только входы и выходы, а также связи между ними. Эта однослойная сеть обучается (с помощью алгоритма быстрого распространения Quickprop [Fahlman, 1988]) так, чтобы свести к минимуму ошибки. Когда уровень ошибок перестает уменьшаться - оценивается эффективность сети. Если ошибка достаточно мала, то мы останавливаемся. В противном случае мы добавляем новые скрытые узлы к сети, чтобы уменьшить остаточную ошибку.

Чтобы создать новые скрытые узлы, мы начинаем с набора узлов-кандидатов, каждый из которых получает на вход сигналы от входов сети и других скрытых узлов, уже присутствующих в сети. Выходы узлов-кандидатов еще не связаны с другими узлами. Мы совершаем несколько проходов по обучающему множеству, и каждый узел-кандидат регулирует свои весовые коэффициенты входящих связей, чтобы увеличить корреляцию между его выходом и остаточной ошибкой сети. Когда оценка корреляции перестает улучшаться, мы выбираем лучшего кандидата, *замораживаем его весовые коэффициенты входящих связей* и добавляем его к сети. Этот процесс называется "завладением". После завладения, кандидат становится полноценным новым узлом сети. Затем мы заново обучаем все коэффициенты на исходящих связях, в том числе и для нового узла. Этот процесс добавления нового узла и переобучения выходного слоя повторяется, пока ошибка не становится незначительной или мы не решаем сдаться и остановиться. Поскольку новые узлы получают входящие соединения от старых, каждый узел эффективно добавляет новый слой к сети. (См. рисунок 1.)

Каскад-корреляция исключает для пользователя необходимость предугадывать размеры сети, ее глубину и топологию. Достаточно малая (но не минимальная) сеть строится автоматически. Поскольку новые узлы после добавления никогда не меняют своих связей и не удаляются, сети могут быть обучены постепенно. Большой набор данных можно разбить на более мелкие "уроки" и увеличение потенциала будет нарастать.

Каскад-корреляция обучает гораздо быстрее, чем сеть с обратным распространением ошибки по нескольким причинам:

В начале только один слой весовых коэффициентов обучается в любой данный момент времени. Никогда не возникает необходимости распространять информацию об ошибках обратно через связи, и мы избегаем драматического замедления, которое является типичным при обучении

многослойных сетей с обратным распространением ошибки. Во-вторых, это "жадный" алгоритм: каждая новая единица захватывает так много оставшейся ошибки, как только может.

В стандартной сети с обратным распространением ошибки, все скрытые узлы меняются одновременно, конкурируя за различные задачи, что должны быть достигнуты — это медленный и иногда ненадежный процесс.

Каскад-корреляция, как и обратное распространение и другие архитектуры с прямой связью, не имеет кратковременной памяти в сети. Выходы в любой данный момент времени являются функцией только от текущих входов и коэффициентов при связях в сети. Конечно, многие реальные задачи требуют распознавания последовательности входов и, в некоторых случаях, создания соответствующей последовательности выходов.

3. Архитектура (Google-переводчик)

Каскадная корреляция [Fahlman, 1990] - это архитектура обучения с учителем, которая строит почти минимальную *многослойную топологию сети* в процессе обучения. Первоначально сеть содержит только входы, выходы и связи между ними. Этот единственный уровень соединений обучается (с использованием алгоритма Quickprop [Fahlman, 1988]), чтобы минимизировать ошибку. Когда не наблюдается дальнейшего улучшения уровня ошибки, оценивается производительность сети. Если ошибка достаточно мала, останавливаемся. В противном случае мы добавляем в сеть новый скрытый блок, пытаемся уменьшить остаточную ошибку.

Чтобы создать новую скрытую единицу, мы начинаем с пула единиц-кандидатов, каждая из которых получает взвешенные соединения от входов сети и от любых скрытых единиц, уже присутствующих в сети. Выходы этих блоков-кандидатов еще не подключены к активной сети. Выполняется несколько проходов через обучающий набор, и каждый блок-кандидат корректирует свои входящие веса, чтобы максимизировать корреляцию

между своим выходом и остаточной ошибкой в активной сети. Когда показатели корреляции перестают улучшаться, мы выбираем лучшего кандидата, *фиксируем его входящие веса* и добавляем его в сеть. Этот процесс называется «владение». После пребывания в должности устройство становится постоянным детектором новых функций в сети. Затем мы повторно обучаем все веса, идущие к выходным единицам, в том числе из нового скрытого блока. Этот процесс добавления нового скрытого блока и повторного обучения выходного слоя повторяется до тех пор, пока ошибка не станет незначительной или мы не сдадимся. Поскольку новый скрытый блок получает соединение со старым, каждый скрытый блок эффективно добавляет новый слой в сеть. (См. Рисунок 1)

Каскадная корреляция избавляет пользователя от необходимости заранее угадывать размер, глубину и топологию сети. Достаточно небольшая (хотя и не минимальная) сеть строится автоматически. Поскольку детектор скрытых объектов, однажды созданный, никогда не изменяется или каннибализируется, сеть можно обучать постепенно. Большой набор данных можно разбить на более мелкие «уроки», и создание функций будет накопительным.

Cascade-Correlation обучается намного быстрее, чем blackprop, по нескольким причинам: во-первых, в любой момент времени обучается только один слой весов. Нет необходимости распространять информацию об ошибках в обратном направлении по соединениям, и мы избегаем резкого замедления, которое типично при обучении многослойных блочных сетей.

Во-вторых, это «жадный» алгоритм: каждая новая единица захватывает как можно больше оставшейся ошибки. В стандартной сети обратного распространения все скрытые элементы меняются одновременно, конкурируя за различные работы, которые необходимо выполнить - медленный и иногда ненадежный процесс.

Каскадная корреляция, как и другие архитектуры с прямой связью, не имеет кратковременной памяти в сети. Выходы в любой момент времени

являются функцией только текущих входов и веса сети. Конечно, многие реальные задачи требуют распознавания последовательности входных данных и, в некоторых случаях, соответствующего производства последовательности выходных данных.

4. Архитектура (переводчик PROMT)

Cascade-Correlation [Fahlman, 1990] - управляемая архитектура обучения, которая строит почти минимальную *многоуровневую сетевую топологию* в ходе обучения. Первоначально сеть содержит только входы, выходные блоки и соединение между ними. Этот одиночный уровень соединений обучается (с использованием алгоритма Quickprop [Fahlman, 1988]) минимизации ошибки. При отсутствии дальнейшего улучшения уровня ошибок оценивается производительность сети. Если ошибка достаточно мала, мы останавливаемся. В противном случае мы добавим новую скрытую единицу в сеть, чтобы уменьшить остаточную ошибку.

Для создания нового скрытого блока мы начинаем с пула блоков-кандидатов, каждый из которых получает взвешенные соединения от входов сети и от любых скрытых блоков, уже присутствующих в сети. Выходы этих блоков-кандидатов еще не подключены к активной сети. Выполняется множество проходов через обучающий набор, и каждый блок-кандидат корректирует свои входящие веса, чтобы максимизировать корреляцию между его выходом и остаточной ошибкой в активной сети. Когда показатели корреляции перестают улучшаться, мы выбираем наилучшего кандидата, *замораживаем его входящие веса* и добавляем его в сеть. Этот процесс называется «пребывание в должности». После пребывания в должности блок становится постоянным детектором новых признаков в сети. Затем мы повторно тренируем все веса, идущие к выходным блокам, включая веса от нового скрытого блока. Этот процесс добавления новой скрытой единицы и повторного обучения выходного уровня повторяется до тех пор, пока ошибка не станет ничтожной или мы не сдадимся. Поскольку новый скрытый блок

получает соединение от старых блоков, каждый скрытый блок эффективно добавляет новый слой в сеть. (См. рисунок 1)

Каскадная корреляция устраняет необходимость предварительного определения размера, глубины и топологии сети. Достаточно небольшая (хотя и не минимальная) сеть создается автоматически. Поскольку детектор скрытых элементов, созданный после его создания, никогда не изменяется или каннибализуется, сеть может быть обучена постепенно. Большой набор данных может быть разбит на более мелкие «уроки», и создание функциональных возможностей будет кумулятивным.

Cascade-Correlation учится гораздо быстрее backprop по нескольким причинам: во-первых, только один слой весов обучается в любое время. Никогда не нужно распространять информацию об ошибках обратно через соединения, и мы избегаем резкого замедления, которое характерно при обучении сетей backprop с множеством уровней.

Во-вторых, это «жадный» алгоритм: каждый новый блок захватывает как можно больше оставшейся ошибки. В стандартной сети backprop все скрытые блоки меняются сразу, конкурируя за различные работы, которые должны быть выполнены - медленный и иногда ненадежный процесс.

Каскадная корреляция, как и обратное распространение и другие архитектуры прямой передачи, не имеет кратковременной памяти в сети. Выходы в любой данный момент времени являются функцией только текущих входов и весов сети. Конечно, многие реальные задачи требуют распознавания последовательности входных данных и, в некоторых случаях, соответствующего производства последовательности выходных данных.

В текстах переводов сохранены стилистические и орфографические ошибки, выданные машинными переводческими системами.

Глобальная коммуникативная цель любого научного произведения – это сообщение о результатах проведенного исследования, а также

объяснение способа их получения, формулировка возникших в процессе новых идей и их обоснование.

Следовательно, стандартное научное изложение состоит в основном из рассуждений, а отличительной чертой функционального стиля научной прозы является не ее насыщенность специальными терминами, а особый формально-логический способ изложения материала.

Далеко не всегда у систем получается успешно опознать термины в тексте. В основном, и система Google, и система PROMT переводят термины, как общеупотребительные слова. Если, например, фраза «multi-layer network topology» переводится как «многослойная топология сети» и «многоуровневая сетевая топология», то фраза «freeze its incoming weights» в переводе звучит бессмысленно: «фиксируем его входящие веса» во варианте Google и «замораживаем его входящие веса» во варианте от компании PROMT. В то время как в переводе О.Г. Василенко фраза приобретает больше смысла, благодаря контексту, который был учтен переводчиком.

Перевод терминов в двух системах представим в таблице 1.

Таблица 1

Перевод терминов в двух системах и в переводе О. Г. Василенко

Оригинал	Перевод системы Google	Перевод системы PROMT	Перевод О. Г. Василенко
Cascade-Correlation	Каскадная корреляция	Cascade-Correlation	Каскад-корреляция
Multi-layer network	Многослойная топология сети	Многоуровневая сетевая топология	Многослойная сеть
Single layer	Единственный уровень соединения	Одиночный уровень соединения	Однослойная сеть

No further improvement is seen in the level of error	Не наблюдается дальнейшего улучшения уровня ошибки	При отсутствии дальнейшего улучшения уровня ошибки	Уровень ошибок перестает уменьшаться
Unit	Блок / единица / устройство	Единица / блок	Узел
Weighted connections	Взвешенные соединения	Взвешенные соединения	На вход сигналы
Active network	Активная сеть	Активная сеть	Другие узлы
Adjust its incoming weights	Корректирует свои входящие веса	Корректирует свои входящие веса	Регулирует свои весовые коэффициенты входящих связей
Freeze its incoming weights	Фиксирует его входящие веса	Замораживает его входящие веса	Замораживает его весовые коэффициенты входящих связей
Tenure	Владение	Пребывание в должности	Завладение
A permanent new feature detector in the net	Постоянный детектор новых функций	Постоянный детектор новых признаков	Полноценный новый узел сети
Cannibalized	Каннибализируется	Каннибализируется	Не удаляются
Feature-building	Создание функций	Создание функциональных возможностей	Увеличение потенциала
Cumulative	Накопительный	Кумулятивный	Нарастать

Back-propagation	-	Обратное распространение	Обратное распространение
The network's weight	Вес сети	Весы сети	Коэффициент при связях в сети

Из таблицы 1 становится понятно, что расхождения в переводе терминов наблюдаются на протяжении всего текста, когда Unit переводится как «блок», «единица», «устройство» и «узел». В переводах PROMT и Google варианты различаются на протяжении всего текста статьи. Данное слово является ключевым для этого текста, и отсутствие единого варианта перевода делает текст разрозненным. При этом большая часть терминов, состоящий из одного слова, полностью совпадают в переводе, в то время как фразовые термины различаются между собой. Задача переводчика в таком случае состоит в грамотном редактировании текста.

Некоторые слова остались непонятны для словарей одной системы PROMT, например, *Cascade-Correlation*, некоторые для обеих систем, например, *blackprop*. Данный факт можно объяснить недостаточностью словарей, установленных в этих системах. Что касается синтаксических особенностей перевода текста, то в обеих системах, как у Google, так и у PROMT предложения, в основном, правильные с грамматической точки зрения, отдельные словосочетания и члены предложения согласованы между собой. Это хорошо заметно на указанных ниже примерах:

Google – *Чтобы создать новую скрытую единицу, мы начинаем с пула единиц-кандидатов, каждая из которых получает взвешенные соединения от входов сети и от любых скрытых единиц, уже присутствующих в сети.*

PROMT – *Для создания нового скрытого блока мы начинаем с пула блоков-кандидатов, каждый из которых получает взвешенные соединения от входов сети и от любых скрытых блоков, уже присутствующих в сети.*

Таким образом можно сделать вывод, что обе системы (Google и PROMT) обладают обширным словарем. Перевод, выполненный при помощи

данных систем, нуждается в меньшем количестве корректировок, чем перевод при использовании сторонних компаний.

2.3. Анализ адекватности перевода научно-технического текста

Несмотря на различия в выборе того, какому переводу отдать свое предпочтение — идиоматическому или буквальному, все переводчики едины во мнении, что перевод должен передавать смысл оригинала. Именно поэтому мы будем использовать как аксиому то, что именно на сохранении смысла оригинала сосредоточено определение степени адекватности перевода. Стоит также учитывать, что поскольку сохранение языковой формы оригинала в большинстве случаев приводит к искажению смысла, то и определение адекватности не должно быть на этом сосредоточено.

Однако при определении адекватности перевода одна из характеристик языковой формы все же является существенной. Языковая форма оригинала в момент его написания была естественной и осмысленной. Она не представляла собой сложную и непонятную грамматическую и лексическую структуру, а являлась естественной совокупностью языковых средств, которые использовались в соответственном стиле речи. Это свойство делало оригинал динамичным, и именно это качество должно соответствовать и переводу.

Следовательно, определение адекватности состоит не только в передаче смысла, но и в упомянутом выше признаке языковой формы. Таким образом, адекватным будет считаться тот перевод, который верно передает смысл и динамику оригинала [37, с. 49]. Под «передает смысл» подразумевается, что перевод доносит до слушателя или читателя точно такую же информацию, как та, что содержится в тексте исходного языка. Заложенное автором сообщение не должно искажаться или изменяться; не должен меняться и объем сообщаемой информации.

«Передает динамику» означает, что, во-первых, в переводе используются естественные и живые языковые структуры, и во-вторых,

адресат перевода без труда понимает смысл сообщения. Перевод должен быть так же естественен и легок для понимания, как и оригинал.

Для анализа адекватности перевода отрывка рассматриваемой нами научно-технической статьи, выполненной как с помощью двух систем машинного перевода, так и О. Г. Василенко составим таблицу 2.

Анализ адекватности перевода научно-технического текста

Оригинал	Перевод системы Google (1)	Перевод системы PROMT (2)	Анализ адекватности	Перевод О. Г. Василенко
<p>Cascade-Correlation [Fahlman, 1990] is supervised learning architecture that builds a near-minimal multi-layer network topology in the course of training.</p>	<p>Каскадная корреляция [Fahlman, 1990] - это архитектура обучения с учителем, которая строит почти минимальную многослойную топологию сети в процессе обучения.</p>	<p>Cascade-Correlation [Fahlman, 1990] - управляемая архитектура обучения, которая строит почти минимальную многоуровневую сетевую топологию в ходе обучения.</p>	<p>1. Перевод термина, как обычного слова. 2. Перевод термина, как обычного слова, отсутствие в программном словаре словосочетания «Cascade-Correlation»</p>	<p>Каскад-корреляция [Fahlman, 1990] является обучаемой системой с учителем, которая строит почти минимальную многослойную сеть во время курса обучения.</p>
<p>Multiple passes through the trainingset are run, and each candidate unit adjusts its incoming weights to maximize the correlation between its output and the residual error in the active net.</p>	<p>Выполняется несколько проходов через обучающий набор, и каждый блок-кандидат корректирует свои входящие веса, чтобы максимизировать корреляцию между своим выходом и остаточной ошибкой в активной сети.</p>	<p>Выполняется множество проходов через обучающий набор, и каждый блок-кандидат корректирует свои входящие веса, чтобы максимизировать корреляцию между его выходом и остаточной ошибкой в активной сети.</p>	<p>1. Неадекватный перевод терминов, несогласование частей предложения, неправильное употребление формы слова «веса» 2. Неадекватный перевод терминов, несогласование частей предложения, неправильное употребление формы слова «веса»</p>	<p>Мы совершаем несколько проходов по обучающему множеству, и каждый узел-кандидат регулирует свои весовые коэффициенты входящих связей, чтобы увеличить корреляцию между его выходом и остаточной ошибкой сети.</p>

<p>After tenure, a unit becomes a permanent new feature detector in the net.</p>	<p>После пребывания в должности устройство становится постоянным детектором новых функций в сети.</p>	<p>После пребывания в должности блок становится постоянным детектором новых признаков в сети.</p>	<p>1. Неправильный перевод слова «tenure», перевод термина, как обычного слова 2. Неправильный перевод слова «tenure», перевод термина, как обычного слова</p>	<p>После завладения, кандидат становится полноценным новым узлом сети.</p>
<p>Cascade-Correlation learns much faster than blackprop for several reasons: First only a single layer of weights is being trained at any given time.</p>	<p>Cascade-Correlation обучается намного быстрее, чем blackprop, по нескольким причинам: во-первых, в любой момент времени обучается только один слой весов.</p>	<p>Cascade-Correlation учится гораздо быстрее blackprop по нескольким причинам: во-первых, только один слой весов обучается в любое время.</p>	<p>1. Отсутствие в программном словаре слов «Cascade-Correlation», «blackprop», предложение звучит незакончено 2. Отсутствие в программном словаре слов «Cascade-Correlation», «blackprop», предложение звучит незакончено</p>	<p>Каскад-корреляция обучает гораздо быстрее, чем сеть с обратным распространением ошибки, по нескольким причинам: в начале только один слой весовых коэффициентов обучается в любой момент времени.</p>

Как можно увидеть в таблице 2, при сопоставительном исследовании оценивается всего 4 отрывка данного перевода, но это не означает, что остальной текст полностью игнорируется. Именно сопоставительный анализ дает нам инструмент для оценки правильности многих аспектов перевода, что является объективным научным обоснованием. Главная проблема при оценке перевода как предмета лингвистического исследования в том, что необходимо исключить все возможные случайности и избежать субъективизма со стороны переводчика. Для этого следует собрать статистические данные.

Мы имеем право считать, что получаем объективное соответствие между формами двух языков, в том случае, если сталкиваемся с одним и тем же явлением перевода неоднократно, в переводах различных авторов, разных переводчиков, при переводе с одного языка и на тот же язык.

Сопоставление текстов, в котором учитываются все возможные изменения в лексико-грамматических формах, логические корреляции и связи, семантические соответствия и экстралингвистическая действительность, позволяет сделать заключение о закономерностях перевода, которые демонстрируют компетенцию лиц, владеющих двумя языками, закономерности, которые также могут служить инструментом для объективной оценки ряда существенных аспектов переводов.

Заключение

К самым традиционным методам, используемым для преодоления языковых барьеров относят обучение разным иностранным языкам и деятельность переводчиков. Но с учетом темпов роста экономических, политических и культурных связей между странами, а также повышение интенсивности информационных потоков по всему миру эти методы стали недостаточными. Появилась необходимость в поиске альтернативного пути решения данной проблемы. Именно одним из них и может считаться создание систем машинного перевода.

Проблемой МП человечество задается уже более полувека. Но до недавнего времени успехи и разработки в данной области были крайне невелики, а итоговая эффективность созданных в те времена систем автоматического перевода была на очень низком уровне. Причиной данного положения стало то, что очень долгое время все ориентировались на первые изобретенные модели, которые основывались на пословном семантико-синтаксическом переводе. При использовании такого подхода основными единицами смысла рассматривались отдельные слова, что создавало проблемы при определении смысла более крупных единиц речи (словосочетаний, предложений, целого текста). Однако стоит упомянуть, что небольшое количество словосочетаний так же допускалось к использованию наряду с отдельными словами, но в основном это были идиоматические выражения, которые сильно уступали по количеству словам в словарях систем автоматического перевода.

В настоящее время мы являемся свидетелями бурного развития компьютерных технологий для перевода текстов. В английской терминологии эти технологии разделяются на два больших класса – МТ (Machine Translation) и САТ (Computer-assisted/aided translation). Первый класс известен в русской лингвистике как автоматический или машинный перевод, то есть, те программы, которые могут

полностью заменить человека-переводчика с разной степенью успеха. В тоже время, системы САТ только лишь автоматизируют и облегчают труд переводчика в различных его аспектах. В первую очередь это программы, открывающие доступ к памяти переводов (translation memory).

МТ сделал возможным недорогой и быстрый перевод, но на данный момент в этой области остается еще много неразрешенных проблем. Перевод сильно зависит от контекста и разнообразных коннотаций слов и словосочетаний. Полный контекст получается предоставить не всегда, и потому машинный перевод ограничен определенными тематиками, которые почти всегда нуждаются в постредактировании. По факту, в настоящее время системы МТ не стремятся полностью заменить переводчиков-людей, вместо этого они скорее выступают в качестве помощников.

В свою очередь, память переводов – это своеобразная лингвистическая база данных, в которой хранятся как исходные тексты, так и их переводы. Тексты в данной программе разделены на сегменты, которые часто совпадают с предложениями. ТМ позволяет облегчить процесс перевода – система «запоминает» переведенные сегменты для дальнейшего использования, благодаря чему переводчику не приходится заниматься переводом одних и тех же предложений во второй раз. Эти программы невероятно полезны для перевода текстов с высоким количеством повторов.

Технология компании PROMT изначально появилась как реакция на потребность в быстром и высококачественном машинном переводе больших объемов текстов в условиях, когда существующие на тот момент системы автоматизированного перевода, основанные в первую очередь на принципах пословного перевода (системы типа PROMT, Сократ или SYSTRAN), не были способны решить данную задачу. Благодаря этому возникла идея интеграции систем пословного машинного перевода с системами типа памяти переводов, которые позволяли сохранять двуязычные фразы, а затем использовать их в

дальнейшем процессе перевода. Однако данная интеграция не позволяет полностью решить поставленную задачу, так как при работе с разнотипными документами вероятность полного совпадения крупных сегментов текстов невероятно низка. Потому системы типа память переводов возможно эффективно использовать исключительно в качестве специального средства автоматизации труда переводчиков при переводе тематически однородных документов, таких как каталоги или однотипная техническая документация.

Проведенный в данной работе сопоставительный анализ и анализ адекватности перевода научно-технического текста показал, что в большинстве случаев результаты МП требуют редактирования. То насколько адекватным можно считать машинный перевод, определяется не только качеством его системы, но также и качеством постредктирования. Системы машинного перевода часто используются специалистами в виде подспорья для быстрого перевода, например, технической документации. В этом случае не возникает к тому же никаких проблем с корректным использованием терминов.

Благодаря интенсивному развитию сети Интернет сейчас наблюдается новый всплеск интереса людей к разнообразным системам МП. В Сети доминирует английский язык, но всегда будут пользователи, которые им не владеют, так же стоит учитывать большое количество разнообразных веб-страниц, которые могут быть написаны не по-английски. В данный момент существует большое количество дополнений к разным браузерам, которые могут осуществить как полный, так и частичный перевод сайтов. Одним из примеров переводчиков такого формата является система Web Trans Site фирмы PROMT.

В данный момент все перспективы развития автоматизированного перевода связаны с дальнейшей разработкой и углублением теории и практики перевода, как машинного, так и «человеческого». Для развития теории важны результаты многих отраслей:

- сопоставительного языкознания,

- общей теории перевода,
- теории закономерных соответствий,
- способов представления знаний,
- оптимизации,
- совершенствования лингвистических алгоритмов

Качество перевода лексических единиц возможно улучшить с помощью новых и более эффективных словарей с расширенной словарной информацией, строгих теорий терминологизации лексики, а также благодаря теории и практике работы с подязыками. Формальные грамматики, ориентированные на перевод, позволят оптимизировать алгоритмы поиска переводных соответствий в предложенной коммуникативной ситуации, которую можно будет описать в рамках соответствующих прикладных теорий представления знаний. Наконец, новые возможности программирования и вычислительной техники также будут вносить свой вклад в улучшение и дальнейшее развитие теории и практики машинного перевода.

Список использованных источников и литературы

1. Андреев, Н. Д. Теория машинного перевода, М., 1981
2. Багриновская Г. П., Кулагина О. С., Ляпунов А. А. О некоторых методологических вопросах, относящихся к машинному переводу// О некоторых вопросах теоретической кибернетики и алгоритмах программирования. - Новосибирск: Изд. СО АН СССР, 1971.
3. Беляева Л. Н., Откупщикова М. И. Автоматический (машинный) перевод. – В сб.: Прикладное языкознание. СПб, 1996
4. Большой англо-русский политехнический словарь: В 2 т. Около 200000 терминов. – М.: РУССО: Лаборатория базовых знаний, 2005.
5. Борисова Л. И. О переводном эквиваленте в научно-техническом переводе // Текст в языке и речевой деятельности: (состав, перевод, автомат. обраб.): сб. науч. тр. / АН СССР, Ин-т языкознания и др.; ред.-сост. Ванников Ю.В., Шахнарович А.М. – М.: ИЯЗ, 1987. - С. 42–43.
6. Виноград Т. Работа с естественными языками. М.: Высшая школа, 1986
7. Виноградов В. С. Перевод: Общие и лексические вопросы: Учебное пособие. - 3-е изд. - М.: КДУ, 2006. - 240 с.
8. Винокуров А. А., Чуканов В. О. Новый метод оценки машинного перевода // Информационные технологии и системы. Hardware Software Security. Тенденции и перспективы - Сборник статей: М., Международная академия информатизации, 1997.
9. Докштейн С. Я., Макаров Е. А., Радоминова С. С. Практический курс перевода [Электронный ресурс] // URL: http://zhurnal.lib.ru/w/wagapow_a_s/dokstein.shtml (01.06.21)
10. Едемский М. Программы автоматического перевода //Мир образования. - 1996. - №11-12. - С.54 – 55.
11. Жолковский А. К., Мельчук И. А. К построению действующей модели языка

- «Смысл – Текст». - В кн.: Машинный перевод и прикладная лингвистика. Вып. 11. - М., 1969
12. Иомдин Л. Л. Уроки машинного перевода для детей и взрослых // Лингвистика для всех. Зимняя лингвистическая школа. – М.: НИИРО, 2004. - С. 56–68.
 13. Иомдин Л. Л. Уроки русско-английского (из опыта работы системы машинного перевода) // Труды Международного семинара Диалог'2002 по компьютерной лингвистике и ее приложениям. Москва, Наука, 2002, Т. 2. С. 234–244.
 14. Климзо Б. Н. Ремесло технического переводчика: об англ. яз., переводе и переводчиках науч.- техн. лит. / Климзо Б. Н. – М.: Валент Р., 2003. – 286 с.
 15. Козеренко Е. Б. Логико-статистические методы представления языковых структур в машинном переводе// Труды Международной конференции «Компьютерная лингвистика и интеллектуальные технологии». - М.: Наука, 2005.
 16. Комиссаров В. Н. Теоретические основы методики обучения переводу / Комиссаров В. Н. - М.: Рема, 1997.
 17. Кулагина О. С. Исследования по машинному переводу. - М.: Наука, 1979. - 320 с.
 18. Кулагина О. С. Машинный перевод: современное состояние // Семиотика и информатика. - М., ВИНТИ, 1989. - Вып. 29
 19. Кулагина О. С. О современном состоянии машинного перевода // Математические вопросы кибернетики, вып. 3, М.: Наука, 1991, стр. 5—50.
 20. Лагунов И. Некоторые проблемы машинного перевода [Электронный ресурс] // Кафедра перевода ФИЯ ЯГУ: [сайт]. - [Б.г.]. – URL: <http://pages.ykt.ru/kperevoda/conf1.html> (04.06.21)
 21. Левин А. Самоучитель работы на компьютере. - 4-е изд., испр. и доп. М.: Нолидж, 1997.
 22. Людсканов А. Селективная стратегия при машинном переводе. - В кн.: Международный семинар по машинному переводу. М., ВЦП, 1975
 23. Ляпунов А. А. Основные проблемы машинного перевода // Вопр. языкознания. – 1956. - № 5. – С. 107-111.

24. Мамедова М. Г. Машинный перевод: эволюция и основные аспекты моделирования / Мамедова М. Г., Мамедова З. Ю.; Нац. акад. наук Азербайджана, Ин-т информ. технологий. - Баку: INFORMASIYA TEXNOLOGIYALARI, 2006. - 158 с.
25. Марчук Ю. Методы моделирования перевода. - М.: Наука, 1985. - 201 с.
26. Марчук Ю. Н. Теория и практика машинного перевода // Рус. филол. вестн. - 1996. - Т. 81.
27. Марчук Ю. Модель «текст-текст» и переводные соответствия в теории машинного перевода // Проблемы компьютерной лингвистики. - Минск: МГЛУ, 1997. - С. 21-29.
28. Марчук Ю. Проблемы машинного перевода. - М.: Наука, 1983. - 231 с.
29. Машинный перевод как информационная и технологическая реальность: науч.-аналит. обзор / АН СССР, ИНИОН. - М.: ИНИОН, 1988. - 36 с.
30. Машинный перевод и прикладная лингвистика. Проблемы создания системы автоматического перевода: [сб. ст.]. - М.: МГПИИЯ, 1986. - 152 с.
31. Мельчук И. А. Машинный перевод, М, 1976
32. Мосавимиянх Т. Общие проблемы англо-персидского машинного перевода // Перевод: традиции и современные технологии. - М., 2002. - С. 76-84.
33. Нелюбин Л. Л. Компьютерная лингвистика и машинный перевод. - М.: ВЦП, 1991. - 151 с.
34. Носов К. Машинный перевод: преодолевая языковые барьеры [Электронный ресурс] // IT Community United: [сайт]. - [2005]. - URL: <http://itc.ua/node/20029/04.06.21>.
35. Петрова Г. Г. Лексико-стилистические трансформации в англо-русских научно-технических переводах // Вопр. гуманитар. наук. - 2004. - № 1. - С. 188-193.
36. Пиотровский Р. Г. Текст, машина, человек. - Л.: Наука, 1975. - 327 с.
37. Пиотровский Р. Г. Инженерная лингвистика и теория языка. - Л.: Наука, 1979. - 112 с.

38. Попова Т. В. Категория определенности-неопределенности как проблема перевода: (на материале рус. и англ. яз.): автореф. дис. ... канд. филол. наук / Попова Т. В.; МГУ им. М. В. Ломоносова. - М., 2001. - 26 с.
39. Проблемы машинного перевода // Компьютерра. - 2002. - № 21. – С. 26-37.
40. Пумпянский А. Л. Чтение и перевод английской научной и технической литературы (лексика, грамматика, фонетика, упражнения). – Минск: Попурри, 1968.
41. Райс К. Классификация текстов и методы перевода // Вопросы теории перевода в зарубежной лингвистике. - М.: Прогресс, 1978. - С. 202-228.
42. Ревзин И. И., Розенцвейг В. Ю., Основы общего и машинного перевода. - М.: Высшая школа, 1964. – 243 с.
43. Роганов В. Р., Роганова С. М., Новосельцева М. Е. Методы искусственного интеллекта для машинного перевода текстов. – Пенза: Изд-во Пензенского ун-та, 2007. – 152 с.
44. Рябцева Н. К. Информационные процессоры и машинный перевод. - М.: Наука, 1986. - 166 с.
45. Хайрутдинов З. Р. Сравнительный анализ лексико-грамматических особенностей оригинальных и переводных текстов английского языка: автореф. дис. ... канд. филол. наук: 10.02.19 / Хайрутдинов З. Р.; Военный ун-т МО РФ. – М., 2008. – 19 с.
46. Цейтин Г. С. Методы синтаксического анализа, использующие предпочтение языковых конструкций (модели и эксперименты). – В кн.: Международный семинар по машинному переводу. – М.: Высшая школа, 1975. - С.131-133
47. Чеповский А. Неразрешимая проблема компьютерной лингвистики // Компьютерра. – 2002. - № 30 ;
48. Шаляпина З. М. Автоматический перевод: эволюция и современные тенденции // Вопросы языкознания. – 1996. - № 2, с. 105 - 117
49. Hutchins W.J. Machine Translation: Past, Present. Future. / Hutchins W. J. - Horwood,

N. Y. Etc. 1986. - 382. p.

50. Knight K. Automating knowledge acquisition for machine translation. AI Magazine 18, pp.81-96, 1997.

Литературные источники

51. Василенко О. Г. Рекуррентная Каскад-Корреляция обучающей архитектуры. 1993 [Электронный ресурс] // URL: http://mathenglish.ru/trs/fahlman1_rus.pdf (24.05.21)
52. Воронович В. В., Конспект лекций для студентов 5-го курса специальности «Современные иностранные языки». [Электронный ресурс] // URL: <https://fsc.bsu.by/wp-content/uploads/2015/12/Mashinny-j-perevod-konspekt-lektsij.pdf> (01.06.21)
53. Fahlman S. E. The Cascaded-Correlation Learning Architecture. School of Computer Science, Carnegie Mellone University, 1991 [Электронный ресурс] // URL: <https://www.semanticscholar.org/paper/The-Recurrent-Cascade-Correlation-Architecture-Fahlman/9e8cf03655d224b0994d0f9d4f5aa80bca07021a> (24.05.21)

The Recurrent Cascade-Correlation Architecture

Scott E. Fahlman May 9, 1991 CMU-CS-91-100

School of Computer Science Carnegie Mellon University Pittsburgh, PA 15213

Abstract

Recurrent Cascade-Correlation (RCC) is a recurrent version of the Cascade-Correlation learning architecture of Fahlman and Lebiere [Fahlman, 1990]. RCC can learn from examples to map a sequence of inputs into a desired sequence of outputs. New hidden units with recurrent connections are added to the network one at a time, as they are needed during training. In effect, the network builds up a finite-state machine tailored specifically for the current problem. RCC retains the advantages of Cascade-Correlation: fast learning, good generalization, automatic construction of a near-minimal multi-layered network, and the ability to learn complex behaviors through a sequence of simple lessons. The power of RCC is demonstrated on two tasks: learning a finite-state grammar from examples of legal strings, and learning to recognize characters in Morse code.

This research was sponsored in part by the National Science Foundation (Contract EET-8716324) and the Defense Advanced Research Projects Agency (Contract F33615-90-C-1465). The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Government or any of its agencies.

The Architecture

Cascade-Correlation [Fahlman, 1990] is supervised learning architecture that builds a near-minimal multi-layer network topology in the course of training. Initially the network contains only inputs, output units and the connection between them. This single layer of connections is trained (using the Quickprop algorithm [Fahlman, 1988]) to minimize the error. When no further improvement is seen in the level of error, the network's performance is evaluated. If the error is small enough, we stop. Otherwise we add a new hidden unit to the network in an attempt to reduce the residual error.

To create a new hidden unit, we begin with a pool of candidate units, each of which receives weighted connections from the network's inputs and from any hidden units already present in the net. The outputs of these candidate units are not yet connected into the active network. Multiple passes through the training set are run, and each candidate unit adjusts its incoming weights to maximize the correlation between its output and the residual error in the active net. When the correlation scores stop improving, we choose the best candidate, freeze its incoming weights, and add it to the network. This process is called «tenure». After tenure, a unit becomes a permanent new feature detector in the net. We then re-train all the weights going to the output units, including those from the new hidden unit. This process of adding a new hidden unit and re-training the output layer is repeated until the error is negligible or we give up. Since the new hidden unit receives connection from the old ones, each hidden unit effectively adds a new layer to the net. (See figure 1)

Cascade-correlation eliminates the need for the user to guess in advance the network's size, depth, and topology. A reasonably small (though not minimal) network is built automatically. Because a hidden-unit feature detector, once built, is never altered or cannibalized, the network can be trained incrementally. A large

data set can be broken up into smaller “lessons,” and feature-building will be cumulative.

Cascade-Correlation learns much faster than backprop for several reasons: First only a single layer of weights is being trained at any given time. There is never any need to propagate error information backwards through the connections, and we avoid the dramatic slowdown that is typical when training backprop nets with many layers.

Second, this is a “greedy” algorithm: each new unit grabs as much of the remaining error as it can. In a standard backprop net, the all the hidden units are changing at once, competing for the various jobs that must be done - a slow and sometimes unreliable process.

Cascade-correlation, like back-propagation and other feed-forward architectures, has no short-term memory in the network. The outputs at any given time are a function only of the current inputs and the network’s weights. Of course, many real-world tasks require the recognition of a sequence of inputs and, in some cases, the corresponding production of a sequence of outputs.

A number of recurrent architectures have been proposed in response to this need. Perhaps the most widely used, at present, is the Elman model [Elman, 1988], which assumes that the network operates in discrete time-steps. The outputs of the network’s hidden units at time t are fed back for use as additional network inputs at time-step $t + 1$. (See figure 2.) These additional inputs can be thought of as state-variables whose contents and interpretation are determined by the evolving weights of the network. In effect, the network is free to choose its own representation of past history in the course of learning.

Recurrent Cascade-Correlation (RCC) is an architecture that adds Elman-style recurrent operation to the Cascade-Correlation architecture. However, some changes were needed in order to make the two models fit together. In the original

Elman architecture there is total connectivity between the state variables (previous Outputs

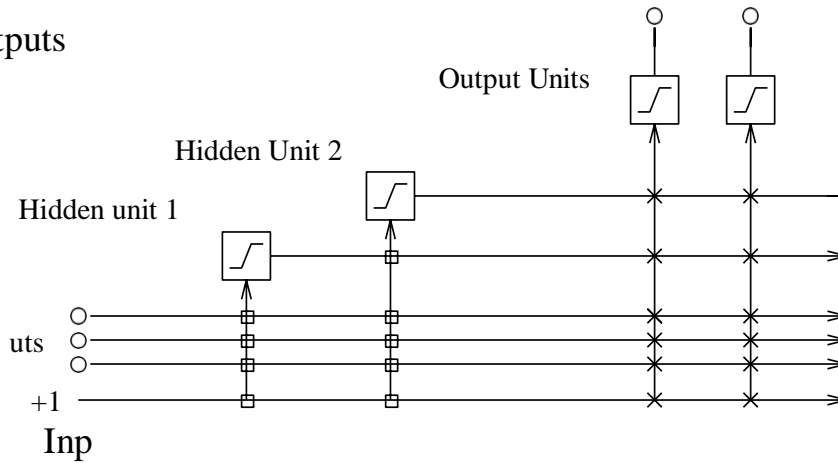


Figure 1: The Cascade-Correlation architecture after two hidden units have been added. The vertical lines sum all incoming activation. Boxed connections are frozen, X connections are trained repeatedly.

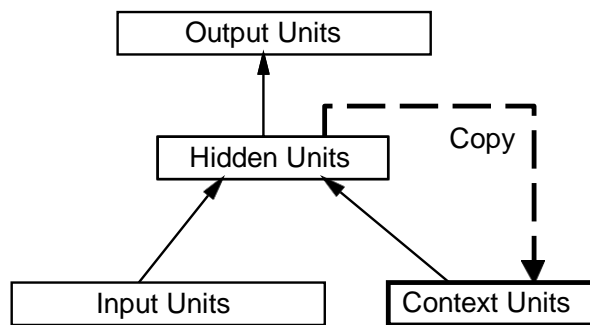


Figure 2: The recurrent network architecture of Elman.

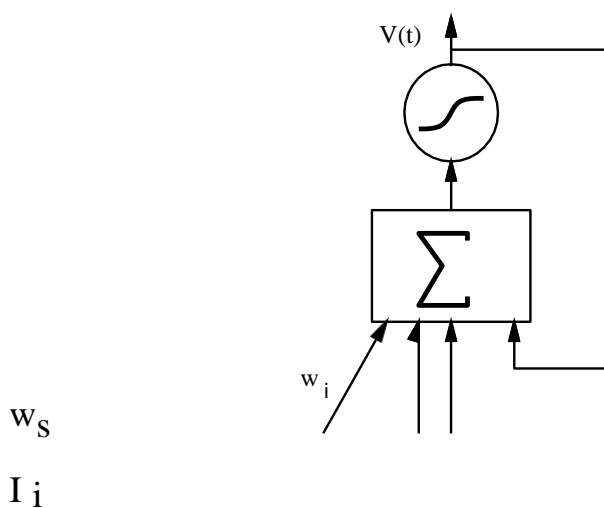


Figure 3: Candidate or hidden unit with a self-recurrent link.

outputs of hidden units) and the hidden unit layer. In Cascade-Correlation, new hidden units are added one by one, and are frozen once they are added to the network. It would violate this concept to insert the outputs from new hidden units back into existing hidden units as new inputs. On the other hand, the network must be able to form recurrent loops if it is to retain state for an indefinite time.

The solution we have adopted in RCC is to augment each candidate unit with a single weighted self-recurrent input that feeds back that unit's own output on the previous time-step (figure 3). That self-recurrent link is trained along with the unit's other input weights to maximize the correlation of the candidate with the residual error. If the recurrent link adopts a strongly positive value, the unit will function as a flip-flop, retaining its previous state unless the other inputs force it to change; if the recurrent link adopts a negative value, the unit will tend to oscillate between positive and negative outputs on each time-step unless the other inputs hold it in place; if the recurrent weight is near zero, then the unit will act as a gate of some kind. When a candidate unit is added to the active network as a new hidden unit, the self-recurrent weight is frozen, along with all the other weights. Each new hidden unit is in effect a single state variable in a finite-state machine that is built specifically for the task at hand. In this use of self-recurrent connections only, the RCC model resembles the "Focused Back-Propagation" algorithm of Mozer [Mozer, 1988].

The output, $V(t)$, of each self-recurrent unit is computed as follows:

$$V(t) = \sum_i I_i(t) w_i + V(t-1) w_s$$

where σ is some non-linear squashing function applied to the weighted sum of inputs I plus the self-weight, w_s , times the previous output. In the studies described here, σ is always the hyperbolic tangent or "symmetric sigmoid" function, with a range from -1 to +1. During the candidate training phase, we adjust the weights w_i and w_s

for each unit so as to maximize its correlation score. This requires computing the derivative of $V(t)$ with respect to these weights:

$$\begin{aligned} \frac{\partial V(t)}{\partial w_i} &= O'(t) I_i(t) + w_s \frac{\partial V(t-1)}{\partial w_i} \\ \frac{\partial V(t)}{\partial w_s} &= O'(t) V(t-1) + w_s \frac{\partial V(t-1)}{\partial w_s} \end{aligned}$$

The rightmost term reflects the influence of the weight in question on the unit's previous state. Since we computed $\frac{\partial V(t-1)}{\partial w}$ on the previous time-step, we can just save this value and use it in the current step. So the recurrent version of the learning algorithm requires us to store a single additional number for each candidate weight, plus $V(t-1)$ for each unit. At $t = 0$ we assume that the unit's previous value and previous derivatives are all zero.

As an aside, the usual formulation for Elman networks treats the hidden units' previous values as independent inputs, ignoring the dependence of these previous values on the weights being adjusted. In effect, the rightmost terms in the above equations are being dropped, though they are not negligible in general. This rough approximation apparently causes little trouble in practice, but it might explain the instability that some researchers have reported when Elman nets are run with aggressive second-order learning procedures such as quickprop. The Mozer algorithm does take these extra terms into account.

Приложения

Приложение № 2

Школа Информационных Технологий

Университет Карнеги-Меллона

Перевод О. Г. Василенко

Аннотация

Рекуррентная Каскад-Корреляция (РКК) является рекуррентной версией Каскад-корреляционной обучающей архитектуры Фальмана и Либьера (Fahlman and Lebiere) [Fahlman, 1990]. РКК может научиться с помощью примеров сопоставлять последовательность входов и желаемую последовательность выходов. Новые скрытые узлы с рекуррентными соединениями добавляются в сеть во время обучения по одному в тот момент, когда они становятся необходимы. По сути, сеть строит конечный автомат специально для текущей задачи. РКК сохраняет преимущества Каскад-Корреляции: быстрое обучение, хорошее обобщение, автоматическое построение практически минимальной многослойной сети, а также возможность изучать сложные поведения через последовательность простых уроков. Мощность РКК демонстрируется на двух задачах: создание конечной грамматики из примеров правильных строк, и обучениераспознавать символы азбуки Морзе.

1. Архитектура

Каскад-корреляция [Fahlman, 1990] является обучаемой системой с учителем, которая строит почти минимальную многослойную сеть во время курса обучения. Изначально сеть содержит только входы и выходы, а также

связи между ними. Эта однослойная сеть обучается (с помощью алгоритма быстрого распространения Quickprop [Fahlman, 1988]) так, чтобы свести к минимуму ошибки. Когда уровень ошибок перестает уменьшаться - оценивается эффективность сети. Если ошибка достаточно мала, то мы останавливаемся. В противном случае мы добавляем новые скрытые узлы к сети, чтобы уменьшить остаточную ошибку.

Чтобы создать новые скрытые узлы, мы начинаем с набора *узлов-кандидатов*, каждый из которых получает на вход сигналы от входов сети и других скрытых узлов, уже присутствующих в сети. Выходы узлов-кандидатов еще не связаны с другими узлами. Мы совершаем несколько проходов по обучающему множеству, и каждый узел-кандидат регулирует свои весовые коэффициенты входящих связей, чтобы увеличить корреляцию между его выходом и остаточной ошибкой сети. Когда оценка корреляции перестает улучшаться, мы выбираем лучшего кандидата, замораживаем его весовые коэффициенты входящих связей и добавляем его к сети. Этот процесс называется "завладением". После завладения, кандидат становится полноценным новым узлом сети. Затем мы заново обучаем все коэффициенты на исходящих связях, в том числе и для нового узла. Этот процесс добавления нового узла и переобучения выходного слоя повторяется, пока ошибка не становится незначительной или мы не решаем сдаться и остановиться. Поскольку новые узлы получают входящие соединения от старых, каждый узел эффективно добавляет новый слой к сети. (См. рисунок 1.)

Каскад-корреляция исключает для пользователя необходимость предугадывать размеры сети, ее глубину и топологию. Достаточно малая (но не минимальная) сеть строится автоматически. Поскольку новые узлы после добавления никогда не меняют своих связей и не удаляются, сети могут быть обучены постепенно. Большой набор данных можно разбить на более мелкие "уроки" и увеличение потенциала будет нарастать.

Каскад-корреляция обучает гораздо быстрее, чем сеть с обратным распространением ошибки по нескольким причинам:

В начале только один слой весовых коэффициентов обучается в любой данный момент времени. Никогда не возникает необходимости распространять информацию об ошибках обратно через связи, и мы избегаем драматического замедления, которое является типичным при обучении многослойных сетей с обратным распространением ошибки. Во-вторых, это "жадный" алгоритм: каждая новая единица захватывает так много оставшейся ошибки, как только может.

В стандартной сети с обратным распространением ошибки, все скрытые узлы меняются одновременно, конкурируя за различные задачи, что должны быть достигнуты — это медленный и иногда ненадежный процесс.

Каскад-корреляция, как и обратное распространение и другие архитектуры с прямой связью, не имеет кратковременной памяти в сети. Выходы в любой данный момент времени являются функцией только от текущих входов и коэффициентов при связях в сети. Конечно, многие реальные задачи требуют распознавания последовательности входов и, в некоторых случаях, создания соответствующей последовательности выходов.

Некоторое количество рекуррентных архитектур было предложены в ответ на эту необходимость. Пожалуй, наиболее широко используемой в настоящее время является модель Эльмана [Elman, 1988], которая предполагает, что сеть оперирует дискретными промежутками времени. Выходы скрытых узлов сети в момент времени T подаются обратно для использования в качестве дополнительных входов сети в момент времени $T + 1$. (См. Рисунок 2). Эти дополнительные входы могут рассматриваться как переменные состояния, содержание и интерпретация которых определяется меняющимися весами сети. По сути, сеть свободна в выборе ее собственного представления прошлой истории в ходе обучения.

Рекуррентная Каскад-корреляция (РКК) - это архитектура, которая добавляет в Каскад-корреляционную архитектуру рекурсивное поведение в стиле Эльмана. Однако, были необходимы некоторые изменения для того, чтобы две модели подходили друг к другу. В оригинальной архитектуре Эльмана существует общая связь между переменными состояниями (предыдущими выходами скрытых узлов) и слоем скрытых узлов.

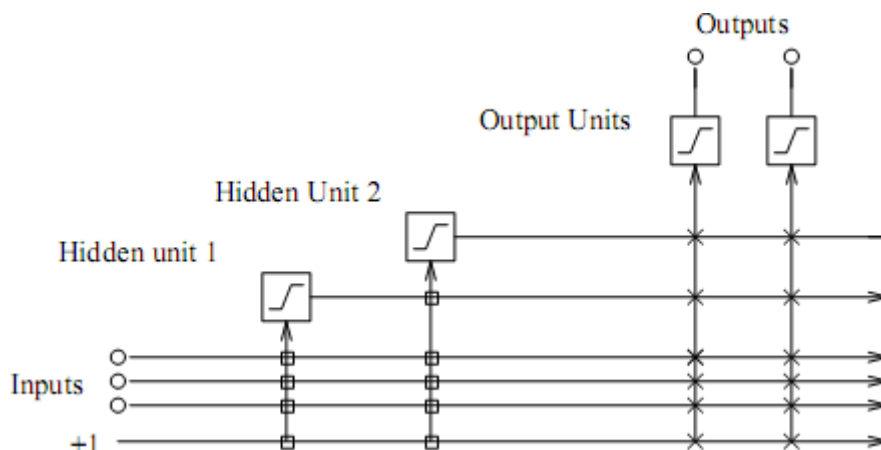


Рисунок 1: Каскадно-корреляционная архитектура, после того как было добавлено два скрытых узла. Вертикальные линии суммируют все входящие импульсы. Соединения в квадратах заморожены, а X-соединения непрерывно обучаются.

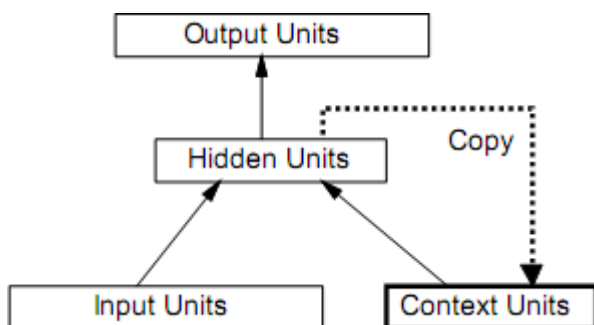


Рисунок 2: Рекуррентная структура сети Эльмана

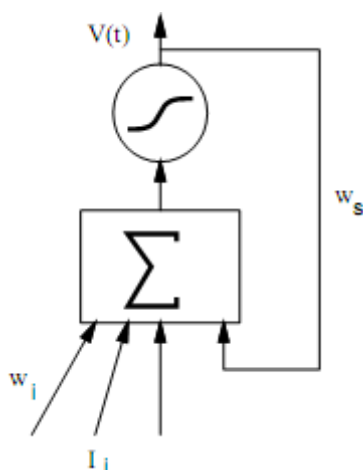


Рисунок 3: Узел-кандидат или скрытый узел с само-рекуррентной связью.

При Каскад-корреляции, новые скрытые узлы добавляются по одному и их состояние замораживается, как только они будут добавлены в сеть. Эта концепция была бы нарушена, если бы выходы новых узлов соединялись со входами уже существующих. С другой стороны, сеть должна иметь возможность создавать рекуррентные петли, если она хочет поддерживать свое состояние в течение неопределенного времени.

Решение, которое мы приняли в РКК: дополнить каждый узел-кандидат одной само-рекуррентной связью с коэффициентом, которая возвращает на вход узла его вывод на предыдущем шаге (рис. 3). Эта само-рекуррентная связь обучается вместе с другими входными коэффициентами узла, чтобы увеличить корреляцию узла с остаточной ошибкой. Если рекуррентная связь примет строго положительное значение — узел будет функционировать в качестве триггера, сохраняя свое прежнее состояние, если другие узлы не заставят его измениться. Если рекуррентная связь примет строго отрицательное значение - узел, как правило, будет колебаться между

положительным и отрицательным значением на выходе в каждый следующий шаг времени, если значение на другом входе не удержит его на месте. А если рекуррентный вес близок к нулю, то узел будет выступать в качестве своеобразного логического вентиля. Когда узел-кандидат добавляется к активной сети как новый скрытый узел, его рекуррентный вес замораживается наряду со всеми другими весами. Каждый новый скрытый узел, по сути, является одной переменной состояния в конечном автомате, создаваемом специально для выполнения этой задачи. При использовании только само-рекуррентных связей модель РКК походит на алгоритм «Сфокусированного обратного распространения» Мозера [Mozer, 1988].

Выход $V(t)$ каждого само-рекуррентного узла рассчитывается следующим образом:

$$V(t) = \sigma \left(\sum I_i(t) w_i + V(t-1) w_S \right)$$

Где σ является некой нелинейной сжимающей функцией, применяемой к взвешенной сумме входов I плюс рекуррентный вес w_S , умноженный на результат предыдущего выхода. В примерах, описанных тут σ всегда является гиперболическим тангенсом или «симметричной сигма-функцией» в пределах $[-1; 1]$. В течение фазы обучения узла- кандидата мы изменяем веса w_i и w_S для каждого узла так, чтобы увеличить его корреляцию. Это требует вычисления производной от $V(t)$ с учетом этих весов:

$$\partial V(t) / \partial w_i = \sigma'(t) I_i(t) w_S \partial V(t-1) / \partial w_i$$

$$\partial V(t) / \partial w_S = \sigma'(t) V(t-1) w_S \partial V(t-1) / \partial w_S$$

Крайне правый член отражает влияние веса в вопросе на предыдущее состояние узла. Т.к. Мы вычислили $\partial V(t-1) / \partial w$ на предыдущем шаге, мы можем просто сохранить это значение и использовать его на текущем шаге. Итак, рекуррентная версия обучающего алгоритма вынуждает нас хранить

одно дополнительное число для веса каждого узла- кандидата, плюс $V \cdot t - 1$ для каждого узла. В момент $t = 0$ мы подразумеваем, что предыдущее значение узла и предыдущие производные все равны нулю.

Как побочный эффект, обычная формулировка сетей Эльмана считает предыдущие значения скрытых узлов *независимыми* входами, игнорируя зависимость этих предыдущих значений от изменения весов. В сущности, крайние правые части уравнений выше отбрасываются, хотя в общем случае они не являются незначительными. Это грубое приближение очевидно причиняет небольшие проблемы на практике, но оно может объяснить ту нестабильность, которую отметили некоторые исследователи, когда сети Эльмана выполняются с агрессивными процедурами обучения второго порядка, такими как quickprop (быстрое распространение). Алгоритм Мозера принимает во внимание эти дополнительные условия.