

Технологии информационного общества**О НЕКОТОРЫХ ОСОБЕННОСТЯХ КОМПЬЮТЕРНОГО ПЕРЕВОДА
ОМОНИМОВ И ПОЛИСЕМАНТИЧНЫХ СЛОВ**

Статья рекомендована к публикации главным редактором Т. В. Ершовой 12.10.2023.

Бозиев Альберд Тахирович

*Кандидат филологических наук, доцент
Кабардино-Балкарский государственный университет
Кабардино-Балкарский научный центр Российской академии наук
Нальчик, Российская Федерация
alberdboziev@mail.ru*

Макоева Дана Гисовна

*Кандидат филологических наук, доцент
Кабардино-Балкарский научный центр Российской академии наук
Нальчик, Российская Федерация
makoevadana@mail.ru*

Аннотация

В статье рассматриваются вопросы, связанные со сложностью лингвистической трактовки полисемичных слов и омонимов. В настоящее время исследователи прилагают большие усилия для достижения такого качества перевода, который не отличался бы от перевода, выполненного человеком. В работе анализируется компьютерный перевод на английский язык некоторых русских полисемичных слов и омонимов. Помимо этого, предлагается метод формализации семантики естественного языка для задач машинного перевода на основе мультиагентных нейрокогнитивных архитектур. Данный метод, основанный на функциональной репрезентации понятий, входящих в состав естественно-языкового высказывания, по мнению авторов, может стать мощным и надежным инструментом для решения ряда переводческих проблем, в том числе при работе с многозначными словами.

Ключевые слова

омонимы; полисемичные слова; компьютерный перевод; машинный перевод; мультиагентная система; когнитивная архитектура; формальная семантика; нейронные сети

Введение

Целью данной статьи является анализ некоторых особенностей компьютерного перевода на английский язык русских предложений, имеющих в своём составе полисемичные слова и омонимы. Предметом исследования послужили примеры из интернет-ресурсов.

Под омонимами в работе понимаются слова, относящиеся к одной части речи и имеющие сходное звучание и написание. Как правило, это лексические омонимы. Полисемантические слова обладают несколькими значениями.

Начиная с 1920-30-х годов проблему соотношения омонимии и полисемии исследуют многие зарубежные и отечественные лингвисты: В. И. Абаев, И. В. Арнольд, О. С. Ахманова, Р. А. Будагов, В. В. Виноградов, Б. А. Ильиш, С. Д. Кацнельсон, Е. Р. Курилович, Л. В. Малаховский, А. . Реформатский, Д. Э. Розенталь, В. Н. Сидоров, А. И. Смирницкий, Ю. П. Сорокин, И. С. Тышлер, E. Buysens, O. espersen, B. Trnka B. U. Weinreich, J. Wells и многие другие.

Лингвистами высказываются противоположные мнения относительно вредности или полезности таких явлений в языке, какими являются омонимы и многозначные слова. По мнению

© Бозиев А. Т., Макоева Д. Г., 2024

Производство и хостинг журнала «Информационное общество» осуществляется Институтом развития информационного общества.

Данная статья распространяется на условиях международной лицензии Creative Commons «Атрибуция — Некоммерческое использование — На тех же условиях» Всемирная 4.0 (Creative Commons Attribution – NonCommercial – ShareAlike 4.0 International; CC BY-NC-SA 4.0). См. <https://creativecommons.org/licenses/by-nc-sa/4.0/legalcode.ru>

https://doi.org/10.52605/16059921_2024_03_118

некоторых из них, данные явления тормозят безмерное раздувание языковых форм (Ф.И. Маулер и Э. Бьюиссанс), другие же считают, что они привносят в язык путаницу и двусмысленность (Э. Эман и А. Реформатский). Интересно в связи с этим мнение известного лингвиста С.О. Карцевского. Говоря об ассиметричном дуализме языкового знака, в случаях, когда план выражения расходится с планом содержания, он подчёркивал, что асимметрия является важным фактором функционирования языка [1, с.87]. Обе точки зрения вполне состоятельны, но трудно признать одну из них главенствующей.

Перед лингвистами остро встаёт вопрос о критериях разграничения омонимов и многозначных слов. Большинство из них признаются следующие критерии:

- этимологический (или исторический) критерий позволяет определить наличие или отсутствие родственных связей происхождения – у омонимов их нет;
- словообразовательный критерий – омонимы имеют различные словообразовательные схемы, а многозначные слова имеют одинаковые или близкие ряды;
- синтаксический критерий позволяет определить сочетаемость – у многозначных слов она совпадает, у омонимов она разная;
- семантический критерий – если разные слова не имеют общих сем, то это омонимы, если имеют – то это полисемное слово;
- контекстный критерий – «окружающий речевой контекст и сама ситуация общения снимают полисемию и достаточно ясно указывают, какое из значений имеется в виду» [1].

Часто приводится и косвенный лексикографический критерий – в толковых словарях омонимы оформляются как отдельные словарные статьи, а многозначное слово как одна единица.

Р.В. Кузьмина пишет, что «...принципиальное различие между омонимией и полисемией заключается в том, что при омонимии мы имеем дело с разными означаемыми, тогда как при полисемии означаемое одно и то же» [3].

Р.В. Кузьмина отмечает, что при полисемии все языковые единицы, имеющие тождественные означающие, связаны деривационными отношениями, тем самым образуя единый словесный знак, при омонимии такие связи полностью отсутствуют [3].

По мнению А. И. Головня, разграничение полисемии и омонимии полезно в трёх смыслах: теоретическом (философском и лингвистическом), прикладном (формализации языка при создании различных интеллектуальных систем), методическом (при обучении русскому языку и составлении словарей) [4]. Данный автор заключает, что в лингвистике остаётся спорной разграничительная линия между полисемией и омонимией [4]. В результате при формализации понятий полисемии и омонимии используется одна и та же методика [4].

Говоря об операциональных процедурах, исследователь И. Ф. Вардуль сводит их к трём приёмам: дистрибутивный (изучение окружающих слов); трансференционный (изучение единиц более высокого уровня); трансформационный (единообразная перестройка рассматриваемых единиц) [цитируется по 4].

Многие лингвисты полагают, что при отграничении омонимов от многозначных слов семантика слова должна быть решающим фактором, но методика этого анализа во многом не ясна. Большинство исследователей русского языка признаётся, что в данном случае необходимо одновременное применение трёх критериев – семантического, лексического и морфологического.

Вот уже несколько десятилетий разработчиками программ машинного перевода уделяется большое внимание вопросам разрешения полисемии и омонимии. Они отмечают, что из-за неоднозначности эти явления представляют наибольшую сложность при обработке естественного языка [5].

Исследователь В. Б. Скромблевич, выделяя ряд операций, входящих в анализ текста, говорит о следующих видах анализа – графематическом, морфологическом, синтаксическом и семантическом. Данный автор указывает на то, что в машинном переводе для разрешения лексической омонимии часто используется метод интерактивного разрешения неоднозначности, эффективность которого доказал научный коллектив Института проблем передачи информации РАН. Суть данного метода заключается в том, что с помощью толковых словарей исходного и целевого языков составляются смысловые дополнения, которые в процессе кодирования позволяют учесть большое количество лексических и грамматических значений. Как отмечает В. Б. Скромблевич, повышение качества машинного перевода – трудоёмкая задача. Её суть в том, чтобы единицей описания стало лексическое значение [5].

1 Материал и методика исследования

Материалом для данного исследования послужили русские и английские предложения, взятые из интернет-ресурсов [6, 7]. Перевод русских предложений на английский осуществлялся с помощью компьютерного переводчика [8]. В работе использовались следующие словари: [9, 10, 11, 12].

В общей сложности было проанализировано свыше 1000 русских примеров на омонимы и многозначные слова. В итоге исследовательский интерес остановился на 17 единицах – омонимах и многозначных словах. К ним относятся следующие слова: свет, коса, ключ, приём, мир, лавка, земля, брак, труба, мат, отделение, перевод, бокс, хвост, шишка, пробка, налет. Полученный сокращенный вариант предложений был переведён на английский язык с помощью компьютерного переводчика. В качестве предмета анализа были выбраны по 500 параллельных примеров из каждого из языков. До начала выполнения практической части работы были рассмотрены спорные теоретические вопросы, касающиеся проблем разграничения омонимов и полисемантических слов.

2 Особенности компьютерного перевода омонимов и полисемантических слов

В словаре С. И. Ожегова [10, с. 523] приводится 3 значения слова «приём» в одной вокабуле:

- отдельное движение,
- способ,
- собрание лиц.

Несмотря на отсутствие объединяющей семы, оно представлено как полисемантическое слово, на самом же деле мы имеем дело с тремя разными словами, которые являются омонимами. Изолированное предъявление для перевода компьютеру лексемы приём даёт английскую лексему *reception*, вероятнее всего в силу высокой статистической рекуррентности. Имеющиеся компьютерные программы и комментарии позволили субъекту перевода безошибочно перевести на английский язык следующие сочетания и фразы из русских предложений: приём в партию – *admission to the party*; прием наглядного обучения – *the method of visual learning*; приём пищи – *eating*; оказан благоприятный приём – *was given a favorable reception*; тактический прием осложнил борьбу войск – *a tactic complicated the struggle of troops*; особый прием наказывать – *a special way of punishing*; какой прием ожидает нас – *what kind of reception awaits us*; приём на работу – *hiring*; педагогический прием – *pedagogical technique*; попасть на приём к – *getting an appointment with*; прием лекарства – *taking medicine*; обучающий приём – *teaching method*; выпить в один приём – *drink in one go*.

В зависимости от контекста различным значениям русского слова приём соответствуют более 10 точных английских эквивалентов. При этом компьютер испытывал затруднения при переводе следующих предложений:

Там были больные, приехавшие из станиц на прием к Василию Федоровичу. – On the wagons lay the sick, who had come from the villages to receive Vasily Fedorovich.

Если из предложения убрать сочетание «из станиц», то получаем: *There were patients who came to see Vasily Fedorovich.*

Экспериментально составленное предложение

Он провёл неожиданный приём, после которого соперник сдался» компьютер перевёл как He held an unexpected reception, after which the opponent surrendered.

Можно ли понять это предложение двояко? Вероятнее всего, да. Видоизменённые предложения имели следующий вид:

Он провёл великолепный приём, после чего соперник сдался. – He had a great reception, after which the opponent surrendered.

Он провёл великолепный борцовский приём, после чего соперник сдался. – He had a great wrestling move, after which the opponent surrendered.

Отличие предложений заключается в одном слове: «борцовский», которого оказалось достаточно, чтобы реализовать задуманное авторами, сработала биграмма.

Рассмотрим слово «ключ». В словаре С. И. Ожегова [10, с. 247] приводится 5 значений слова «ключ» в прямом и переносном значении: отмычка, разгадка, музыкальный термин, а также его омоним (отдельной вокабулой) со значением «источник воды». Ю. Яесс [9] даёт 20 значений данного слова в одной словарной статье. Приведём примеры:

- ключ от сейфа – the key to the safe;
- повернул ключ в замке зажигания – turned the key in the ignition;
- прошел сквозь запертую на ключ дверь – walked through a locked door;
- схватив тяжелый ключ – grabbing a heavy key;
- ключ к объяснению – the key to explaining;
- дверь он закрыл на ключ – he closed the door with a key;
- употребляются три ключа: скрипичный (или ключ G [«соль»]), басовый (или ключ F [«фа»]) и альтовый (или ключ C [«до»]) – three clefs are used: treble (or G clef [“sol”]), bass (or F clef [“fa”]), and alto (or C clef [“to”]);
- ключ от сейфа – the key to the safe.

Компьютер успешно справился с переводом различных значений слова ключ – в прямом и переносном значении, но испытывал затруднения с переводом омонима со значением источник:

Действовали как горный ключ на изнывающего от жажды путника. – Acted like a mountain key on a thirsty traveler.

Я был отправлен на ключ Алмазный. – I was sent to the Diamond key.

Ключ находился где-то в южной части Тихого океана – The key was somewhere in the South Pacific.

Экспериментальные предложения компьютер перевёл следующим образом:

Я выпил воды из ключа – I drank water from the key.

Ключ как источник воды – key as a source of water.

Ожидаемое слово spring появилось только в атрибутивной фразе: ключевая вода – spring water. Если фразу довести до предложения, то получим: ключевая вода стоит дорого – spring water is expensive.

Рассмотрим слово «коса». В словаре С. И. Ожегова [10, с. 264] приводится 3 разных слова с написанием «коса» со следующими значениями:

- сплетённая прядь волос,
- металлическое орудие,
- удалённая полоса земли.

Ю. Яесс [9] даёт 20 значений данного слова в одной словарной статье.

- там шла песчаная коса – there was a sandy spit;
- пусть у нее и коса своя, и румянец свой – let her have her own braid and her blush;
- его орудиями были не плуг, не коса и не серп... his tools were not a plow, not a scythe and not a sickle...
- густая, длинная коса – a thick, long braid;
- серп, коса, мыло, спички – a sickle, a scythe, soap, matches;
- жестокая коса войны – the cruel scythe of war...

Субъект перевода испытал затруднения в примерах:

Значит, коса крепкая, косить ею будет легко. – This means that the braid is strong, it will be easy to mow with it.

Ее красота и коса не остались незамеченными. – Her beauty and scythe did not go unnoticed.

После повторного обращения к переводчику второе предложение приняло правильный вид:

Her beauty and braid did not go unnoticed.

Последующие обращения каждый раз давали такой же результат. Можно предположить, что положительный эффект дала сочинительная связь и расширенное семантическое поле: «красота» – «коса». Эксперименты с первым предложением дали следующие результаты:

Значит, крепкая коса, косить ею будет легко. – So, strong scythe, it will be easy to mow with it.

При разделении «крепкая» и «коса» получаем:

So, a strong braid, it will be easy to mow with it.

Примечательно, что нет никакой реакции на правую часть предложения. Непонятно, какую роль играет слитное или раздельное написание данного словосочетания.

Рассмотрим слово «свет». С. И. Ожегов [10, с. 624] даёт два разных слова «свет» – в физическом смысле, прямом и переносном (4 значения) и второе слово со значением «мир», «земля» (2 значения).

- на свет появлялась – was born;
- зажечь свет – to turn on the light;
- с момента своего первого выхода в свет – from its first publication;
- свет фар скоро был замечен – the headlights were soon seen;
- свет не без добрых людей – the world is not without good people;
- затем она вышла в свет на многих языках мира, даже на японском – then it was published in many languages of the world, even in Japanese.

Переводчик затруднялся в переводе предложений, в которых слово «свет» употреблено в переносном смысле:

- позволили этой книге увидеть свет – allowed this book to see the light of day;
- появился на Божий свет – I came into the light of God.

Рассмотрим слово «мир». По С. И. Ожегову [10, с. 312] слово «мир» представлено 2 омонимами: в значении «Земля» в прямом и переносном значении данное слово имеет 5 ЛСВ, в значении «покой» 2 значения.

- мир секретности и мир открытости – the world of secrecy and the world of openness;
- для меня открылся совершенно новый мир – a completely new world opened up for me;
- вынужден был пойти на мир – was forced to make peace;
- красота помогает понять мир – beauty helps to understand the world;
- приблизить реальный мир к миру иному – to bring the real world closer to another world;
- мир и единение – the peace and unity;
- это были демонстрации за мир – these were demonstrations for peace.

Все значения омонимичной пары компьютер перевёл верно.

Рассмотрим слово «лавка». Согласно С. И. Ожегову [10, с. 279], «лавка» (скамейка) и «лавка» (небольшой магазин) – два разных слова.

- открылась лавка – a shop opened;
- лавка моего отца – my father's shop;
- у них была овощная лавка – they had a greengrocer's shop;
- виднелась запорошенная снегом речная лавка – a snow-covered bench could be seen;
- книжная лавка – a bookstore;
- тут лавка со спичками, солью, мылом... - there is a shop with matches, salt, soap...
- здесь есть мясная лавка, булочная, почта ... - there is a butcher shop, a bakery, a post office...

Только в одном случае компьютер допустил ошибку:

Росло огромное тенистое дерево, вокруг которого была фигурная лавка. – Grew a huge shady tree, around which was a figured shop.

Переводчик оказался чувствителен к небольшому добавлению к тексту справа:

Вокруг которого была фигурная лавка из дуба – around which was a figured bench made of oak.

Рассмотрим слово «земля». Согласно С. И. Ожегову [10, с. 206], слово «земля» имеет 6 значений (полисемантическое слово).

В темноте скрывались земля и люди. – In the darkness hid the earth and the people.

Земля переходила из рук в руки. – The land passed from hand to hand.

Самое мощное на планете Земля. – The most powerful on the planet Earth.

Вахтенный на носу закричал «земля!» – The watchman on the bow shouted “land!”

Земля снова стала бедной. – The land again became poor.

Относим к классу «земля-земля». – Refer to the ground-to-ground class.

Рыхлая земля превратилась в грязевую топь. – The loose earth turned into a mud swamp.

Эта земля была богата различными плодами. – This land was rich in various fruits.

... Или земля сотворена в семь дней? – Or was the earth created in seven days?

... Чтобы земля снова приносила урожай. – ... To make the land bear fruit again.

Кто-то полагал, будто Земля плоская. – Someone believed that the Earth was flat.

Компьютер успешно справился с переводом различных значений слова «земля».

Рассмотрим слово «брак». В словаре С. И. Ожегова [10, с. 34] приводится два разных слова (омонима) с написанием брак: относящийся к взаимоотношениям между людьми и отступающий от норм. В данном случае Ю. Яесс [9] даёт одну вокабулу с двумя не связанными между собой значениями.

...И этот брак оказался весьма прочным. – ... And this marriage turned out to be very strong.

Брак его племянницы несчастен. – His niece's marriage was unhappy.

Он вступал в брак по законам церкви. – He entered into marriage according to the laws of the church.

Это был брак по расчету, стратегический брак. – It was a marriage of convenience, a strategic marriage.

Переводчик не проводит различия между омонимичной парой.

Он начал допускать много брака в работе, его уволили. – He began to allow a lot of marriage in his work, he was fired.

Брак в работе говорит о многом. – Marriage at work speaks volumes.

Эксперименты с предложениями не дали результатов: для компьютера брак – это только женитьба или замужество, но не дефект.

Рассмотрим слово «труба». И в словаре С. И. Ожегова [10, с. 722], и в словаре Ю. Яесса [9] «труба» – многозначное слово, которое имеет 4 и 5 значений соответственно.

... Вечно раскаленная труба из кухни. – ... An ever-hot pipe from the kitchen...

Труба была инструментом симфонического оркестра. – The trumpet was a symphony orchestra instrument.

Снова прозвучала труба... – The trumpet sounded again...

Его голос разнесся по палубе, как труба. – His voice resounded across the deck like a trumpet.

Потом в мастерской лопнула труба. – Then a pipe burst in the workshop.

Альтовая труба висела на стене. – The alto trumpet hung on the wall.

Ключевым словом при переводе слова «труба» является английское pipe:

А вот и труба, над которой вьется дымок. – And here is the pipe, over which smoke curls.

Трансформированное предложение переводится как:

А вот и печная труба, над которой вьется дымок. – And here is the chimney, over which smoke curls.

Переводчик не справляется со сленговым оборотом:

У него трубы горят. – His pipes are on fire.

Рассмотрим слово «мат». В словаре С. И. Ожегова [10, с. 303] приводится четыре разных слова (омонима) с написанием «мат» – поражение в шахматной игре, плетёный половик, неприличная брань, шероховатость. В данном случае Ю. Яесс [9] даёт одну вокабулу с восемью как связанными, так и не связанными между собой значениями. Компьютер успешнее всего справился с переводом предложений, в которых «мат» – часть шахматной игры:

Объявил мат королю противника. – Declared checkmate to the opponent's king.

Такая программа уже смогла поставить мат. – Such a program has already been able to checkmate.

Он ставил мне мат – He checkmated me.

Самым неудачным является перевод синонимического ряда: бранные слова, мат, ругань, сплетни – swear words, swearing, swearing, gossip. Псевдоошибкой может показаться перевод следующего предложения:

Мат должен быть музыкой – Mat must be music.

Данное предложение может быть понято неоднозначно любым носителем языка.

Рассмотрим слово «отделение». В словаре С. И. Ожегова [10, с. 414] слово «отделение» – многозначное слово (6 значений). Ю. Яесс [9] даёт одну вокабулу с семью как связанными, так и не связанными между собой значениями.

Поступил на литературное отделение. – Entered the literary department.

Наше отделение отстрелялось вполне прилично. – Our squad fought back pretty well.

Экстренно перевели в реанимационное отделение. – Was urgently transferred to the intensive care unit.

Пока кончилось первое отделение программы... – Before the end of the first part of the program.

Отделение от рода есть и отделение от мира. – Separation from the clan is also separation from the world.

Второе отделение было отдано ему. – The second section was given to him

Перевели в отделение выздоравливающих. – Was even transferred to the convalescent unit.

Как видно из примеров, русскому слову отделение соответствует 7 английских точных эквивалентов.

Рассмотрим слово «перевод». В словаре С. И. Ожегова [10, с. 442] слово «перевод» встречается два раза: осуществление работы над текстом, доставка и перенос в пространстве (2 омонима). Ю. Яесс [9] даёт одну вокабулу с одиннадцатью как связанными, так и не связанными между собой значениями.

- последовал первоначальный перевод на шестнадцать тысяч – followed by the initial transfer of sixteen thousand;
- полный перевод не сохранился – complete translation has not been preserved;
- просит сделать перевод каких-то статей – asks for a translation of some articles;
- перевод в новую должность – to be transferred to a new position;
- уточняющие перевод текста – to clarify the translation of the text;
- получить денежный перевод – to receive a money order.

Ошибка допущена только в одном случае:

Перевод не означал повышения – This translation in no way meant an increase.

При повторном предъявлении переводчик даёт правильный перевод:

Transfer did not mean promotion.

Экспериментальная фраза «перевод на должность» переводится как *transfer to the post*.

Рассмотрим слово «бюкс». Согласно С. И. Ожегову [10, с. 51], в русском языке три слова с написанием «бюкс» (3 омонима): причёска, спорт, изолятор. Ю. Яесс [9] в одной вокабуле приводит 9 как связанных, так и не связанных между собой значений.

- считали бюкс хорошим средством против хронических головных болей – considered boxing a good remedy for chronic headaches;
- заглянул ко мне в бюкс – she looked into my box;
- выбор пал на бюкс – the choice fell on boxing;
- бюкс для ремонта машин – a car repair box;
- в разные спортивные секции: кто в бюкс, кто в шахматы – in different sports sections: some in boxing, some in chess;
- не досмотрев бюкс – without watching boxing.

Экспериментальная фраза «спортивная причёска “бюкс”» переводится как *sports hairstyle boxing*.

Рассмотрим слово «хвост». Хвост по словарю С.И. Ожегова [10, с.995] – одно многозначное слово, которое имеет 8 значений в прямом и переносном смысле. Ю. Яесс [9] приводит 31 значение данной лексемы в одной вокабуле.

Дымный хвост сигнальной ракеты. – The smoky tail of the signal flare.

У него был хвост трубы. – He had a tail pipe.

Приподнимали хвост и разворачивали всю машину. – Lifted the tail and turned the whole car around.

Хвост поезда. – The tail of the train.

Переводчик буквально переводит следующие предложения:

У меня образовалось два хвоста. – I had two tails.

Почувствовал за собой хвост. – He felt his tail.

Рассмотрим слово «шишка». В соответствие со словарём С. И. Ожегова [10, с.797], в русском языке два созвучных слова «шишка» – одно означает плод, выпуклость, а второе – большой начальник. Ю. Яесс [9] приводит 23 значения данной лексемы в одной вокабуле.

На голове вскочила шишка. – A bump popped up on his head.

Образовалась большая шишка. – A large bump formed there.

Шишка достигла размеров яйца. – The lump had reached the size of an egg.

Еловая шишка лежала в том же месте. – A fir cone lay in the same place.

Собирали шишки для самовара. – Collected cones for a samovar.

Большая шишка... упала с дерева. – A big cone fell from a tree.

Только в одном случае переносное значение слова «шишка» переведено правильно:

А тут сидит большая шишка. – And here a big shot sits.

Все шишки обрушатся на него – All the bumps would fall on him (буквальный перевод).

Рассмотрим слово «пробка». По мнению С. И. Ожегова [10, с. 536], пробка – многозначное слово, которое имеет 4 значения: наружная часть коры, закупорка для бутылки, переносное значение «затор» и «предохранитель». Ю. Яесс говорит об 11 значениях [9].

Была пробка на дороге. – There was a traffic jam.

Образовалась пробка. – A traffic jam formed.

Пробка вылетела из бутылки. – The cork flew.

Пробка изображала девушку. – The cork depicted a girl.

Компьютер затруднился перевести только фразеологизм: Вылетел как пробка из бутылки – Flew out of there like a traffic jam.

Рассмотрим слово «налёт». С.И. Ожегов [10, с. 336] считает, что слово «налёт» имеет два значения: нападение и слой на поверхности. Ю. Яесс [9] говорит о 8 значениях.

Пока не закончился воздушный налёт. – Until the air raid is over.

Налёт был настолько мощный... – The raid was so powerful...

Сделав огневой налет... – Having made a fire raid...

Общий налёт. – Total flight time.

На зубах появился налёт. – Plaque appeared on the teeth.

Все значения переданы верно.

3 Мультиагентная нейрокогнитивная архитектура для систем автоматического компьютерного перевода

С целью создания интеллектуальной системы для решения задачи автоматического перевода предлагается использовать мультиагентную систему (МАС).

МАС представляет собой набор интеллектуальных агентов, обладающих распределенными знаниями. Суть представляемого подхода состоит в том, что ни один агент не может в полной мере функционировать и/или решить некую задачу в одиночку, т. к. не обладает необходимым набором знаний. Таким образом решение задач основывается на взаимодействии агентов внутри системы или за ее пределами посредством пересылки сообщений с запросом искомой информации или ее предоставлением. Как и в реальной жизни мы не можем предугадать все возможные сценарии развития событий, в связи с этим заранее запрограммированная модель поведения агентов не всегда справляется с решением сложных задач. Вместо этого агенты должны найти решение самостоятельно, используя обучение. Такой подход схож с аналогичными процессами в любом человеческом сообществе, знания и опыт людей увеличиваются посредством общения. Подобно социальным группам, агенты в мультиагентных системах взаимодействуют друг с другом с целью обмена знаниями и обучения навыкам [12].

Далее следует пояснить понятие «интеллектуальный агент». В информационных технологиях под интеллектуальным агентом понимается программа, способная не только строго выполнять задание, поставленное пользователем, но и самостоятельно принимать решение о

целесообразности своих действий и их последствиях. Интеллектуальность агентов определяется их способностью собирать информацию через систему сенсоров и осуществлять влияние на внешний мир через систему актуаторов.

Еще одним понятием, используемым в данном подходе, является нейрокогнитивная архитектура. Как известно мышление возможно за счет электрического тока, передающегося от нейрона к нейрону у нас в мозге. Для моделирования процессов мышления человека агенты в рассматриваемом подходе называются и представляются в виде нейронов головного мозга. За счет наличия большого количества нейронов-агентов в системе достигается ее мультиагентность. С целью визуализации процессов, происходящих в системе, применяются инструменты имитационного моделирования, позволяющего отследить процесс создания и угасания нейрона, а также процессы создания и распада нейронных ансамблей (группа нейронов, связанная между собой). Все агенты-нейроны в системы упорядочены особым образом, в зависимости от своих типов и функции. Такое функционально детерминированное расположение агентов являет собой когнитивную архитектуру. Когнитивная архитектура – это среда функционирования интеллектуальных агентов, с ее помощью мы пытаемся моделировать не только поведение, но и структурные свойства моделируемой системы, а именно мозга человека. На рис. 1 представлена 3D модель мультиагентной нейрокогнитивной архитектуры.

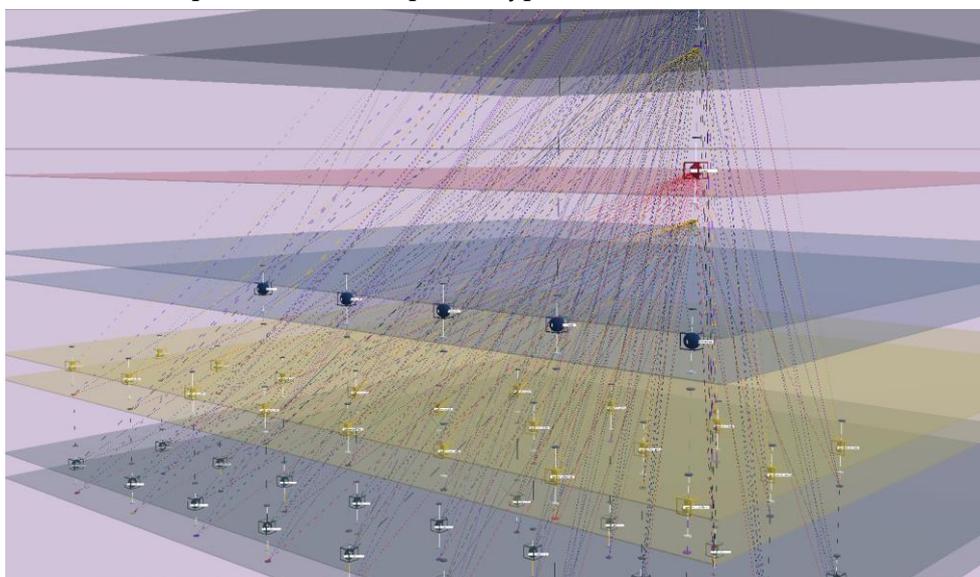


Рис. 1. 3D модель мультиагентной нейрокогнитивной архитектуры

В то время как мультиагентные системы могут использоваться в различных областях, включая робототехнику, распределенное управление, телекоммуникации, экономику и т. д., объектом нашего особого интереса является имитационное моделирование семантики естественного языка. Как показано в работе [13] центральным сдерживающим фактором в моделировании «семантики естественного языка является ... проблема формализации смысла естественно-языковых высказываний» [13 с. 48], что является ключевой проблемой в системах распознавания, понимания, синтеза речи, а также в задачах автоматического компьютерного перевода.

Одна из гипотез, выдвинутых в [13], которой мы придерживаемся, состоит в том, что интеллектуальный агент должен быть погружен в социальную среду, с которой ему придется взаимодействовать посредством естественного языка. Такое погружение сродни проблеме освоения естественного языка (Language Acquisition Problem) детьми, которые изначально не знают никакого языка, но будучи погруженными в языковую среду своей семьи, овладевают естественным языком за пару лет. Данный подход позволяет смоделировать «естественный» способ построения глубинной структуры естественно-языкового высказывания.

В парадигме данного подхода понимание высказываний – сложный процесс, протекающий в мозге человека, который мы планируем смоделировать с помощью мультиагентной нейрокогнитивной архитектуры.

В разрабатываемой нами системе представлены агенты разных лингвистических типов: морфологический, синтаксический, семантический, лексический. Для репрезентации значения в системе необходимы два вида агентов: агенты-слова и соответствующие им агенты-понятия.

Агенты-слова являются хранилищем фонетической, парадигматической и синтагматической информации. Агенты-понятия хранят в своих базах знаний описание объекта, обозначаемого этим словом. Между двумя агентами, хранящими разную информацию об одной и той же единице языка, устанавливаются связи. Активация одного из них влечет за собой активацию второго. Подробнее о механизме мультиагентной репрезентации элементов естественного языка говорится в работе [14], механизм обоснования символов и мультиагентные нейрокогнитивные модели семантики естественного языка подробно описаны в [13, 15, 16].

В процессе понимания и распознавания речи задействована функциональная репрезентация данного высказывания в мультиагентной системе, т. е. слова не запоминаются отдельно друг от друга. Иными словами, появление в системе нового существительного непременно инициирует появление связанного с ним глагола. Далее следуя логике грамматики зависимостей, новое существительное будет пытаться объединиться в нейронную связь с теми прилагательными, с которыми оно может сочетаться. Глагол, в свою очередь, будет искать те наречия, которые могут быть использованы для его идентификации. «На уровне моделей неосознаваемый процесс имеет место интерпретация высказываний в терминах элементов графа проблемной ситуации, а на уровне имитационных моделей – идентификация этих процессов на предмет полноты и целостности» [13 с. 48].

Рассмотрим программную реализацию данного процесса. В систему через клавиатуру вводится слово «ключ», т. е. ввод был осуществлен с клавиатуры, создается лексический агент типа существительное. Данный агент типа существительное инициирует создание соответствующего семантического/концептуального агента типа объект, между ними устанавливаются контрактные отношения, которые гарантируют узнавание себя в будущем. Как упоминалось ранее, мы считаем, что объекты не запоминаются отдельно от контекста, т. е. объект или новое слово не отражаются в памяти без соответствующего действия/глагола и т. д.

Таким образом, автоматический перевод должен начинаться с обучения системы. Рассмотрим процесс обучения, погружения в языковую среду системы через один из сенсоров, клавиатуру. Введем слово «яблоко». Созданный агент существительное начнет поиск своего глагола. Предположим, что в системе нет ни одного агента типа глагол, тогда в соответствии со своей базой знаний агент существительное задаст вопрос пользователю: «Земля что делает»? В ответ на этот вопрос пользователь может ввести любой глагол, который сочетается с этим существительным. Далее существительное будет искать прилагательные, с которыми может связаться, и при необходимости может задать вопросы пользователю; такой же процесс будет выполнен и глаголом. В результате мы можем получить некий ансамбль нейронов, состоящий, например, из существительного «ключ», связанного с глаголами «поворачивать» (на данный момент используется только нефинитная форма глагола), с прилагательным «золотой» или «старый». В следующий раз, когда мы введем слово «ключ», система найдет нейронную связь «золотой-ключ-поворачивать» и сможет задать нам вопросы: Ключ поворачивать? Ключ золотой? В этот раз мы хотим ввести новое значение ключа как источника воды и отвечаем: «бить из-под земли», на второй вопрос даем ответ: «ледяной».

Мы предполагаем, что в перспективе система сможет обучаться не только с помощью пользователя, а использовать интернет как источник примеров различных естественно-языковых высказываний, будет строить их функциональные репрезентации и определять сочетаемостные особенности слов, что, в свою очередь, позволит строить мультиагентную модель различных коллокаций полисемичных и амонимичных слов.

Заключение

Таким образом, проведенное исследование позволило сделать определённые выводы. В теоретическом (лингвистическом) плане существуют некоторые разногласия, касающиеся трактовки таких феноменов, как полисемантность и омонимия. Многие исследователи отмечают размытость между данными явлениями – начало одного и конец другого трудно идентифицируемы. Несмотря на то, что разработаны надёжные операционные методики (морфологические, лексические, синтаксические и лексические), толкование некоторых лексем наталкивается на трудности в плане синхронии и диахронии. Нерешённость данной лингвистической проблемы вызывает дополнительные сложности коллективам, которые занимаются вопросами машинного перевода, особенно тем, кто пытается внедрить и обучать

нейросети и интеллектуальных агентов. Очевидна трудоёмкая ручная работа, которая стоит за программами и комментариями. Анализ компьютерных переводов полисемичных слов и омонимов с русского языка на английский показал, что большие усилия разработчиков и исследователей дают свои положительные результаты. Анализ позволяет заключить следующее:

1. В некоторых случаях переводчик выполняет перевод, на который способен только человек.
2. В переводах максимально учитывается сочетаемость, валентность, семантические поля, информация из грамматики и морфологии.
3. Перевод получается успешным, если он опирается на биграммы.

Нет надёжных данных, говорящих о том, что леммы опираются на информацию справа от себя. Имеются проблемы с переводом переносных значений и устойчивых (фразеологических) единиц и рядом расположенных групп синонимов. Ближайшая цель исследователей – применить мультиагентную нейрокогнитивную архитектуру для проведения семантического анализа элементов естественного языка для снятия омонимии и полисемии в задаче автоматического компьютерного перевода.

Литература

1. Карцевский С.О. Об асимметричном дуализме лингвистического знака. М.: Аспект Пресс, 2001. – С. 76-81.
2. Бирюк И. Б. Омонимия и полисемия как формы языковой асимметрии. URL: https://ffl.msu.by/info/faculty/foreign/nauch_rab/konf/2018.12.0317/doklady/40_biruk_ib.pdf (дата обращения: 11.01.2023).
3. Кузьмина Р.В. К вопросу о разграничении явлений омонимии и полисемии и об особенностях их представления в орфоэпическом словаре. URL: <https://main.isuct.ru/files/publ/vgf/2006/01/129.html> (дата обращения: 6.07.2022).
4. Головня АИ. Проблема разграничения полисемии и омонимии в русском языке. URL: <https://elib.bsu.by/bitstream/123456789/50202/> (дата обращения: 16.08.2022).
5. Скроблевич В. Б Разрешение полисемии и омонимии при машинном переводе. URL: <http://elib.mitso.by/bitstream/edoc/3054/1/Теория167.pdf> (дата обращения: 19.09.2022).
6. Авдина А. И. Вопросы классификации омонимов (на материале словарей омонимов) // Вестник ЮУрГУ. Серия: Лингвистика. 2022. №2. URL: <https://cyberleninka.ru/article/n/voprosy-klassifikatsii-omonimov-na-materiale-slova-rey-omonimov> (дата обращения: 03.05.2023).
7. Словарь синонимов русского языка URL: <https://sinonim.org/>: [<https://sinonim.org/>] (дата обращения: 19.09.2022).
8. Google Переводчик URL: www.translate.google.ru (дата обращения: 19.09.2022).
9. Яесс Ю. Словарь омонимов и многозначных слов русского язык URL: <https://proza.ru/2020/04/17/951> (дата обращения: 19.09.2022).
10. Ожегов С.И. Словарь русского языка. М: Рус.яз.,1982. - 816 с.
11. Hornby A.S. Oxford Advanced Learner's Dictionary of Current English. М: Издательство «Русский язык», 1982. Т-1.544 с.
12. Bennane A. Tutoring and Multi-Agent Systems: Modeling from Experiences // Informatics Educ. 2010. Vol. 9. Pp. 171-184.
13. Нагоев З. В., Нагоева О. В. Обоснование символов и мультиагентные нейрокогнитивные модели семантики естественного языка Обоснование символов и мультиагентные нейрокогнитивные модели семантики естественного языка. Нальчик: Издательство КБНЦ РАН, Нальчик: Издательство КБНЦ РАН, 2022. – 150 с.
14. Makoeva D., Nagoeva O., Gurtueva I. Formal Representation of natural language elements in multi-agent system based of self-organization of distributed neurocognitive architectures // Procedia Computer Science, 2022, volume 213, p. 631-635 DOI: 10.1016/j.procs.2022.11.114
15. Анчечков М.И., Бжихатлов К.Ч., Канкулов С.А., Нагоев З.В., Нагоева О.В. Мультиагентный алгоритм обоснования символов конвенционального языка на основе ситуативно обусловленного развития нейрокогнитивной архитектуры // Известия Кабардино-Балкарского научного центра РАН. – 2022. – № 6(110). – С. 48-60. – DOI 10.35330/1991-6639-2022-6-110-48-60. – EDN SBIEVA.

16. Анчеков М.И., Бжихатлов К.Ч., Нагоев З.В., Пшенокова И.А. Онтоэписоциофилогенетическое развитие систем общего искусственного интеллекта на основе мультиагентных нейрокогнитивных архитектур // Известия Кабардино-Балкарского научного центра РАН. – 2022. – № 6(110). – С. 61-75. – DOI 10.35330/1991-6639-2022-6-110-61-75. – EDN PRPNCF.

ON SOME PECULIARITIES OF COMPUTER TRANSLATION OF HOMONYMS AND POLYSEMANTIC WORDS

Boziev, Alberd Takhirovich

*PhD in philology, associate professor
Kabardino-Balkarian State University
Kabardino-Balkarian Scientific Center of the RAS
Nalchik, Russian Federation
alberdboziev@mail.ru*

Makoeva, Dana Gisovna

*PhD in philology, associate professor
Kabardino-Balkarian Scientific Center of the RAS
Nalchik, Russian Federation
makoevadana@mail.ru*

Abstract

The article deals with issues related to the complexity of the linguistic interpretation of polysemic words and homonyms. Contradictory interpretations of these phenomena to a certain extent complicates the work of developers of programs and comments for computer translation. Currently, researchers are making great efforts to achieve a quality of translation that would not differ from human translation.

Keywords

homonyms; lexical homonyms; polysemic words; computer translation; homonymy elimination; machine translation; neural networks

References

1. Karcevskij S.O. Ob asimmetrichnom dualizme lingvisticheskogo znaka. M.: Aspekt Press, 2001. – S. 76-81.
2. Biryuk I. B. Omonimiya i polisemiya kak formy yazykovoj asimmetrii. URL: https://fll.msu.by/info/faculty/foreign/nauch_rab/konf/2018.12.0317/doklady/40_biruk_ib.pdf (accessed on: 11.01.2023).
3. Kuz'mina R.V. K voprosu o razgranichenii yavlenij omonimii i polisemii i ob osobennostyah ih predstavleniya v orfoepicheskom slovare. URL: <https://main.isuct.ru/files/publ/vgf/2006/01/129.html>] (accessed on: 6.07.2022).
4. Golovnya AI. Problema razgranicheniya polisemii i omonimii v russkom yazyke. URL: <https://elib.bsu.by/bitstream/123456789/50202/> (accessed on: 16.08.2022).
5. Skromblevich V. B Razreshenie polisemii i omonimii pri mashinnom perevode. URL: <http://elib.mitso.by/bitstream/edoc/3054/1/Teoriya167.pdf> (data obrashcheniya: 19.09.2022).
6. Avdina A. I. Voprosy klassifikatsii omonimov (na materiale slovarej omonimov) // Vestnik YUUrGU. Seriya: Lingvistika. 2022. №2. URL: <https://cyberleninka.ru/article/n/voprosy-klassifikatsii-omonimov-na-materiale-slovarej-omonimov> (data obrashcheniya: 03.05.2023).
7. Slovar' sinonimov russkogo yazyka URL: <https://sinonim.org/>: [<https://sinonim.org/>] (accessed on: 19.09.2022).
8. Google Perevodchik URL: www.translate.google.ru (data obrashcheniya: 19.09.2022).
9. Yaess Yy. Slovar' omonimov i mnogoznachnyh slov russkogo yazyk URL: <https://proza.ru/2020/04/17/951> (accessed on: 19.09.2022).
10. Ozhegov S.I. Slovar' russkogo yazyka. M: Rus.yaz.,1982. - 816 s.
11. Hornby A.S. Oxford Advanced Learner's Dictionary of Current English. M: Izdatel'stvo «Russkij yazyk», 1982. T-1.544 s.
12. Bennane A. Tutoring and Multi-Agent Systems: Modeling from Experiences // Informatics Educ. 2010. Vol. 9. Pp. 171-184.
13. Nagoyev Z. V., Nagoyeva O. V. Obosnovaniye simvolov i mul'tiagentnyye neyrokognitivnyye modeli semantiki yestestvennogo yazyka Obosnovaniye simvolov i mul'tiagentnyye

- neyrokognitivnyye modeli semantiki yestestvennogo yazyka. Nal'chik: Izdatel'stvo KBNTS RAN, Nal'chik: Izdatel'stvo KBNTS RAN, 2022. 150 s.
14. Makoeva D., Nagoeva O., Gurtueva I. Formal Representation of natural language elements in multi-agent system based of self-organization of distributed neurocognitive architectures // *Procedia Computer Science*, 2022, volume 213, p. 631-635 DOI: 10.1016/j.procs.2022.11.114
 15. Anchekov M.I., Bzhikhatlov K.CH., Kankulov S.A., Nagoyev Z.V., Nagoyeva O.V. Mul'tiagentnyy algoritm obosnovaniya simvolov konventsional'nogo yazyka na osnove situativno obuslovlennogo razvitiya neyrokognitivnoy arkhitektury // *Izvestiya Kabardino-Balkarskogo nauchnogo tsentra RAN*. 2022. № 6(110). S. 48-60. DOI 10.35330/1991-6639-2022-6-110-48-60. EDN SBIEVA.
 16. Anchekov M.I., Bzhikhatlov K.CH., Nagoyev Z.V., Pshenokova I.A. Ontoepisotsiofilogeneticheskoye razvitiye sistem obshchego iskusstvennogo intellekta na osnove mul'tiagentnykh neyrokognitivnykh arkhitektur // *Izvestiya Kabardino-Balkarskogo nauchnogo tsentra RAN*. – 2022. – № 6(110). – S. 61-75. – DOI 10.35330/1991-6639-2022-6-110-61-75. EDN PRPNCF.