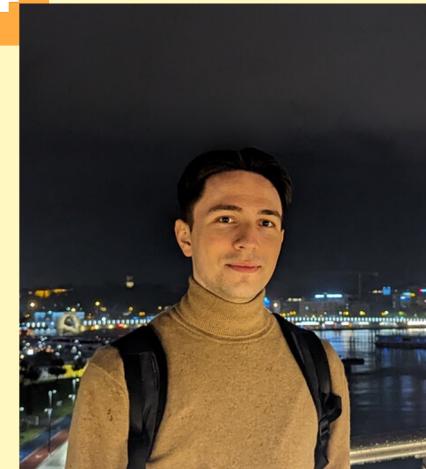


Машинный перевод

Николай Карпачёв

руководитель группы базового качества
перевода, Яндекс



Содержание

1. Задача перевода
2. Статистический word-based и фразовый перевод
3. Перевод с помощью нейросетей
4. Оценка качества
5. Машинный перевод в 2024 году

Задача перевода

- Получить функцию преобразования текстов на исходном языке в эквивалентные по смыслу тексты на целевом языке
- “Текстами” считаем любую последовательность символов из алфавита

$$f : X \rightarrow Y$$

или

$$f : x^{\|\infty\|} \rightarrow y^{\|\infty\|}$$

Задача перевода

Не обязательно перевод между языками

The screenshot shows a user interface for translating code between different programming languages. At the top, there's a "Target Program Language" dropdown set to "Bash" and a "Your Program Language" dropdown set to "Python".

Target Program Language (Bash):

```
#!/bin/bash
sum=0
for ((i = 1; i <= 10; i++)); do
    ((sum += i))
done
echo "The sum is: $sum."
```

Your Program Language (Python):

```
int sum = 0;
for (int i = 1; i <= 10; i++) {
    sum += i;
}
System.out.println("The sum is: " + sum);
```

Задача перевода

Не обязательно перевод между языками

АНГЛИЙСКИЙ ЭМОДЗИ

cooking book on sale

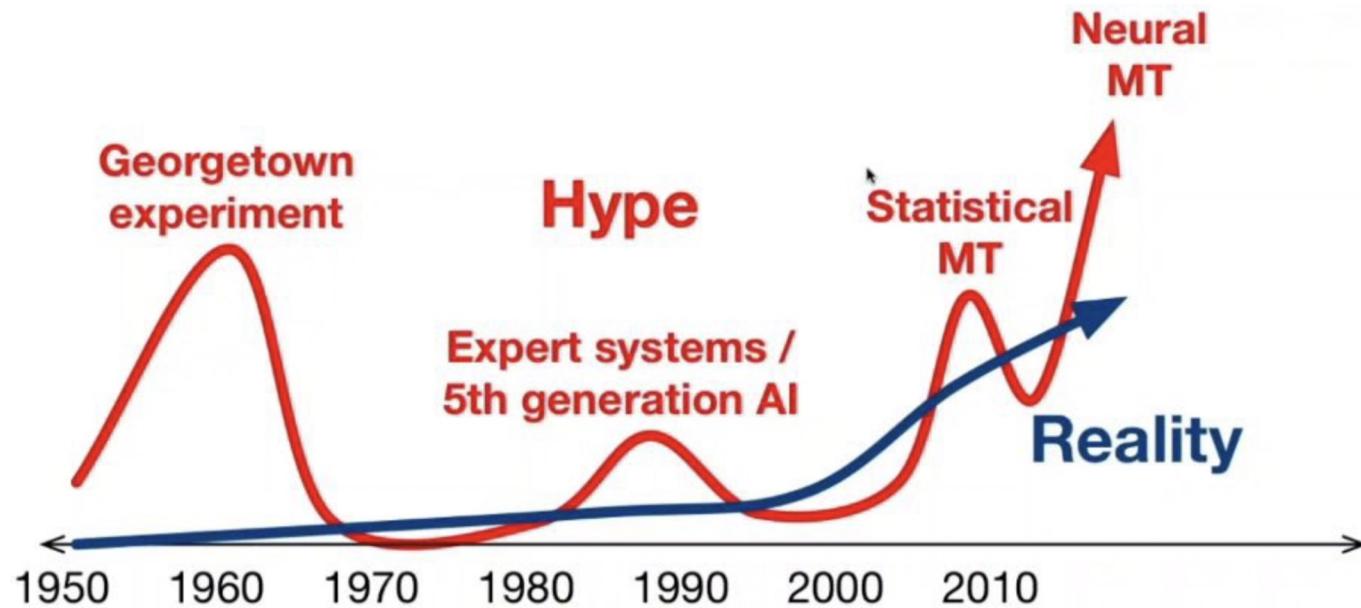
X

🔍📚↔️ \$

Перевести в Google

Bookmark Share Like Dislike

Задача перевода



Задача перевода

- Джорджтаунский эксперимент (1954)
 - Перевод русский -> английский
 - 6 грамматических правил
 - Mainframe IBM 701

Задача перевода

- Джорджтаунский эксперимент (1954)
 - Перевод русский -> английский
 - 6 грамматических правил
 - Mainframe IBM 701

Система работала только на малом числе возможных src
Нельзя называть автоматическим переводом

Задача перевода

- Машинный перевод на системах правил (rule-based MT)
 - Словари + алгоритмические отображения
 - Качество крайне плохое
 - Много эвристик и ad-hoc решений
 - Громоздкие системы

Задача перевода

- Статистический перевод
 - Порождающие вероятностные модели
 - Инференс правил на основе массива примеров переводов
 - Качество - приемлемое (на текстах средней сложности или ниже)

На основе SMT появились первые коммерческие системы перевода

Задача перевода

- Нейросетевой машинный перевод (NMT)
 - end-to-end обучение одной модели отображения текста в перевод
 - почти полное отсутствие induction bias в модели, все закономерности выучиваются из данных
 - нужны гигантские массивы примеров переводов
 - слабая интерпретируемость

На high-resource языковых парах качество - высокое

Статистические модели перевода

Пословный перевод

- Лексические словари
- Простейшее решение - переводить по словно согласно словарю

das	Haus	ist	klein
the	house	is	small



Пословный перевод

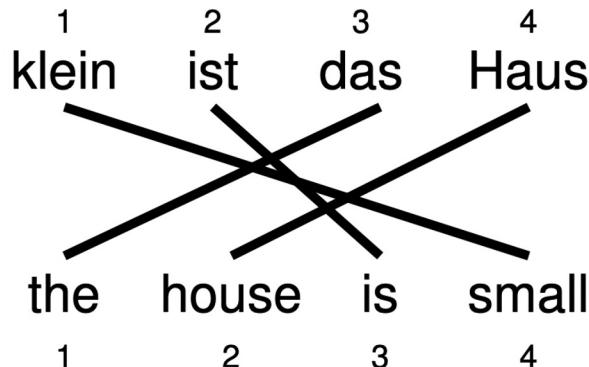
- Лексические словари
- Простейшее решение - переводить по словно согласно словарю

das	Haus	ist	klein
the	house	is	small

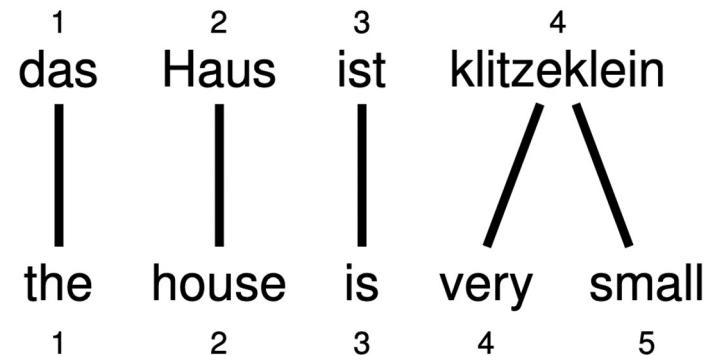
Очевидные проблемы: выравнивание, многозначность, контекст

Пословный перевод

- Изменение порядка слов



- One-to-many выравнивание



Word-based statistical translation

Строим порождающую вероятностную модель

$$p(\mathbf{e}, \mathbf{a} | \mathbf{f}) = \frac{\epsilon}{(l_f + 1)^{l_e}} \prod_{j=1}^{l_e} t(e_j | f_{a(j)})$$

Diagram illustrating the components of the word-based statistical translation model:

- Arrows point from the components of the formula to their corresponding labels:
 - From $p(\mathbf{e}, \mathbf{a} | \mathbf{f})$ to *перевод* (translation).
 - From $(l_f + 1)^{l_e}$ to *исходный текст* (source text).
 - From $t(e_j | f_{a(j)})$ to *пословная вероятностная модель* (word-level probabilistic model).
 - From ϵ to *модель выравнивания* (alignment model).

Word-based statistical translation

Строим порождающую вероятностную модель

das	
<i>e</i>	$t(e f)$
the	0.7
that	0.15
which	0.075
who	0.05
this	0.025

Haus	
<i>e</i>	$t(e f)$
house	0.8
building	0.16
home	0.02
household	0.015
shell	0.005

ist	
<i>e</i>	$t(e f)$
is	0.8
's	0.16
exists	0.02
has	0.015
are	0.005

klein	
<i>e</i>	$t(e f)$
small	0.4
little	0.4
short	0.1
minor	0.06
petty	0.04

$$\begin{aligned} p(e, a|f) &= \frac{\epsilon}{4^3} \times t(\text{the}| \text{das}) \times t(\text{house}| \text{Haus}) \times t(\text{is}| \text{ist}) \times t(\text{small}| \text{klein}) \\ &= \frac{\epsilon}{4^3} \times 0.7 \times 0.8 \times 0.8 \times 0.4 \\ &= 0.0028\epsilon \end{aligned}$$

Word-based SMT: обучение

Нужно: оценить оптимальные параметры порождающей вероятностной модели, имея параллельный корпус текстов

$$p(\mathbf{e}, a | \mathbf{f}) = \frac{\epsilon}{(l_f + 1)^{l_e}} \prod_{j=1}^{l_e} t(e_j | f_{a(j)})$$

В чем сложность:

- Для оценки параметров t нужны веса модели выравнивания a
- Для оценки параметров a нужны веса модели перевода

Word-based SMT: обучение

Нужно: оценить оптимальные параметры порождающей вероятностной модели, имея параллельный корпус текстов

$$p(\mathbf{e}, a | \mathbf{f}) = \frac{\epsilon}{(l_f + 1)^{l_e}} \prod_{j=1}^{l_e} t(e_j | f_{a(j)})$$

В чем сложность:

- Для оценки параметров t нужны веса модели выравнивания a
- Для оценки параметров a нужны веса модели перевода

EM-алгоритм для оценки параметров t и a

Word-based SMT: Noisy channel

- Пословный перевод + выравнивание никак не учитывают гладкость итогового текста
- Можем явно построить вероятностную модель с учетом частоты перевода с т.зр. целевого языка

Word-based SMT: Noisy channel

- Пословный перевод + выравнивание никак не учитывают гладкость итогового текста
- Можем явно построить вероятностную модель с учетом частоты перевода с т.зр. целевого языка

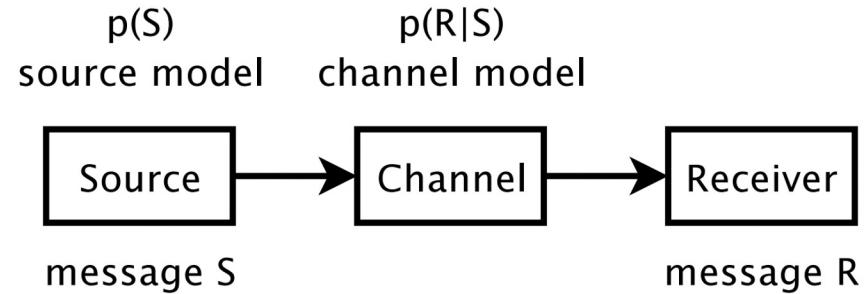
n-граммная языковая модель (n-gram LM)

$$\begin{aligned} p(\mathbf{e}) &= p(e_1, e_2, \dots, e_n) \\ &= p(e_1)p(e_2|e_1)\dots p(e_n|e_1, e_2, \dots, e_{n-1}) \\ &\simeq p(e_1)p(e_2|e_1)\dots p(e_n|e_{n-2}, e_{n-1}) \end{aligned}$$

Word-based SMT: Noisy channel

По формуле Байеса

$$\begin{aligned}\operatorname{argmax}_e p(e|f) &= \operatorname{argmax}_e \frac{p(f|e) p(e)}{p(f)} \\ &= \operatorname{argmax}_e p(f|e) p(e)\end{aligned}$$



Модель “зашумленного канала”

PBMT (Phrase-Based Machine Translation)

Похоже на word-based SMT, но

- вместо словаря - фразовая таблица
- более сложная модель выравнивания
- декодинг

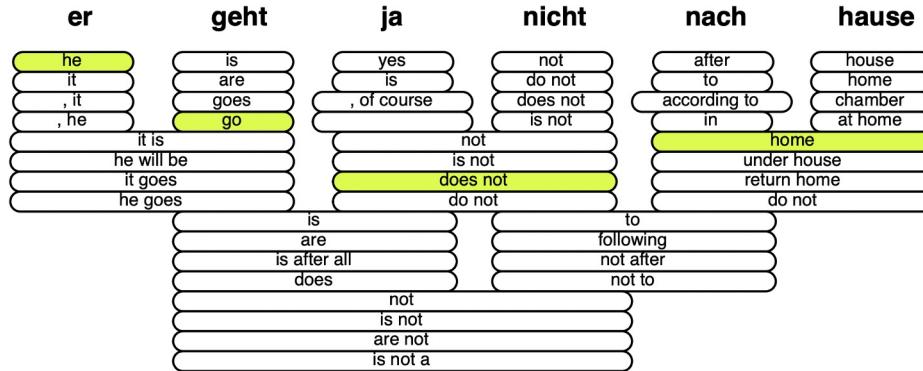
PBMT (Phrase-Based Machine Translation)

N-Gram phrase table

er	geht	ja	nicht	nach	hause
he	is	yes	not	after	house
it	are	IS	do not	to	home
, it	goes	, of course	does not	according to	chamber
, he	go	,	is not	in	at home
it is				home	
he will be			not		under house
It goes			is not		return home
he goes			does not		
			do not		do not
				to	
					following
					not after
					not to
			not		
			is not		
			are not		
			is not a		

PBMT (Phrase-Based Machine Translation)

Decoding



- Задача поиска оптимального разбиения - NP-полная
- Приближенный декодинг
- Эвристический beam search

РВМТ - проблемы

1. Неучет контекста
2. “Склейка” перевода из кусочков фраз
3. Громоздкий пайплайн обучения, много компонент

На практике:

- буквальные переводы
- большие проблемы со сложными зависимостями, идиомами

Нейросетевой перевод

Sequence to sequence models

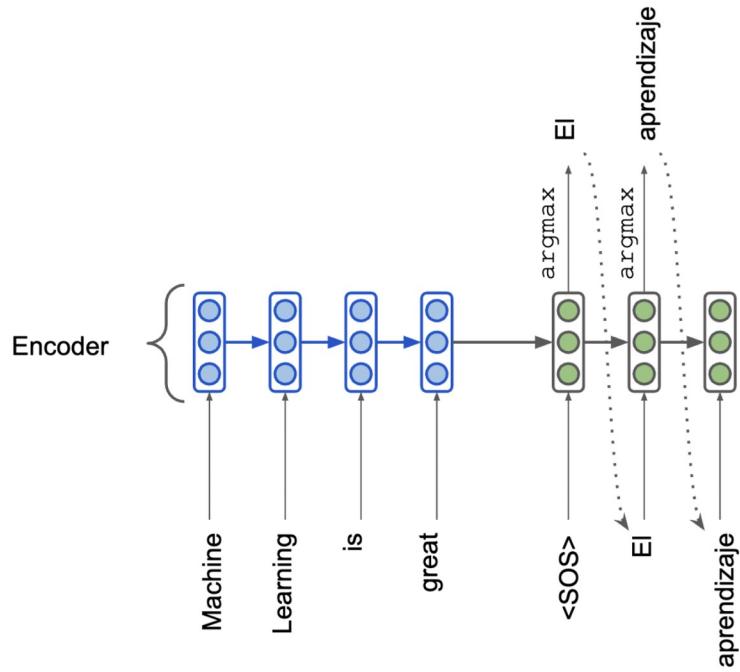
- Текст делится на токены
- Авторегрессионное предсказание следующего токена

$$P(y|x) = P(y_2|y_1, x)P(y_3|y_1, y_2, x)\dots \underbrace{P(y_T|y_1, y_2, \dots, x)}$$

- Модель - это одна дифференцируемая функция
- Обучение end-to-end

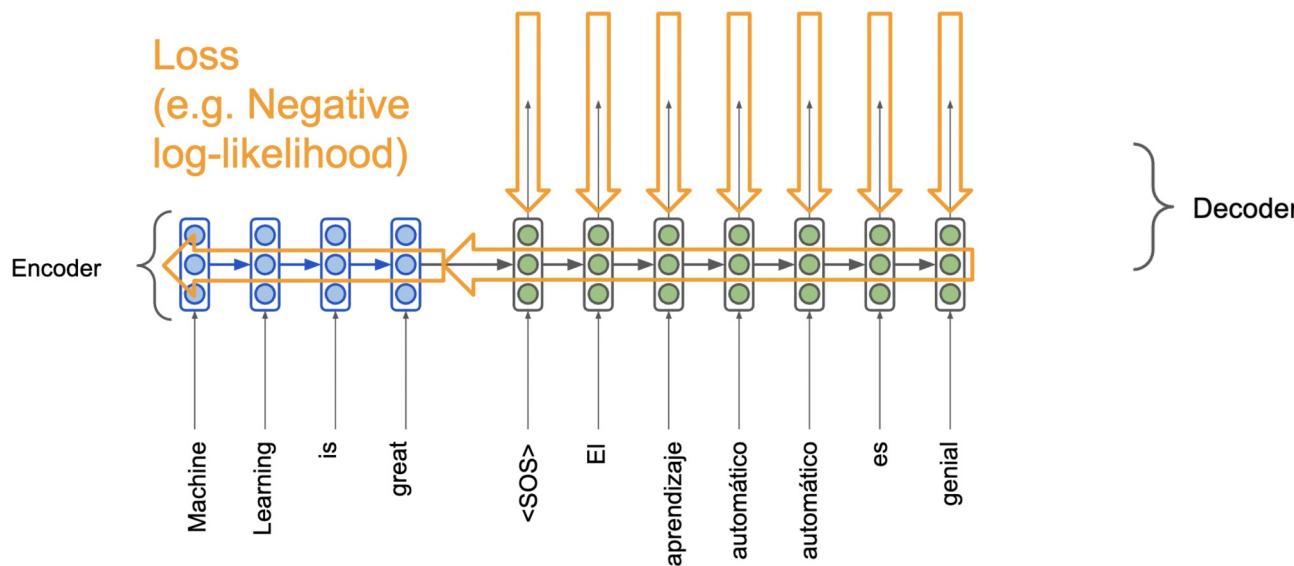
Encoder-decoder

Sutskever et.al. (2014)



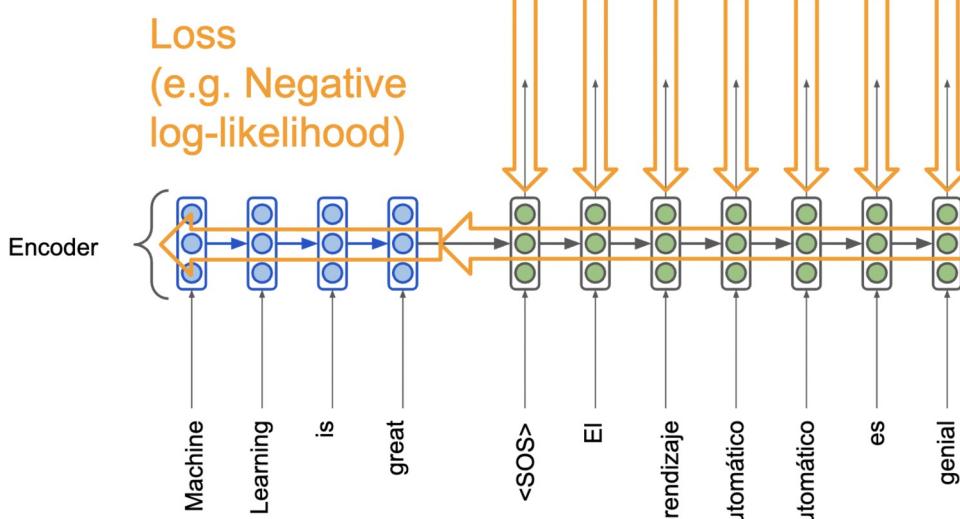
Encoder-decoder

Sutskever et.al. (2014)



Encoder-decoder

Sutskever et.al. (2014)



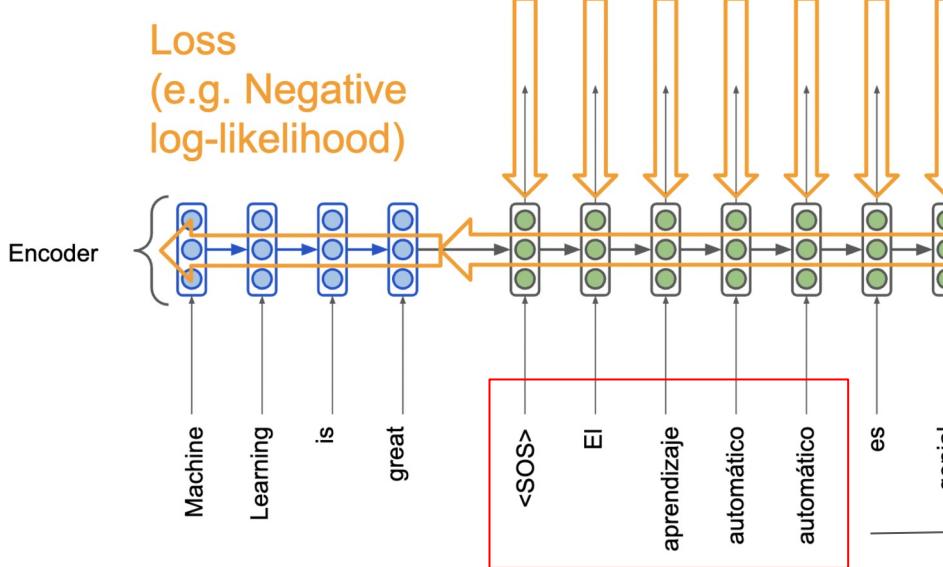
$$f(s)_i = \frac{e^{s_i}}{\sum_j^C e^{s_j}}$$

$$CE = - \sum_i^C t_i \log(f(s)_i)$$

Функция потерь - категориальная кроссэнтропия
потокенная

Encoder-decoder

Sutskever et.al. (2014)



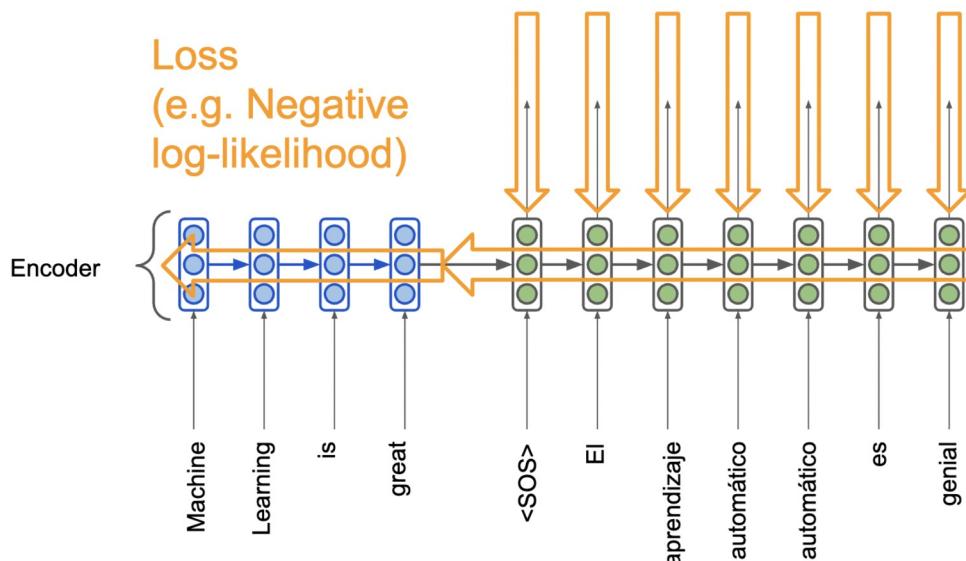
$$f(s)_i = \frac{e^{s_i}}{\sum_j^C e^{s_j}}$$
$$CE = - \sum_i^C t_i \log(f(s)_i)$$

Функция потерь - категориальная кроссэнтропия
потокенная

“teacher forcing” - при обучении
используем префикс эталонного
перевода

Encoder-decoder

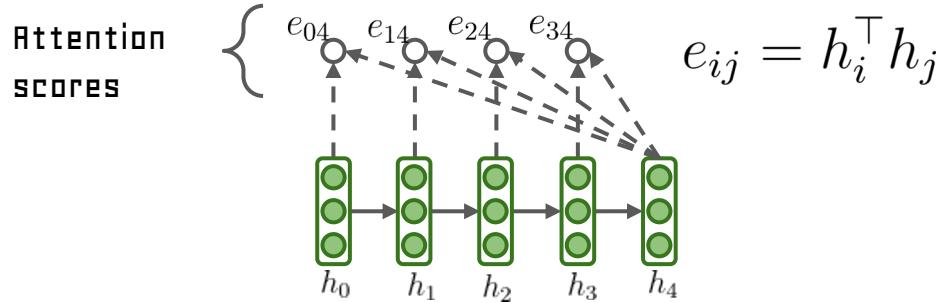
Sutskever et.al. (2014)



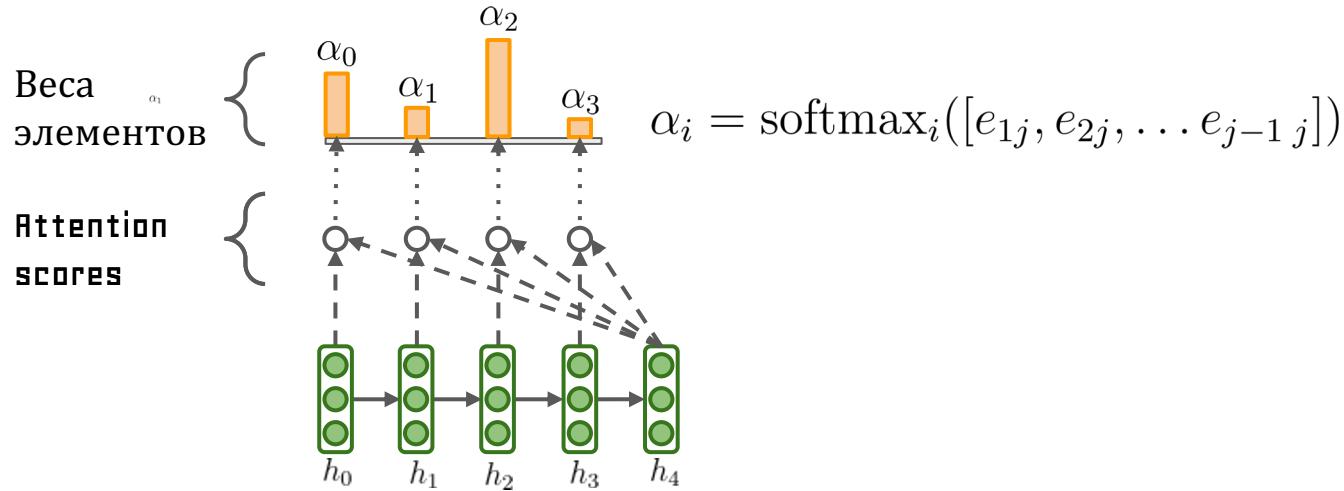
Проблемы seq2seq

1. Сжатие всей информации в один вектор *hidden_state*
2. Качество сильно деградирует на длинных src
3. Затухающие градиенты

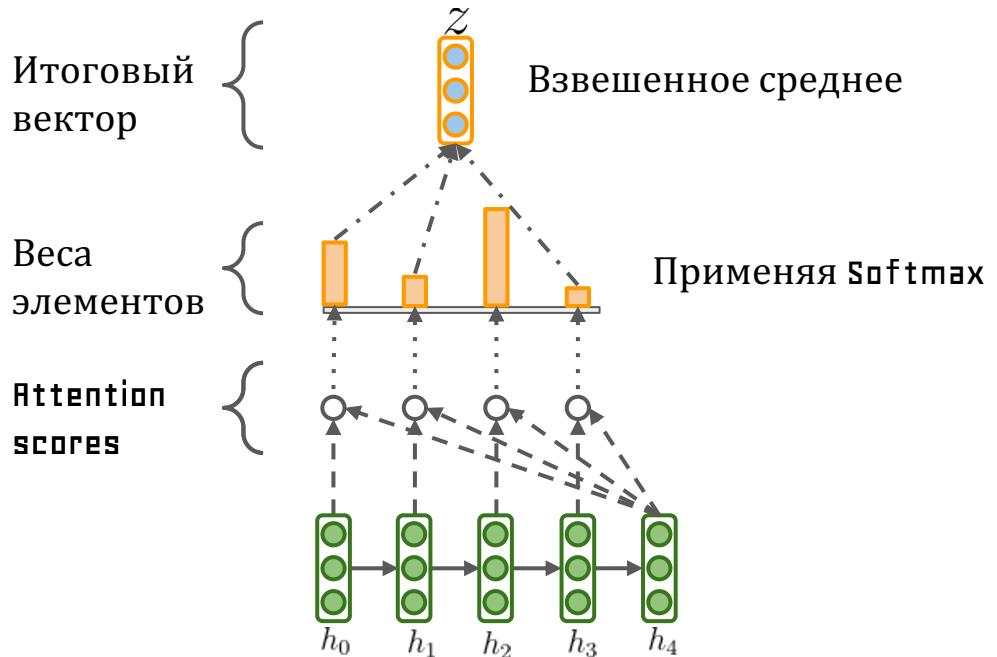
Attention [Bahdanau et.al. 2014]



Attention [Bahdanau et.al. 2014]



Attention [Bahdanau et.al. 2014]

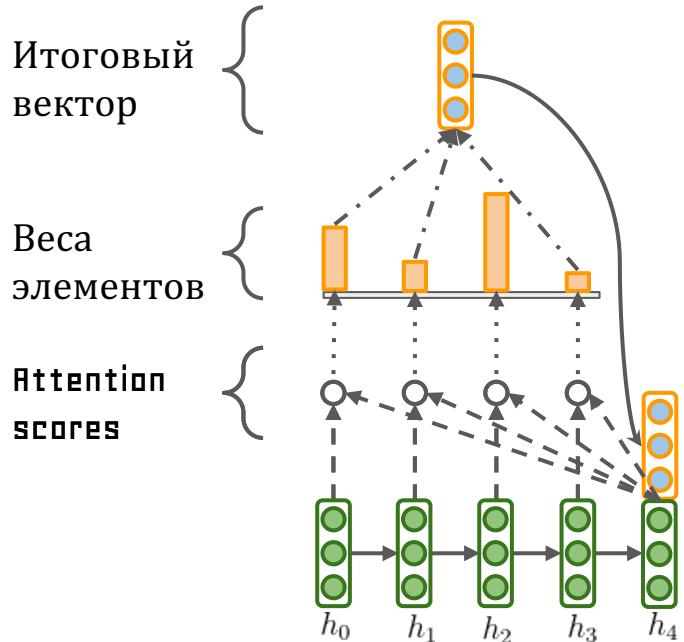


$$z = \sum_{i=1}^T \alpha_i h_i$$

$$\sum_{i=1}^T \alpha_i = 1$$

$$\forall i \quad \alpha_i \in [0; 1]$$

Attention [Bahdanau et.al. 2014]

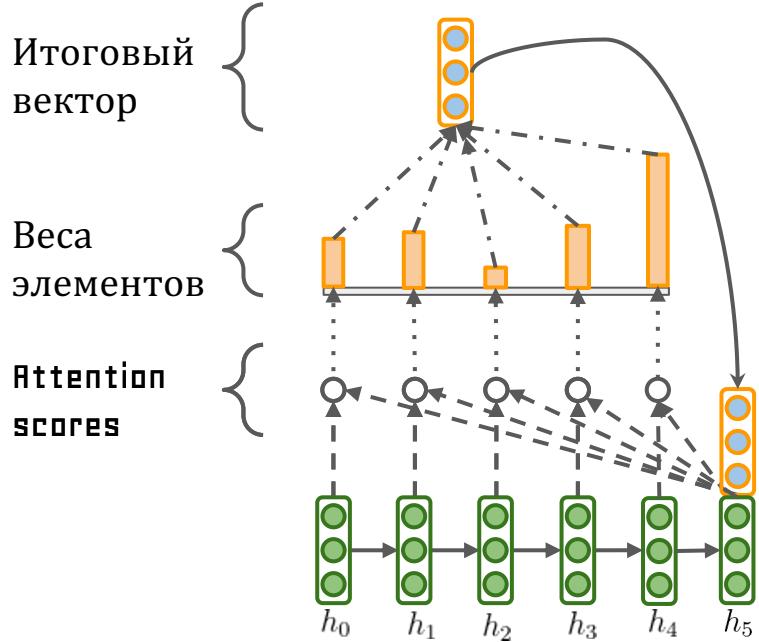


$$z = \sum_{i=1}^T \alpha_i h_i$$

$$\sum_{i=1}^T \alpha_i = 1$$

$$\forall i \quad \alpha_i \in [0; 1]$$

Attention [Bahdanau et.al. 2014]

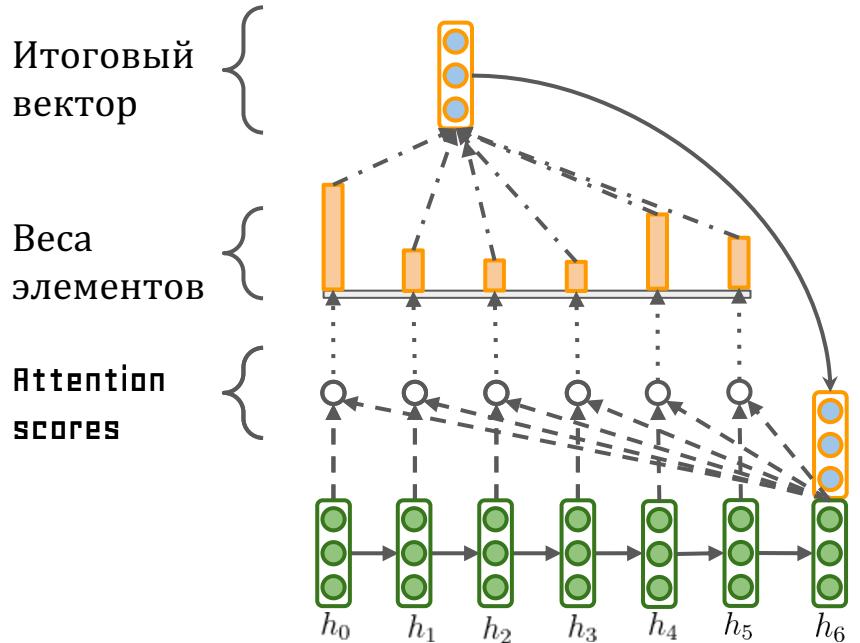


$$z = \sum_{i=1}^T \alpha_i h_i$$

$$\sum_{i=1}^T \alpha_i = 1$$

$$\forall i \quad \alpha_i \in [0; 1]$$

Attention [Bahdanau et.al. 2014]



$$z = \sum_{i=1}^T \alpha_i h_i$$

$$\sum_{i=1}^T \alpha_i = 1$$

$$\forall i \quad \alpha_i \in [0; 1]$$

Transformer [Vasvani et.al. 2017]

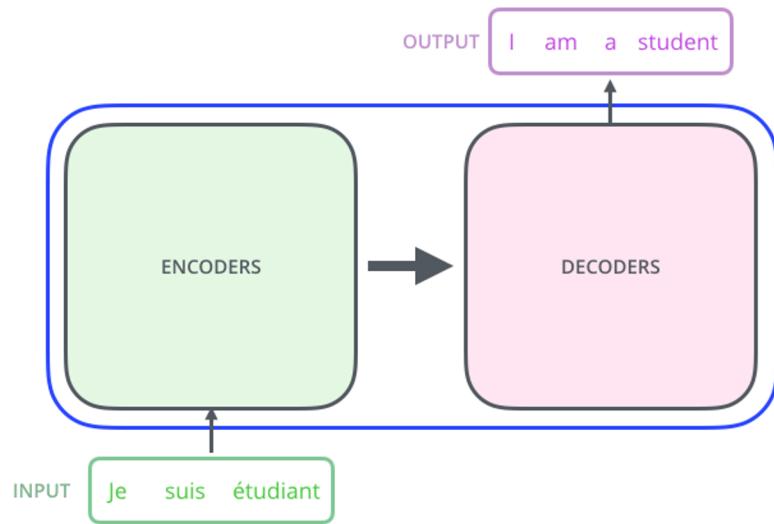
Ключевая идея:

- для учета контекстных зависимостей не нужна рекуррентность
- достаточно модели попарных связей между токенами через механизм внимания



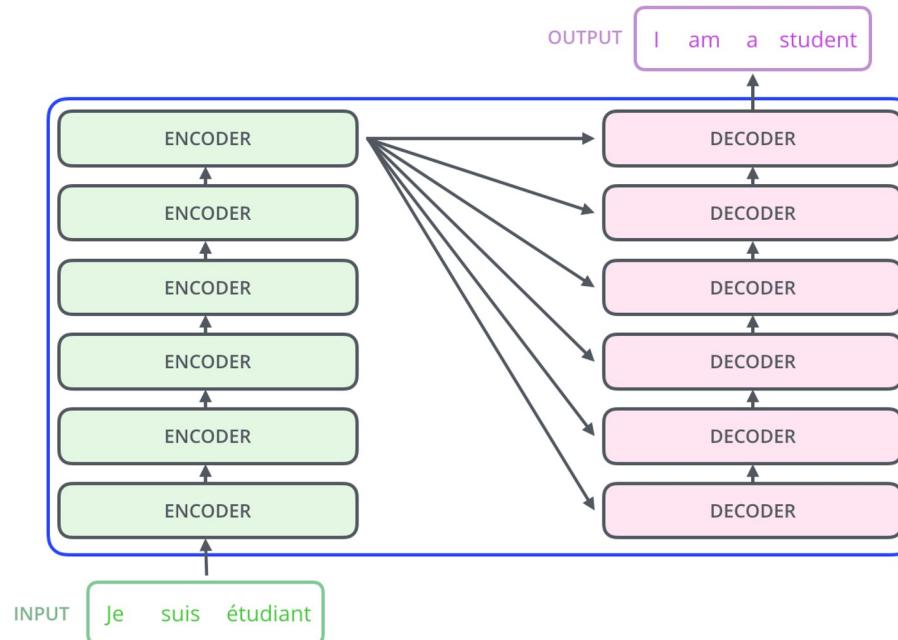
Источник изображения: <http://jalammar.github.io/illustrated-transformer/>

Transformer [Vasvani et.al. 2017]



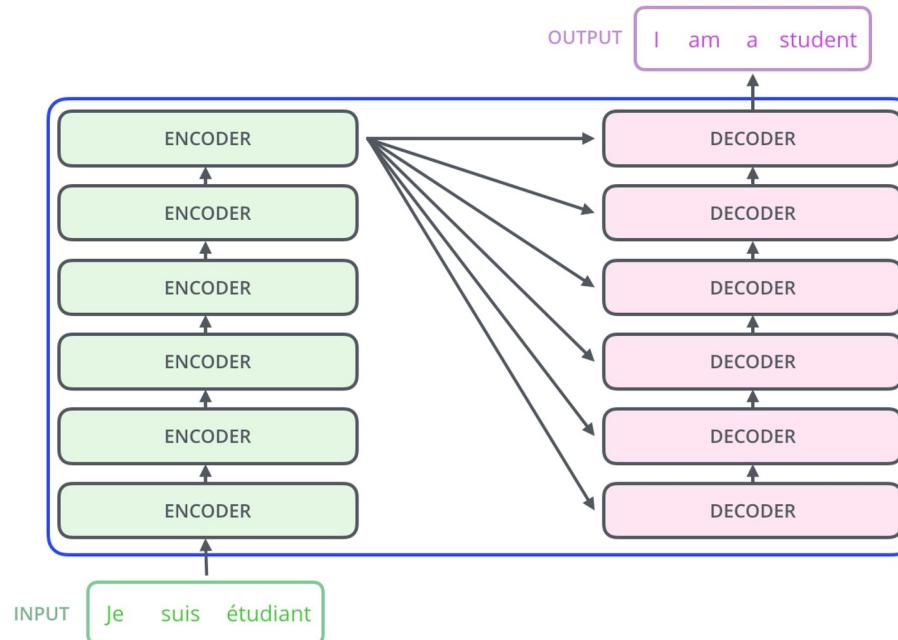
Источник изображения: <http://jalammar.github.io/illustrated-transformer/>

Transformer [Vaswani et.al. 2017]



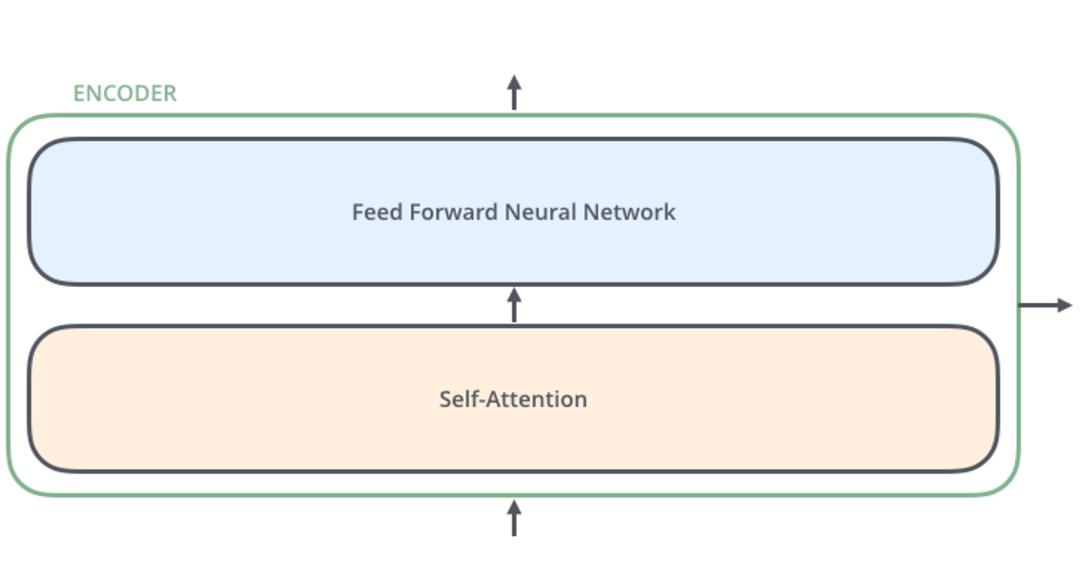
Источник изображения: <http://jalammar.github.io/illustrated-transformer/>

Transformer [Vaswani et.al. 2017]



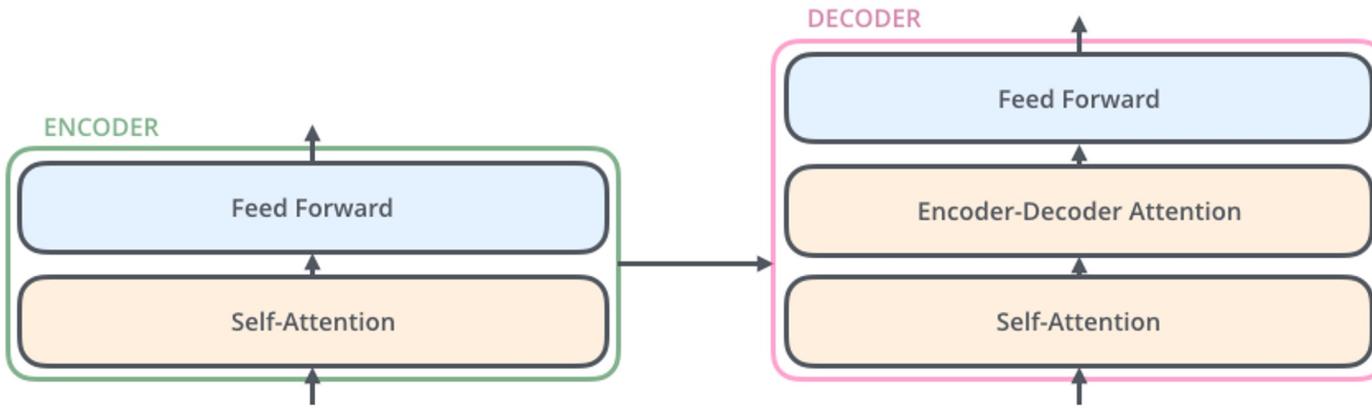
Источник изображения: <http://jalammar.github.io/illustrated-transformer/>

Transformer [Vaswani et.al. 2017]



Источник изображения: <http://jalammar.github.io/illustrated-transformer/>

Transformer [Vasvani et.al. 2017]



Источник изображения: <http://jalammar.github.io/illustrated-transformer/>

Обучение

- Обучающая выборка - “параллельные” корпуса

Обучение

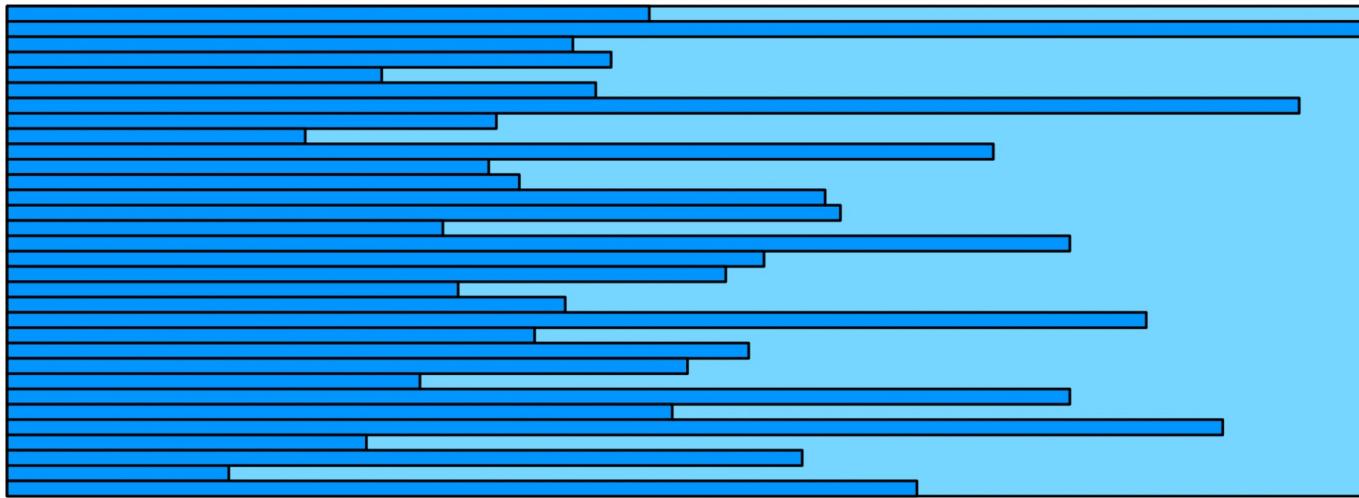
Обучающая выборка - “параллельные” корпуса

Откуда брать обучающие данные?

1. Ручные переводы аннотаторами
2. Автоматический сбор данных
 - a. Переводы книг / статей + выравнивание
 - b. Поиск похожих по смыслу предложений и абзацев среди неструктурированных данных

Обучение

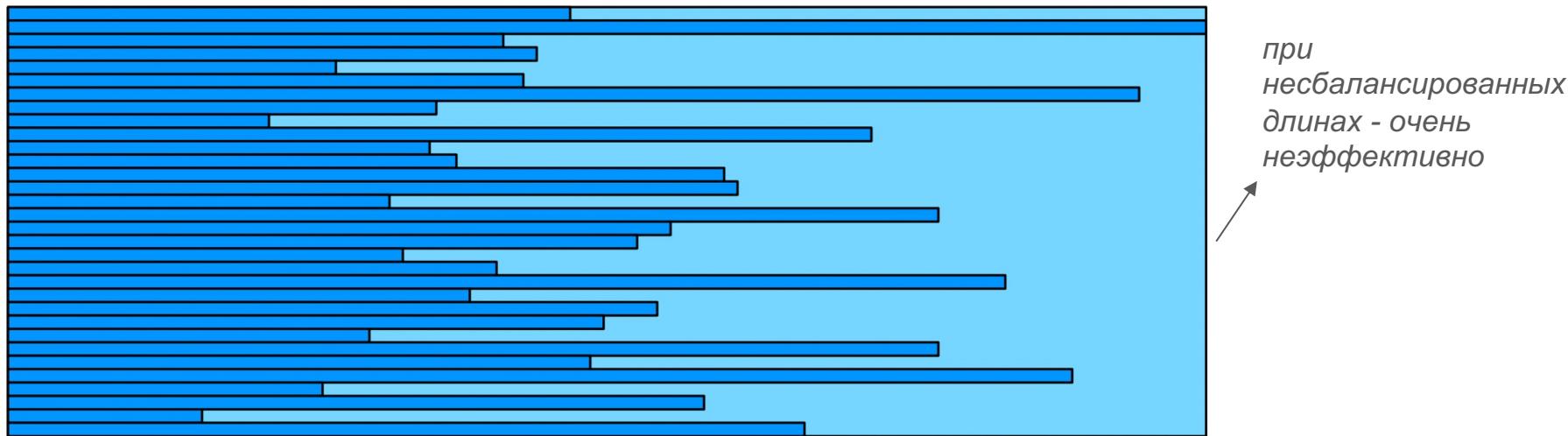
- Обучающие данные (предложения) сильно различаются по длине



- при упаковке данных в мини-батчи необходимо выровнять длину путем добавления спец.токенов паддинга
- по этим токенам не считается функция потерь

Обучение

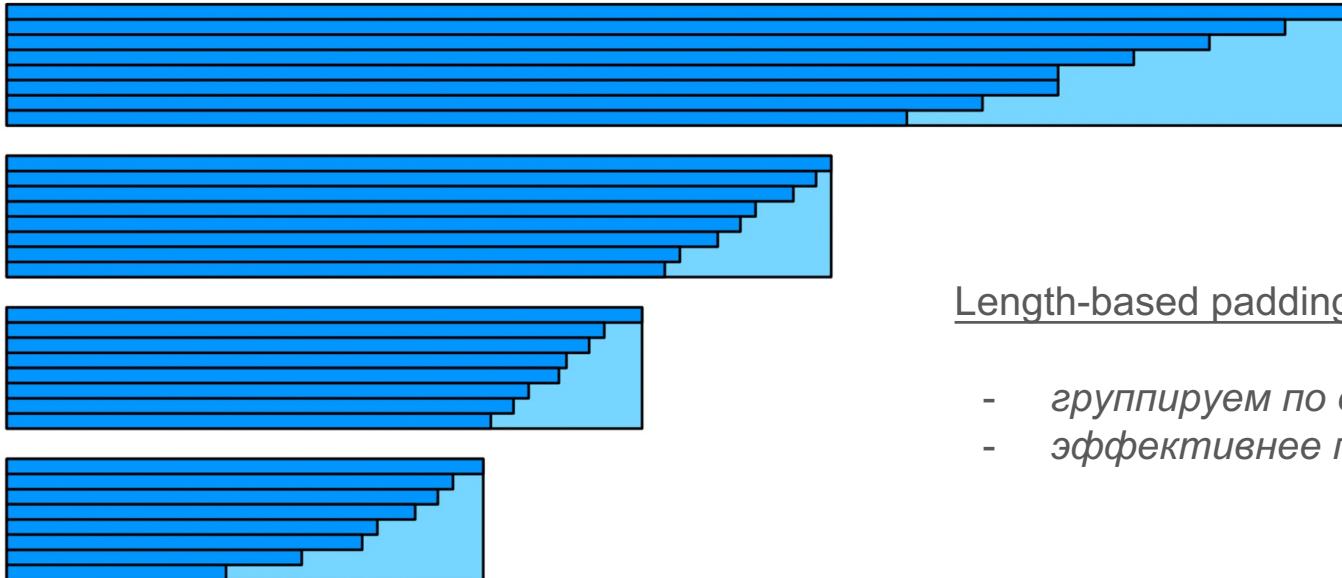
- Обучающие данные (предложения) сильно различаются по длине



- при упаковке данных в мини-батчи необходимо выровнять длину путем добавления спец.токенов паддинга
- по этим токенам не считается функция потерь

Обучение

- Обучающие данные (предложения) сильно различаются по длине



- группируем по длине
- эффективнее по времени

Оценка качества перевода

Почему оценивать качество перевода - сложно?

У одного предложения может быть несколько допустимых переводов

English: “He left the bank”

Русский:

- “Он покинул банк”
- “Он ушел из банка”
- “Он ушел с берега”
- “Он оставил банк”
- “Он отошел от берега”

Почему оценивать качество перевода - сложно?

У одного предложения может быть несколько допустимых переводов

English: “He left the bank”

Русский:

- “Он покинул банк”
- “Он ушел из банка”
- “Он ушел с берега”
- “Он оставил банк”
- “Он отошел от берега”

Недостаточно просто сравнивать с эталонным переводом (accuracy)

Малейшее несовпадение будет штрафоваться так же сильно, как и абсолютно несвязанный текст

Потокенные overlap-based метрики

- Разбиваем текст на токены
- Смотрим на долю совпадений между токенами перевода и токенами референса

Precision	$\frac{\text{correct}}{\text{output-length}}$	F-мера	$\frac{\text{precision} \times \text{recall}}{(\text{precision} + \text{recall})/2}$
Recall	$\frac{\text{correct}}{\text{reference-length}}$	Расстояние Левенштейна (нормированное)	$\text{WER} = \frac{\text{substitutions} + \text{insertions} + \text{deletions}}{\text{reference-length}}$

Потокенные overlap-based метрики

- Разбиваем текст на токены
- Смотрим на долю совпадений между токенами перевода и токенами референса

Precision

$$\frac{\text{correct}}{\text{output-length}}$$

F-мера

$$\frac{\text{precision} \times \text{recall}}{(\text{precision} + \text{recall})/2}$$

Recall

$$\frac{\text{correct}}{\text{reference-length}}$$

Расстояние Левенштейна (нормированное)

$$\text{WER} = \frac{\text{substitutions} + \text{insertions} + \text{deletions}}{\text{reference-length}}$$

проблемы?

BLEU

- смотрим на n-граммную точность
- “какая доля n-грамм перевода присутствует в референсе?”

$$\text{BLEU} = \min \left(1, \frac{\text{output-length}}{\text{reference-length}} \right) \left(\prod_{i=1}^4 precision_i \right)^{\frac{1}{4}}$$

BLEU

- смотрим на n-граммную точность
- “какая доля n-грамм перевода присутствует в референсе?”

$$\text{BLEU} = \min \left(1, \frac{\text{output-length}}{\text{reference-length}} \right) \left(\prod_{i=1}^4 \text{precision}_i \right)^{\frac{1}{4}}$$



очень короткие переводы по построению имеют высокую точность
явным образом учитываем длину

Нейросетевые автометрики

- BLEURT | Трансформерные encoder-only модели, предсказывают меру близости к референсу
- COMET | Аналогично COMET, но в режиме QE
Т.е. вычисляет “качество в абсолюте”, а не близость
- GEMBA | Дообучение GPT под выделение ошибок

Human evaluation

- Более точная оценка - ручная аннотация человеком

Human evaluation

Pointwise

оценка перевода по инструкции

Side-by-side

сравнение переводов одного и того же текста

Side by Side метрики

Исходный текст:

Boethius engages with the nature of predestination and free will, the problem of evil, human nature, virtue, and justice. He speaks about the nature of free will and determinism when he asks if God knows and sees all, or does man have free will. On human nature, Boethius says that humans are essentially good, and only when they give in to "wickedness" do they "sink to the level of being an animal." On justice, he says criminals are not to be abused, but rather treated with sympathy and respect, using the analogy of doctor and patient to illustrate the ideal relationship between prosecutor and criminal.

Показать контекст

Перевод 1:

Боэций занимается природой предопределения и свободы воли, проблемой зла, человеческой природой, добродетелью и справедливостью. Он говорит о природе свободы воли и детерминизма, когда задается вопросом, знает ли Бог и видит ли все или у человека есть свобода воли. О человеческой природе Боэций говорит, что люди по сути добры, и только когда они поддаются "злу", они "понижаются до уровня животного". О справедливости он говорит, что преступников не следует обижать, а скорее относиться к ним с сочувствием и уважением, используя аналогию врача и пациента для иллюстрации идеальных отношений между обвинителем и преступником.

Перевод 2:

Боэций затрагивает вопросы природы предопределения и свободы воли, проблему зла, человеческой природы, добродетели и справедливости. О природе свободы воли и детерминизма он говорит, задаваясь вопросом, знает ли Бог и видит ли он все, или же человек обладает свободной волей. О природе человека Боэций говорит, что человек по сути своей добр, и только когда он поддается "пороку", он "опускается до уровня животного". Что касается правосудия, то он говорит, что с преступниками не следует жестоко обращаться, а лучше относиться к ним с сочувствием и уважением, используя аналогию врача и пациента для иллюстрации идеальных отношений между обвинителем и преступником.

Выберите из пары переводов текстов лучший:

- 1 Первый 2 Второй Q Одинаковы по качеству E Ошибка

Side by Side метрики

Исходное слово или фраза

I found this originally on Mr breakfast.com and tweaked it a bit to omit the sugar.

Переводы

Изначально я нашла этот рецепт на сайте Mr breakfast.com и немного изменила его, опустив сахар.



Я нашел это первоначально на Mr. breakfast.com и немного подправил его, чтобы не добавлять сахар.



Pointwise метрики

Исходное предложение

(yes, I know it sucks, but you will have to restart the PC again after installation).

Перевод

(да, я знаю, что это плохо, но вам придется снова перезагрузить компьютер после установки).

1 Плохо 2 Хорошо

Sometimes even if one just passes by the mysterious washerwoman by the side of the road in the daytime and have not heard the wail, that, too, may be all the warning one gets of an impending death in the family.

Иногда даже если кто-то просто проходит мимо таинственной прачки на обочине дороги днем и не слышит вопля, это тоже может быть единственным предупреждением о неминуемой смерти в семье.

64

Очень плохо

Очень хорошо

Проблемы этих метрик

sbs

Сравнение систем,
а не оценка перевода

Метрика выделения ошибок

Number box

Represents a control that can be used to display and edit numbers. This supports validation, increment stepping, and computing inline calculations of basic equations, such as multiplication, division, addition, and subtraction.

Is this the right control?

You can use a NumberBox control to capture and display mathematic input. If you need an editable text box that accepts more than numbers, use the TextBox control. If you need an editable text box that accepts passwords or other sensitive input, see PasswordBox. **If you need a text box to enter search terms, see AutoSuggestBox.** If you need to enter or edit formatted text, see RichEditBox.

Recommendations

- "Text" and "Value" make it easy to capture the value of a NumberBox as a String or as a Double without needing to convert the value between types. When programmatically altering the value of a NumberBox, it is recommended to do so through the "Value" property. "Value" will overwrite "Text" in initial set up. After the initial set up, changes to one will be propagated to the other, but consistently making programmatic changes through "Value" helps avoid any conceptual misunderstanding that NumberBox will accept non-

Поле для ввода номера

Представляет собой элемент управления, который можно использовать для отображения и редактирования числовых данных, а также для проверки, пошагового увеличения или уменьшения значений и выполнения встроенных вычислений по основным формулам, таким как умножение, деление, сложение и вычитание.

Это та самая кнопка, которая нужна?

Для ввода и отображения математических данных можно использовать элемент управления NumberBox. Если вам нужно создать редактируемое текстовое поле, в которое можно вводить не только числа, но и другие данные, используйте элемент управления TextBox. Для создания редактируемого текстового поля, в котором можно вводить пароли и другие конфиденциальные данные, см. раздел PasswordBox. Для ввода или редактирования форматированного текста см. раздел RichEditBox.

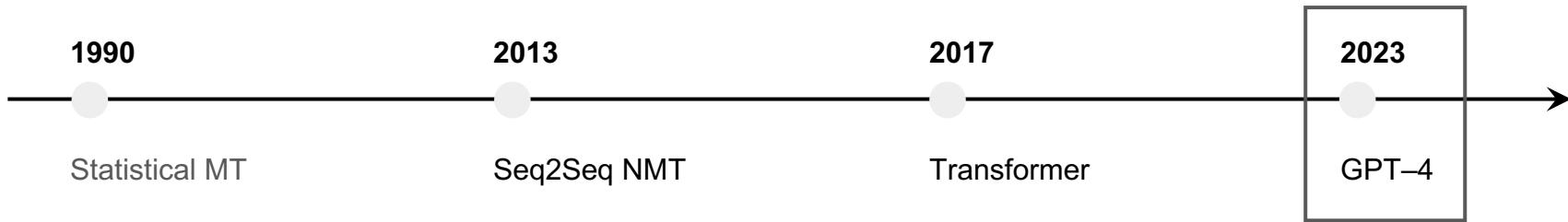
Рекомендации

- Свойства «Текст» и «Значение» позволяют легко считывать значение элемента управления «Числовое поле» как строку или как число с плавающей запятой без необходимости преобразования значения между типами данных.

E	major
R	minor
<hr/>	
C	Выделение
ALT +C	Всё
ALT +X	В переводе нет ошибок

Машинный перевод в 2024 году

Машинный перевод в 2024 году



- › end2end обучение одной вероятностной модели перевода
- › Сильный буст fluency
- › Проблемы с out-of-domain параллельного корпуса

Задача перевода решена??

Машинный перевод в 2024 году

WMT (Workshop on Machine Translation)

- › Ежегодный бенчмарк качества open-source систем
- › General MT Task - перевод «обычных текстов»

Машинный перевод в 2024 году

**Findings of the 2023 Conference on Machine Translation (WMT23):
LLMs Are Here But Not Quite There Yet**

Rank	Ave. ↑	System (En-De)	MQM ↓
1-5	89.0	GPT4-5shot	2.96
1-5	88.8	ONLINE-B	3.72
1-4	88.3	ONLINE-W	3.95
2-6	88.1	ONLINE-A	4.71
4-6	88.0	ONLINE-Y	5.64
1-6	87.7	Human-refA	5.67
7-8	86.7	ONLINE-M	6.57
7-8	85.5	ONLINE-G	6.94
9	84.0	Lan-BridgeMT	8.67
10	82.7	LanguageX	9.25
11-12	76.8	NLLB_MBR_BLEU	9.54
11-12	75.7	NLLB_Greedy	10.79
13	73.6	AIRC	14.23



Интегральный скор качества



Число ошибок на сегмент

На high-resource языковых парах NMT на уровне человека (но содержит чуть больше смысловых ошибок)

Машинный перевод в 2024 году

Перевод на уровне предложений

- › Классическая постановка задачи
- › Много данных, в т.ч. автоматически собранных
- › Качество близко к человеческому

Перевод полных текстов

- › Очень мало high-quality документных переводов
- › Задача объективно сложнее для моделирования
- › Очень много проблем с контекстом и гладкостью

Машинный перевод в 2024 году

Original:

«90 min: Off the line by Olmo! Palmer's outswinging corner is met by Rice, whose powerful header is pushed away by Unai Simon»



Sent-level:

«90 минут: **снято с линии** Ольмо! Навес Палмера с углового **отбивает Райс**, но Унаи Симон **отразил удар головой**»

Ideal:

«90 минута: **вынос с ленточки** от Ольмо!
Уходящий навес Палмера с углового **принял Райс**,
чей мощный удар головой отражен Унаи Симоном»

Машинный перевод в 2024 году

Original:

«90 min: Off the line by Olmo! Palmer's outswinging corner is met by Rice, whose powerful header is pushed away by Unai Simon»



Sent-level:

«90 минут: **снято с линии** Ольмо! Навес Палмера с углового **отбивает Райс**, но Унаи Симон **отразил удар головой**»

Для практических приложений
задача больше неинтересна

Ideal:

«90 минута: **вынос с ленточки** от Ольмо!
Уходящий навес Палмера с углового **принял Райс**,
чей мощный удар головой отражен Унаи Симоном»

Хотим делать контекстный
перевод крупных фрагментов
с качеством на уровне человека

Документный перевод: challenges

1. Данные для обучения

Переводы, сделанные человеком

Есть две опции:

- › Литературный перевод - хорошо, мало, плохое выравнивание
- › Перевод “близко к тексту” - не дает гладкость на стороне dst

Automated crawlers and augmentations

Для сбора таких данных уже нужна документная модель

Документный перевод: challenges

2. Учет длинного контекста

- › Маленьким enc-dec моделям не хватает capacity для моделирования сложного контекста
- › Мультимодальный контекст
- › Консистентность и гладкость в переводе

Как дообучать LLM под перевод?

Вариант 1: Full Finetune на переводные данные

- › Улучшили формат
- › Нет галлюцинаций
- › Модель максимально похожа на seq2seq обучение с нуля



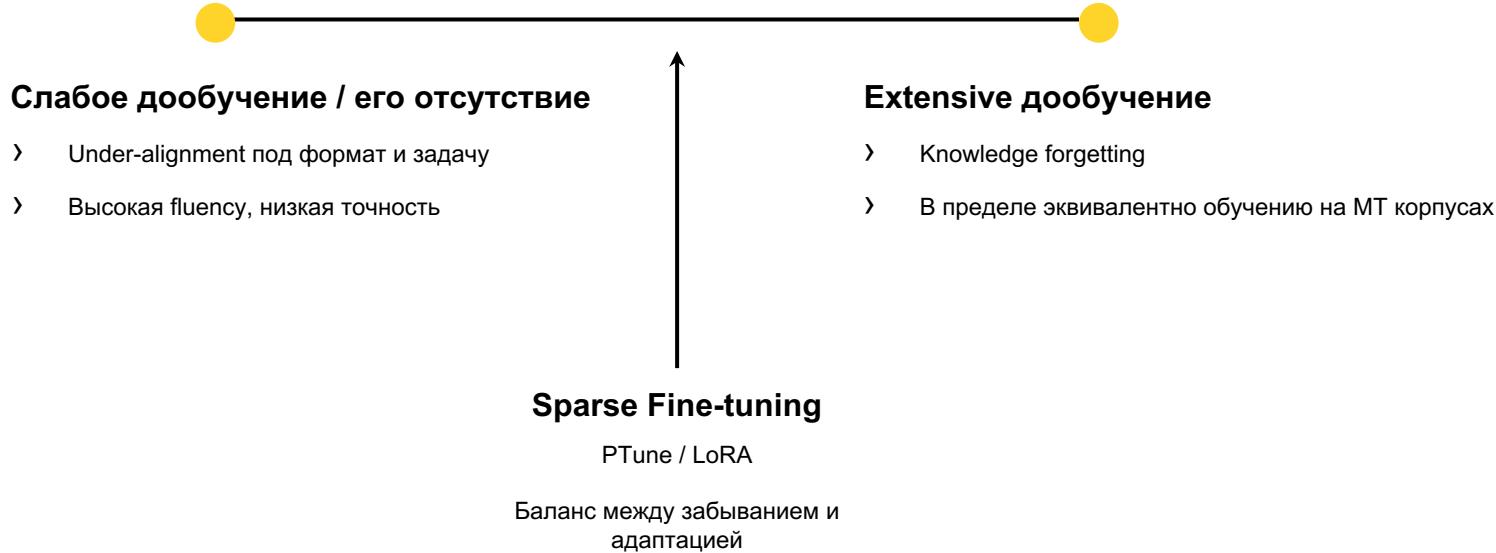
Слабое дообучение / его отсутствие

- › Under-alignment под формат и задачу
- › Высокая fluency, низкая точность

Extensive дообучение

- › Knowledge forgetting
- › В пределе эквивалентно обучению на MT корпусах

Как дообучать LLM под перевод?



Машинный перевод в 2024 году: что дальше?

Мультимодальный перевод

- › Перевод с учетом контекста изображения / видео
- › Учет более сложных эмбеддингов содержимого (тематика книги, формальность текста)

Базовое качество

Ошибок стало мало. Но уровень человека не достигнут на сложных / плохо покрытых доменах

Контролируемый перевод:

- › Стиль, сложность, формальность
- › Гибкий режим с фидбеком от пользователя (интерактивный переводчик)

Длинный контекст:

Единовременный перевод очень большого документа (книга)

Спасибо за внимание



Y&Z Y