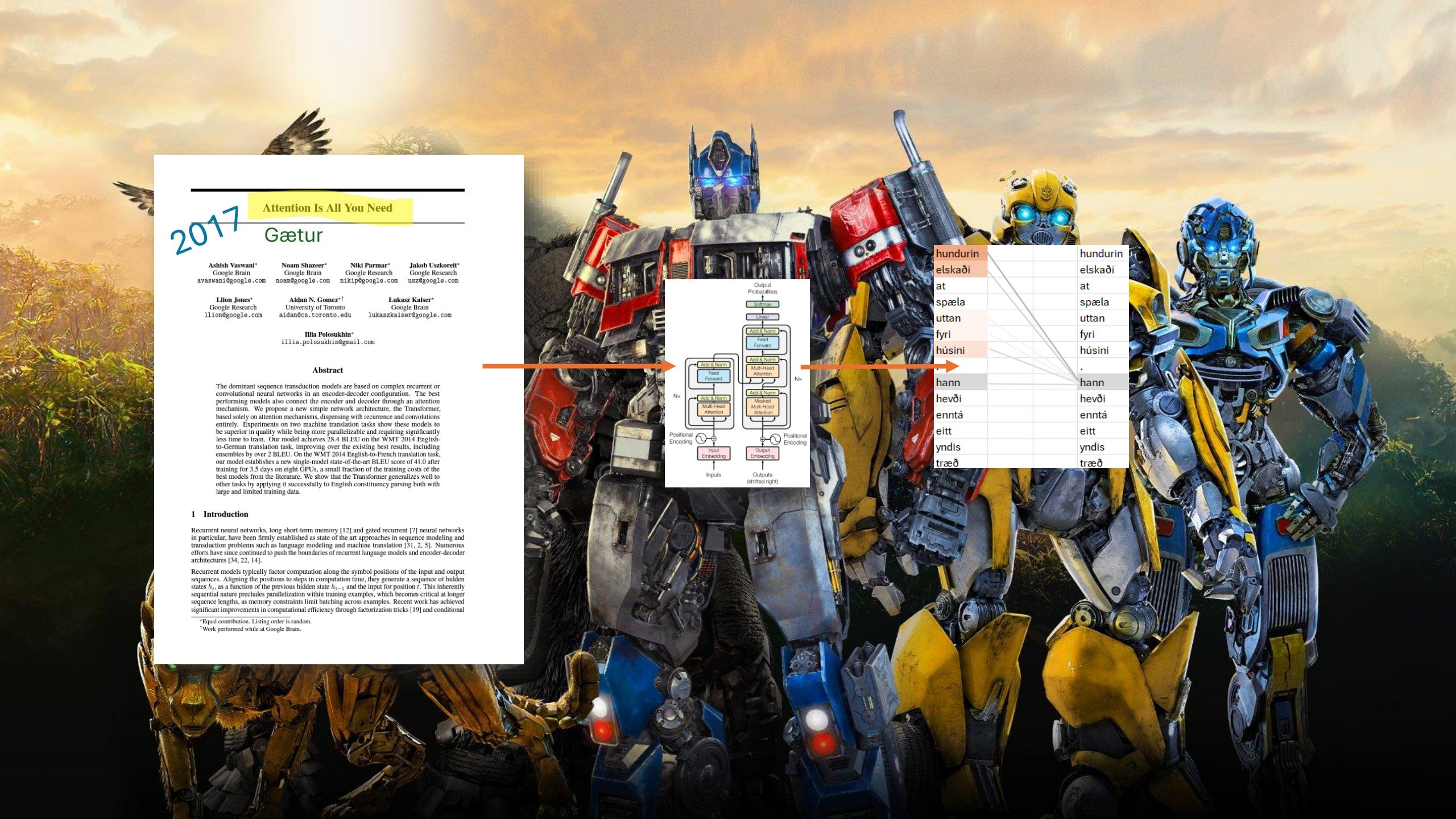


Vitlíki til Tala-til-Tekst málþókni til føroykst

Dávid í Lág, M.Sc. Teldufrøði, ph.d.-ví sindastarv



FRÓÐSKAPARSETUR FØROYA
UNIVERSITY OF THE FAROE ISLANDS



2017

Attention Is All You Need

Gætur

Ashish Vaswani*
Google Brain
avaswani@google.com

Noam Shazeer*
Google Brain
noam@google.com

Niki Parmar*
Google Research
nikip@google.com

Jakob Uszkoreit*

Google Research
usz@google.com

Llion Jones*
Google Research
lliond@google.com

Aidan N. Gomez*†
University of Toronto
aidan@cs.toronto.edu

Lukasz Kaiser*
Google Brain
lukasz.kaiser@google.com

Illia Polosukhin*
illia.polosukhin@gmail.com

Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks in an encoder-decoder configuration. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.0 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.

1 Introduction

Recurrent neural networks, long short-term memory [12] and gated recurrent [7] neural networks in particular, have been firmly established as state of the art approaches in sequence modeling and transduction problems such as language modeling and machine translation [31, 2, 5]. Numerous efforts have since continued to push the boundaries of recurrent language models and encoder-decoder architectures [34, 22, 14].

Recurrent models typically factor computation along the spatial positions of the input and output sequences. Aligning the positions to steps in computation time, they generate a sequence of hidden states h_t as a function of the previous hidden state h_{t-1} and the input for position t . This inherently sequential nature precludes parallelization within training examples, which becomes critical at longer sequence lengths, as memory constraints limit batching across examples. Recent work has achieved significant improvements in computational efficiency through factorization tricks [19] and conditional

*Equal contribution. Listing order is random.

†Work performed while at Google Brain.

Alheims mál og tøkni

- 6-7.000 mál eru til í verðini
- 3.000 mál hava eitt skriftmál
- 900+ mál hava eitt knappaborð
- 140 mál hava Talu-til-Tekst tøkni

Gransking í Tala-til-Tekst

- Hví granska í Tala-til-Tekst tøkni?
 - Økja um atkomu og inklusjón
 - Gagnnýta tí nýggjastu vitlíkistøknina
 - Áhugi fyrir føroyska málinum
- Endamálið við granskningini
 - Gera tað gjørligt hjá teldum at skilja, og skriva, føroyska talu
 - Gera føroyskt tøkt í tí talgilda heiminum
 - Dríva tøkniliga nýskapan

Hvat er Tala-til-Tekst?

- Umskapa talað mál til skrivaðan tekst
 - Lætt hjá menniskjum, trupult hjá teldum
- Tøknin kann t.d. brúkast til
 - Tekst frá tímanum ella fundinum
 - Undirtekstir á YouTube
 - Voice commands/search
- Avbjóðingin
 - Tøknin riggar væl á enskum
 - Riggar ikki væl á fôroyskum

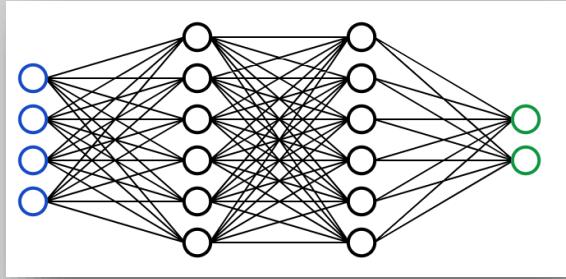


Hví er føroyskt ein avbjóðing?

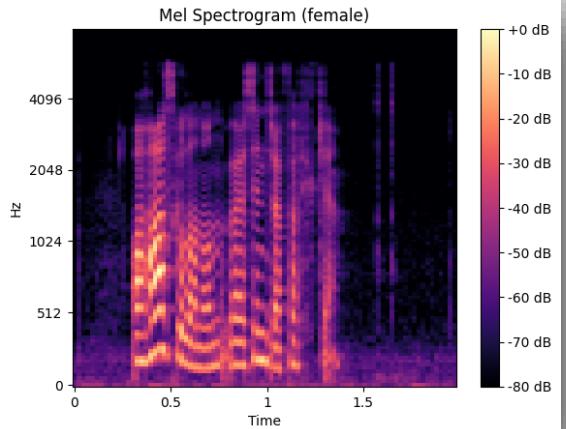
- Lítið mál (“low-resource”)
 - Bara 54.000+ fólk tosa føroyskt
 - Ikki nógvar dátur
 - Manglandi atgongd til dátur
 - “Code-switching”
- Kompleksiteturin av talu
 - Dialektir, aldur, kyn, talumátar, larmur á upptøkum, v.m.

Tala-til-Tekst tøkni

- Maskinlæra við Neuralum Netverkum at finna mynstur í talu
- Mappa talu til eini umboðan av talu
- Brúka dátur at trena modellir at skilja føroyskt
- Málmodellið ger forsøgn um orð ella setningar



"Á háskúlanum var eitt sindur av fóldi"



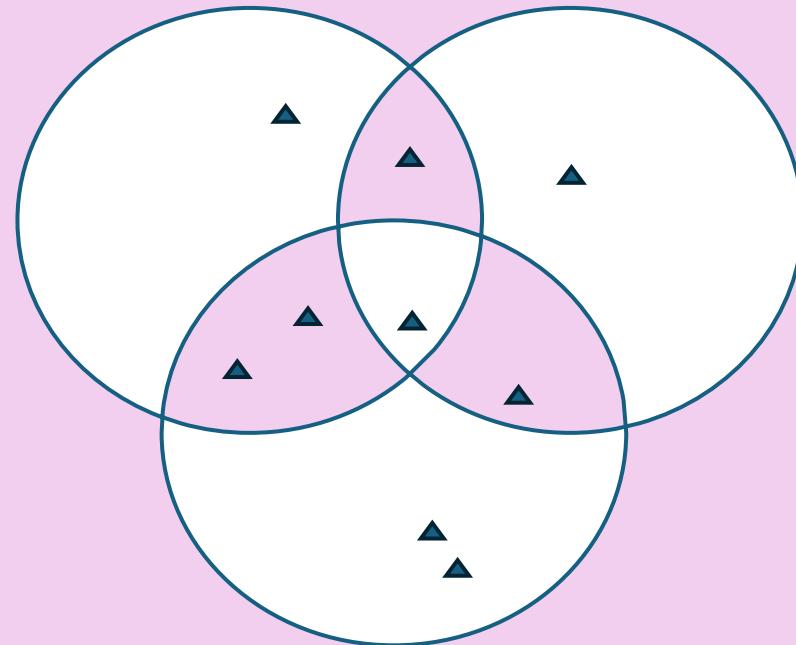
Vitanarflyting millum mál

- Old Nordiskt sum felags byrjan
- Vitanarflyting frá stórrri málum til smærri
- Ymsar fjarstøður millum føroyskt og hini málini
- Menna nýggj málmodellir sum innihalda dátur frá málum við nógvari vitan um føroyskt



Granskningin

- Fleirmálmodellir læra talaðar representációir
- Menna nýggj málmodell
 - Meta: *wav2vec 2.0*
 - 56 mál, 56.000 tímar, 317M
 - Lærur um mál uttan tekstu við at fjala partar talu undir trening
 - OpenAI: *Whisper*
 - 99 mál, 660.000 tímar, 1.5B
 - Lærur mál við myndum og tilhoyrandi tekstu
- Distillatión av málmodellum
 - Minka um stødd og stytta um roknutíð
- Dátu øking
 - Samla dátur sum longu finnast
 - Gera nýggjar syntetiskar dátur út frá tí sum er til
 - Gera nýggjar dátur við Tekst-til-Talu tøkni



Hvat siga dátur?



Ávirkan av granskningini í heiminum

- Vísindaliga ávirkan: Betri avrik við færri dátum fyri lítil mál
- Tekniska ávirkan: Gera nýggj Talu-til-Tekst modell tøk fyri øll at brúka
- Sosiala ávirkan: Betri atgongd til góða Talu-til-Tekst tøkni fyri øll
 - Fáa tekst úr føroyeskari talu
 - Hjálp til fólk við niðursettari ella ongari hoyrn, orðblind, v.fl.
 - Tekst til alt hjá KVF, sjónvarp og útvarp
 - Løgtingið, beinleiðis og savnið
 - Brúka tekstin til greining/granskning
 - Til tey sum vilja læra føroykst

Til endans

- Nútímans málþøkni at fáa føroyskt inn í tann talgilda heimin
- Áhugavert øki sum er í stórari menning
- Granskigin kemur við nýggjari vitan

Takk fyri meg

davidl@setur.fo

huggingface.co/davidilag

