BLEU： 译文中有多少n-gram在标准译文中出现 **（准确率）**

ROUGE： 译文中有多少词(1-gram)在标准译文中出现 **（召回率）**

METTOR： BLEU和ROUGE的调和平均，**考虑了同义词和时态（比如like和likes)**

SARI： 同时考虑的输入和标准译文，**希望输出在输入的基础上做出改变**

Unified Medical Language System (UMLS)

A set of files and software that brings together many health and biomedical vocabularies and standards to enable interoperability between computer systems.

Consumer Health Vocabulary (CHV)

CHV is designed to complement the existing framework of the UMLS and to aid the needs of consumer health applications. It enables these applications to translate technical terms to consumer friendly language.

Leveraging Social Media for Medical Text Simplification

SIGIR, 2020

H5 = 75

■ 网络包含大量的生物医学信息：包括研究、患者信息、健康记录、临床实验

■ 这些信息主要由医学专家撰写编辑，缺乏相关知识的普通读者难以使用

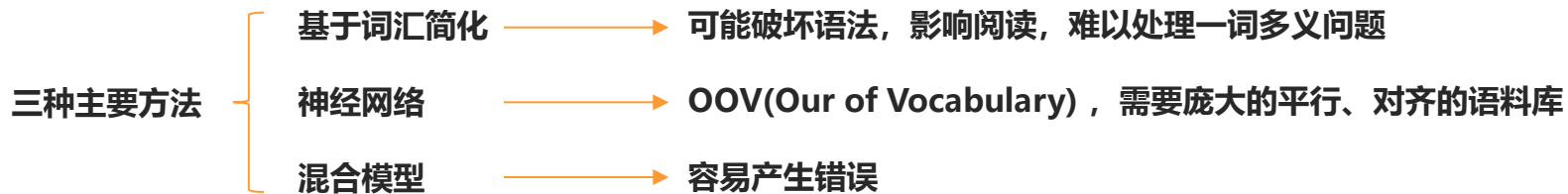医疗文本简化(Medical Text Simplification, MTS)
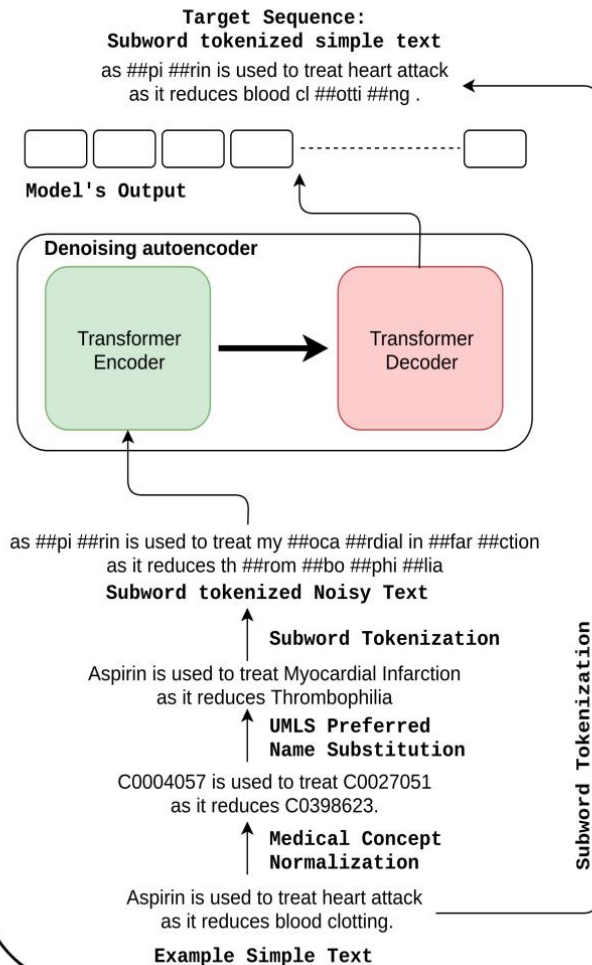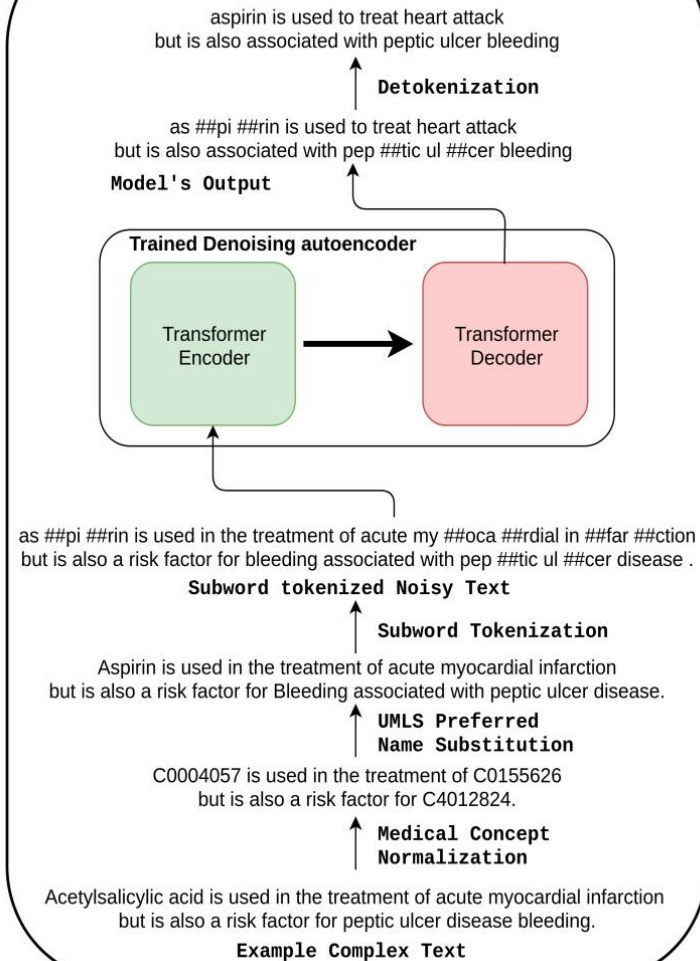
目标： 减少语言的复杂性、翻译医学术语
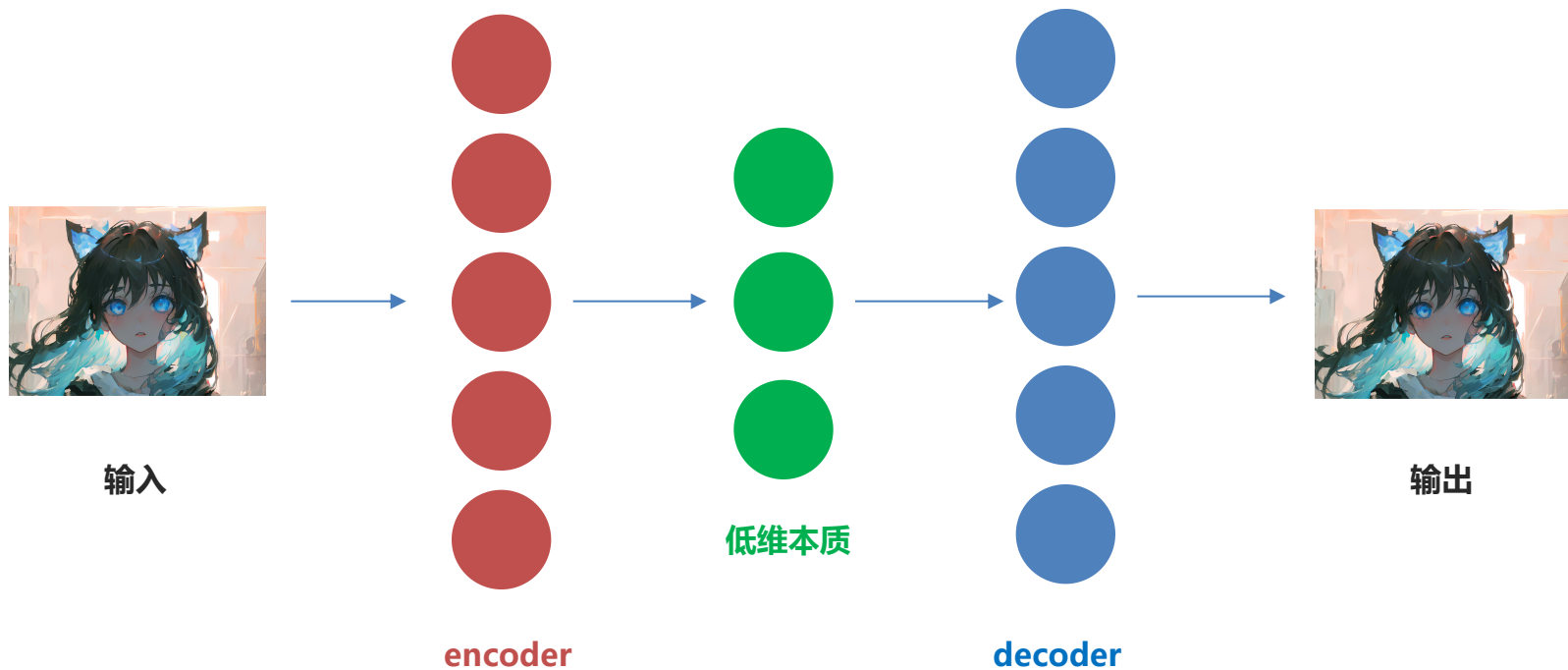
三个主要任务
- 词汇简化
- 句法简化
- 句子压缩

三种主要方法

基于词汇简化 ⟶ 可能破坏语法，影响阅读，难以处理一词多义问题

神经网络 ⟶ OOV(Our of Vocabulary)，需要庞大的平行、对齐的语料库

混合模型 ⟶ 容易产生错误

Training phase

Target Sequence:
Subword tokenized simple text

as ##pi ##rin is used to treat heart attack
as it reduces blood cl ##otti ##ng .

Model's Output

Denoising autoencoder

Transformer Encoder → Transformer Decoder

as ##pi ##rin is used to treat my ##oca ##rdial in ##far ##ction
as it reduces th ##rom ##bo ##phi ##lia
Subword tokenized Noisy Text

↑ **Subword Tokenization**

Aspirin is used to treat Myocardial Infarction
as it reduces Thrombophilia

↑ **UMLS Preferred Name Substitution**

C0004057 is used to treat C0027051
as it reduces C0398623.

↑ **Medical Concept Normalization**

Aspirin is used to treat heart attack
as it reduces blood clotting.

Example Simple Text

Subword Tokenization

Inference phase

aspirin is used to treat heart attack
but is also associated with peptic ulcer bleeding

↑ **Detokenization**

as ##pi ##rin is used to treat heart attack
but is also associated with pep ##tic ul ##cer bleeding

Model's Output

Trained Denoising autoencoder

Transformer Encoder → Transformer Decoder

as ##pi ##rin is used in the treatment of acute my ##oca ##rdial in ##far ##ction
but is also a risk factor for bleeding associated with pep ##tic ul ##cer disease .
Subword tokenized Noisy Text

↑ **Subword Tokenization**

Aspirin is used in the treatment of acute myocardial infarction
but is also a risk factor for Bleeding associated with peptic ulcer disease.

↑ **UMLS Preferred Name Substitution**

C0004057 is used in the treatment of C0155626
but is also a risk factor for C4012824.

↑ **Medical Concept Normalization**

Acetylsalicylic acid is used in the treatment of acute myocardial infarction
but is also a risk factor for peptic ulcer disease bleeding.

Example Complex Text

自编码器 Aotu-Encoder

输入　　　　　　encoder　　　　　低维本质　　　　decoder　　　　　　输出

Strength

▶ 使用子词分词

▶ 社交媒体文本+自编码器

▶ UMLS词汇替换，添加噪声

▶ 在推理阶段做同样的处理

▶ 解决了OOV问题，模型学习到了词映射

▶ 不需要对齐的数据

▶ 没有减少信息且语法正确

▶ 减少两个阶段中句子风格的差异

	Training Dataset	Medical Paper-Blog				Medical EW-SEW			
		BLEU	SARI	ROUGE	METEOR	BLEU	SARI	ROUGE	METEOR
NTS	EW-SEW	29.40	36.07	53.49	26.97	49.14	28.45	64.08	36.06
NTS	Newsela	11.64	31.09	32.39	13.19	19.41	28.01	41.22	17.98
Transformer	EW-SEW	12.44	30.52	33.30	14.66	18.75	27.13	39.87	18.06
Transformer	Newsela	7.24	27.93	25.92	9.92	8.50	23.28	26.66	10.24
NTS + R	EW-SEW	30.77	36.07	53.49	26.97	**49.14**	28.45	64.08	36.06
NTS + R	Newsela	11.64	31.09	32.39	13.19	19.41	28.01	41.22	17.98
Transformer + R	EW-SEW	12.44	30.52	33.30	14.66	18.76	27.13	39.87	18.06
Transformer + R	Newsela	7.24	27.93	25.92	9.92	8.50	23.28	26.66	10.24
DAE	Medical Blogs	**33.60**$^{\circ\bullet\star}$	**41.15**$^{\circ\bullet\star}$	**65.15**$^{\circ\bullet\star}$	**34.20**$^{\circ\bullet\star}$	38.05	**33.15**$^{\circ\bullet\star}$	**64.96***	**37.09**$^{\circ\bullet\star}$

Table 2: Comparison between all the baselines and our Denoising Autoencoder (DAE) approach. '+R' indicates replacement. Significant differences ($p < 0.05$, paired t-test) compared to the NTS, NTS + R, and Transformer are denoted by $\circ$, $\bullet$ and $\star$ respectively.

	Medical Paper-Blog				Medical EW-SEW				i2b2				MIMIC			
	G	M	S	E	G	M	S	E	G	M	S	E	G	M	S	E
NTS	3.86	3.72	-0.4	3.52	4.52	4.12	-0.06	2.12	4.52	4.00	-0.02	2.82	3.88	3.62	-0.34	4.10
NTS + R	3.60	3.64	-0.54	3.36	4.06	3.92	-0.22	2.52	4.02	3.68	-0.10	2.72	3.52	3.44	-0.46	3.84
Transformer	2.12	1.94	-0.94	5.06	3.44	2.74	-0.28	3.92	2.86	2.36	-0.62	4.54	2.32	2.02	-1.18	5.50
Transformer + R	2.10	1.92	-1.00	5.06	3.18	2.68	-0.36	4.00	2.76	2.24	-0.92	4.62	2.24	1.94	-1.20	5.44
DAE	**4.68**	**4.6**	**0.12**	**1.88**	**4.32**	**4.38**	**0.06**	**2.20**	**4.34**	**4.34**	**0.22**	**1.92**	**4.32**	**4.28**	**-0.10**	**2.12**

Table 3: Comparison of the average scores assigned by human evaluators to all the models for Grammar (G), Meaning preservation (M), Simplicity (S) and Error type (E). '+R' indicates replacement.

Original Sentence	DAE	NTS	NTS + Replacement
it was felt that the patient should have an augmented medical regimen and as an outpatient undergo an exercise tolerance test with viability evaluation to determine whether he would benefit from revascularization.	it was felt that the patient should have an **increased** medical regimen and as an outpatient undergo an **exercise test** with **viable assessments** to determine whether he would benefit from revascularization.	it was felt that the patient should have an augmented medical regimen.	it was felt that the patient should have an **increased by** medical regimen.
other concerns during his stay was his kidney function , due to his renal transplant in 2003.	other concerns during his stay was his kidney function , due to his **kidney transplants** in 2003.	other concerns during his stay was his kidney function , due to his renal transplant in 2003.	other concerns during his stay was his entire kidney function , due to his **renal transplanted tissue** in 2003.
change in lv mass was positively associated with systolic blood pressure and body mass index and negatively associated with treated hypertension and high-density lipoprotein cholesterol level.	changes in **left ventricular** mass was positively associated with systolic blood pressure and body mass index and negatively associated with **treating** hypertension and **hdl cholesterol levels**.	it is associated with systolic blood pressure and body mass index.	it is associated **in addition to systole**, **function** blood pressure and body **observation of a** mass index.
surgery alone or surgery combined with pelvic radiation is typically used to treat vaginal cancer.	surgery alone or surgery combined with **hip radiation** is typically used to treat vaginal cancer.	it is used to treat vaginal cancer.	it is used to treat **into vagina ca - cancer**.
this type of color blindness is usually a sex-linked condition.	this type of color blindness is usually a **gender - linked** condition.	this type of color blindness is usually a sex-linked condition.	this type of color **loss of vision** is usually a sex-linked condition.

Table 4: Sample outputs of different MTS models. The changes made by the models are indicated using boldfaced font.

Two Parents, One Child:
Dual Transfer for Low-Resource
Neural Machine Translation

ACL, 2021

H5 = 169

Introduction and Background
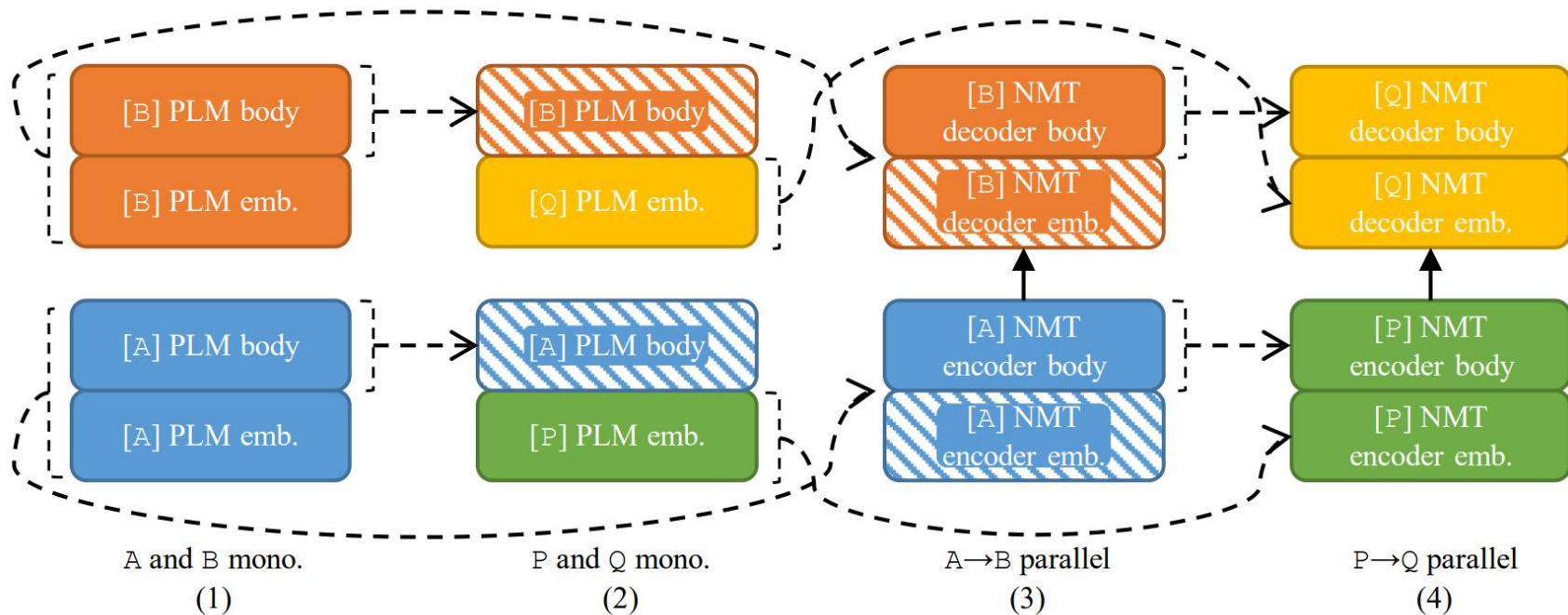
平行语料难以获得，但单语语料不少

能够使用高资源语料辅助低资源翻译任务的训练，即转移学习

Figure 1: Dual transfer from PLM and high-resource A→B NMT to low-resource P→Q NMT. Dashed lines represent initialization. Parameters in striped blocks are frozen in the corresponding step, while other parameters are trainable. Different colors represent different languages. Data used in each step is also listed.

在单语语料中训练高语言资源A和B的语言模型（Bert)

使用A和B的语言模型，初始化除了embedding之外的P和Q的语言模型参数，并将这些参数固定，训练P和Q的embedding
将A和P，B和Q放在相同的语义空间中

训练A -> B的翻译模型： (Transformer)
使用预训练的语言模型A的编码器和B的解码器进行初始化，固定embedding

训练完毕后，初始化P和Q的编码器、解码器

Methods

Transformer

Bert

BERT

Transformer Encoder

Transformer Encoder

Transformer Encoder

向量表示

输入文本
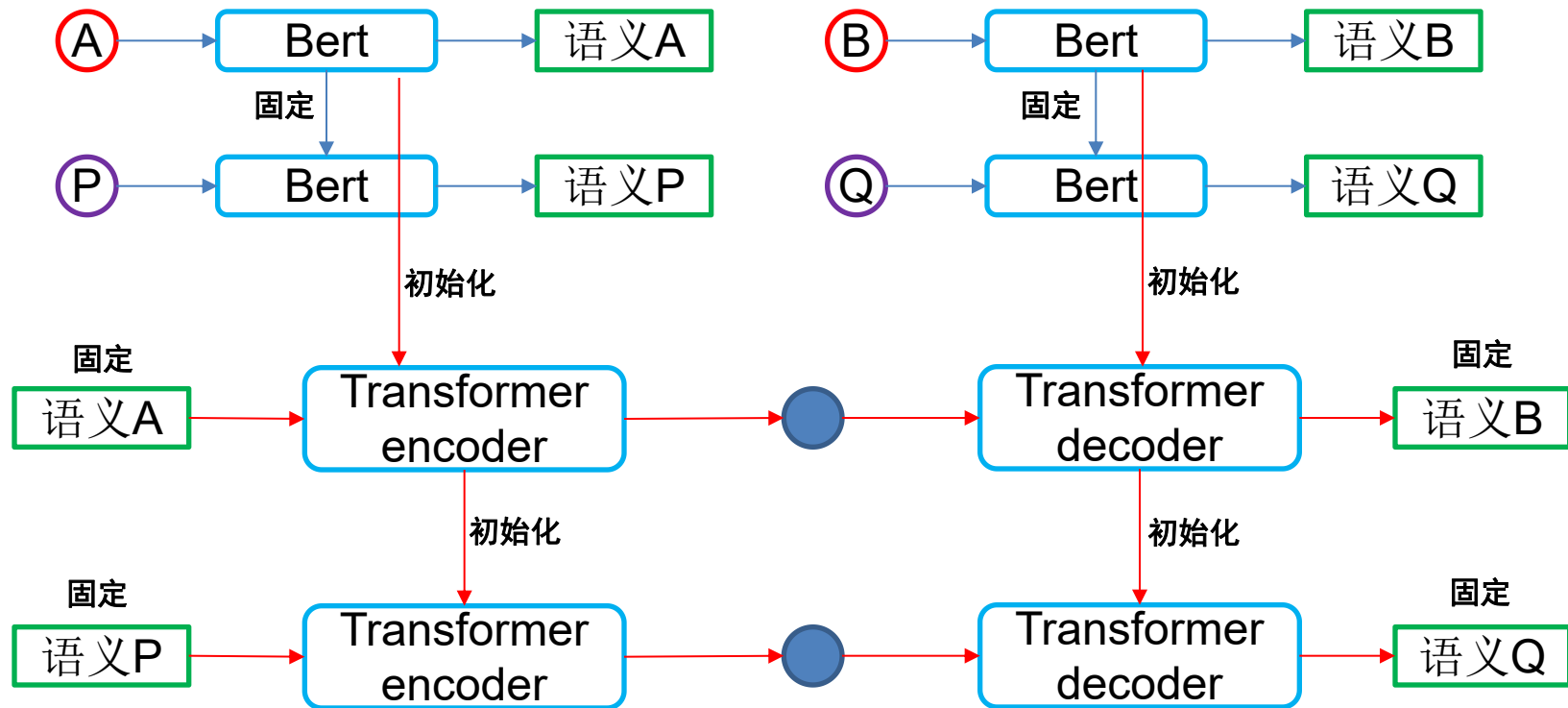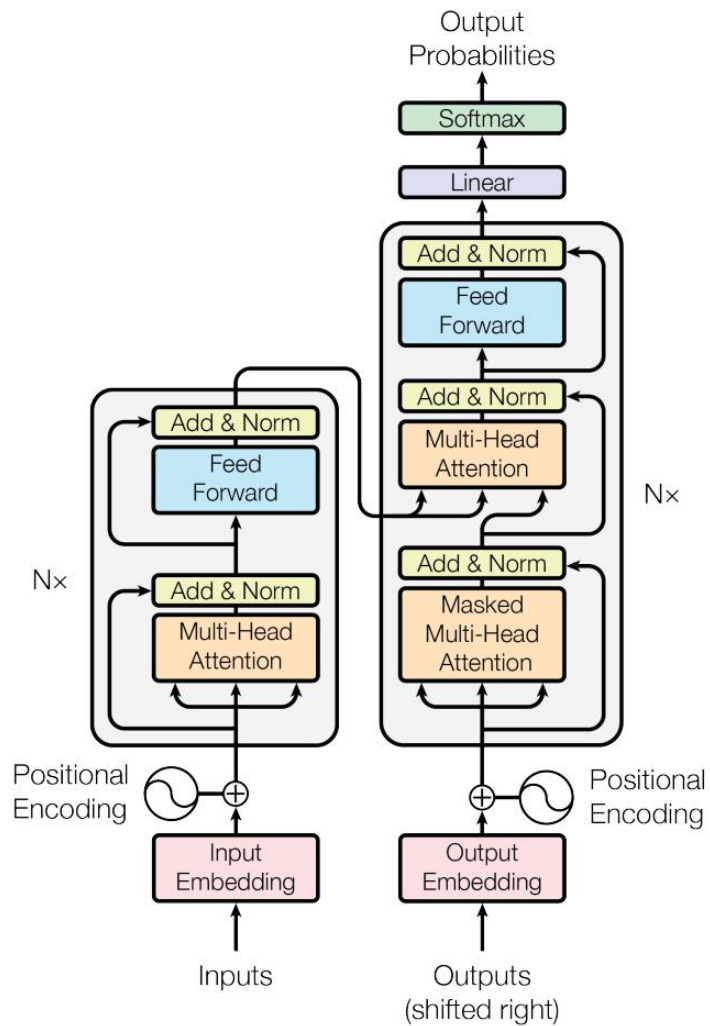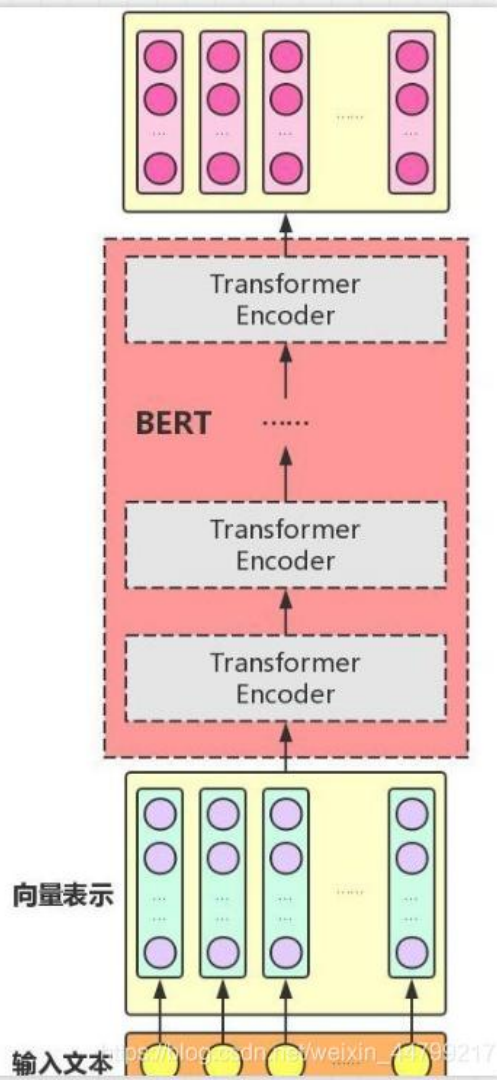
Transformer

approach	V	BLEU
no transfer	✓	21.76
(Zoph et al., 2016)	✓	21.07
(Kim et al., 2019)	✓	22.25
BERT2RND	✓	22.89
BERT2BERT	✓	23.44
(Kocmi and Bojar, 2018)	✗	23.58
BBERT2BBERT	✗	23.90
BBERT transfer	✗	24.03
dual transfer (word)	✓	**24.81**
dual transfer (word+position)	✓	24.28

Table 3: BLEU on et→en, with the best in bold. "✓" in the "V" column indicates independent vocabulary, while "✗" means the approach relies on shared vocabulary. Our approach (dual transfer) has two variants, with or without position embeddings in the transfer parameters.

▶ Transformer

parallel data size ($\times 10^3$)	0	1	5	10	50	100	500	1000
dual transfer (word)	0.43	9.06	11.74	12.97	17.44	18.84	22.10	23.72
+freezing parent NMT encoder	**6.20**	8.82	11.58	12.76	16.62	18.50	21.69	23.59

Table 4: BLEU on et→en. Freezing the parent NMT encoder helps our approach to perform zero-shot translation.

approach	tr→en	en→et	en→tr	fr→es
no transfer	15.44	16.29	9.63	10.59
BERT2BERT	19.73	17.36	11.78	18.26
dual transfer (word)	**21.12**	**19.41**	**13.18**	**22.28**
dual transfer (word+position)	20.29	18.79	13.16	-

Table 5: BLEU on translation directions shown in columns, grouped by shared target transfer, shared source transfer, and general transfer. "-" means the experiment was not carried out.

Context-Aware Automatic Text Simplification of Health Materials in Low-Resource Domain

ACL, 2020

H5 =169

Introduction

理解网络上的健康信息所需的能力超过了一般患者的健康素养

老年人通常是和卫生保健系统打交道最多但健康素养最低的人群

低健康素养会导致差的健康结果

Dr.Babel Fish

第一步：通过词的**使用频率**判断该词是否难以理解

第二步：为需要简化的**复杂词**生成对应的**一系列简单词**，并赋予一个**替换概率**

第三步：为对应的**简单词**生成**语言模型**概率，选择替换词

Dr.Babel Fish

"**巴别鱼**，很可能是宇宙中最奇异的事物。它靠接收脑电波的能量为生，并且不是从其携带者身上接收，而是从周围的人身上。它从这些脑电波能量中吸收所有未被人察觉的精神频率，转化成营养。然后它向携带者的思想中排泄一种由被察觉到的精神频率和大脑语言中枢提供的神经信号混合而成的心灵感应矩阵。所有这些过程的实际效果就是，**如果你把一条巴别鱼塞进耳朵，你就能立刻理解以任何形式的语言对你说的任何事情。** 你所听到的解码信号就是巴别鱼向你的思想提供的脑电波矩阵。"

——《银河系漫游指南》

Step 1: 识别难以理解的词

Ct = 面向一般用户的健康相关文本

统计Ct中词的词频，认为词频越高，词越容易被理解

设置一个阈值ft，使用频率低于ft的被认为是难以被理解的词语

Step 2: 产生候选的替换词

UMLS: 为一个医学词提供一系列同义词，可以输入词进行检索，得到概念,认为检索结果中靠前的有更大的概率是正确的概念

为一个概念产生i个候选概念，候选概念和$1/r^i$成正比

Step 3: 使用语言模型选择简单词

在面向普通用户的文本集中训练出一个6-gram的语言模型，使用6-gram的语言模型计算候选词的概率

一个句子可能有多个词需要替换，每一个词都有一个替换概率和语言模型生成的概率，将这两个概率相乘，并找到使得整个句子概率最大的词语组合

MetaMap+CHV

使用MetaMap找到可对应到UMLS概念的短语，如果短语中包含难以理解的词，且该UMLS概念能被CHV中的词进行替换，则替换，否则不变

NMT:

LSTM+注意力

Pointer-Generator

选择医疗机构的health records，由专业医生进行简化

收集了五个面向普通用户的医学文本资源网站，用于统计词频和训练语言模型

使用多项选择题测试用户对简化前后的句子的理解程度

Results

Method	Supervision Type	PINC	SARI
No Change	N/A	0.00	32.83
MetaMap+CHV	Knowledge-Aware	25.84	45.64
DBF	Knowledge-Aware	19.61	55.33
Pointer-Generator	Direct Supervision	32.25	54.75
Seq2Seq-w-Att	Direct Supervision	50.81	**79.26**

Table 5: Performance of the simplification systems on all medication instructions

Method	Supervision Type	PINC	SARI
No Change	N/A	0.00	39.29
MetaMap+CHV	Knowledge-Aware	26.32	54.35
DBF	Knowledge-Aware	21.52	**56.51**
Pointer-Generator	Direct Supervision	36.34	40.01
Seq2Seq-w-Att	Direct Supervision	78.35	48.27

Table 6: Performance of the simplification systems on the free-text subset of the medication instructions

PINC measures how many n-grams differ between the two sentences

Results

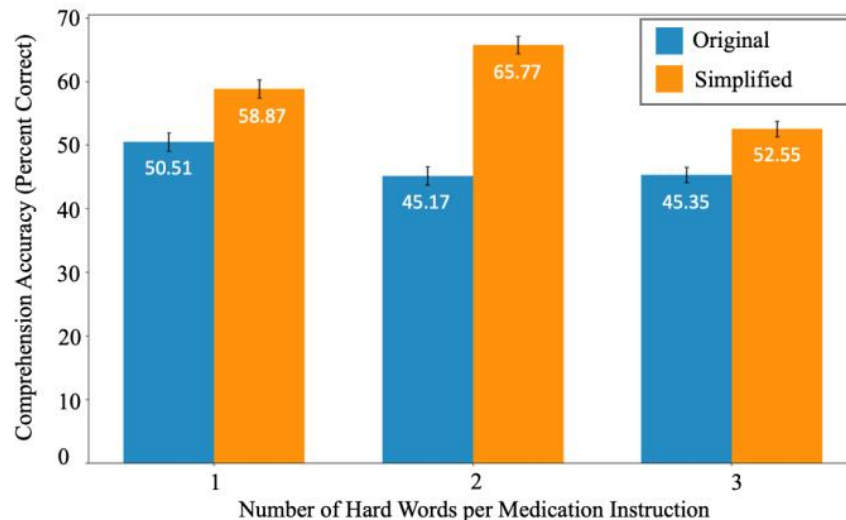Figure 3: Example questions from the online human subjects study

Figure 4: Impact of DBF on the different hardness levels of medication instructions