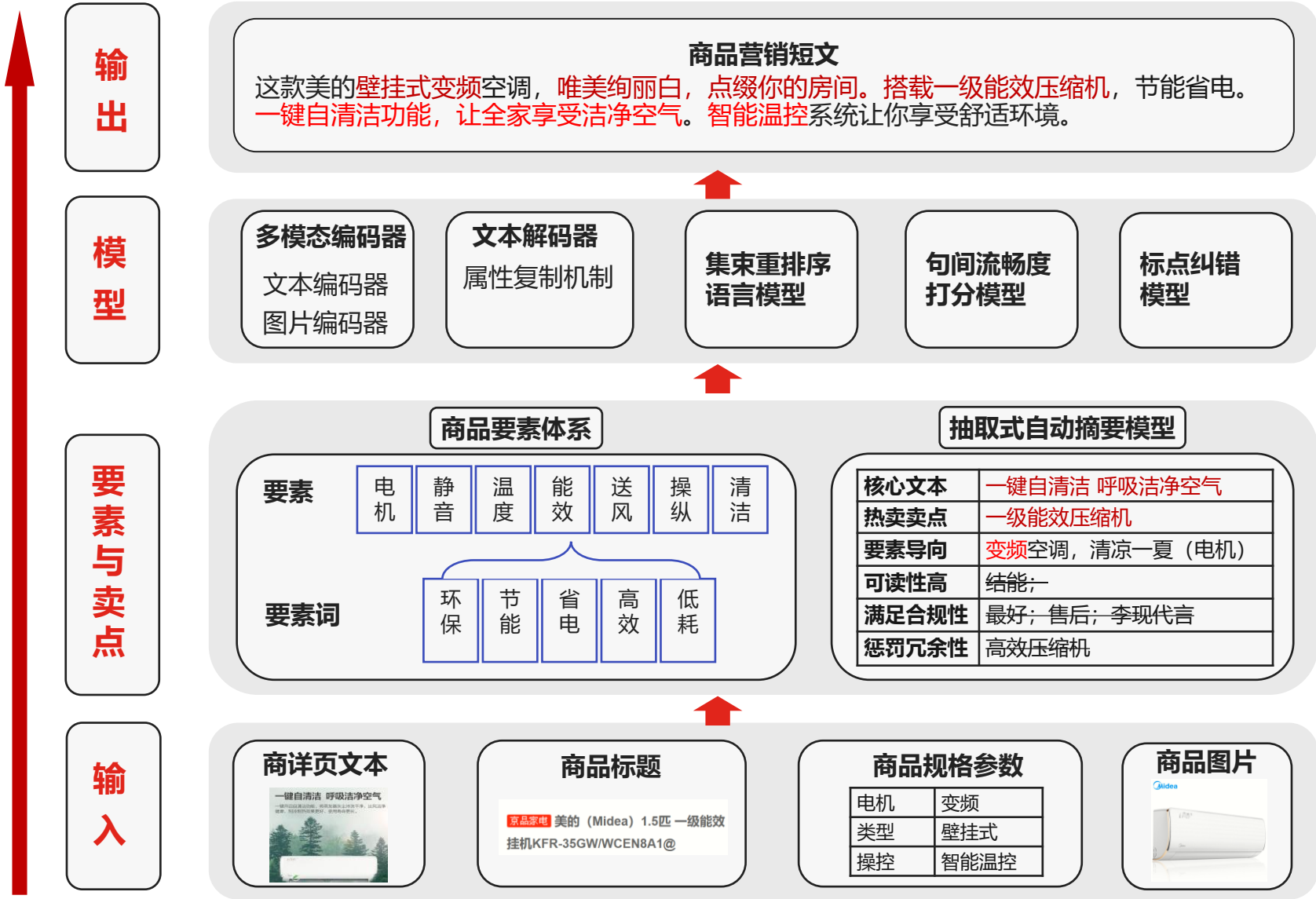


商品营销文案自动生成



◆ 京东发现好货频道



◆ 京东智能客服



◆ 京东搭配购



◆ 社交电商（京东电商广告联盟APP：京粉）



◆ 辅助写作

1. 输入商品ID

2. 生成营销短文

3. 卖点提取与分类

4. 同类卖点短句展示

5. 卖点一键复制与编辑

京东智联云
云 智 联 世 界

工作台

退出登录

62680022197

创作

卖点短文

可爱的维尼熊图案，吸引宝宝的注意力，让孩子爱上吃饭。采用PP材质，安全无异味，使用更放心。硅胶的指环，柔软舒适，不易脱落。人体工程学的设计，方便夹取食物。

引用

外观-可爱

可爱的维尼熊图案，吸引宝宝的注意力，让孩子爱上吃饭。

引用

更多...

外观-可爱

叉勺采用了可爱的印花图案，能够吸引宝贝的注意力，从而让他爱上吃饭，妈妈更省心。

引用

可爱的3D卡通造型，让孩子爱不释手，爱上吃饭，养成良好的饮食习惯。

引用

可爱的凯蒂猫造型，吸引宝宝的注意力，让孩子爱上吃饭。

引用

可爱的小黄鸭造型，充满童趣，宝宝更喜欢。

一键复制

↩ ↪

🖼️ 📏

Ω 😊 😄

🔗 🔗 🔗 🔗

B *I* U ~~S~~ ^{x₂} ^{x³} A A *I*_x

格式 字体 大小

A Neural Attention Model for Sentence Summarization

Alexander M. Rush
Facebook AI Research /
Harvard SEAS
srush@seas.harvard.edu

Sumit Chopra
Facebook AI Research
spchopra@fb.com

Jason Weston
Facebook AI Research
jase@fb.com

Abstract

Summarization based on text extraction is inherently limited, but generation-style abstractive methods have proven challenging to build. In this work, we propose a fully data-driven approach to abstractive sentence summarization. Our method utilizes a local attention-based model that generates each word of the summary conditioned on the input sentence. While the model is structurally simple, it can easily be trained end-to-end and scales to a large amount of training data. The model shows significant performance gains on the DUC-2004 shared task compared with several strong baselines.

1 Introduction

Summarization is an important challenge of natural language understanding. The aim is to produce a condensed representation of an input text that

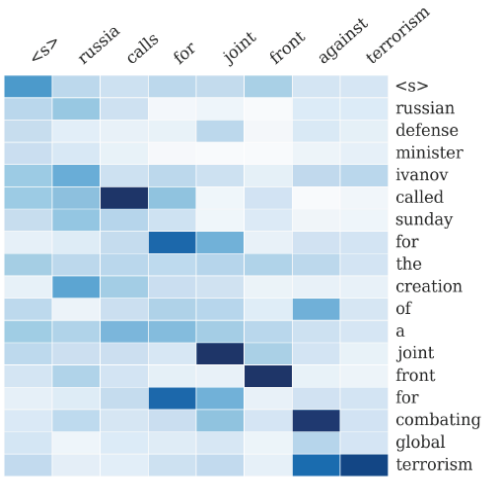
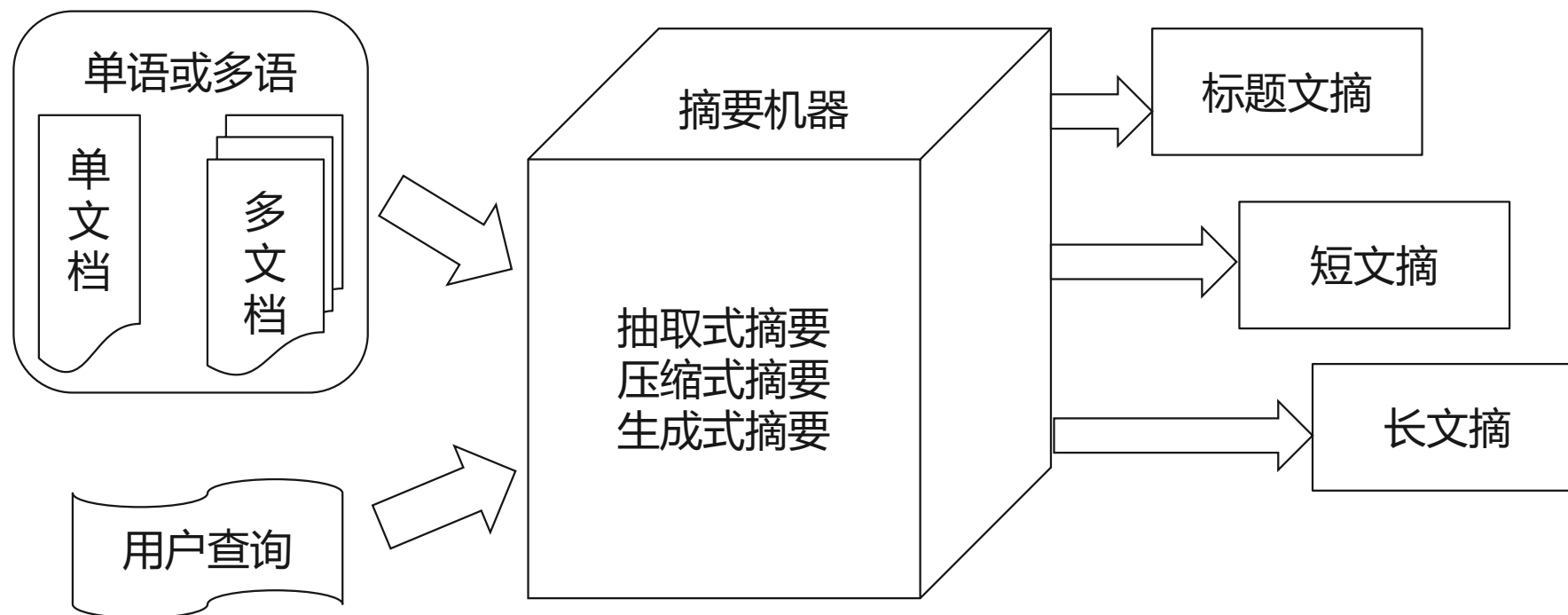


Figure 1: Example output of the attention-based summarization (ABS) system. The heatmap represents a soft alignment between the input (right) and the generated summary (top). The columns represent the distribution over the input after generating each word.

Lapata, 2008; Woodsend et al., 2010). These approaches are described in more detail in Section 6.



◆ 抽取式摘要

- ◆ 直接从原文中抽取已有的句子组成摘要
- ◆ 简单易实现

◆ 压缩式摘要

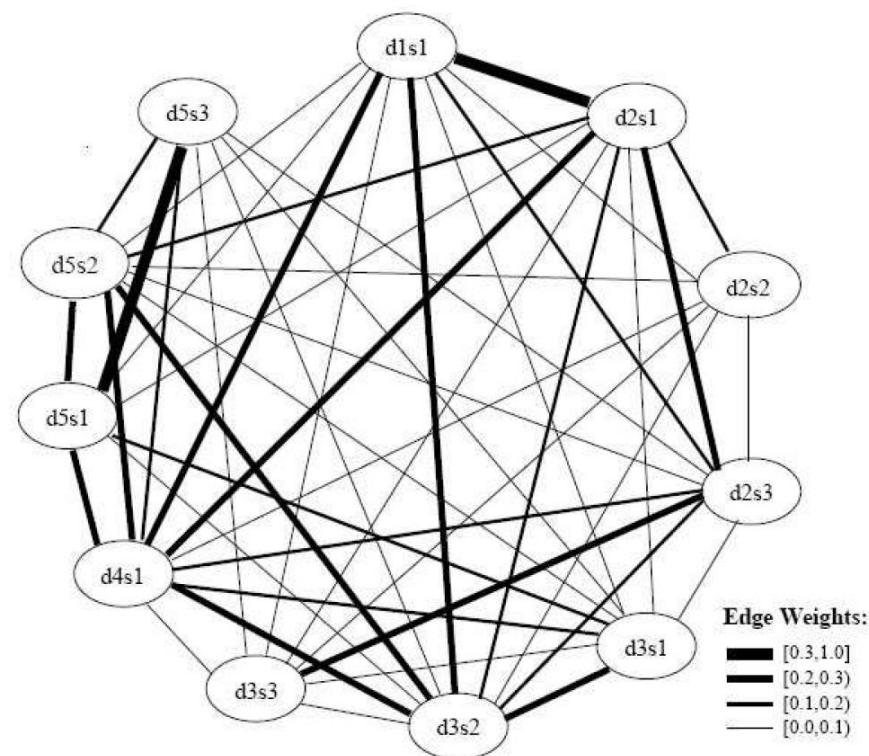
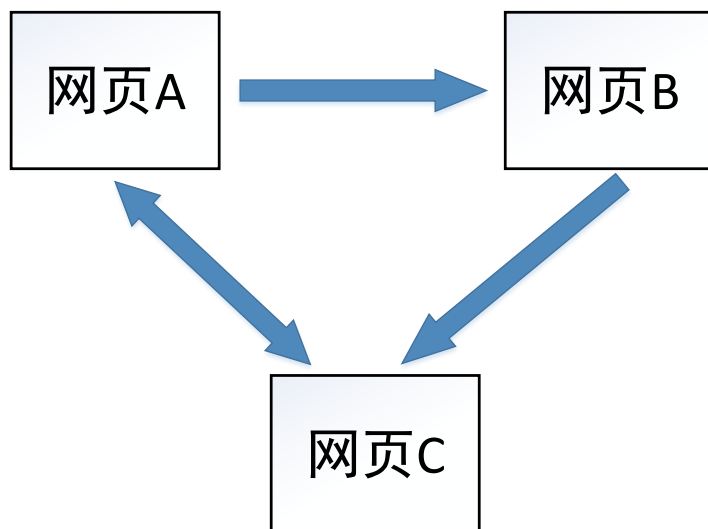
- ◆ 抽取并简化原文中的重要句子构成文摘
 - ◆ 长征五号B运载火箭首次飞行任务的圆满成功 → 长征五号B火箭首飞成功
- ◆ 可读性难以保证

◆ 生成式摘要

- ◆ 改写或重新组织原文内容形成最终文摘
- ◆ 可控性差

◆ 抽取式摘要

- ◆ 基于词频的方法
- ◆ 基于句子聚类的方法
- ◆ 基于图模型的方法
 - ◆ PageRank
 - ◆ TextRank

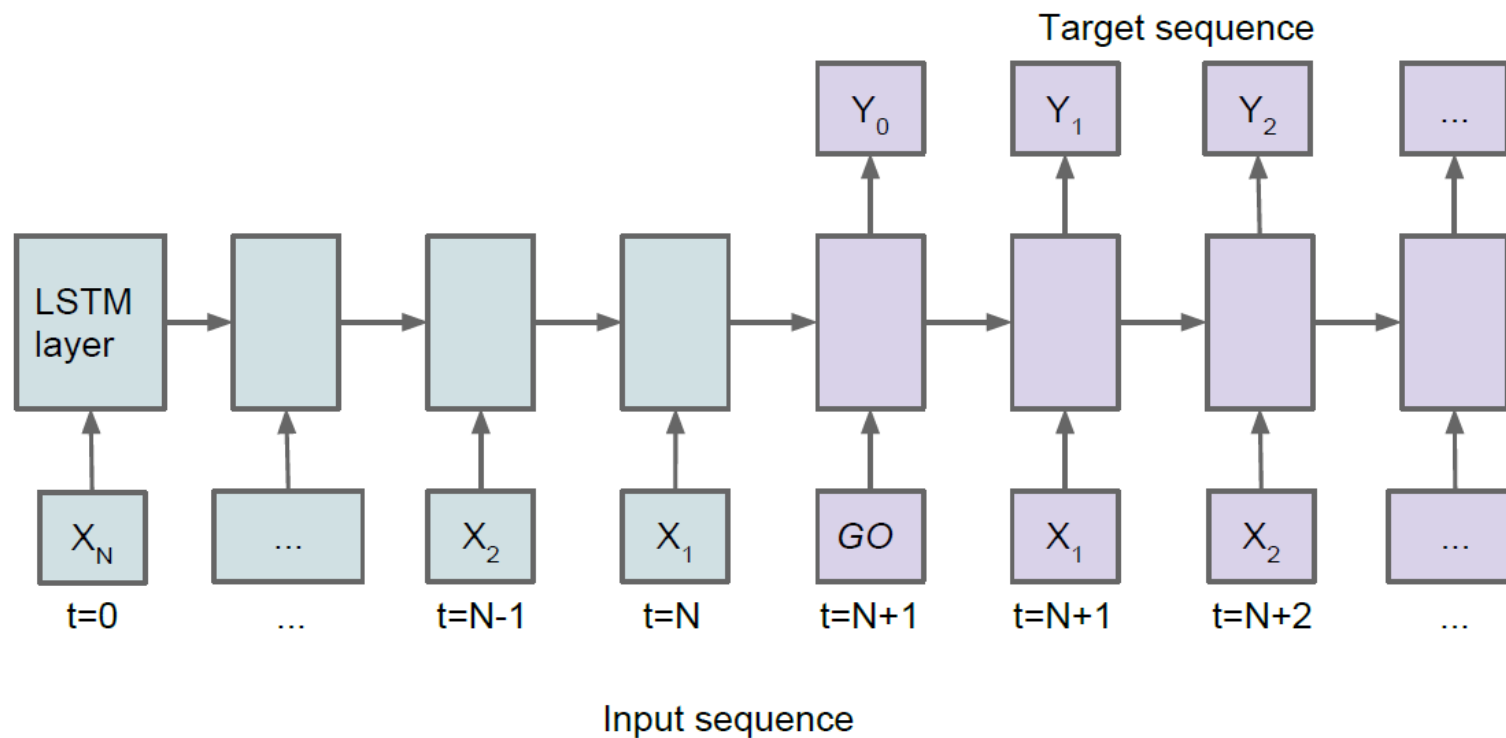


$$\begin{cases} s_1 = a_{12}s_2 + a_{13}s_3 \\ s_2 = a_{21}s_1 + a_{23}s_3 \\ s_3 = a_{31}s_1 + a_{32}s_2 \end{cases}$$

TextRank: Bringing order into text. Rada Mihalcea and Paul Tarau. EMNLP 2004.

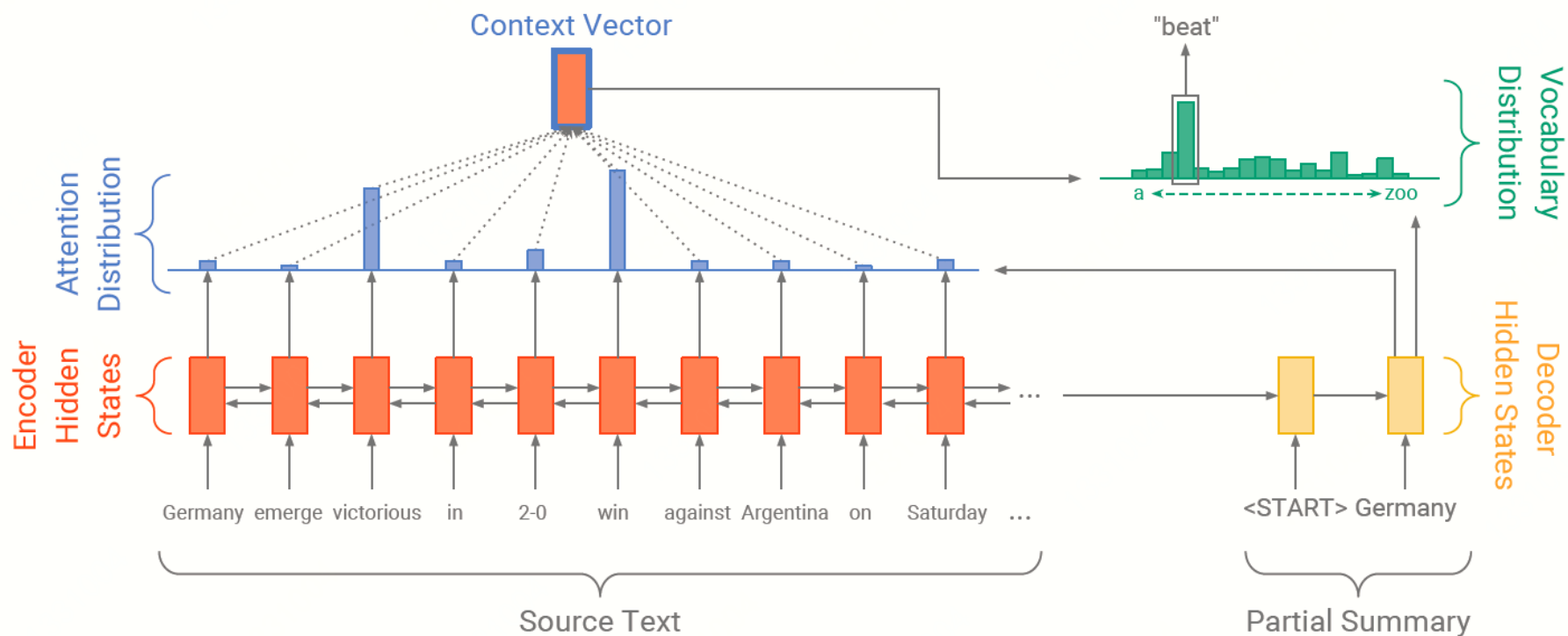
◆ 生成式摘要

◆ “编码器-解码器” 框架



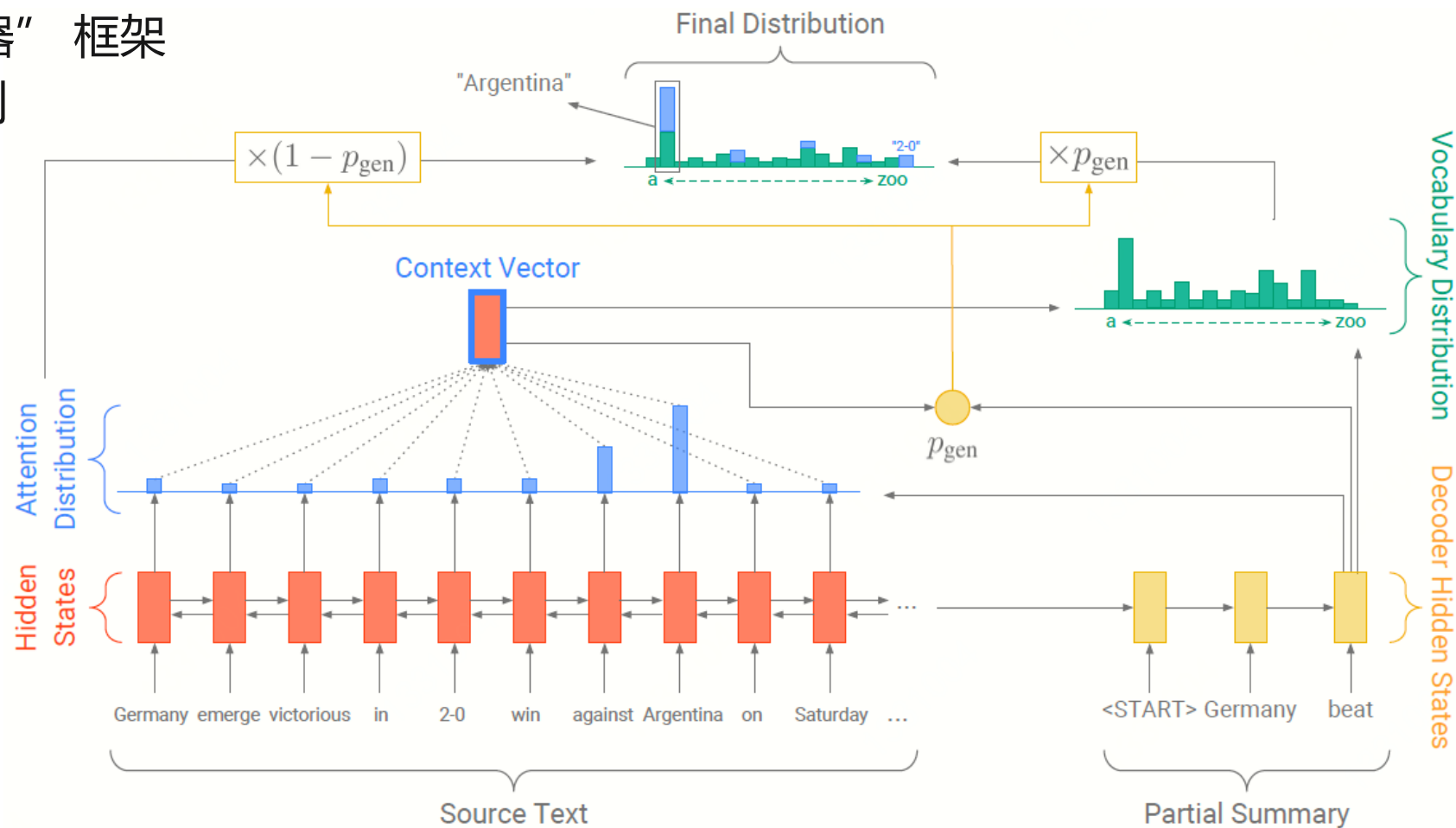
◆ 生成式摘要

- ◆ “编码器-解码器” 框架
- ◆ “注意力” 机制



◆ 生成式摘要

- ◆ “编码器-解码器” 框架
- ◆ “注意力” 机制
- ◆ “复制” 机制



■ Get To The Point: Summarization with Pointer-Generator Networks. Abigail See, Peter J. Liu, Christopher D. Manning. ACL 2017.

◆ 生成式摘要

- ◆ “编码器-解码器” 框架
- ◆ “注意力” 机制
- ◆ “复制” 机制
 - ◆ 长征五号B运载火箭首次飞行任务的圆满成功 → 长征五号B火箭首飞成功
 - ◆ $P(\text{火箭}) = \lambda * P_{gen}(\text{火箭} \in \text{词典}) + (1 - \lambda) * P_{copy}(\text{火箭})$

◆ 生成式摘要

- ◆ “编码器-解码器” 框架
- ◆ “注意力” 机制
- ◆ “复制” 机制
- ◆ “覆盖度” 机制
 - ◆ 惩罚重复的注意力

◆ 基本要求：像人写的摘要

- ◆ 重要
- ◆ 流畅
- ◆ 凝练
- ◆ 正确

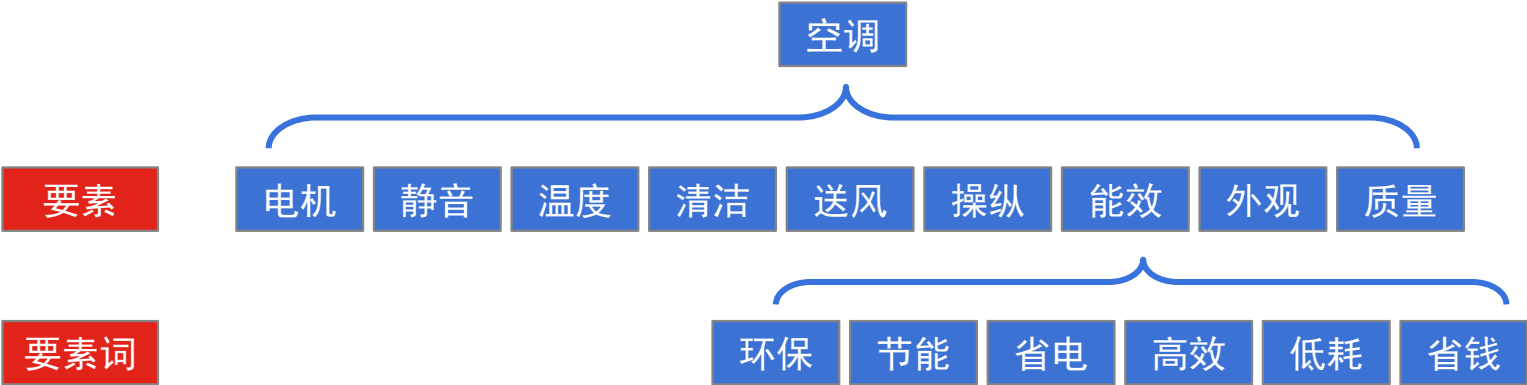
◆ 进阶：人们满意的摘要

- ◆ 写人们感兴趣的内容
- ◆ 模拟人写摘要的方式
- ◆ 探索模型的可解释性

◆ 任务定义

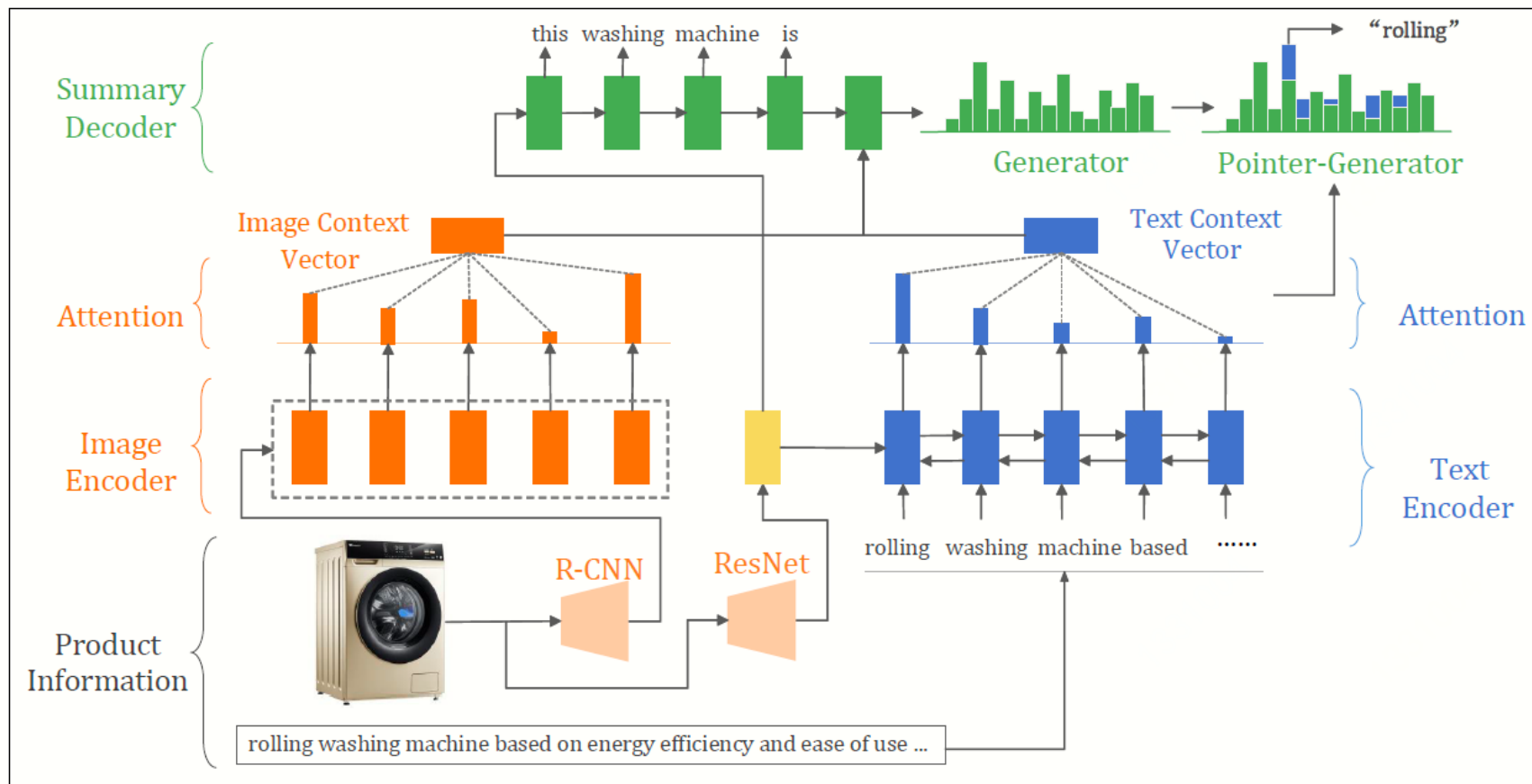


◆ 商品要素定义




- ◆ 商品摘要重在用户感兴趣
 - ◆ 感性的人：
典型：颜值控，外观与众不同
 - ◆ 理性的人：
典型：技术宅，技术领先、功能实用

◆ 多模态摘要满足感性用户



◆ 多模态摘要满足感性用户

	<p>PGNet: 时尚的两件套设计，穿出不同种类的风格，选用优质的针织面料，手感细腻，具有良好的亲肤效果，穿着舒适不紧绷，配以甜美的喇叭袖，丰富了整体的视觉效果。</p> <p>(Fashionable two-piece design wearing in different styles. It uses high-quality knitted fabric, feeling delicate, skin friendly, comfortable but not tight. Sweet trumpet sleeve enrich the overall visual effect.)</p> <p>PGNet+EncInit+MMAtt: 一款来自欧芮儿很时尚的套装，以橙色为主色调的条纹设计，靓丽吸睛，让你穿出不一样的时髦造型。丰富视觉效果，凸显层次美感，木耳边的喇叭袖，增添潮流看点。</p> <p>(A very fashionable suit from Orel. The stripe design with orange as the main color, which is beautiful and attractive, making you in different fashionable shapes. Rich visual effects, highlight the level of aesthetic feeling, and the horn sleeve beside the wood ear, add trend points.)</p>
---	--

- ◆ 基于商品要素的摘要满足理性用户
 - ◆ 重要 → 重要的商品要素
 - ◆ 流畅 → 流畅的商品要素
 - ◆ 凝练 → 凝练的商品要素
 - ◆ 正确 → 正确的商品要素

◆ 基于商品要素的摘要满足理性用户

- ◆ 重要 → 重要的商品要素
 - ◆ 奖励函数：商品要素准确率
- ◆ 流畅 → 流畅的商品要素
- ◆ 凝练 → 凝练的商品要素
- ◆ 正确 → 正确的商品要素

Ground-Truth Summary:

This washer has the *volume* of 10 kilogram. Remote *operation* system facilitates using, especially for elderly people. The advanced brushless *motor* makes it more reliable.



Ground-Truth Aspects:

Capacity, *Control*, *Motor*



Rewarded Aspects:

Capacity, *Control*, *Motor*

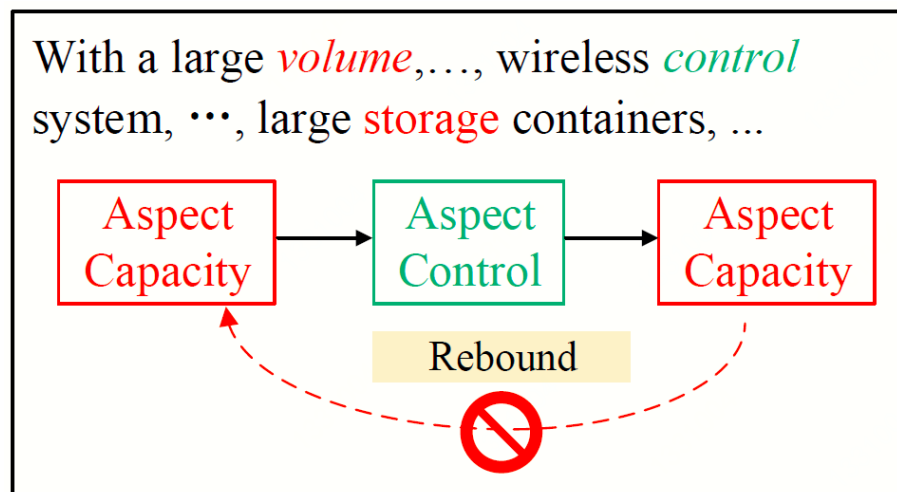


Generated Summary:

With a large *volume*,..., wireless *control* system, ..., it is driven by a stable *motor*,...

◆ 基于商品要素的摘要满足理性用户

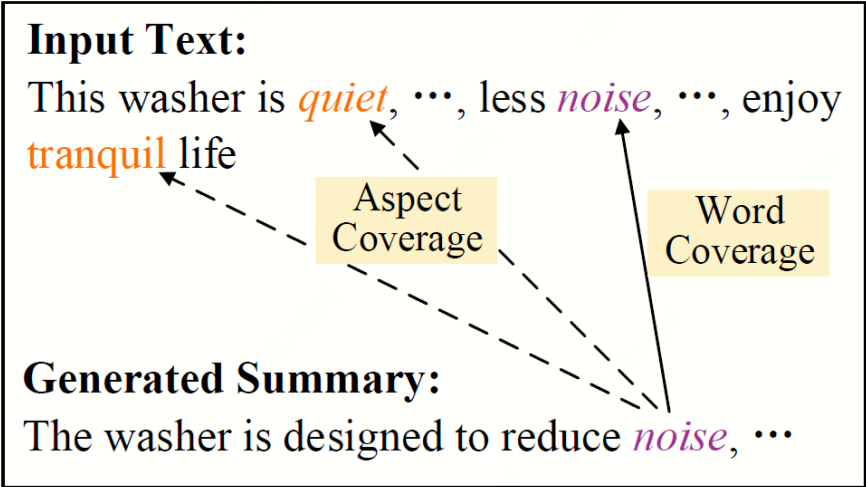
- ◆ 重要 → 重要的商品要素
- ◆ 流畅 → 流畅的商品要素
 - ◆ 商品要素防回跳机制
- ◆ 凝练 → 凝练的商品要素
- ◆ 正确 → 正确的商品要素



$P(w) = 0, w$ is related to Capacity

◆ 基于商品要素的摘要满足理性用户

- ◆ 重要 → 重要的商品要素
- ◆ 流畅 → 流畅的商品要素
- ◆ 凝练 → 凝练的商品要素
 - ◆ 商品要素覆盖度机制
- ◆ 正确 → 正确的商品要素



$$\begin{aligned} c_{t,i} &= \sum_{\tau=1}^{t-1} \alpha_{\tau,i} \\ e_{t,i} &= u_c^T \tanh(W_c s_{t-1} + V_c h_i + w_c c_{t,i}) \\ \mathcal{L}^w &= \sum_{t=1}^T \sum_{i=1}^N \min(\alpha_{t,i}, c_{t,i}) \\ \mathcal{L} &= \mathcal{L} + \mathcal{L}^w \end{aligned}$$


词级别的覆盖度机制



$$\begin{aligned} \alpha_{t,a_j} &= \sum_{x_k \in a_j} \alpha_{t,k} \\ c_{t,a_j} &= \sum_{\tau=1}^{t-1} \alpha_{\tau,a_j} \end{aligned}$$

要素级别的覆盖度机制

◆ 基于商品要素的摘要满足理性用户

- ◆ 重要 → 重要的商品要素
- ◆ 流畅 → 流畅的商品要素
- ◆ 凝练 → 凝练的商品要素
- ◆ 正确 → 正确的商品要素（至关重要！）
 - ◆ “属性词仅允许复制” 机制
 - ◆ $P(\text{定频}) = \lambda * P_{copy}(\text{定频} \in \text{输入}) + (1 - \lambda) P_{gen}(\text{定频} \in \text{词典})$ 
 - ◆ $P(\text{定频}) = P_{copy}(\text{定频} \in \text{输入}) = 0$



美的 (Midea) 1.5匹 变频 智弧 冷暖 智能壁挂式卧室空调挂机 KFR-35GW/WDAA3@

【五一狂欢Zui后2天！7日价格上涨，后患无期！】3期免息！高密度银离子滤网呵护全家健康，还空调干净的“肺”！ECO节能！一键强劲制冷！[查看一级高配爆款](#)

京 东 价 **¥1799.00** 降价通知

累计评价
87万+

选购指数
6.2

优 惠 券 **满2000减30**

促 销 **换购** 购买1件可优惠换购热销商品 立即换购 >>

◆ 实验效果

	Home App.		Clothing		Cases&Bags	
	R-1	R-2	R-1	R-2	R-1	R-2
Seq2seq	21.57	7.18	23.05	6.84	23.18	6.94
PG	31.31	10.93	29.11	9.24	31.11	10.27
MMPG+EncInit (E)	32.67	11.54	30.14	9.85	32.05	11.36
MMPG+DecInit (D)	32.88	11.88	30.73	10.29	32.69	11.78
MMPG+MMAtt (M)	32.76	11.67	30.67	10.13	32.59	11.58
MMPG+D+M	32.87	11.88	30.72	10.31	32.66	11.74
MMPG+E+D+M	32.89	11.89	30.71	10.30	32.68	11.77

	Home Appliances			Clothing			Cases&Bags		
	R-1	R-2	R-L	R-1	R-2	R-L	R-1	R-2	R-L
Lead	21.97	9.54	12.79	19.83	8.39	13.56	21.49	9.37	14.19
LexRank	24.06	10.01	18.19	26.87	9.01	17.76	27.09	9.87	18.03
Seq2seq	21.57	7.18	17.61	23.05	6.84	16.82	23.18	6.94	17.29
MASS	28.19	8.02	18.73	26.73	8.03	17.72	27.19	9.03	18.17
PG	31.31	10.93	21.11	29.11	9.24	19.92	31.11	10.27	21.79
MMPG	32.88	11.88	21.96	30.73	10.29	21.25	32.69	11.78	22.27
MMPG+Aspect	33.97	12.43	22.21	31.81	10.87	21.32	33.67	12.44	22.31
MMPG+Aspect+Consistency	34.36	12.52	22.35	31.93	11.09	21.54	33.78	12.51	22.43

- ◆ 人写摘要的方式
 - ◆ 一半摘抄，一半创作

Observation:

Input sentence: France and Germany called on world leaders Monday to take rapid action to press for the closure of Ukraine 's Chernobyl nuclear plant , site of the world 's worst ever nuclear disaster .

Reference summary: World leaders urged to back Chernobyl closure plan .

Solution:

Step1. Extracting keywords: world leaders closure Chernobyl

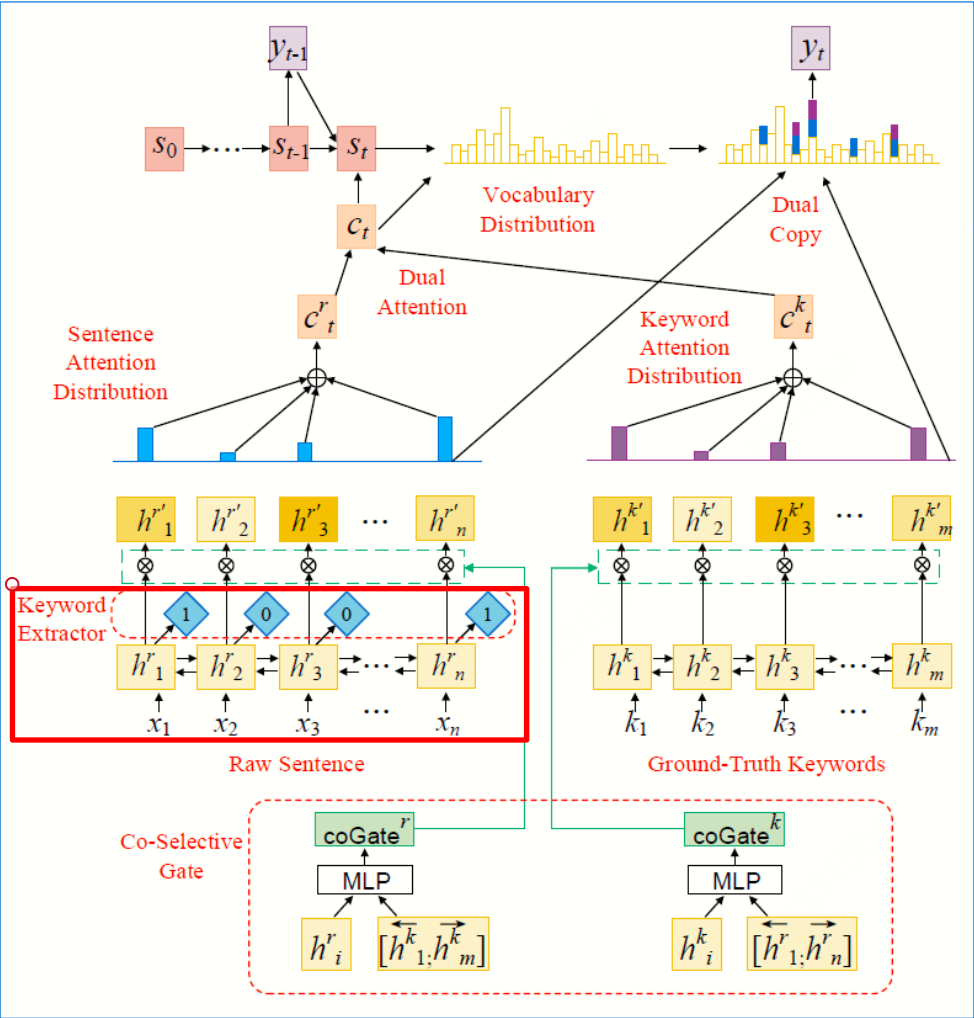
Step2. Generating summary guided by the keywords: World leaders called for action on Chernobyl closure .

◆ 模拟人写摘要的方式

- ◆ 步骤一：抽取ground-truth卖点词/短语
- ◆ 步骤二：卖点识别模型+基于ground-truth卖点的自动文摘模型（多任务学习）
- ◆ 步骤三：基于预测的卖点的自动文摘模型

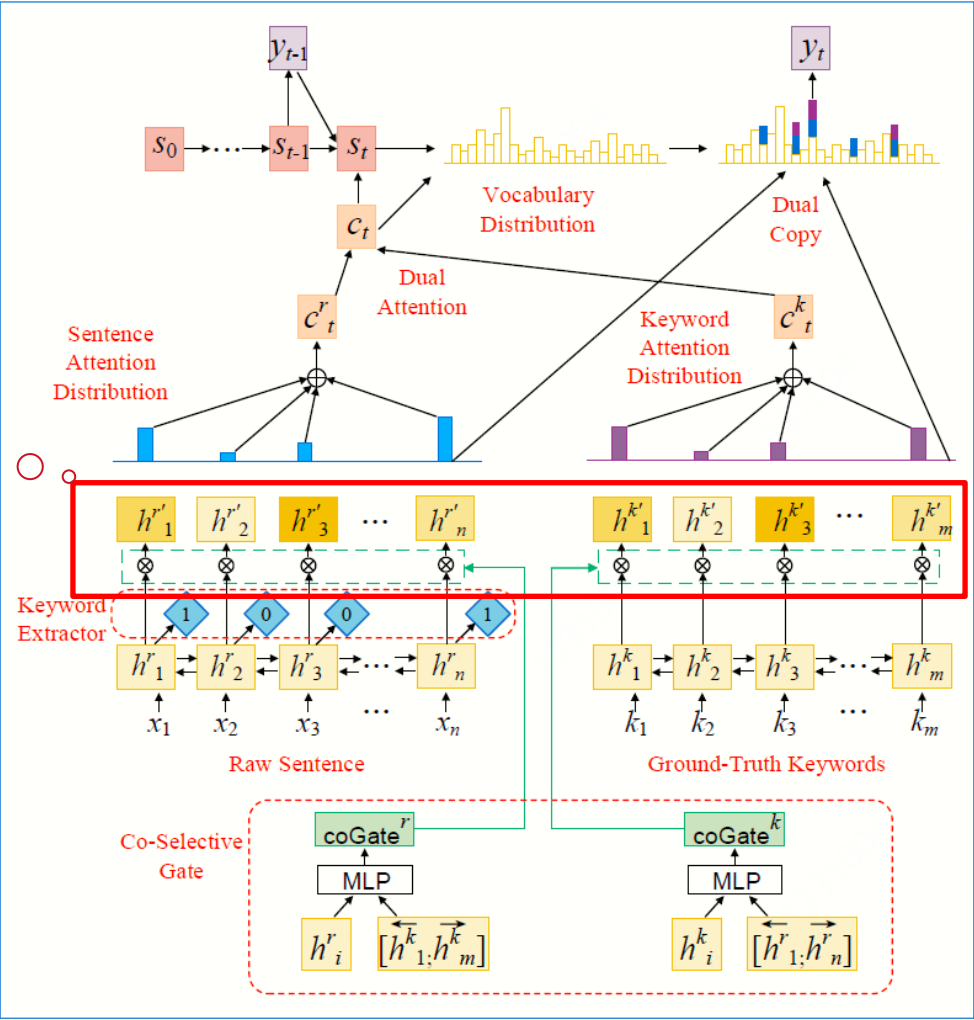
模型细节

卖点识别
模块



◆ 模型细节

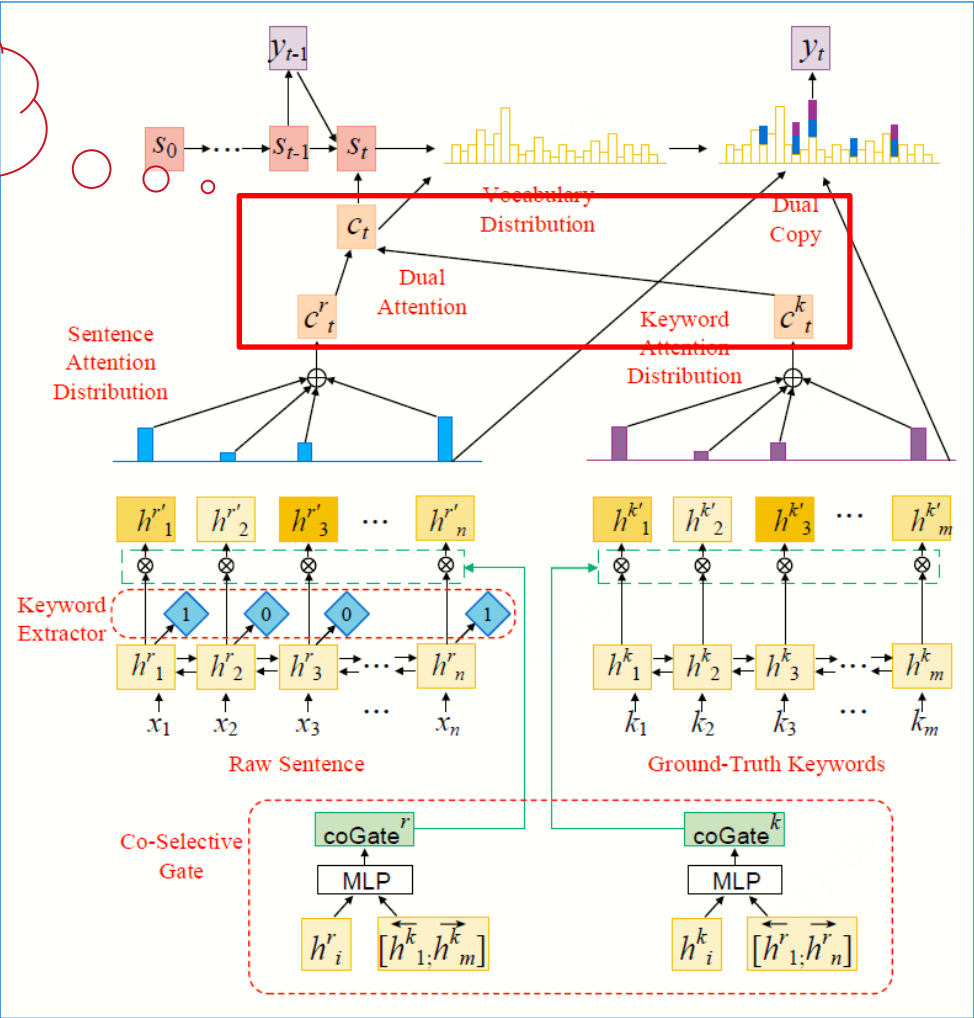
卖点选择性
编码模块



$$coGate_i^r = \sigma(\mathbf{W}_p h_i^r + \mathbf{U}_p a^k)$$
$$coGate_i^k = \sigma(\mathbf{W}_q h_i^k + \mathbf{U}_q a^r)$$
$$h_i^{r'} = h_i^r \odot coGate_i^r$$
$$h_i^{k'} = h_i^k \odot coGate_i^k$$

◆ 模型细节

双重注意力模块



I. 直连

$$c_t = [c_t^r; c_t^k]$$

II. 门控机制融合

$$g_t = \sigma(\mathbf{W}_g c_t^r + \mathbf{U}_g c_t^k)$$

$$c_t = g_t \odot c_t^r + (1 - g_t) \odot c_t^k$$

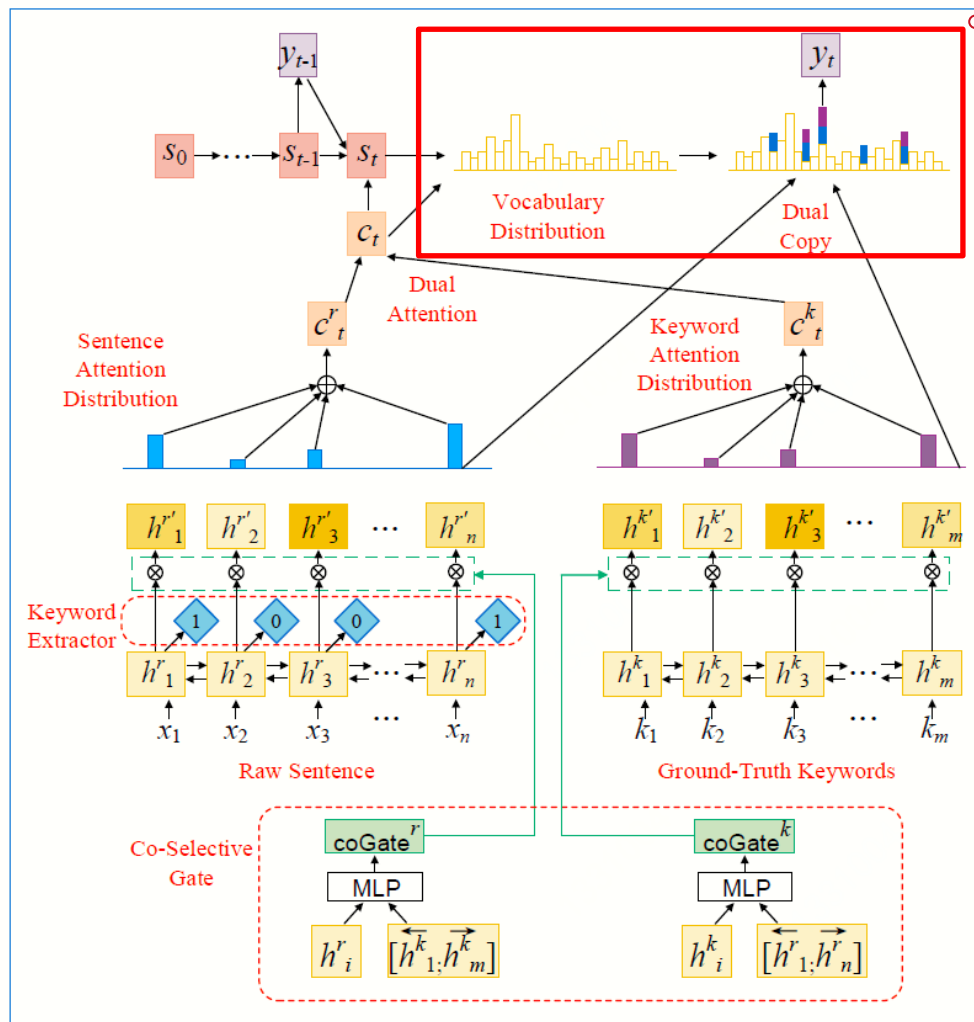
III. 层次化融合

$$\beta_t^r = \sigma(\mathbf{U}_\beta s_{t-1} + \mathbf{W}_\beta c_t^r)$$

$$\beta_t^k = \sigma(\mathbf{U}_\beta s_{t-1} + \mathbf{W}_\beta c_t^k)$$

$$c_t = \beta_t^r c_t^r + \beta_t^k c_t^k$$

◆ 模型细节



双重复制
模块

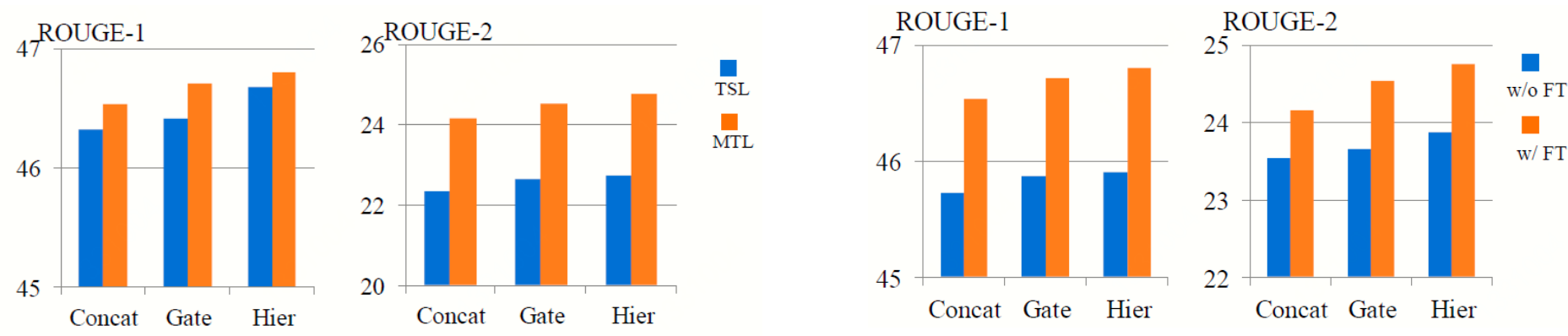
$$P_{copy_s}(w) = \sum_{i:x_i=w} \alpha_{t,i}^r$$

$$P_{copy_k}(w) = \sum_{i:k_i=w} \alpha_{t,i}^k$$

$$\lambda_t = \text{sigmoid}(w_a^T c_t + u_g^T s_t + v_g^T y_{t-1} + b_g)$$

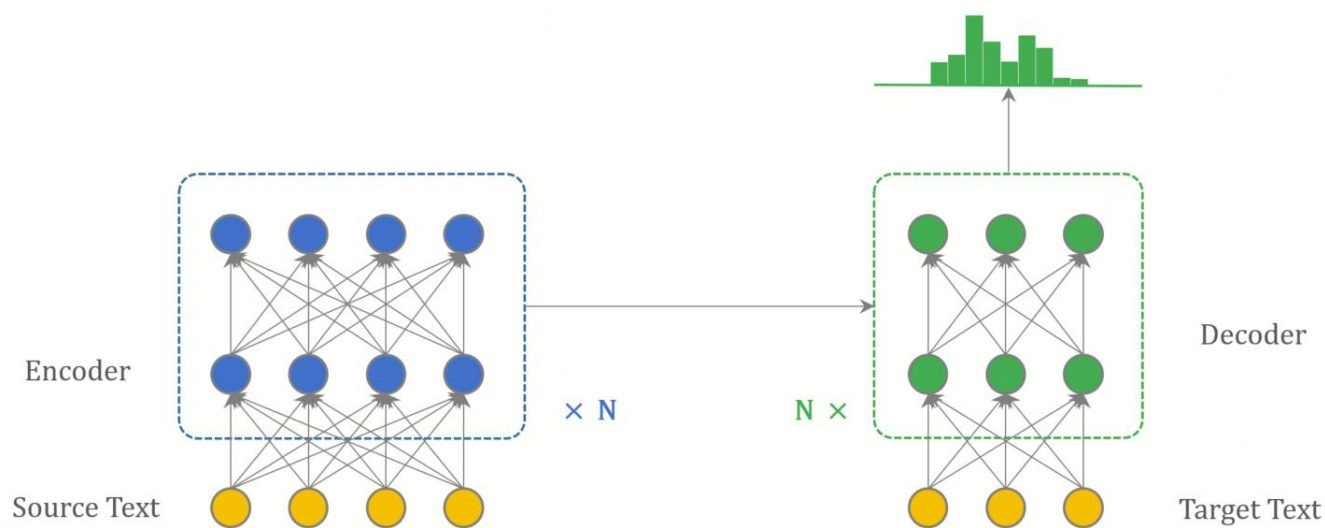
$$P(w) = \lambda_t P_{gen}(w) + \frac{1}{2}(1 - \lambda_t)(P_{copy_s}(w) + P_{copy_k}(w))$$

◆ 实验效果



Model	Attribute Keyword Coverage
CopyNet	41.73
DualCopyNet	44.53
DualCopyNet + Copy Loss	48.13
AttrCopyNet	49.31

◆ Transformer (self-attention)



我	爱	北京	天安门
$\begin{bmatrix} 0.1 \\ 0.2 \\ 0.4 \\ 0.3 \\ 0.2 \end{bmatrix}$	$\begin{bmatrix} 0.7 \\ 0.3 \\ 0.6 \\ 0.1 \\ 0.2 \end{bmatrix}$	$\begin{bmatrix} 0.4 \\ 0.5 \\ 0.2 \\ 0.6 \\ 0.1 \end{bmatrix}$	$\begin{bmatrix} 0.3 \\ 0.9 \\ 0.1 \\ 0.3 \\ 0.7 \end{bmatrix}$

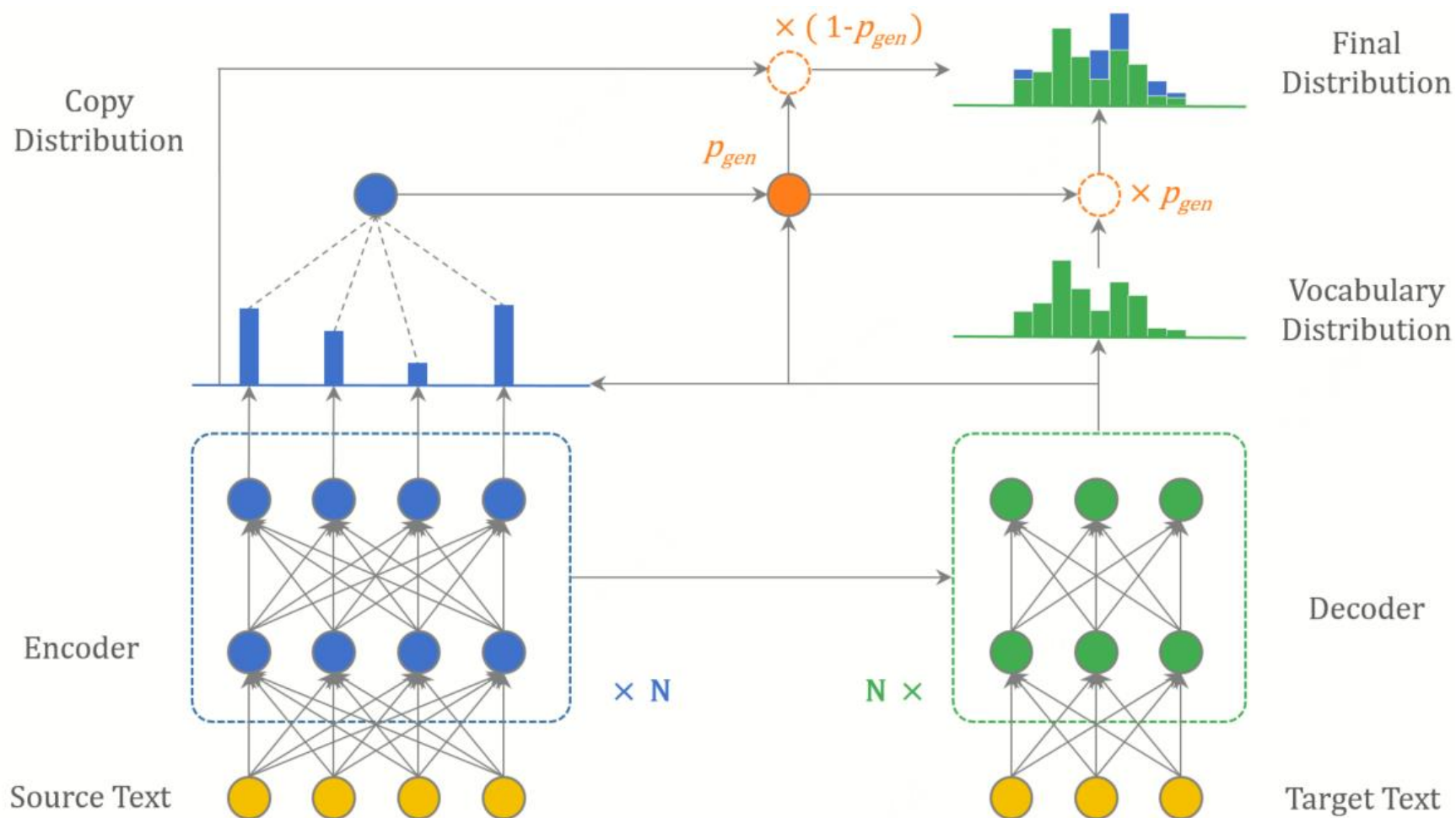
$q = [0.4, 0.5, 0.2, 0.6, 0.1]$ 北京

$k = v = \begin{bmatrix} 0.1, 0.2, 0.4, 0.3, 0.2 \\ 0.7, 0.3, 0.6, 0.1, 0.2 \\ 0.4, 0.5, 0.2, 0.6, 0.1 \\ 0.3, 0.9, 0.1, 0.3, 0.7 \end{bmatrix}$ 我
爱
北京
天安门

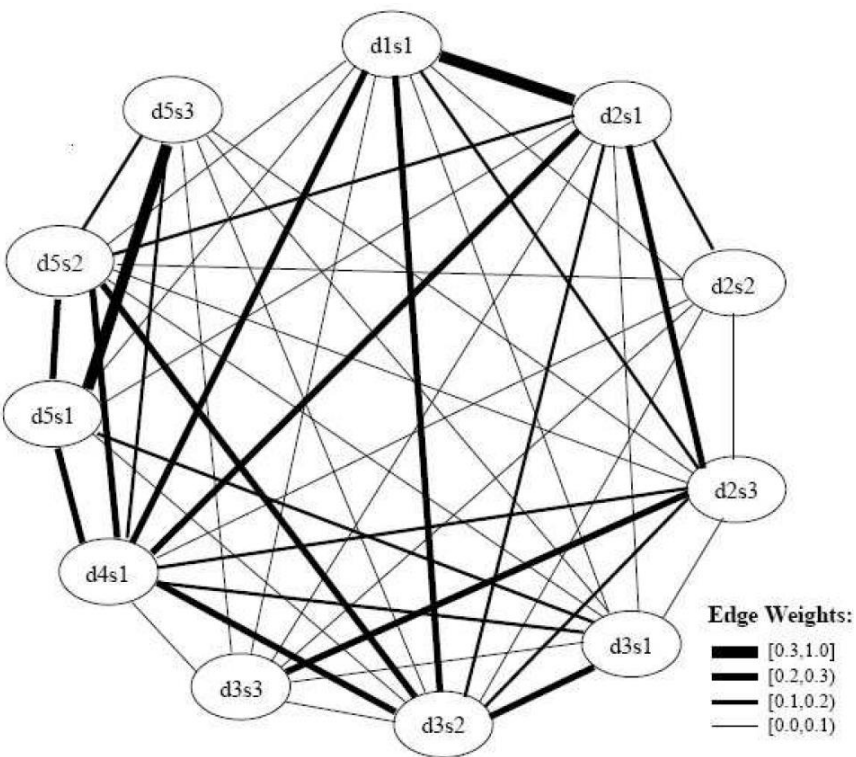
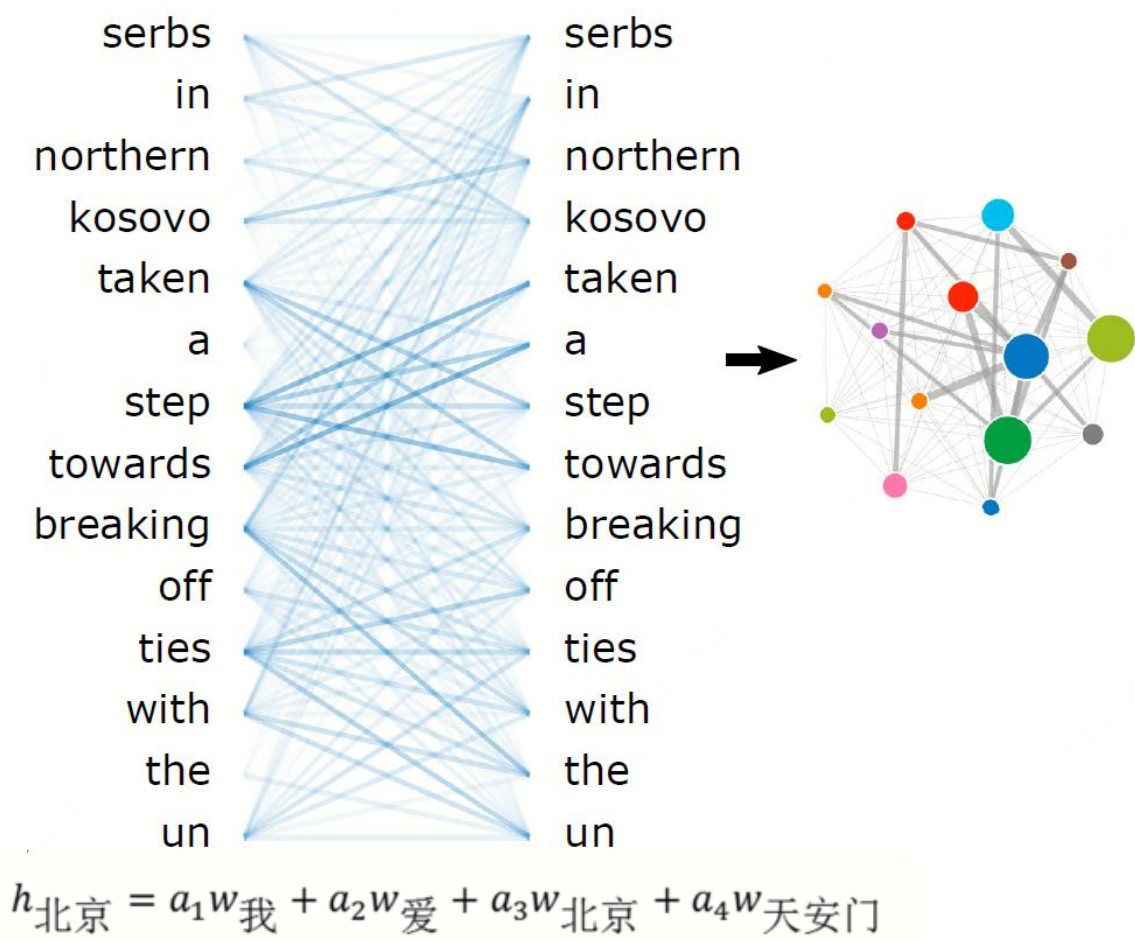
$$Attn(K, Q, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V$$

$$h_{\text{北京}} = a_1 w_{\text{我}} + a_2 w_{\text{爱}} + a_3 w_{\text{北京}} + a_4 w_{\text{天安门}}$$

◆ Transformer+复制机制

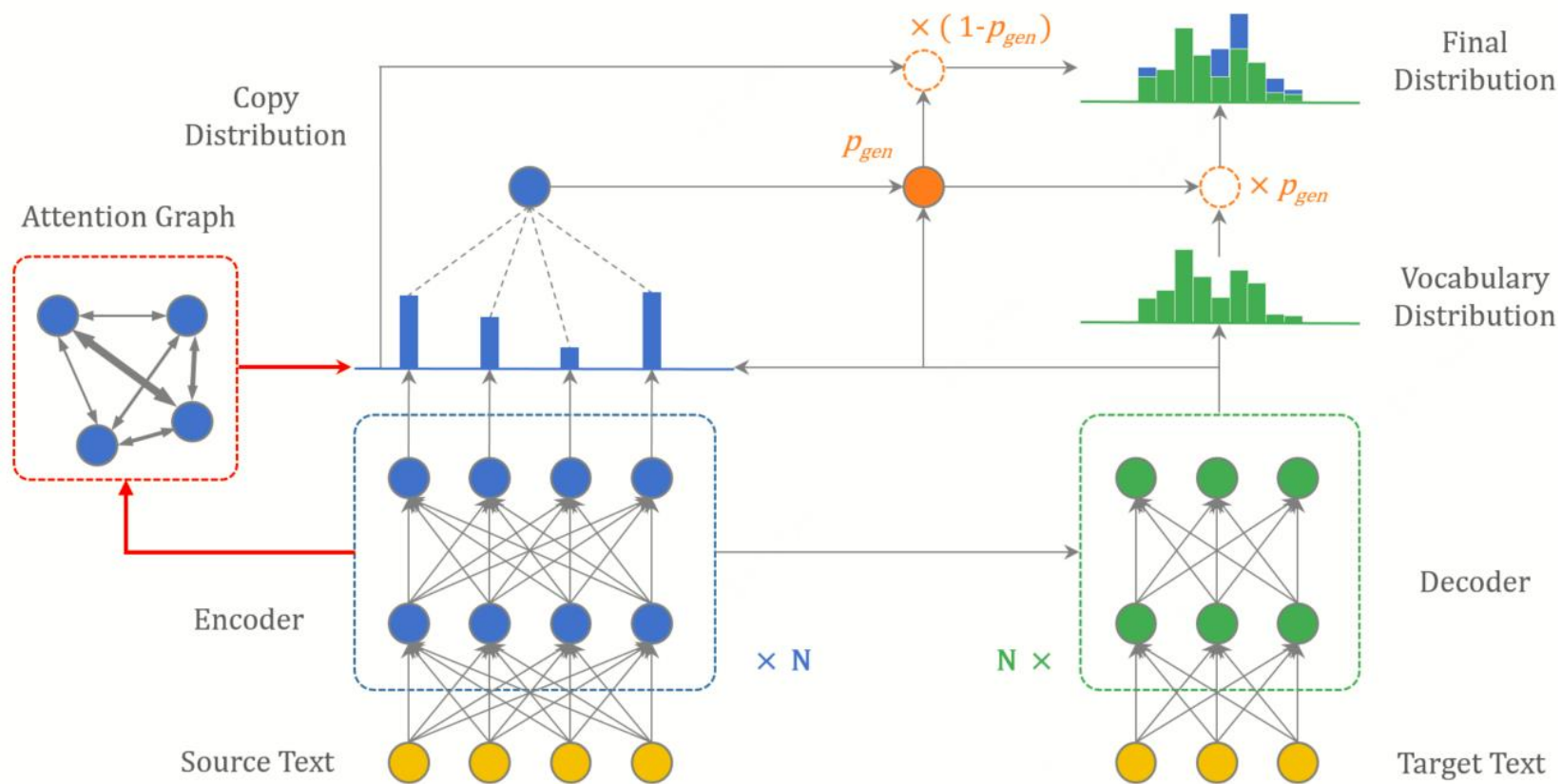


◆ 自注意力全连接图



$$\begin{cases} s_1 = a_{12}s_2 + a_{13}s_3 \\ s_2 = a_{21}s_1 + a_{23}s_3 \\ s_3 = a_{31}s_1 + a_{32}s_2 \end{cases}$$

◆ Transformer+受到自注意力指导的复制机制



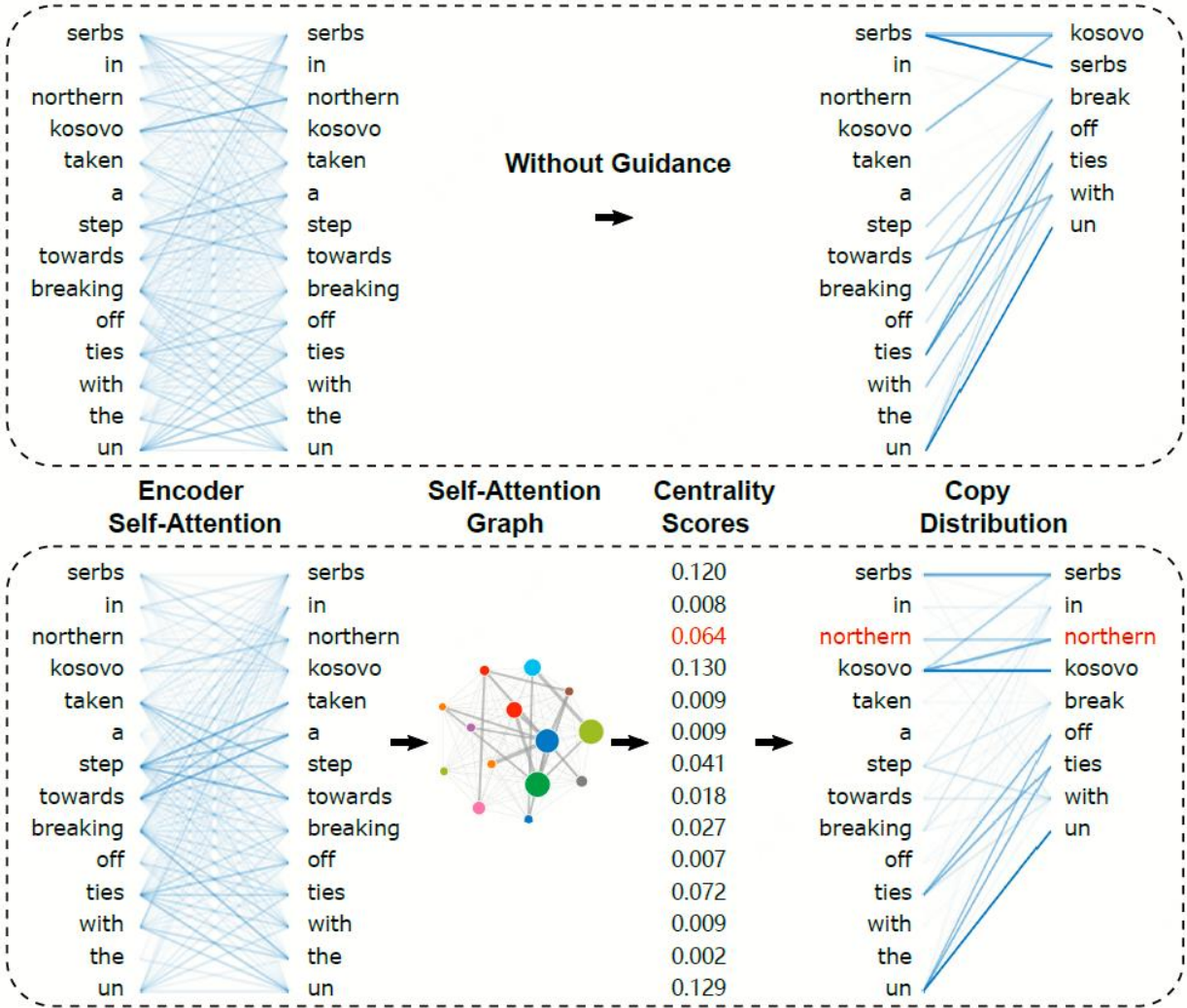
◆ 自注意力指导的复制机制

$$\alpha_{t,i} = \text{softmax}\left(\frac{(W_s s_t)^T W_h h_i}{\sqrt{d_k}}\right) \quad \rightarrow \quad \text{softmax}\left(\frac{(W_s s_t)^T (W_h h_i + w_p \text{score}_i)}{\sqrt{d_k}}\right)$$

α : 编码器-解码器间的注意力分布, 复制概率分布
score: 重要性得分

$$\mathcal{L} = -\frac{1}{T} \sum_t \log P(y_t) + \lambda \text{KL}\left(\frac{1}{T} \sum_t \alpha_t, \text{score}\right)$$

◆ 自注意力指导流程示例



◆ 实验结果

Lead-3*	40.34	17.70	36.57
PGNet*	39.53	17.28	36.38
Bottom-Up*	41.22	18.68	38.34
MASS	41.38	19.11	38.42
MASS+Copy	41.71	19.41	38.66
SAGCopy Outdegree	42.53	19.92	39.44
SAGCopy Indegree-1	42.30	19.75	39.23
SAGCopy Indegree-2	42.56	19.89	39.40
SAGCopy Indegree-3	42.34	19.72	39.29

Table 2: ROUGE F_1 scores on the CNN/Daily Mail

◆ 实验结果

Models	RG-1	RG-2	RG-L
ABS*	29.55	11.32	26.42
ABS+*	29.76	11.88	26.96
SEASS*	36.15	17.54	33.63
SeqCopyNet*	35.93	17.51	33.35
MASS*	38.73	19.71	35.96
MASS+Copy	38.53	19.93	35.86
SAGCopy Outdegree	38.86	19.91	36.06
SAGCopy Indegree-1	38.84	20.39	36.27
SAGCopy Indegree-2	38.70	20.16	36.09
SAGCopy Indegree-3	38.69	19.83	35.98

Table 3: Experimental result on the Gigaword dataset.

谢谢