

Get To The Point: Summarization with Pointer-Generator Networks

ADVISOR: JIA-LING, KOH

SOURCE: ACL-2017

SPEAKER: WEI, LAI

DATE: 2020/4/6

OUTLINE

Introduction

Method

Experiment

Conclusion

OUTLINE

Introduction

Method

Experiment

Conclusion

Introduction

lagos, nigeria (cnn) a day after winning nigeria's presidency, *muhammadu buhari* told cnn's christiane amanpour that he plans to aggressively fight corruption that has long plagued nigeria and go after the root of the nation's unrest. *buhari* said he'll "rapidly give attention" to curbing violence in the northeast part of nigeria, where the terrorist group boko haram operates. by cooperating with neighboring nations chad, cameroon and niger, he said his administration is confident it will be able to thwart criminals and others contributing to nigeria's instability. for the first time in nigeria's history, the opposition defeated the ruling party in democratic elections. *buhari* defeated incumbent goodluck jonathan by about 2 million votes, according to nigeria's independent national electoral commission. the win comes after a long history of military rule, coups and botched attempts at democracy in africa's most populous nation.

Long Text



Input
400 tokens/text



Summary



Output
100 tokens/text (train)
120 tokens/text (test)

muhammadu buhari says he plans to aggressively fight corruption that has long plagued nigeria. he says his administration is confident it will be able to thwart criminals. the win comes after a long history of military rule, coups and botched attempts at democracy in africa's most populous nation.

Introduction

Extractive

Aim to select salient snippets, sentences or passages from documents.

Abstractive

Aim to concisely paraphrase the information content in the documents.

Introduction

1. Inaccurately reproducing factual details
2. Inability to deal with out-of-vocabulary (OOV) words

Baseline Seq2Seq + Attention: UNK UNK says his administration is confident it will be able to **destabilize nigeria's economy**. UNK says his administration is confident it will be able to thwart criminals and other **nigerians**. **he says the country has long nigeria and nigeria's economy.**

3. Repeating themselves

Pointer-Gen: *muhammadu buhari* says he plans to aggressively fight corruption **in the northeast part of nigeria**. he says he'll "rapidly give attention" to curbing violence **in the northeast part of nigeria**. he says his administration is confident it will be able to thwart criminals.

OUTLINE

Introduction

Method

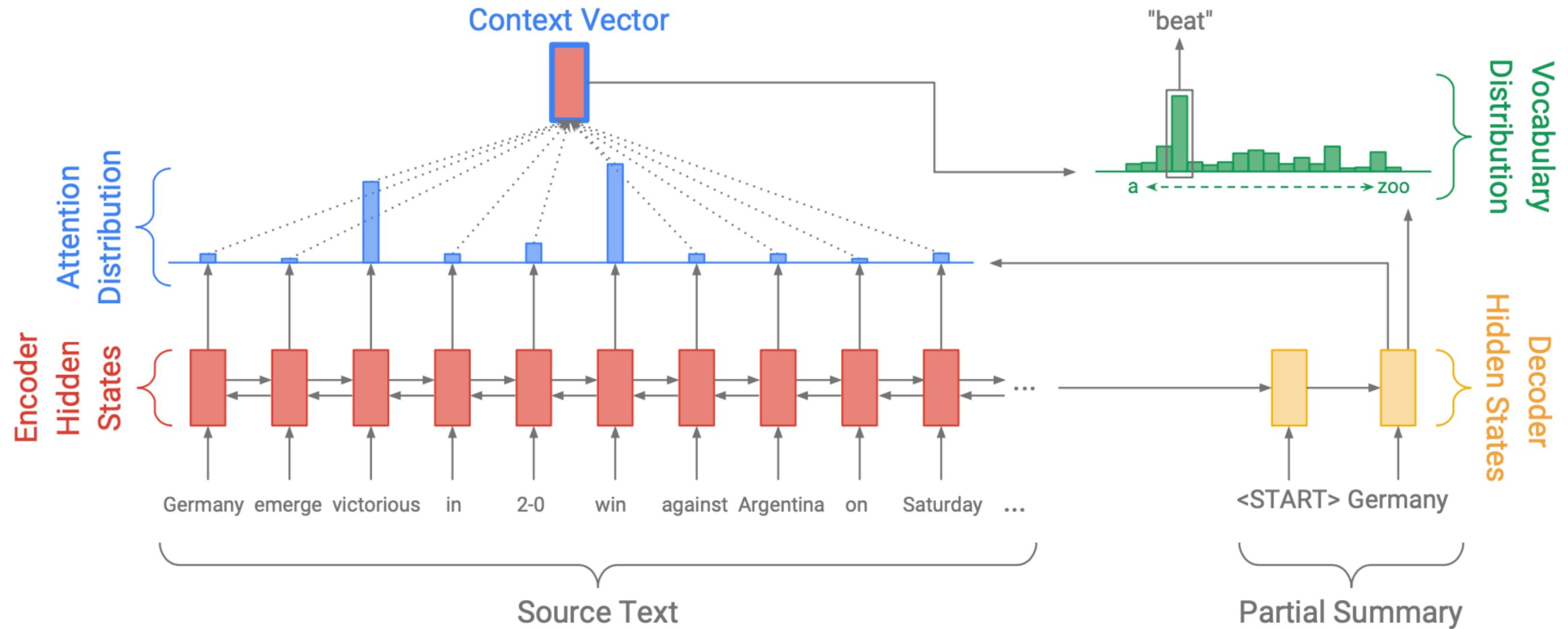
Experiment

Conclusion

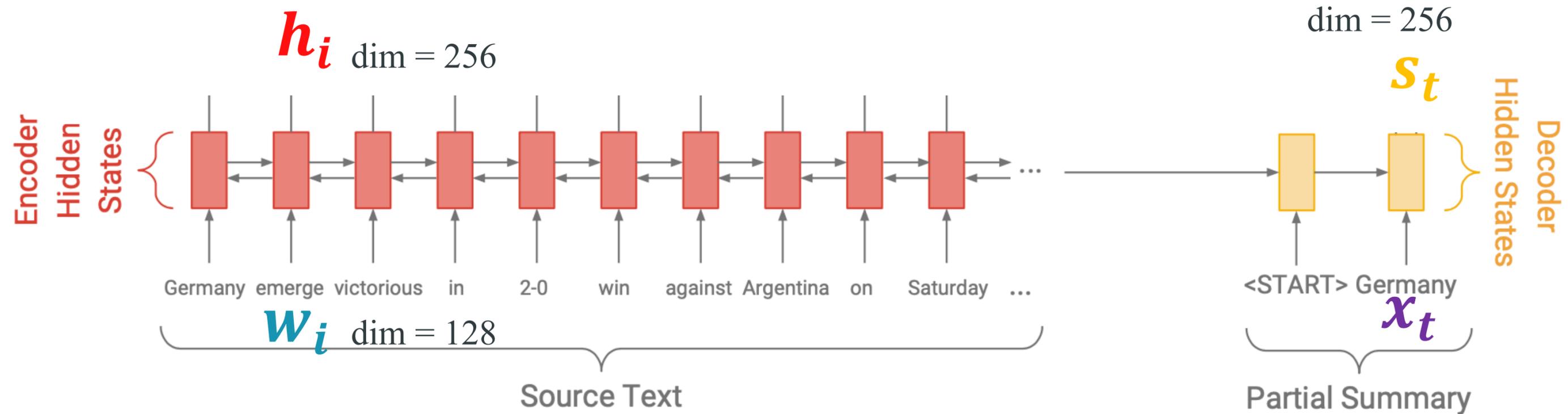
Baseline

Sequence-to-sequence attentional model

Sequence-to-sequence attentional model



Encoder/Decoder Hidden State



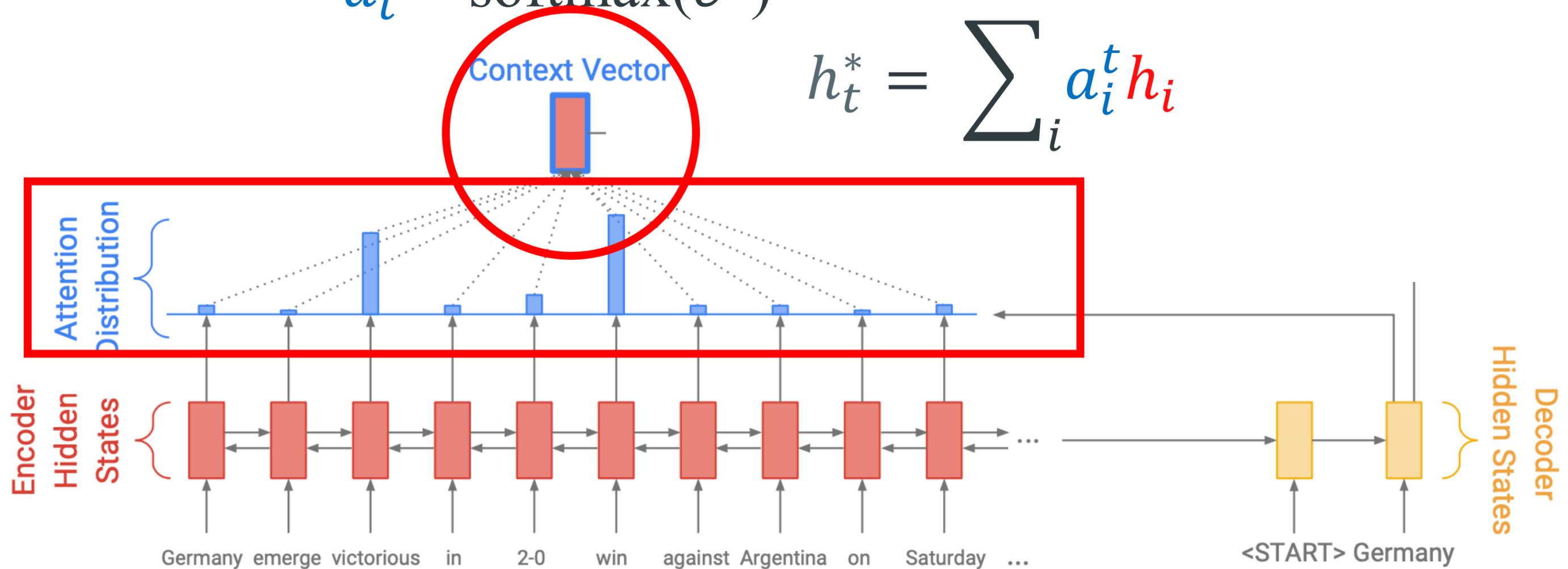
Attention distribution

$$e_i^t = v^T \tanh(W_h h_i + W_s s_t + b_{attn})$$

dim = 256 dim = 256

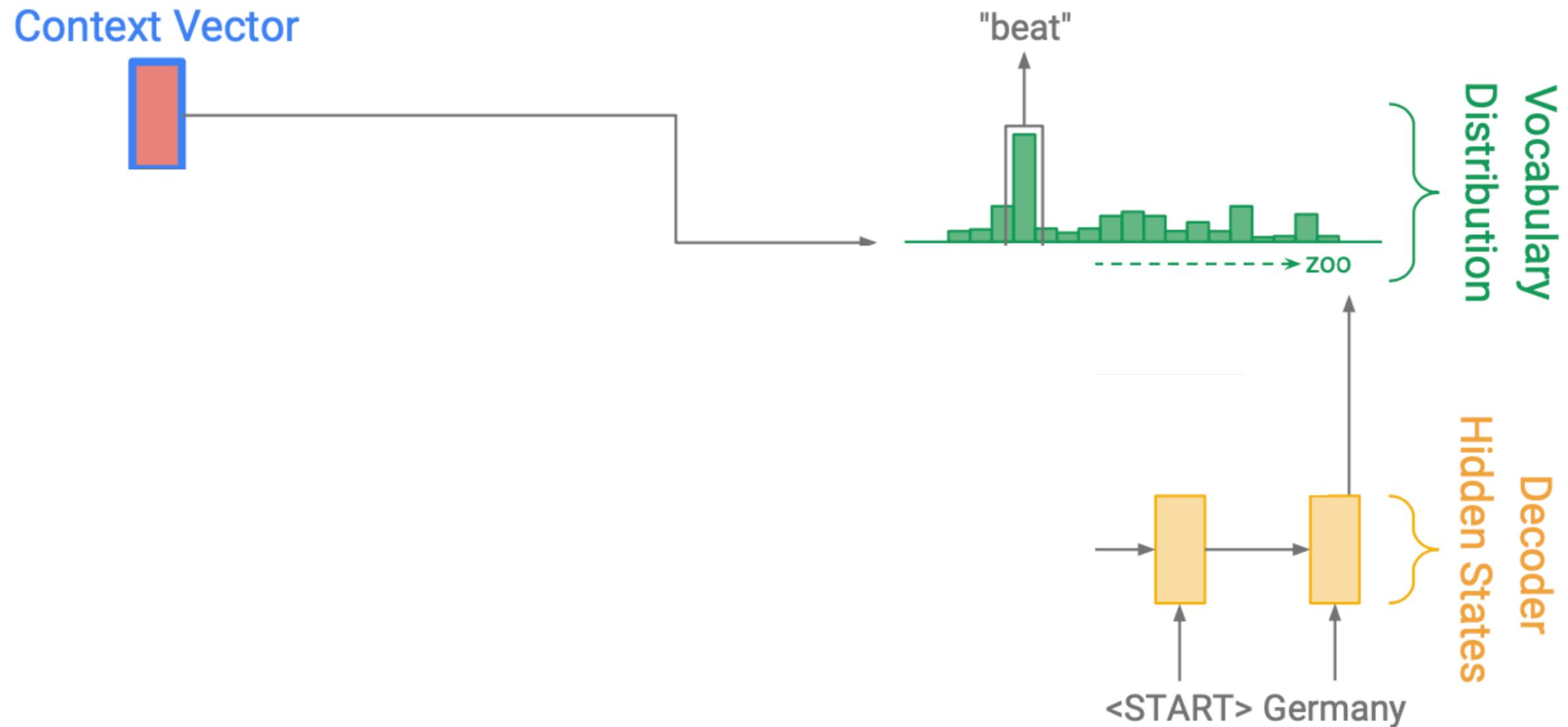
$$a_t = \text{softmax}(e^t)$$

$$h_t^* = \sum_i a_i^t h_i$$



Vocabulary Distribution

$$P_{vocab} = \text{softmax}(V'(V[s_t, h_t^*] + b) + b')$$



Loss Function

$$P(w) = P_{vocab}(w)$$

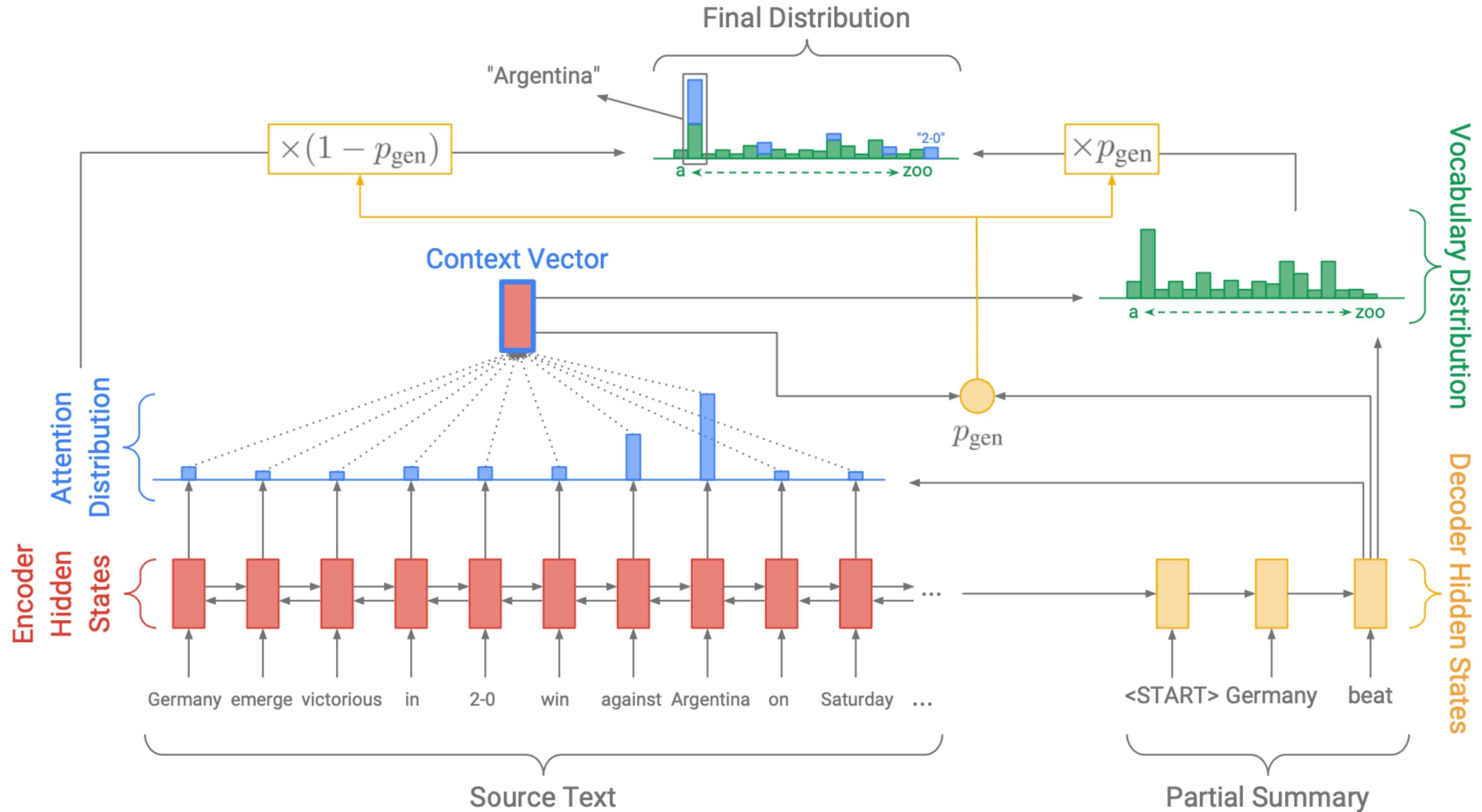
target word

$$loss_t = -\log P(w_t^*)$$

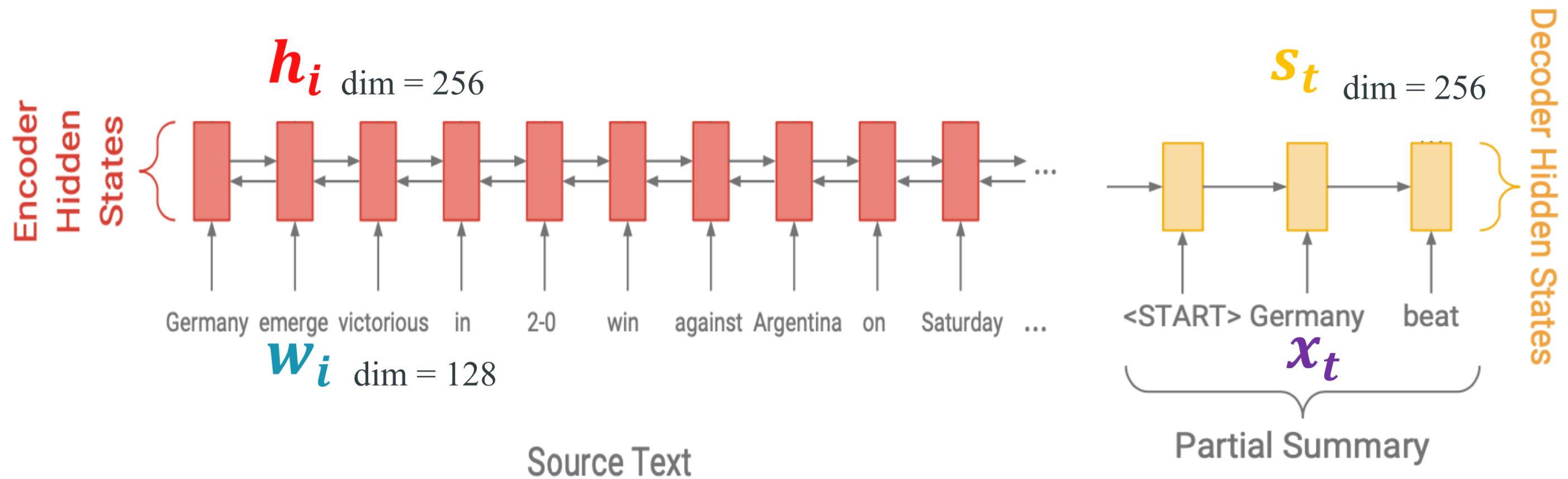
$$loss = \frac{1}{T} \sum_{t=0}^T loss_t$$

Pointer-generator network

Pointer-generator Model



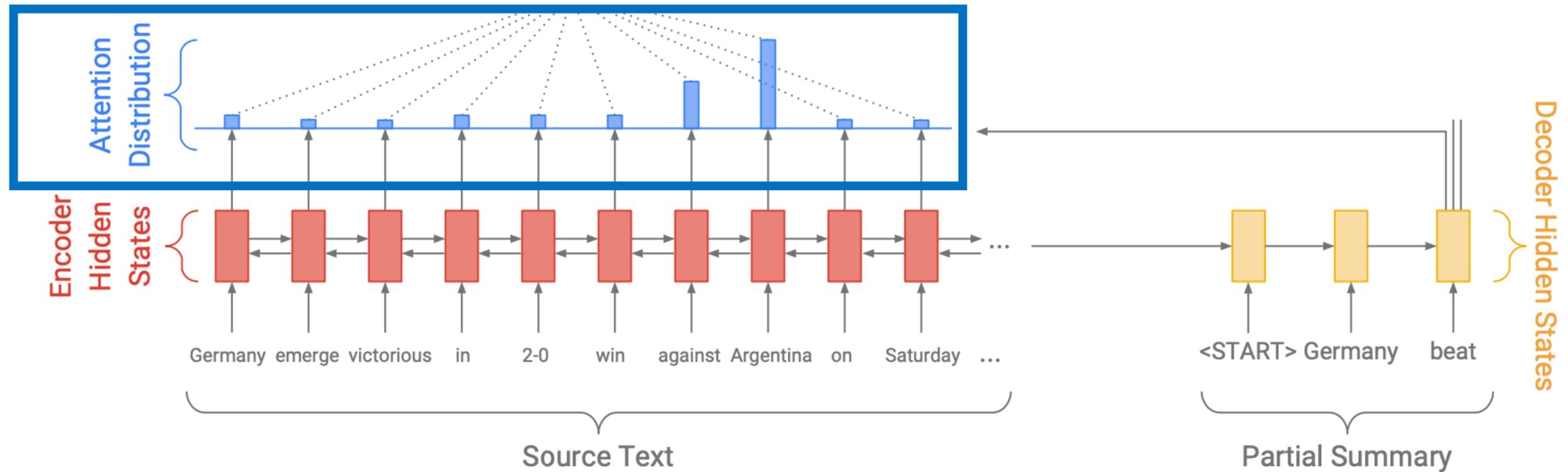
Encoder/Decoder Hidden State



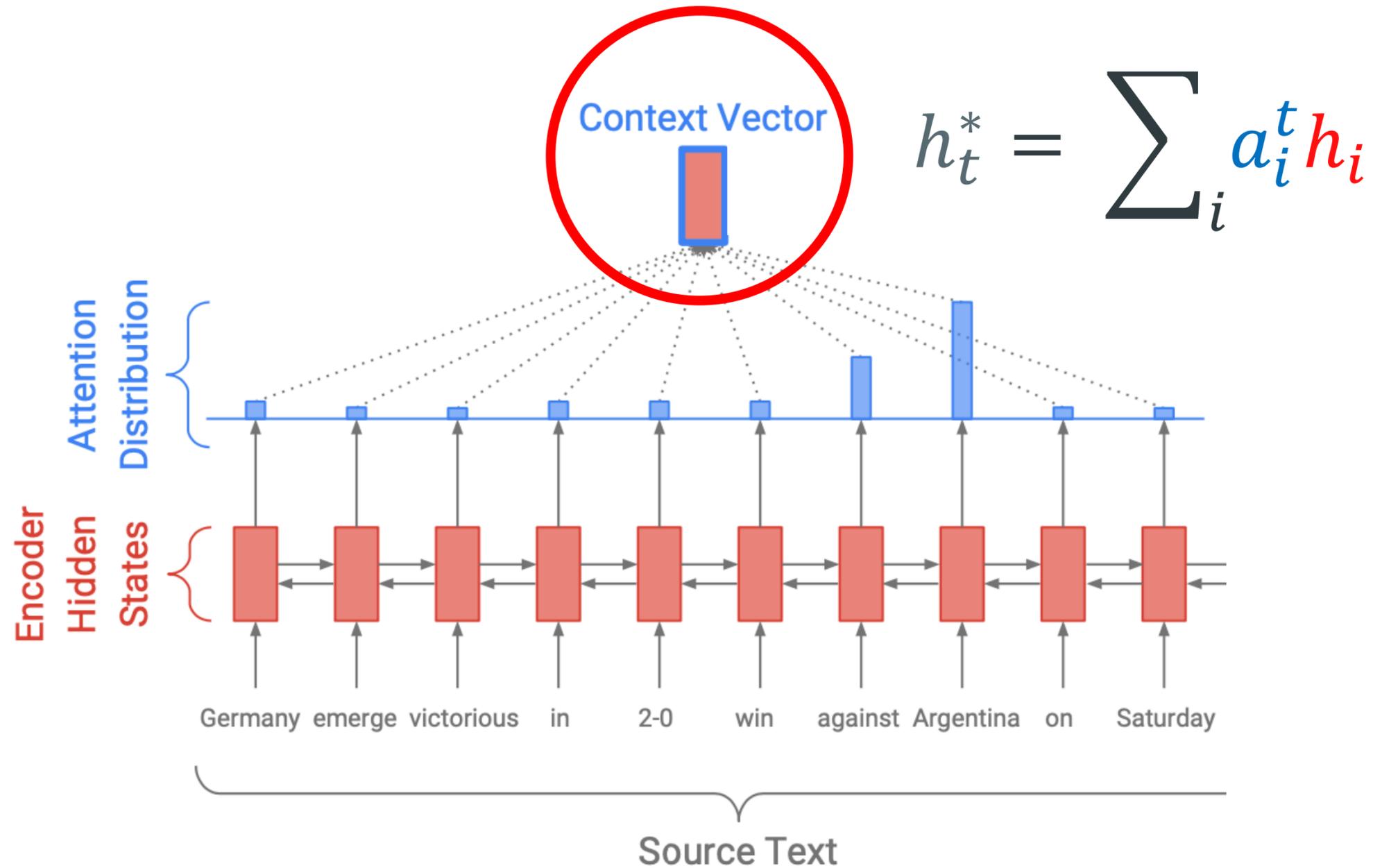
Attention Distribution

$$e_i^t = v^T \tanh(W_h \overset{\text{dim} = 256}{h_i} + W_s \overset{\text{dim} = 256}{s_t} + b_{\text{attn}})$$

$$a_t = \text{softmax}(e^t)$$

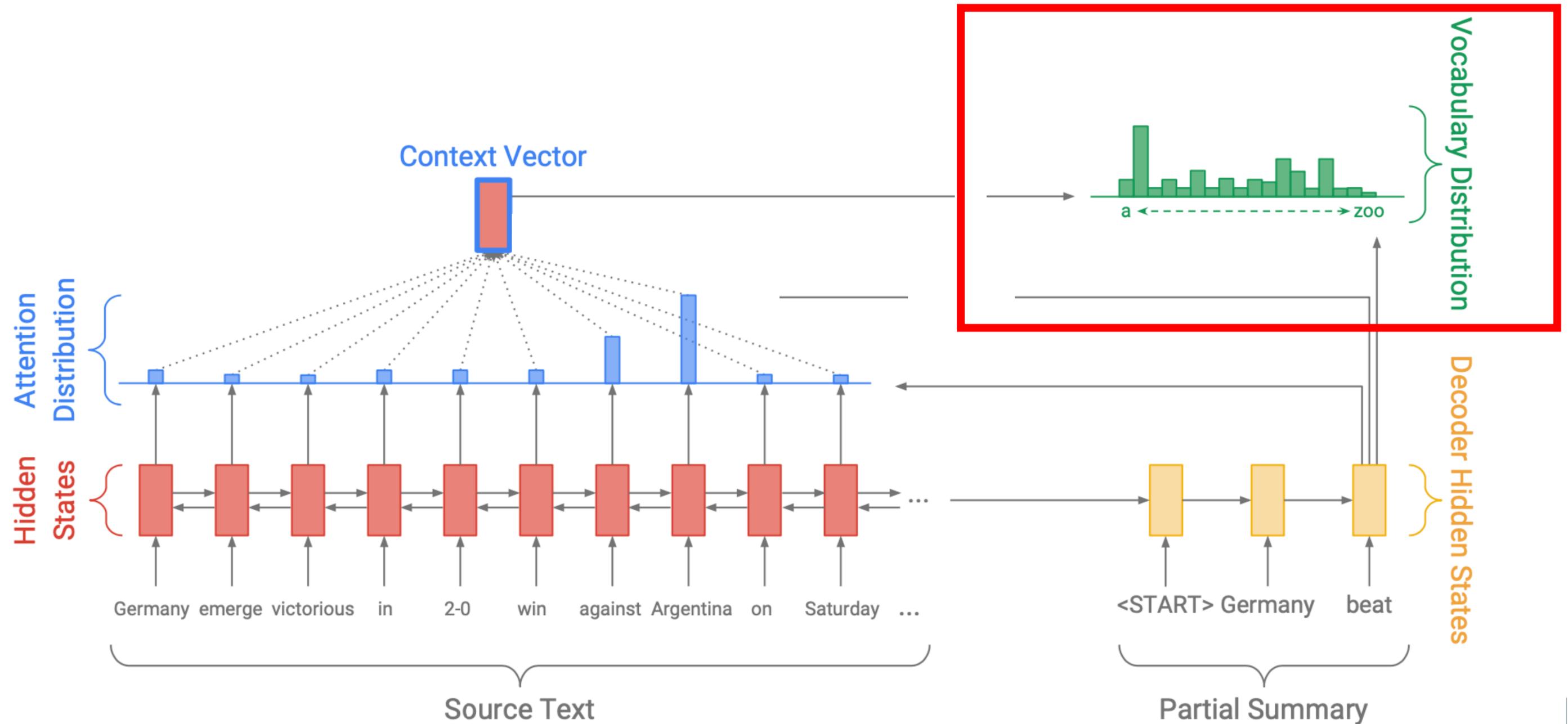


Context Vector

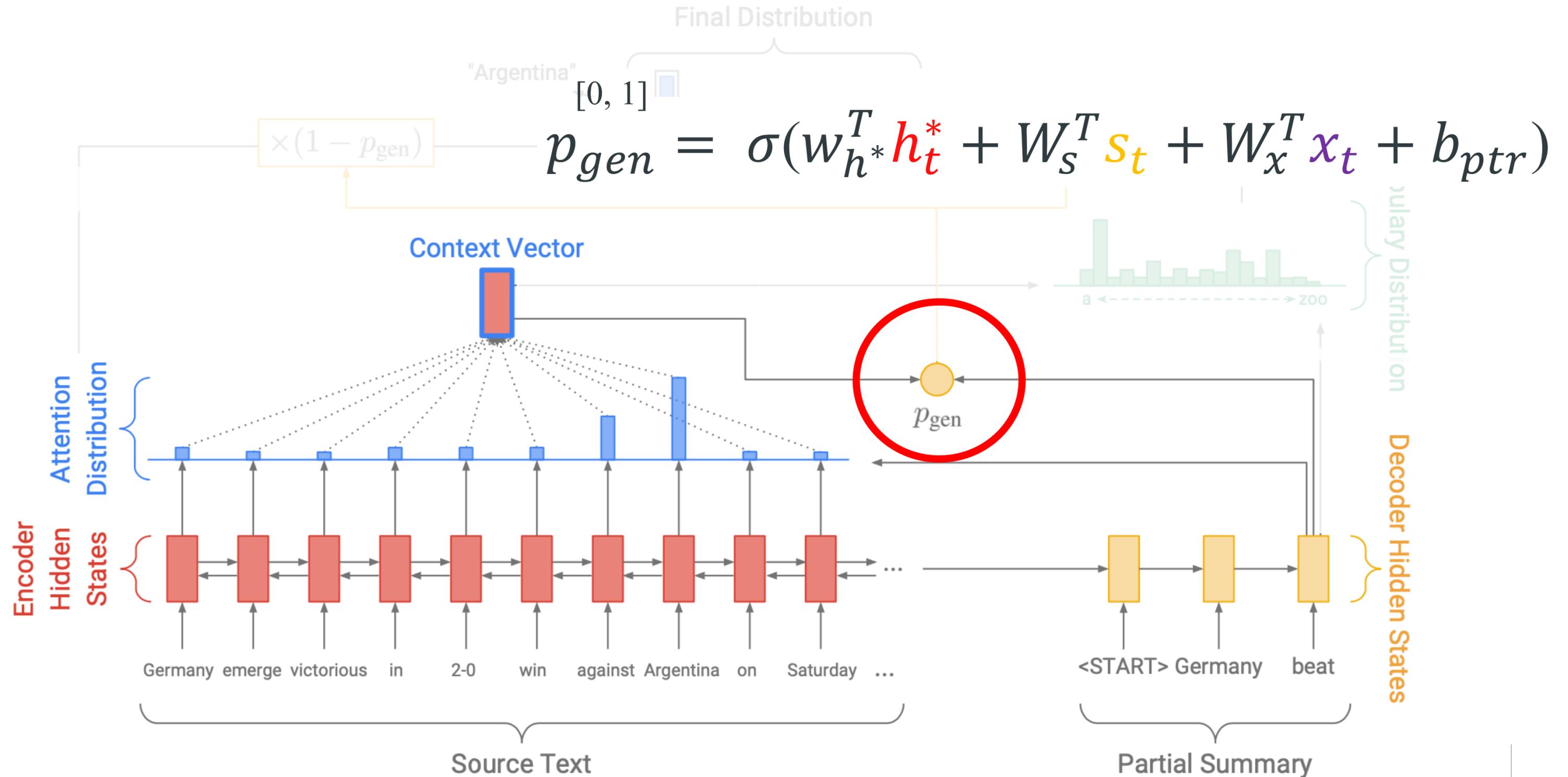


Vocabulary Distribution

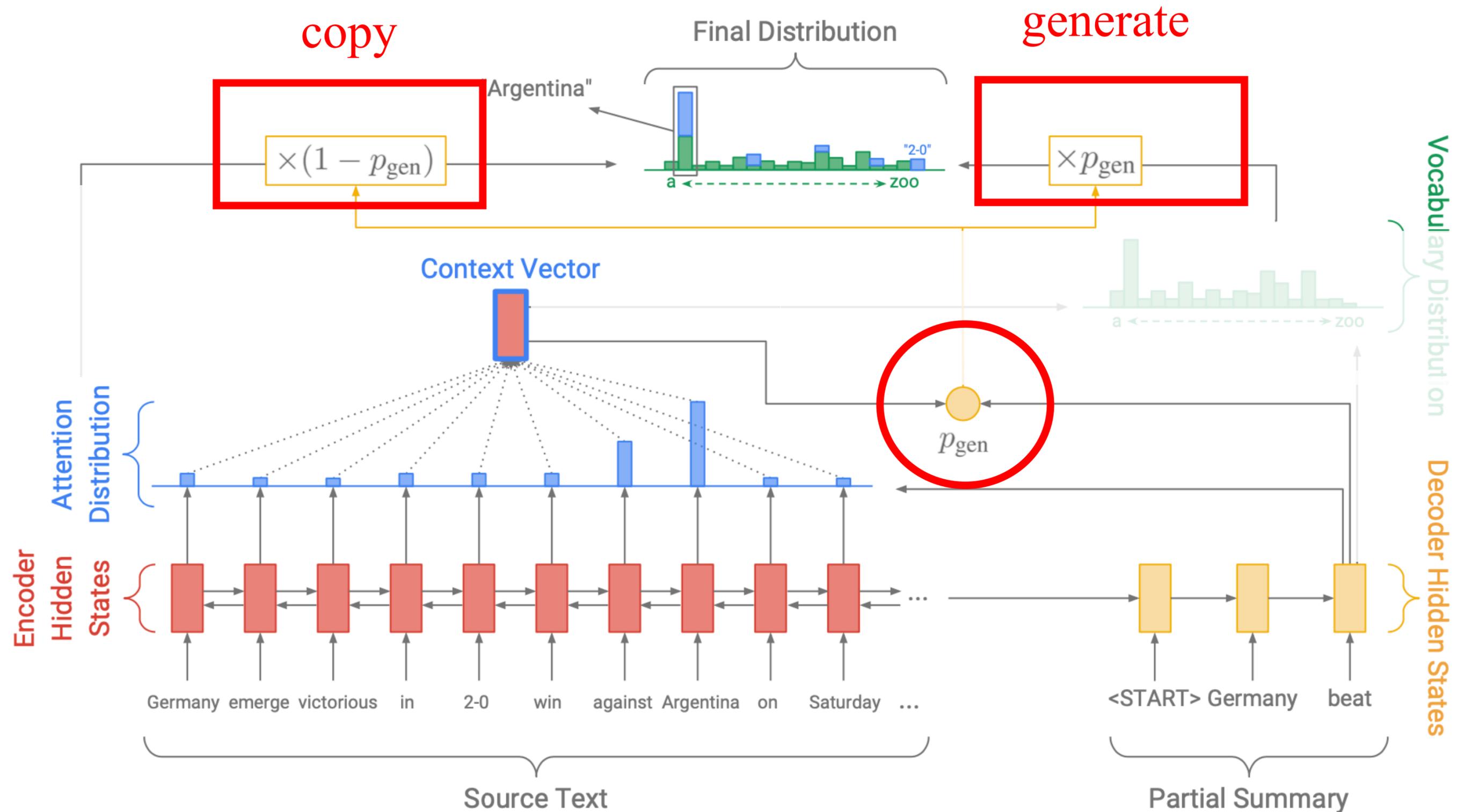
$$P_{vocab} = \text{softmax}(V'(V[s_t, h_t^*] + b) + b')$$



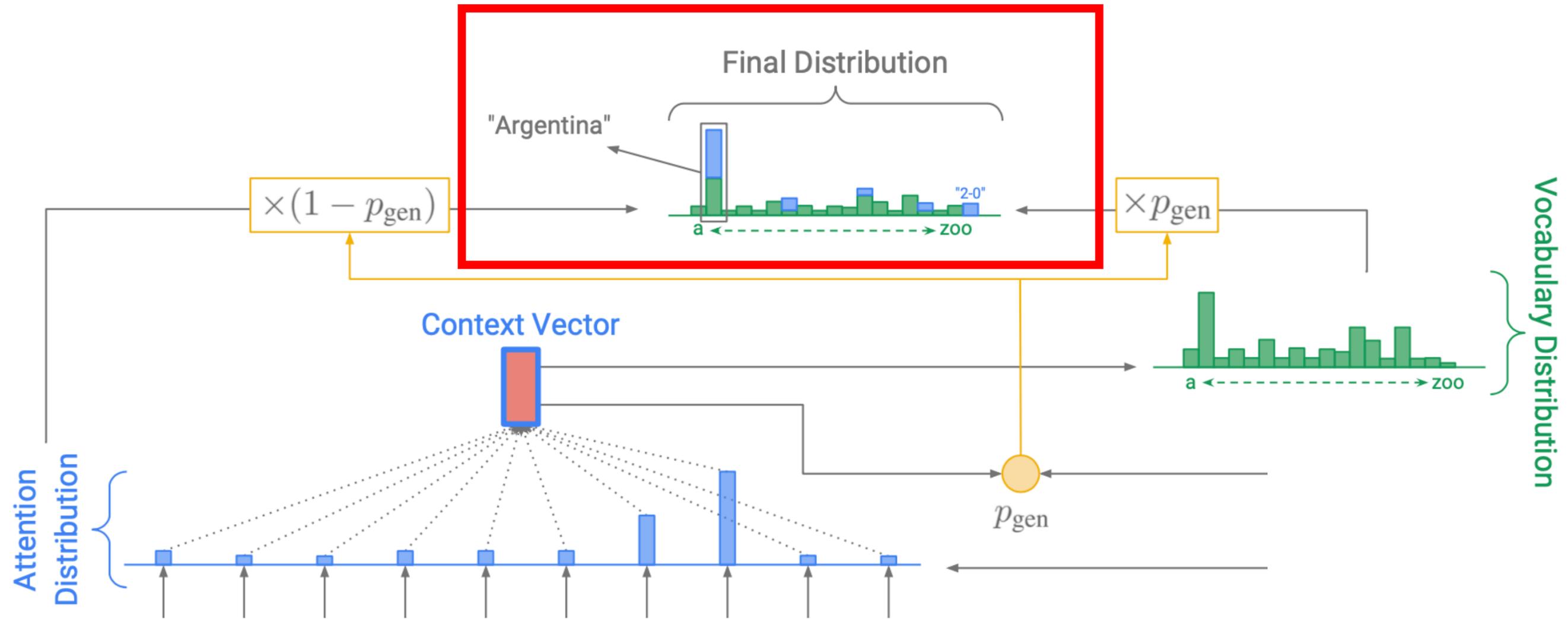
Generation Probability



Generation Probability



Generation Probability



$$P(w) = p_{gen}P_{voc}(w) + (1 - p_{gen}) \sum_{i:w_i=w} a_i^t$$

Loss Function

$$loss_t = -\log P(w_t^*)$$

$$loss = \frac{1}{T} \sum_{t=0}^T loss_t$$

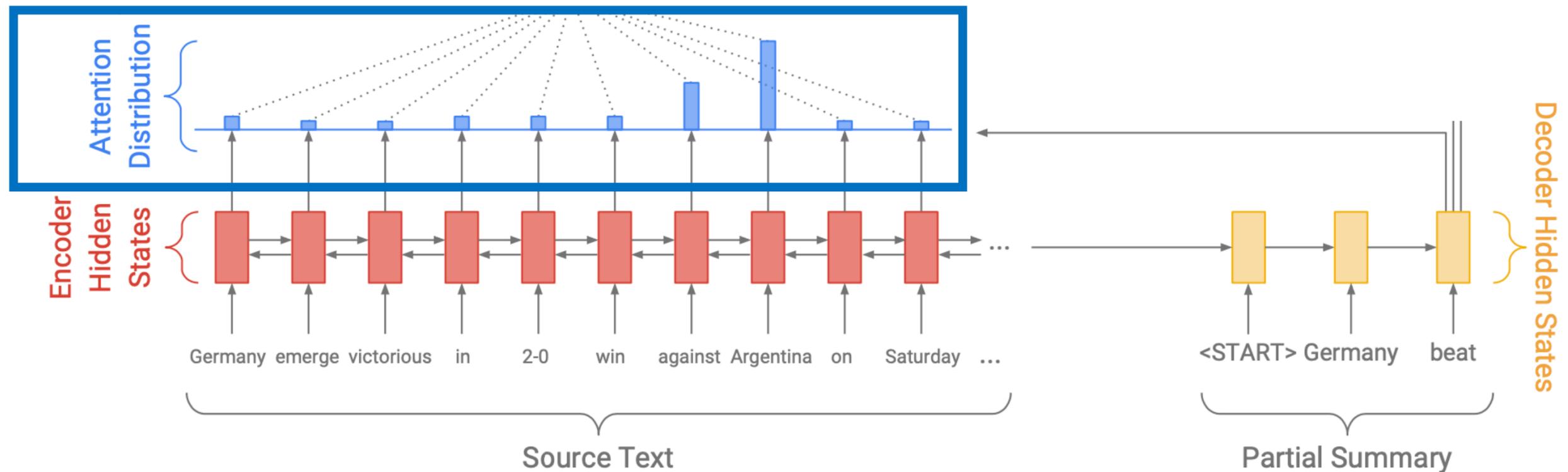
Pointer-generator network + Coverage mechanism

Coverage mechanism

$$c^t = \sum_{t'=0}^{t-1} a^{t'}$$

$$e_i^t = v^T \tanh(W_h h_i + W_s s_t + w_c c_i^t + b_{attn})$$

$$a_t = \text{softmax}(e^t)$$



Loss Function

$$loss_t = -\log P(w_t^*) + \lambda \sum_i \min(a_i^t, c_i^t)$$

Coverage loss

$$loss = \frac{1}{T} \sum_{t=0}^T loss_t$$

OUTLINE

Introduction

Method

Experiment

Conclusion

Dataset

Deep Mind Q&A Dataset	# of articles
CNN	Approximately 90k
Daily Mail	Approximately 197k

781 tokens/article
56 tokens/summary

Set	# of pairs
Train	287,226
Validation	13,368
Test	11,490

ROUGE and METEOR

	ROUGE			METEOR		
	1	2	L	exact match	+ stem/syn/para	
<u>abstractive model (Nallapati et al., 2016)*</u>	35.46	13.30	32.65	-	-	abtractive
<u>seq-to-seq + attn baseline (150k vocab)</u>	30.49	11.17	28.08	11.65	12.86	
<u>seq-to-seq + attn baseline (50k vocab)</u>	31.33	11.81	28.83	12.03	13.20	
<u>pointer-generator</u>	36.44	15.66	33.42	15.35	16.65	
<u>pointer-generator + coverage</u>	39.53	17.28	36.38	17.32	18.72	
<u>lead-3 baseline (ours)</u>	40.34	17.70	36.57	20.48	22.21	extractive
<u>lead-3 baseline (Nallapati et al., 2017)*</u>	39.2	15.7	35.5	-	-	
<u>extractive model (Nallapati et al., 2017)*</u>	39.6	16.2	35.3	-	-	

Table 1: ROUGE F₁ and METEOR scores on the test set. M

Coverage Elimination

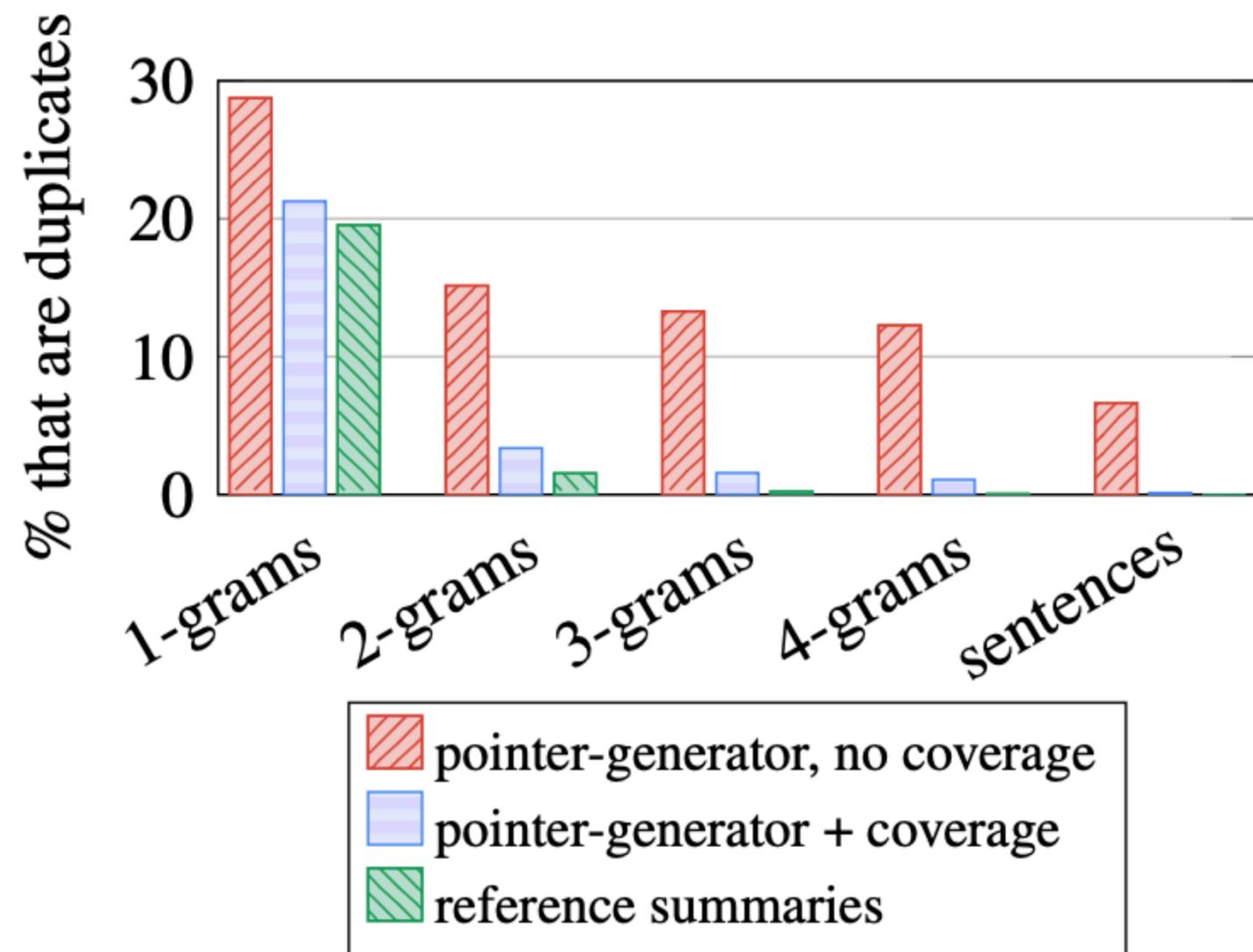


Figure 4: Coverage eliminates undesirable repetition.

Novel n-grams Comparison

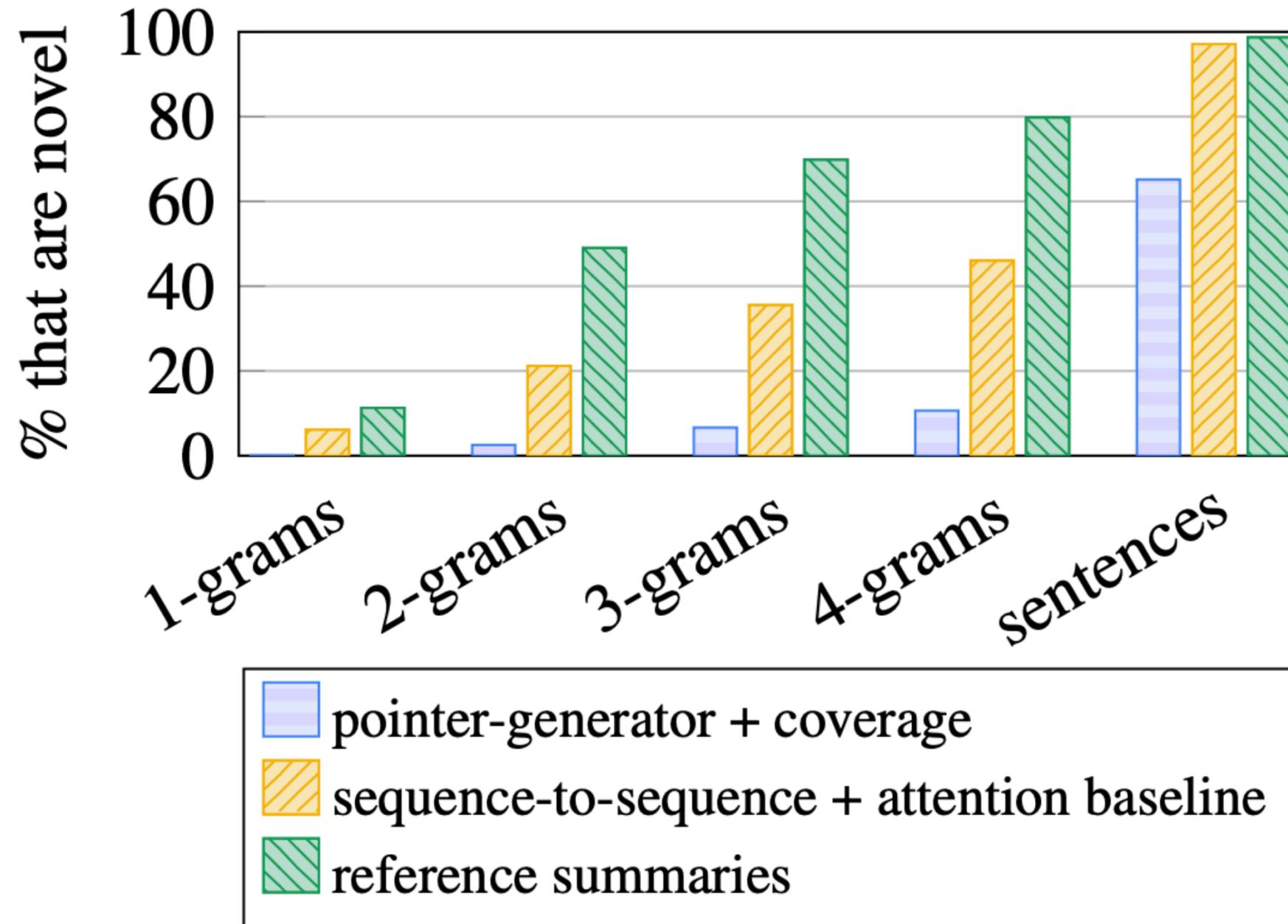


Figure 6

Examples

Article: andy murray (...) is into the semi-finals of the miami open , but not before getting a scare from 21 year-old austrian dominic thiem, who pushed him to 4-4 in the second set before going down 3-6 6-4, 6-1 in an hour and three quarters. (...)

Summary: andy murray **defeated** dominic thiem 3-6 6-4, 6-1 in an hour and three quarters.

Article: (...) wayne rooney smashes home during manchester united 's 3-1 win over aston villa on saturday. (...)

Summary: manchester united **beat** aston villa 3-1 at old trafford on saturday.

Figure 7: Examples of abstractive summaries produced by our model (**bold** denotes novel words).

Examples

Original Text (truncated): lagos, nigeria (cnn) a day after winning nigeria's presidency, *muhammadu buhari* told cnn's christiane amanpour that **he plans to aggressively fight corruption that has long plagued nigeria** and go after the root of the nation's unrest. *buhari* said he'll "rapidly give attention" to curbing violence in the northeast part of nigeria, where the terrorist group boko haram operates. by cooperating with neighboring nations chad, cameroon and niger, **he said his administration is confident it will be able to thwart criminals** and others contributing to nigeria's instability. for the first time in nigeria's history, the opposition defeated the ruling party in democratic elections. *buhari* defeated incumbent goodluck jonathan by about 2 million votes, according to nigeria's independent national electoral commission. **the win comes after a long history of military rule, coups and botched attempts at democracy in africa's most populous nation.**

Baseline Seq2Seq + Attention: UNK UNK says his administration is confident it will be able to **destabilize nigeria's economy**. UNK says his administration is confident it will be able to thwart criminals and other **nigerians**. **he says the country has long nigeria and nigeria's economy.**

Pointer-Gen: *muhammadu buhari* says he plans to aggressively fight corruption **in the northeast part of nigeria**. he says he'll "rapidly give attention" to curbing violence **in the northeast part of nigeria**. he says his administration is confident it will be able to thwart criminals.

Pointer-Gen + Coverage: *muhammadu buhari* says he plans to aggressively fight corruption that has long plagued nigeria. he says his administration is confident it will be able to thwart criminals. the win comes after a long history of military rule, coups and botched attempts at democracy in africa's most populous nation.

OUTLINE

Introduction

Method

Experiment

Conclusion

Conclusion

1. We presented a hybrid pointer-generator architecture with coverage, and showed that it reduced inaccuracies and repetition.
2. We applied our model to a new and challenging long text dataset.
3. Our model exhibits many abstractive abilities, but attaining higher levels of abstraction remains an open research question.