

# **Advanced NLP**

**Summer 2023**

**Anoop Sarkar**

# NEURAL MACHINE TRANSLATION BY JOINTLY LEARNING TO ALIGN AND TRANSLATE

**Dzmitry Bahdanau**

Jacobs University Bremen, Germany

**KyungHyun Cho    Yoshua Bengio\***

Université de Montréal

ICLR 2015

<https://arxiv.org/abs/1409.0473>

# Machine Translation

$$\arg \max_{\mathbf{y}} p(\mathbf{y} \mid \mathbf{x})$$

Target language

Source language

# Neural Machine Translation

$$\arg \max_{\mathbf{y}} p(\mathbf{y} \mid \mathbf{x}) = \prod_{t=1}^T p(y_t \mid \{y_1, \dots, y_{t-1}\}, c),$$

$$\mathbf{y} = (y_1, \dots, y_{T_y}).$$

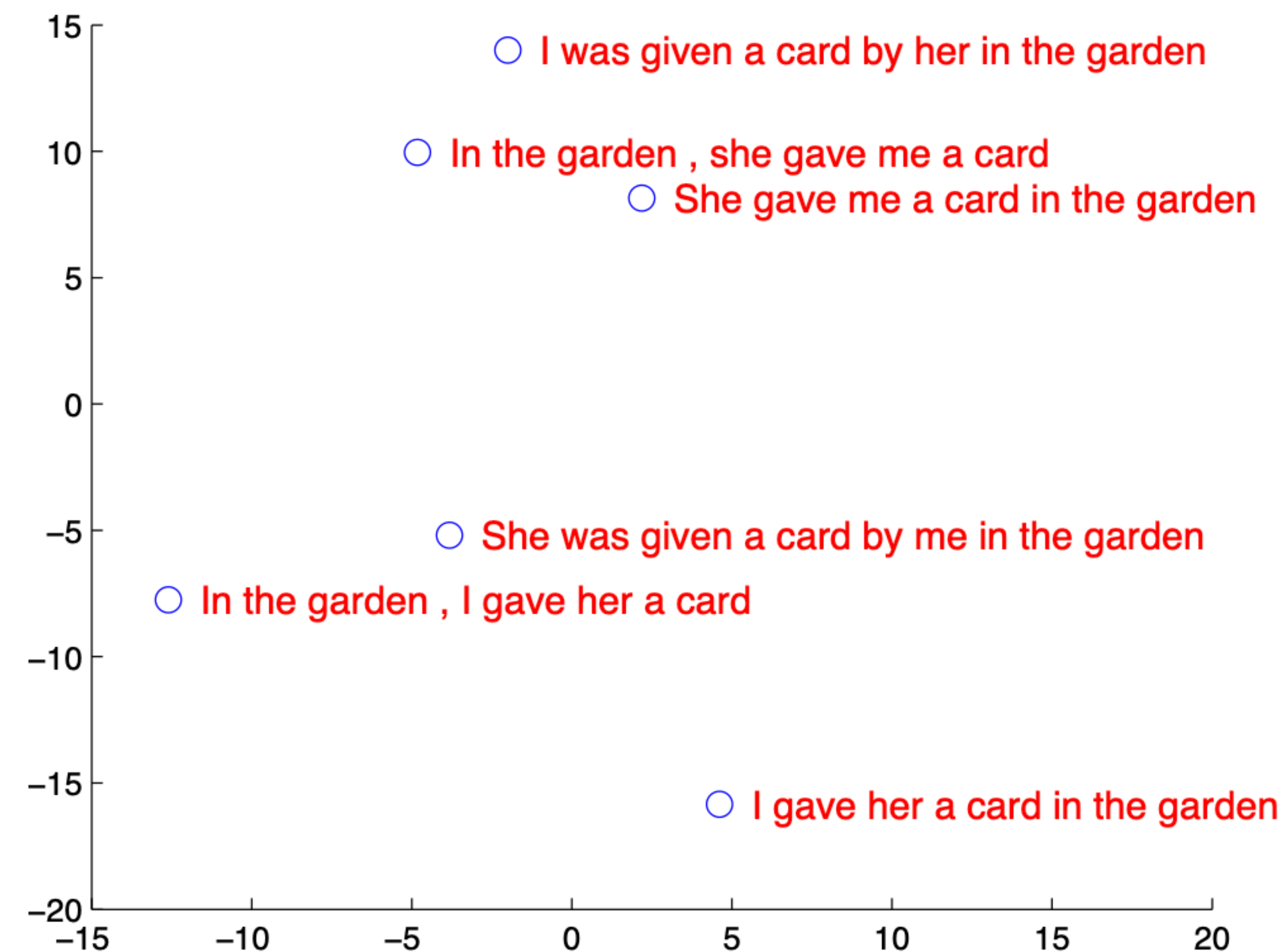
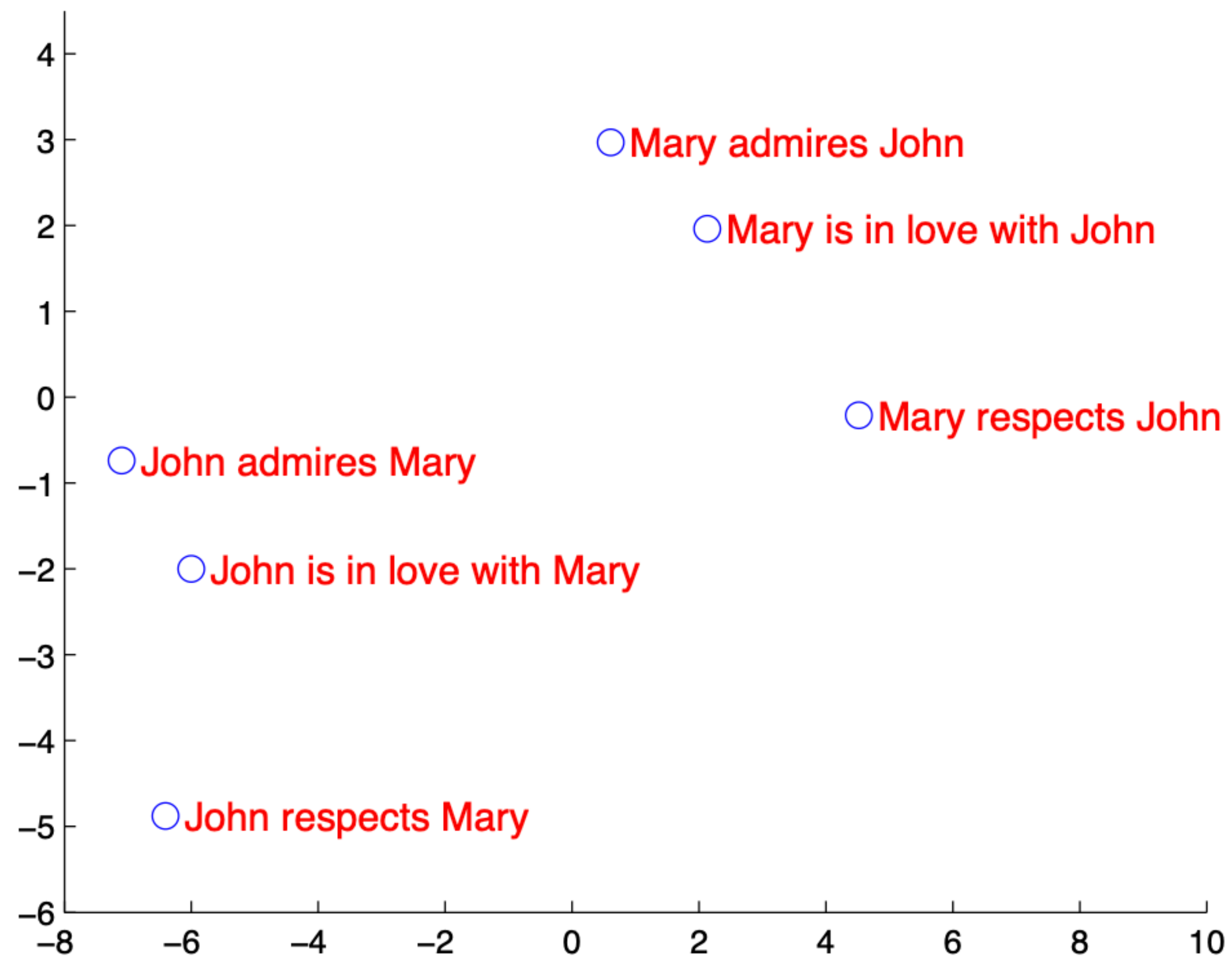
$$c = q(\{h_1, \dots, h_{T_x}\})$$

$$\mathbf{x} = (x_1, \dots, x_{T_x}).$$

$h_i$  is a vector representation of  $x_i$

$c$  is a vector assembled from all the  $h_i$  vectors

$s_i$  is a vector representation of  $y_i$



Sequence to Sequence Learning with Neural Networks <https://arxiv.org/abs/1409.3215>

# Attention Networks in Neural Machine Translation

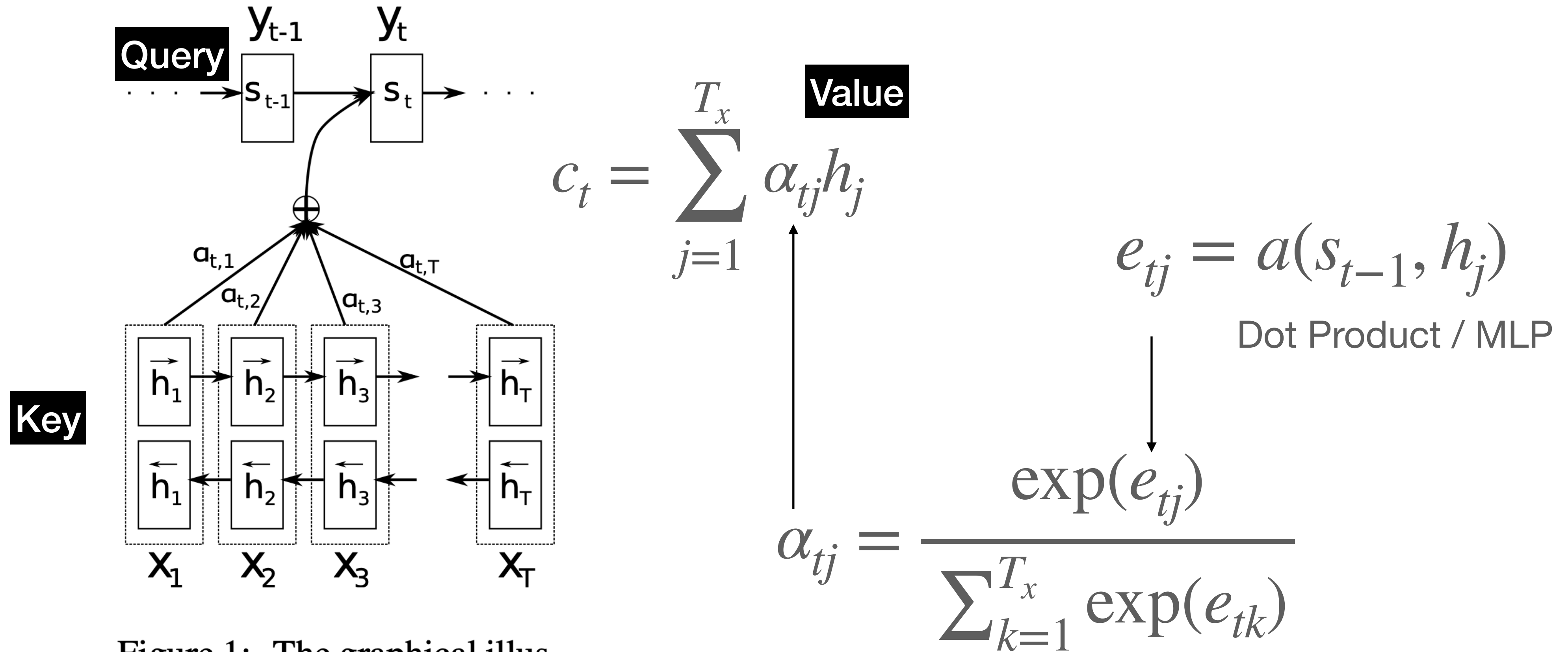
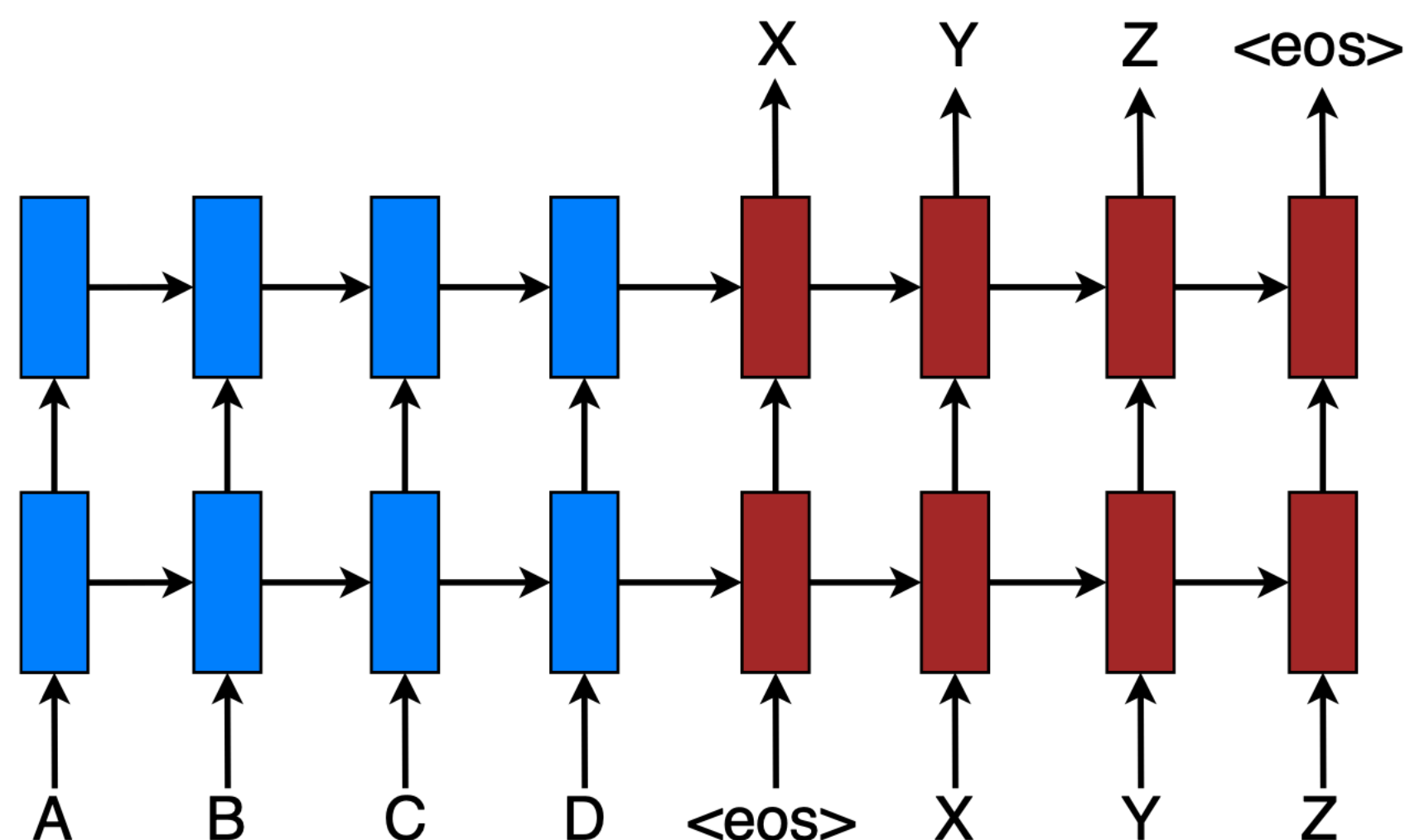


Figure 1: The graphical illustration of the proposed model trying to generate the  $t$ -th target word  $y_t$  given a source sentence  $(x_1, x_2, \dots, x_T)$ .

# Effective Approaches to Attention-based Neural Machine Translation

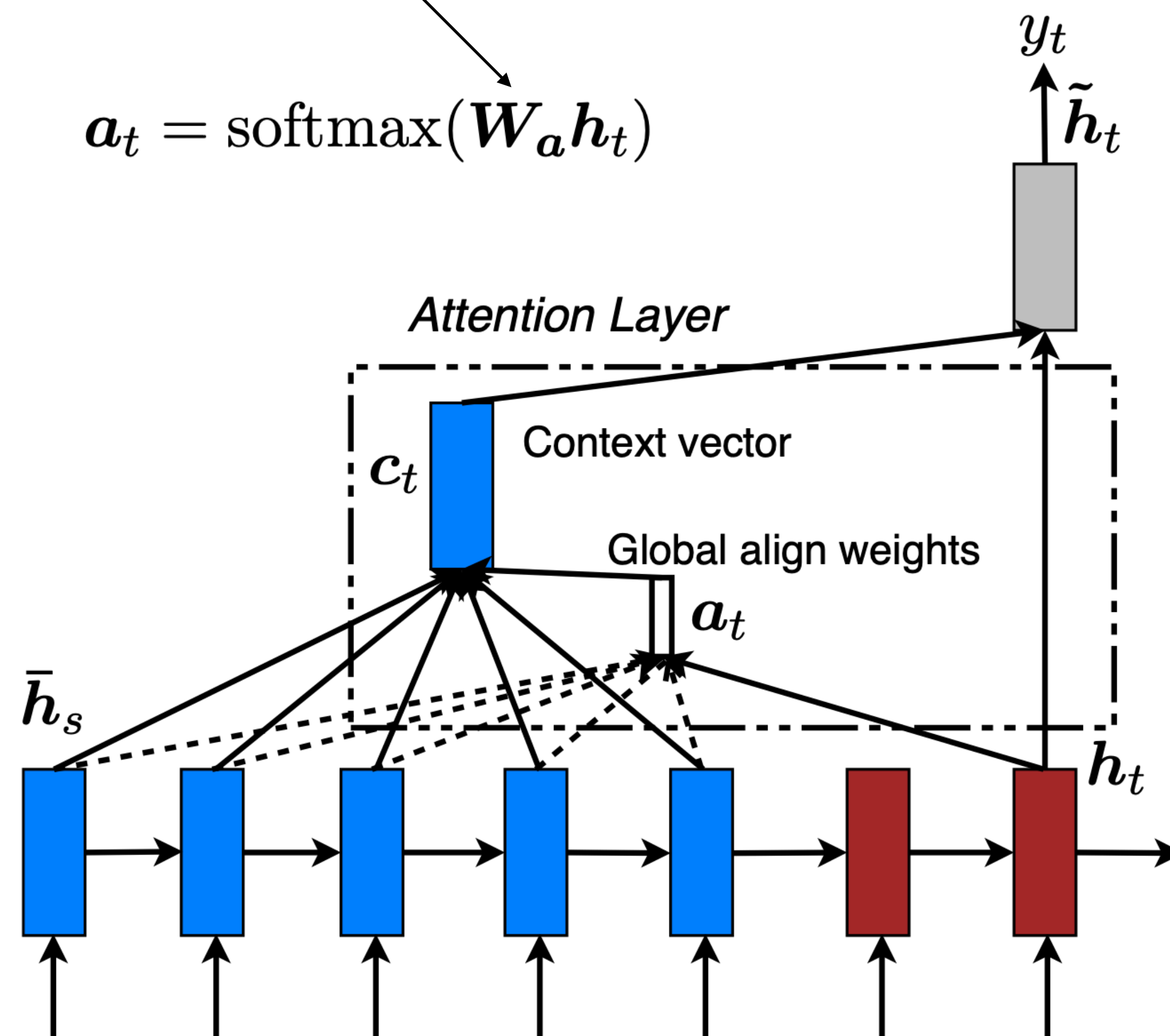


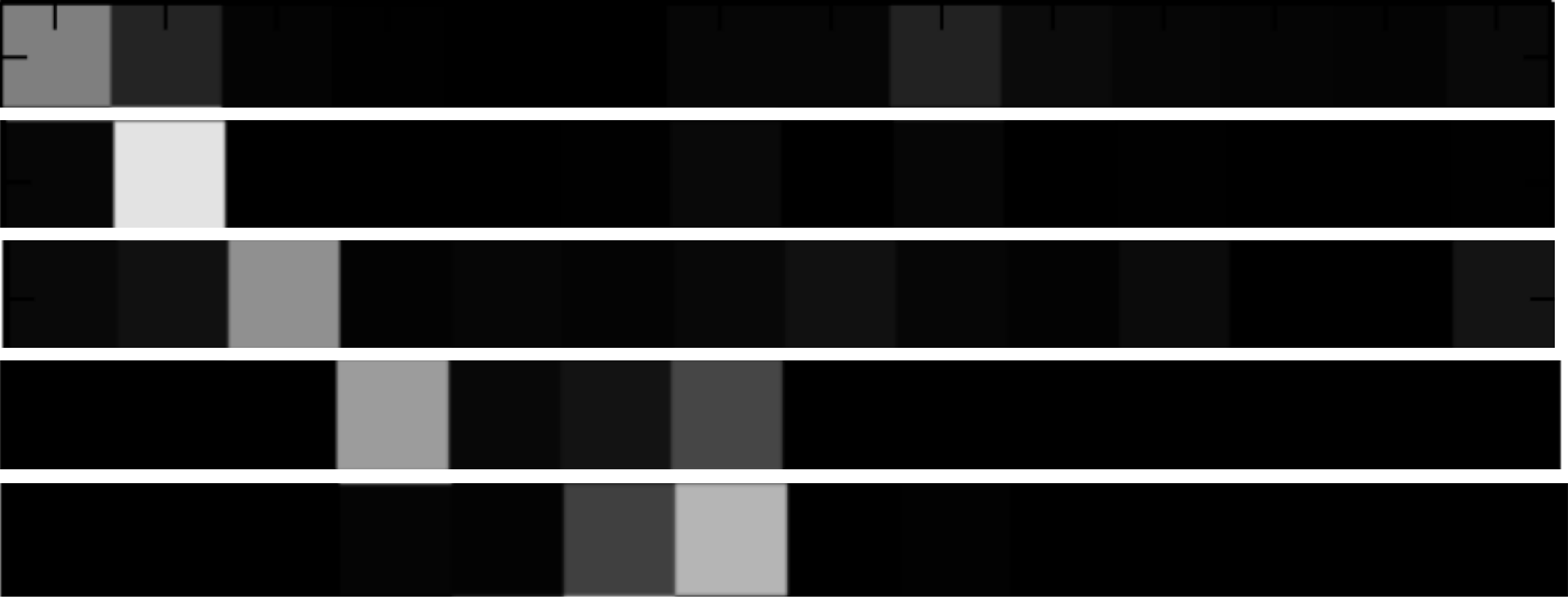
$$\text{score}(\mathbf{h}_t, \bar{\mathbf{h}}_s) = \begin{cases} \mathbf{h}_t^\top \bar{\mathbf{h}}_s & \text{dot} \\ \mathbf{h}_t^\top \mathbf{W}_a \bar{\mathbf{h}}_s & \text{general} \\ \mathbf{v}_a^\top \tanh(\mathbf{W}_a [\mathbf{h}_t; \bar{\mathbf{h}}_s]) & \text{concat} \end{cases}$$

<https://arxiv.org/abs/1508.04025>

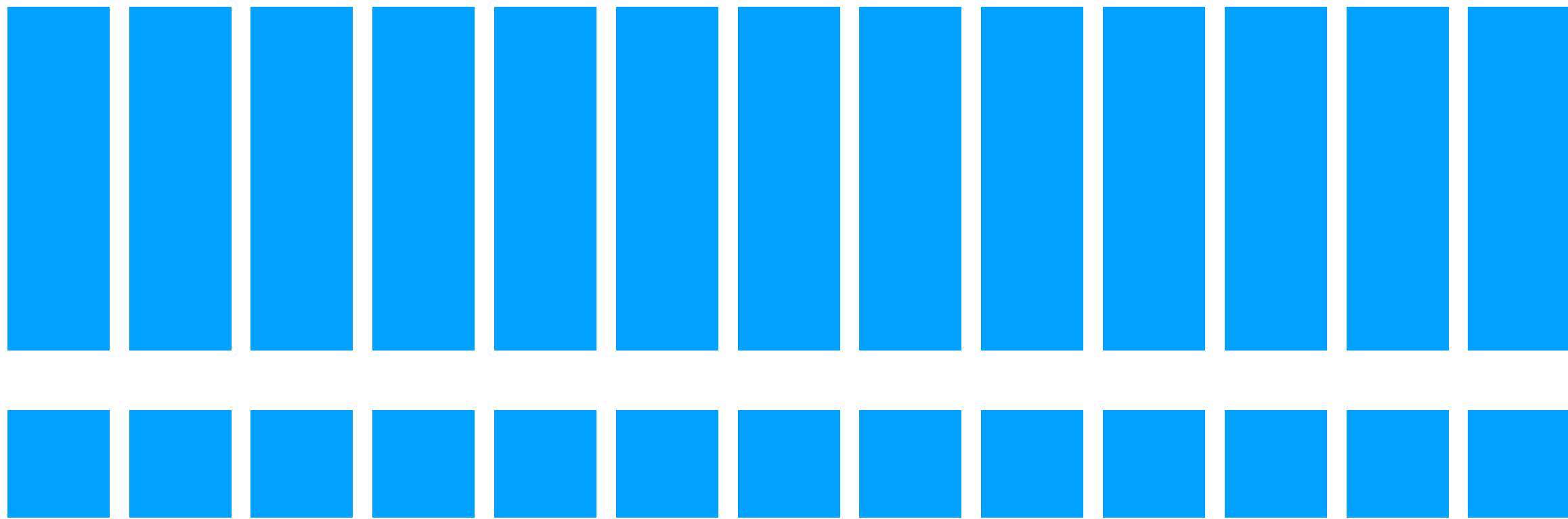
$$\frac{\exp(\text{score}(\mathbf{h}_t, \bar{\mathbf{h}}_s))}{\sum_{s'} \exp(\text{score}(\mathbf{h}_t, \bar{\mathbf{h}}_{s'}))}$$

$$\mathbf{a}_t = \text{softmax}(\mathbf{W}_a \mathbf{h}_t)$$

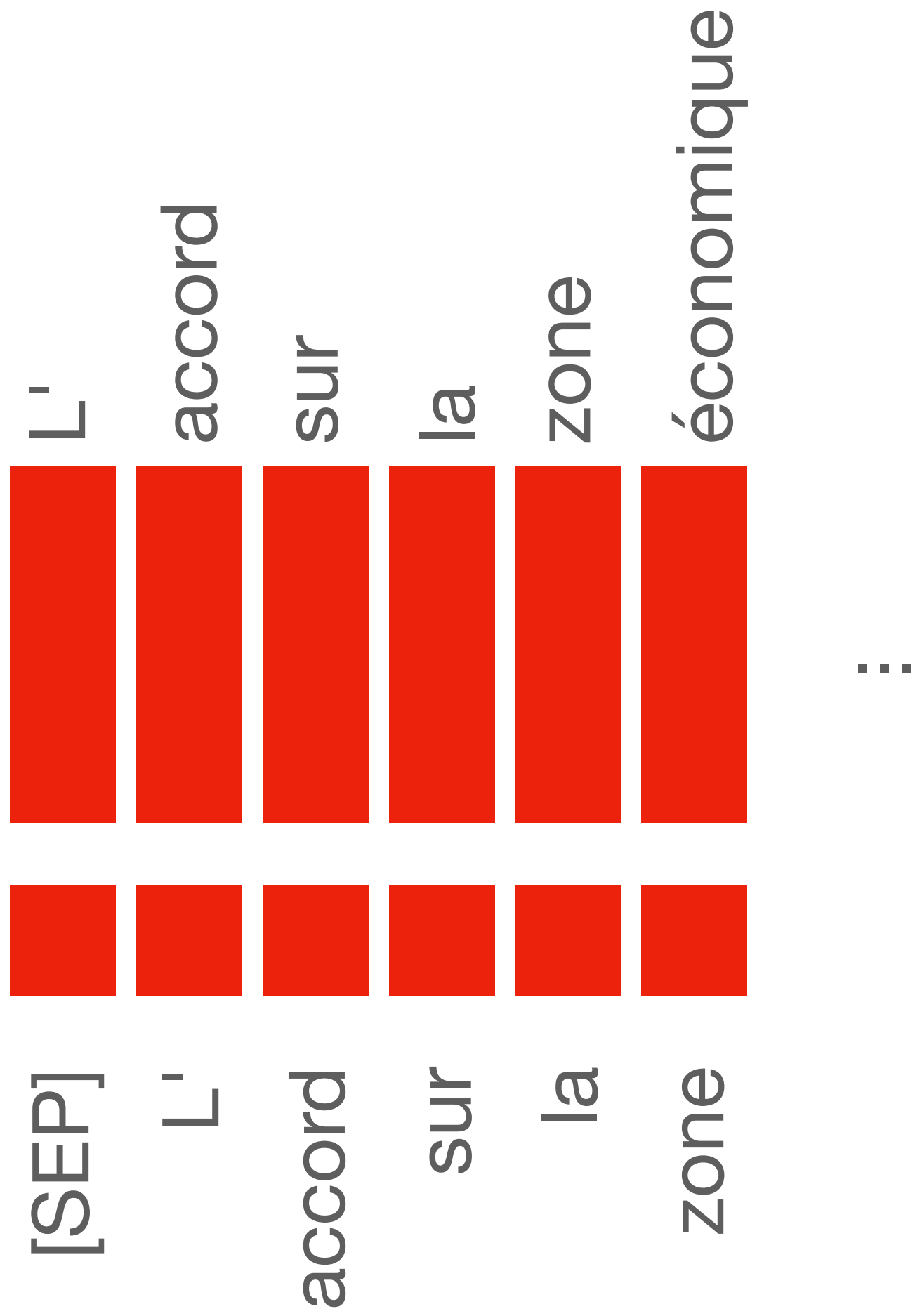




L'  
accord  
sur  
la  
zone



The  
agreement  
on  
the  
European  
Economic  
area  
was  
signed  
in  
August  
1992  
.



[SEP]  
L'  
accord  
sur  
la  
zone  
...











**Source**

*An admitting privilege is the right of a doctor to admit a patient to a hospital or a medical centre to carry out a diagnosis or a procedure, based on his status as a health care worker at a hospital.*

**No attention**

*Un privilège d'admission est le droit d'un médecin de reconnaître un patient à l'hôpital ou un centre médical d'un diagnostic ou de prendre un diagnostic en fonction de son état de santé.*

**With attention**

*Un privilège d'admission est le droit d'un médecin d'admettre un patient à un hôpital ou un centre médical pour effectuer un diagnostic ou une procédure, selon son statut de travailleur des soins de santé à l'hôpital.*

**Source**

*This kind of experience is part of Disney's efforts to "extend the lifetime of its series and build new relationships with audiences via digital platforms that are becoming ever more important," he added.*

**No attention**

*Ce type d'expérience fait partie des initiatives du Disney pour "prolonger la durée de vie de ses nouvelles et de développer des liens avec les lecteurs numériques qui deviennent plus complexes.*

**With attention**

*Ce genre d'expérience fait partie des efforts de Disney pour "prolonger la durée de vie de ses séries et créer de nouvelles relations avec des publics via des plateformes numériques de plus en plus importantes", a-t-il ajouté.*