Practice 07

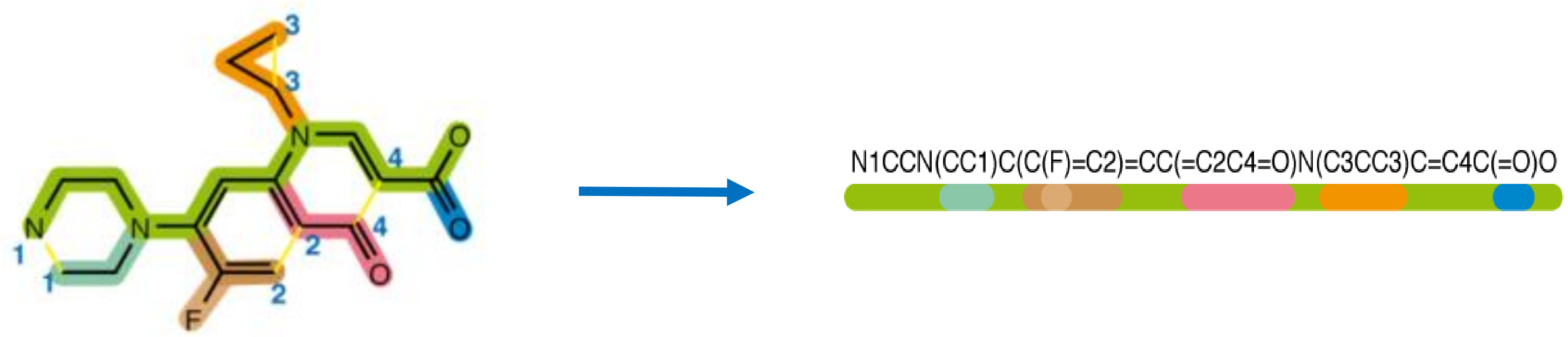
SMILES and RNN

Seongok Ryu
Department of Chemistry, KAIST

Contents

- SMILES
- Applications of RNN
- Modern advanced RNN
- RNN to predict logP
- Assignment #6

SMILES

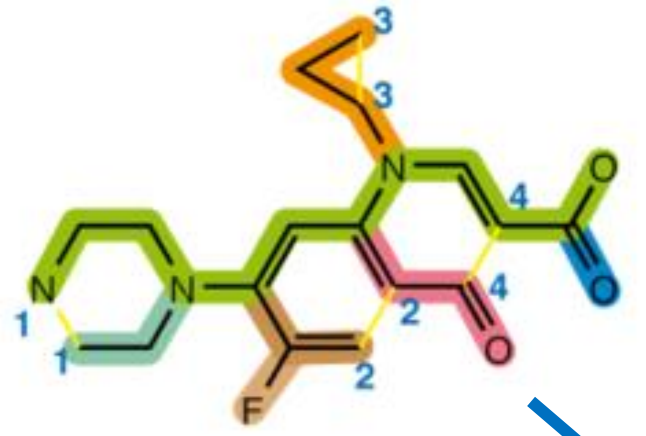


*“The **simplified molecular-input line-entry system** (**SMILES**) is a specification in form of a line notation for describing the structure of chemical species using short ASCII strings. SMILES strings can be imported by most molecule editors for conversion back into two-dimensional drawings or three-dimensional models of the molecules.”*

https://en.wikipedia.org/wiki/Simplified_molecular-input_line-entry_system

KAIST

SMILES



N	1	C	C	N	(	C	C	1	)	C	(	C	(	F	)	=	C	2	)	=	C	C	(	=	C	2	C	4	=	O	)	N	(	C	3	C	C	3	)	C	=	C	4	C	(	=	O	)	O

shape = [#batch, #characters]

SMILES

N	1	C	C	N	(	C	C	1	)	C	(	C	(	F	)	=	C	2	)	=	C	C	(	=	C	2	C	4	=	O	)	N	(	C	3	C	C	3	)	C	=	C	4	C	(	=	O	)	O

One-hot encoding

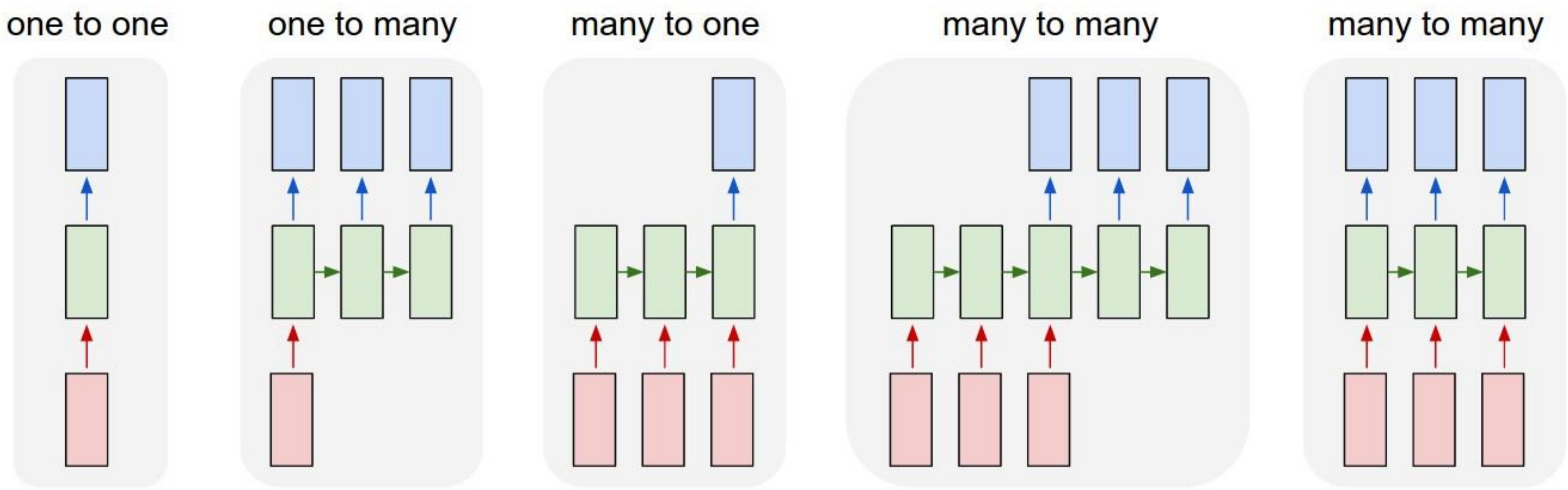
C	0	0	1	1	0	0	1	1	0	0	1	0	1	0	0	0	0	1	0	0	0	1	1	0	0	1	0	1	0	0	0	0	0	0	1	0	1	1	0	0	1	0	1	0	1	0	0	0	0	0
N	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
F	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
=	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0
…	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

shape = [#batch, #characters, #types]

Applications of RNN

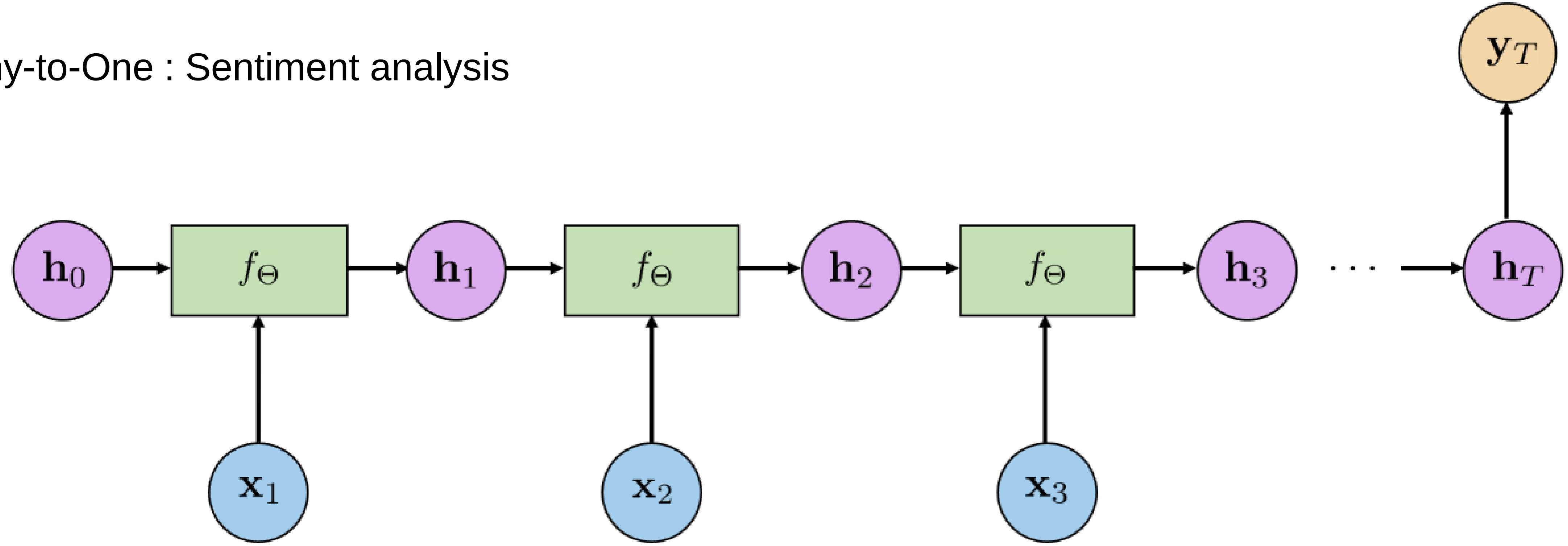
Applications of RNN

Depending on the architecture, many different tasks can be done using RNN



Applications of RNN

Many-to-One : Sentiment analysis



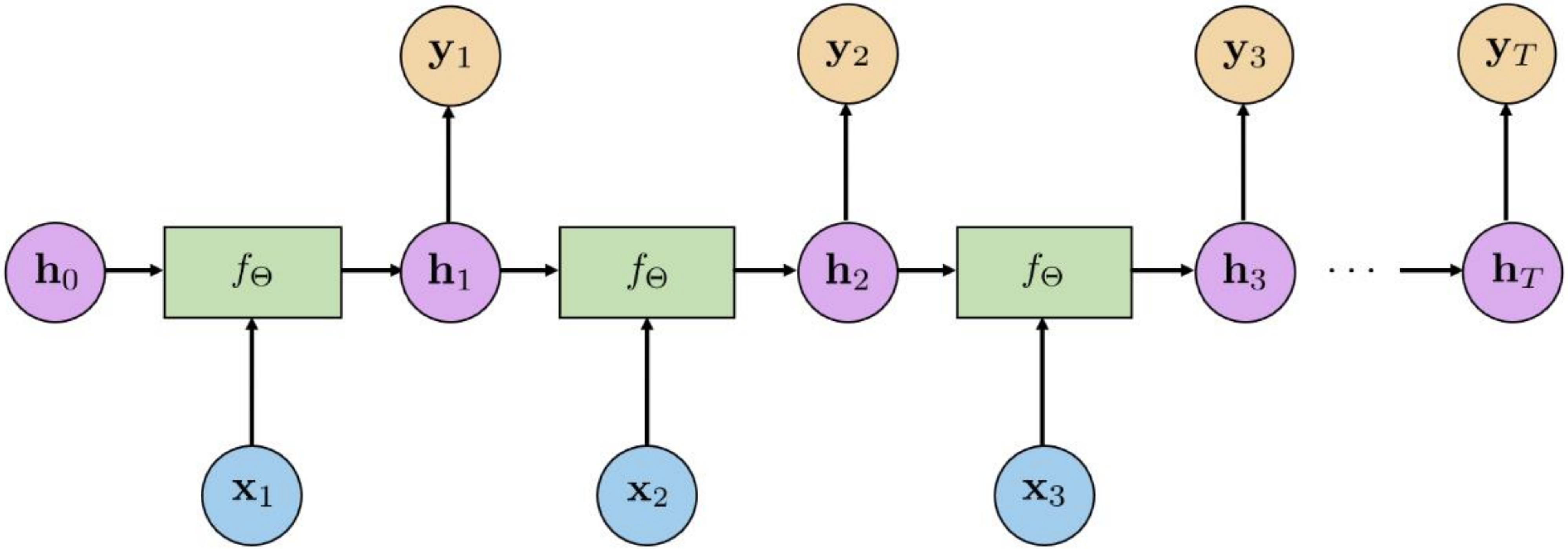
e.g., **Sentiment Classification**
(Sequence of words → sentiment)

→ Good paper or not?

Applications of RNN

Many-to-Many : Neural machine translation

http://alinlab.kaist.ac.kr/resource/Lec1_Introduction_to_NN.pdf



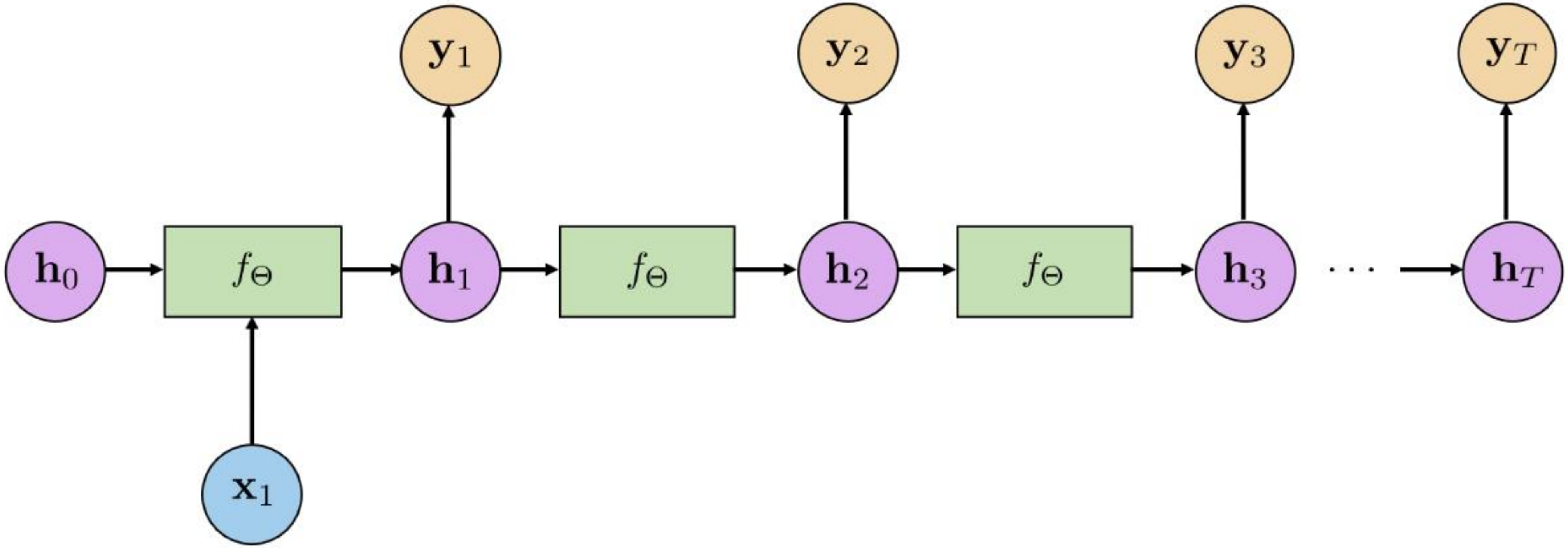
e.g., **Machine Translation**
(Sequence of words → Sequence of words)

Input sentence:	*Translation (PBMT):*	*Translation (GNMT):*	*Translation (human):*
李克強此行將啟動中加總理年度對話機制，與加拿大總理杜魯多舉行兩國總理首次年度對話。	Li Keqiang premier added this line to start the annual dialogue mechanism with the Canadian Prime Minister Trudeau two prime ministers held its first annual session.	Li Keqiang will start the annual dialogue mechanism with Prime Minister Trudeau of Canada and hold the first annual dialogue between the two premiers.	Li Keqiang will initiate the annual dialogue mechanism between premiers of China and Canada during this visit, and hold the first annual dialogue with Premier Trudeau of Canada.

Applications of RNN

One-to-Many : Image captioning

http://alinlab.kaist.ac.kr/resource/Lec1_Introduction_to_NN.pdf



e.g., **Image Captioning**
(Image → sequence of words)

A white teddy bear sitting in the grass | *A man in a baseball uniform throwing a ball* | *A woman is holding a cat in her hand*

Modern advanced RNN

Modern advanced RNN

Attention is all you need

Recall Graph attention

Attention Is All You Need

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
llion@google.com

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

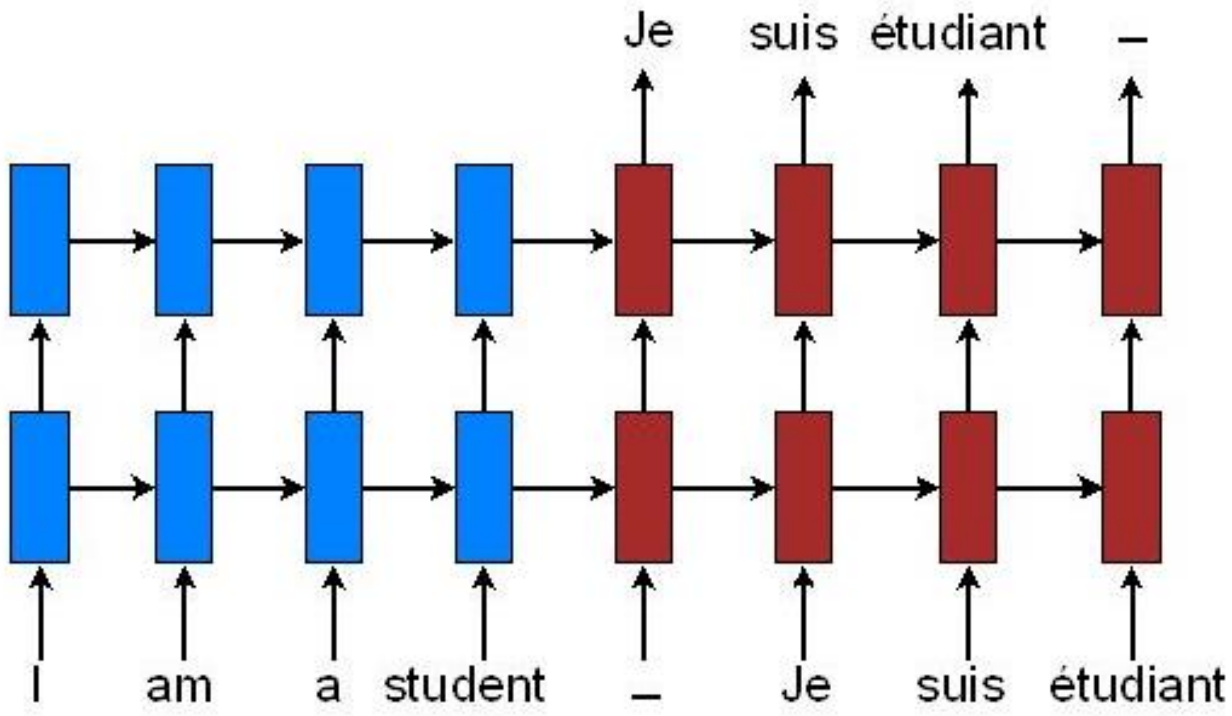
Łukasz Kaiser*
Google Brain
lukaszkaiser@google.com

Illia Polosukhin* ‡
illia.polosukhin@gmail.com

31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

Modern advanced RNN

What is the **attention**?



Basic structure of neural translational machine

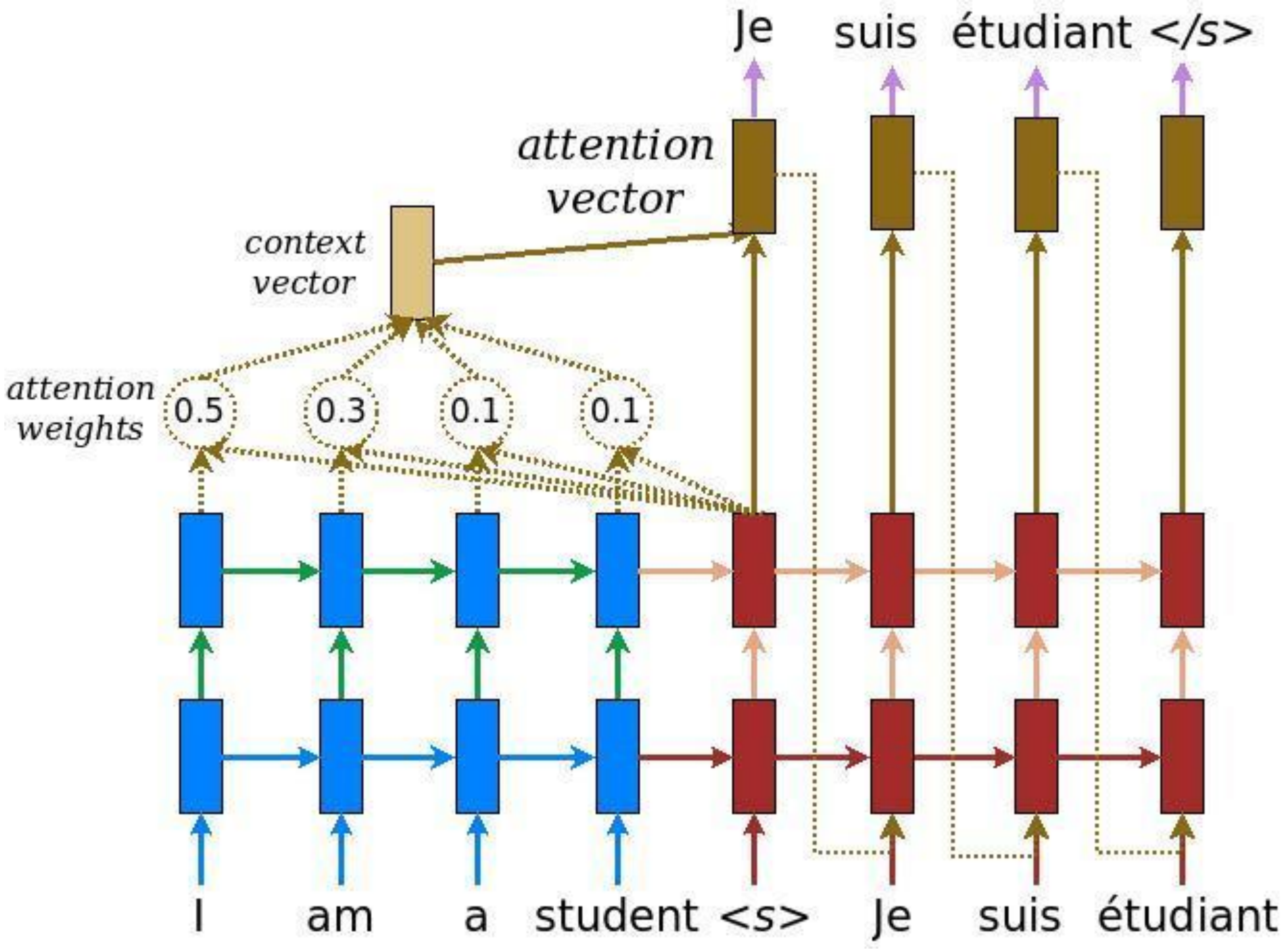
Luong, Minh-Thang, Hieu Pham, and Christopher D. Manning. "Effective approaches to attention-based neural machine translation." *arXiv preprint arXiv:1508.04025* (2015).

Modern advanced RNN

What is the **attention**?

- However, each output from decoder has different importances in relation with input words.

- Attention mechanism grasps the importance between the words and highlights (larger weight) on more important words.



Luong, Minh-Thang, Hieu Pham, and Christopher D. Manning. "Effective approaches to attention-based neural machine translation." *arXiv preprint arXiv:1508.04025* (2015).

Modern advanced RNN

What is the **attention**?

The attentional vector $\tilde{\boldsymbol{h}}_t$ is provided from the context vector $\boldsymbol{c}_t$ and hidden state $\boldsymbol{h}_t$.

$$\tilde{\boldsymbol{h}}_t = \tanh(\boldsymbol{W}_c[\boldsymbol{c}_t, \boldsymbol{h}_t])$$

Then, the predictive output word is given by

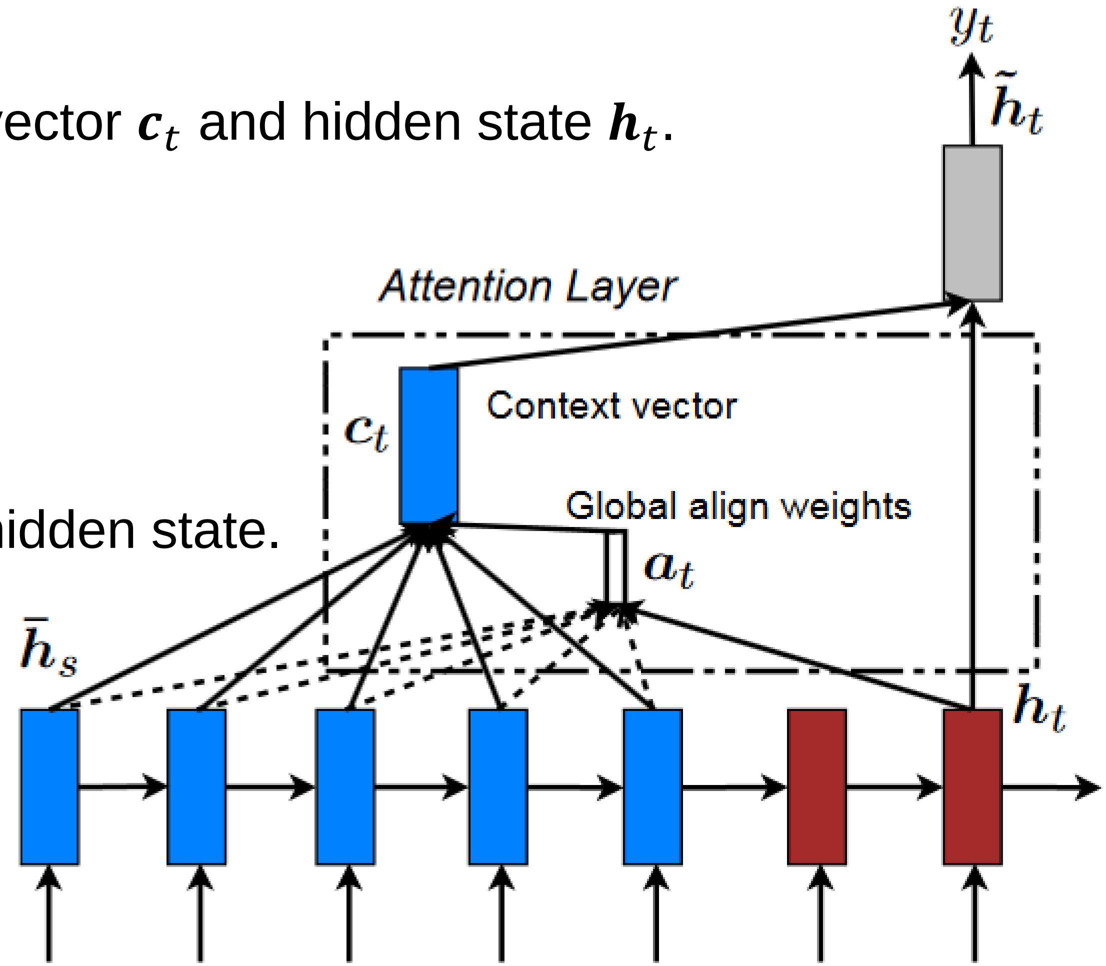
$$p(y_t|y_{<t}, x) = \mathrm{softmax}\left(\boldsymbol{W}_s \tilde{\boldsymbol{h}}_s\right)$$

The context vector is weighted average of the source hidden state.

$$\boldsymbol{c}_t = \boldsymbol{a}_t(s) \cdot \bar{\boldsymbol{h}}(s)$$

$$a_t(s) = \frac{\exp\left(\mathrm{score}(\boldsymbol{h}_t, \bar{\boldsymbol{h}}_s)\right)}{\sum_{s'} \exp\left(\mathrm{score}(\boldsymbol{h}_t, \bar{\boldsymbol{h}}_{s'})\right)}$$

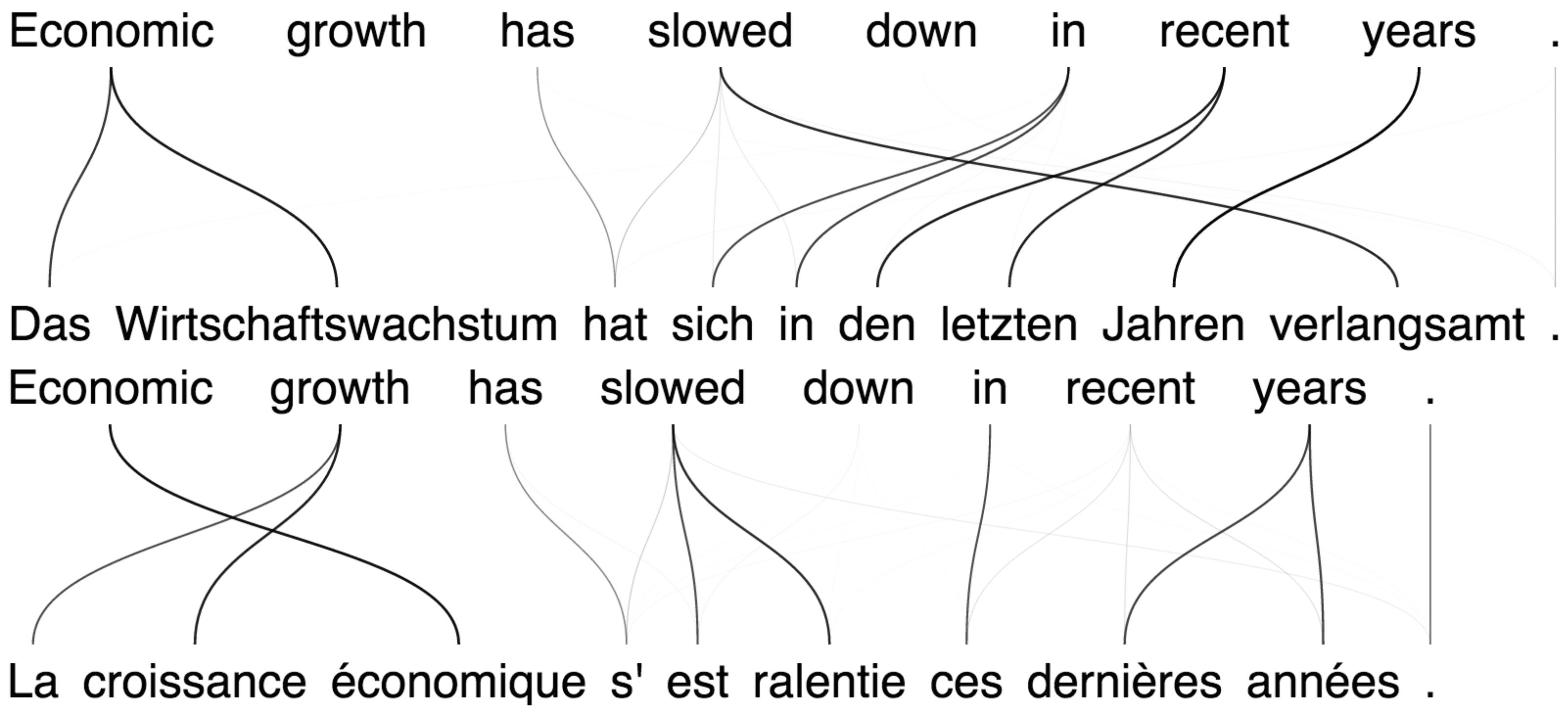
$$\mathrm{score}\left(\boldsymbol{h}_t, \bar{\boldsymbol{h}}_{s'}\right) = \boldsymbol{v}_a^T \tanh\left(\boldsymbol{W}_a\left[\boldsymbol{h}_t; \bar{\boldsymbol{h}}_s\right]\right)$$

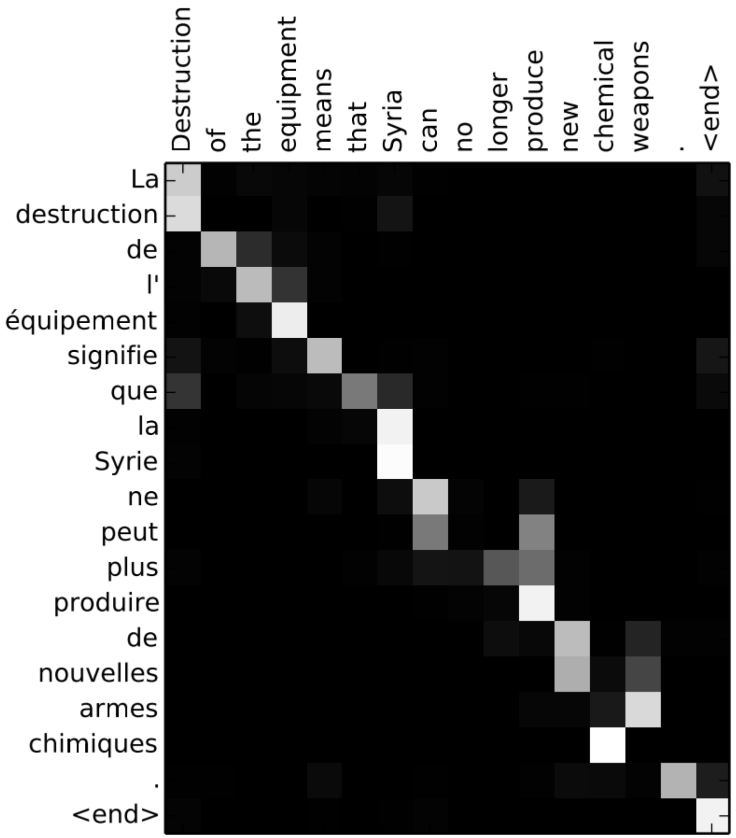


Luong, Minh-Thang, Hieu Pham, and Christopher D. Manning. "Effective approaches to attention-based neural machine translation." *arXiv preprint arXiv:1508.04025* (2015).

Modern advanced RNN

Results of the attention mechanism





https://devblogs.nvidia.com/introduction-neural-machine-translation-gpus-part-3/figure6_sample_translations-2/

Modern advanced RNN

What is the **attention**?

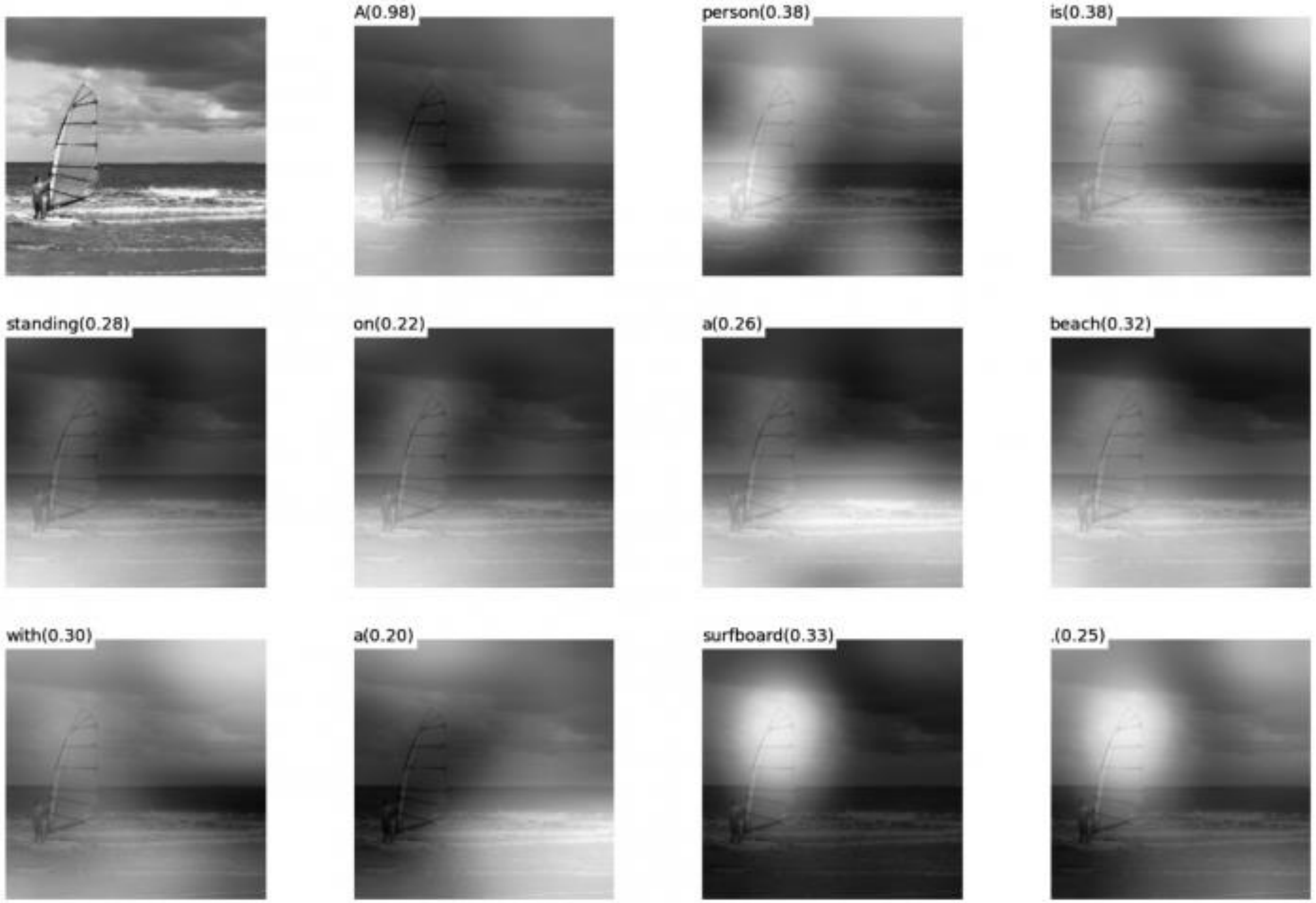
While attention is typically thought of as an orienting mechanism for perception, its “spotlight” can also be focused internally, toward the contents of memory. This idea, a recent focus in neuroscience studies, has also inspired work in AI. In some architectures, **attentional mechanisms have been used to select information to be read out from the internal memory of the network.** This has helped provide recent successes in machine translation and led to important advances on memory and reasoning tasks. These architectures offer a novel implementation of content-addressable retrieval, **which was itself a concept originally introduced to AI from neuroscience.**

Neuroscience-Inspired Artificial Intelligence

Demis Hassabis,[1,2,*] Dharshan Kumaran,[1,3] Christopher Summerfield,[1,4] and Matthew Botvinick[1,2]
[1]DeepMind, 5 New Street Square, London, UK
[2]Gatsby Computational Neuroscience Unit, 25 Howland Street, London, UK
[3]Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London, UK
[4]Department of Experimental Psychology, University of Oxford, Oxford, UK

Modern advanced RNN

Results of the attention mechanism



(b) A person is standing on a beach with a surfboard.

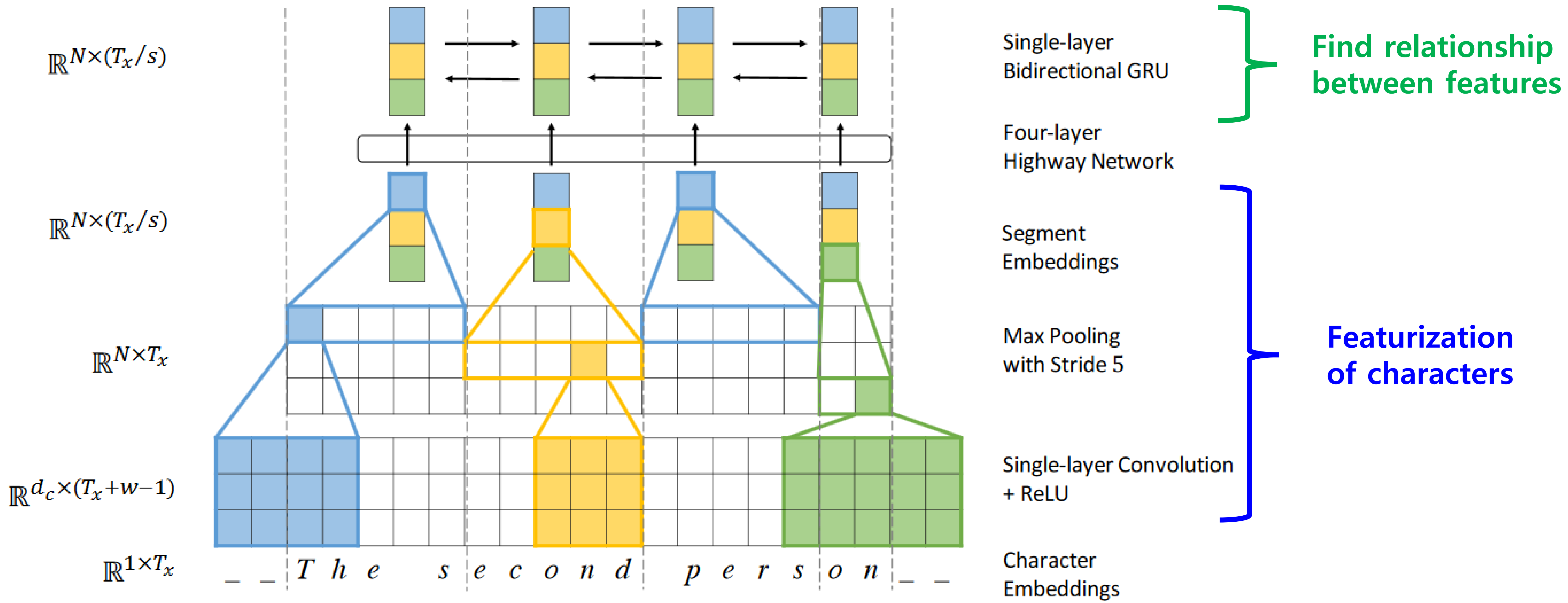
Modern advanced RNN

Results of the attention mechanism

A woman is throwing a frisbee in a park. A dog is standing on a hardwood floor. A stop sign is on a road with a mountain in the background.

Modern advanced RNN

Character level NMT

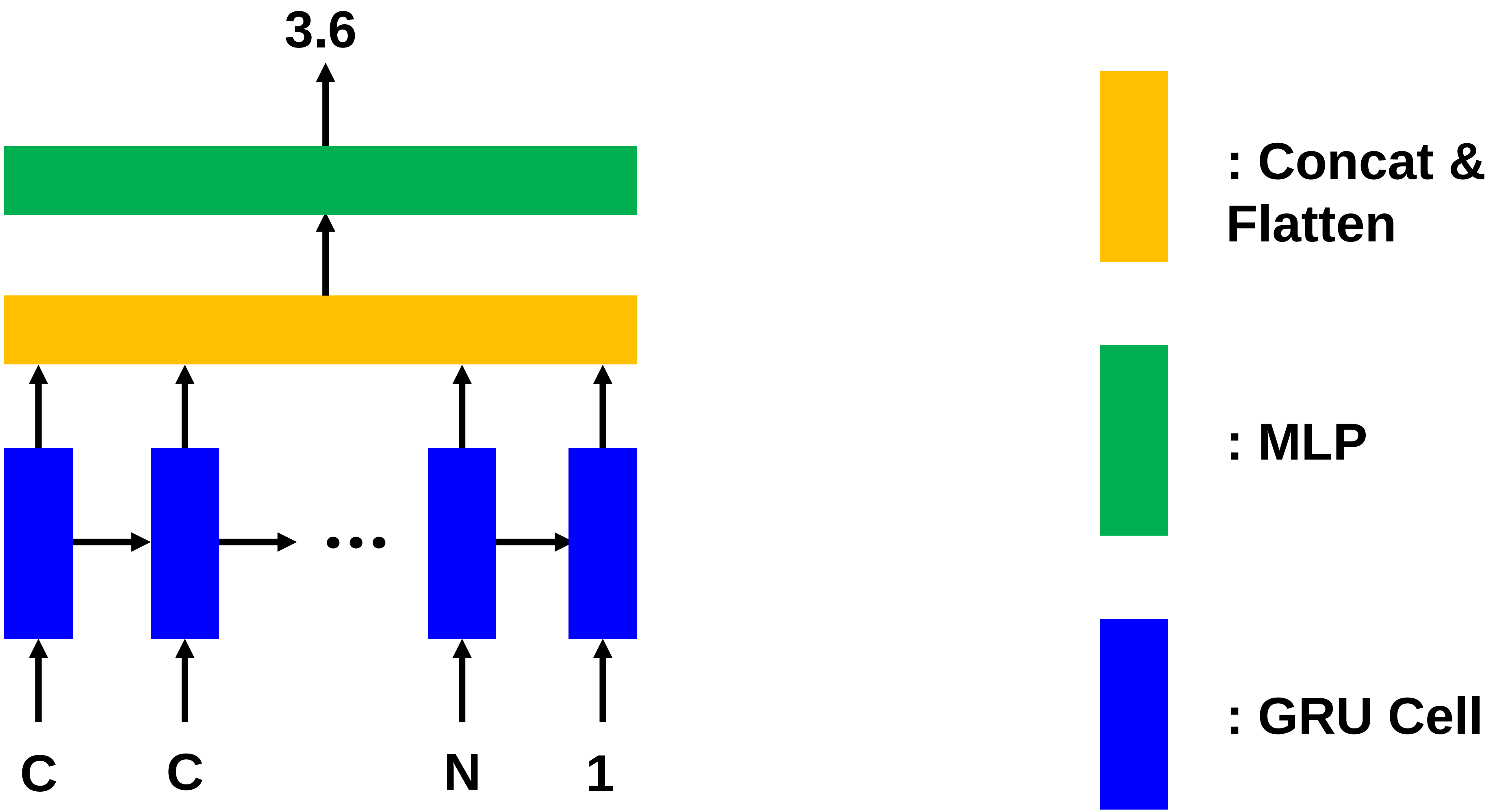


Lee, Jason, Kyunghyun Cho, and Thomas Hofmann. "Fully character-level neural machine translation without explicit segmentation." *arXiv preprint arXiv:1610.03017* (2016).

RNN to predict logP

RNN to predict logP

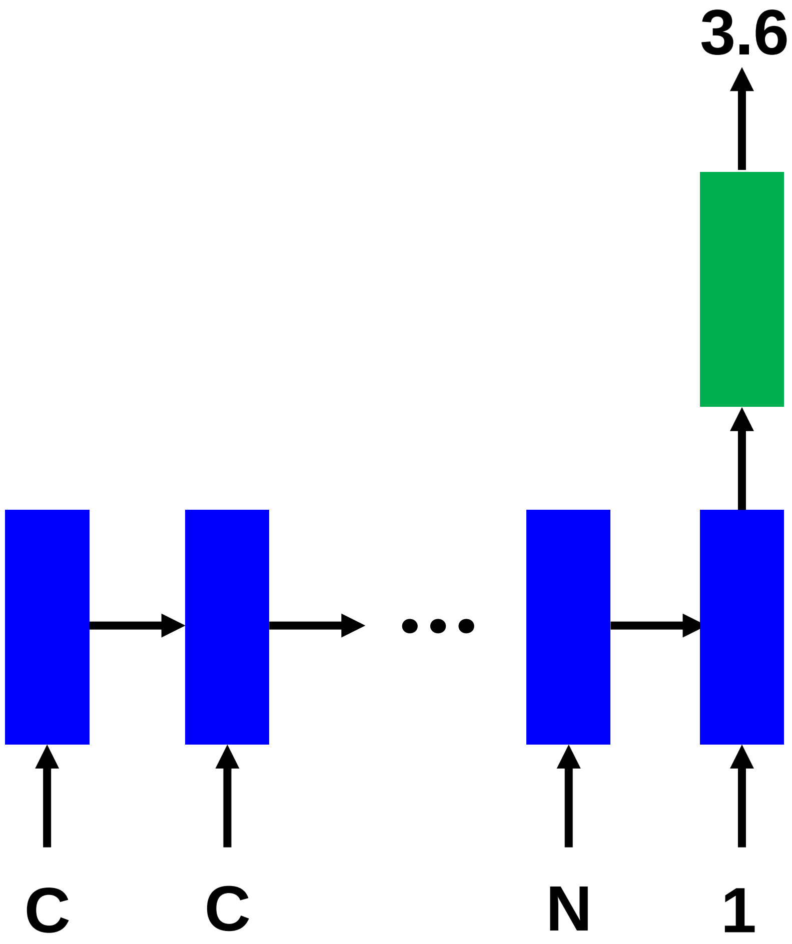
Model 1)

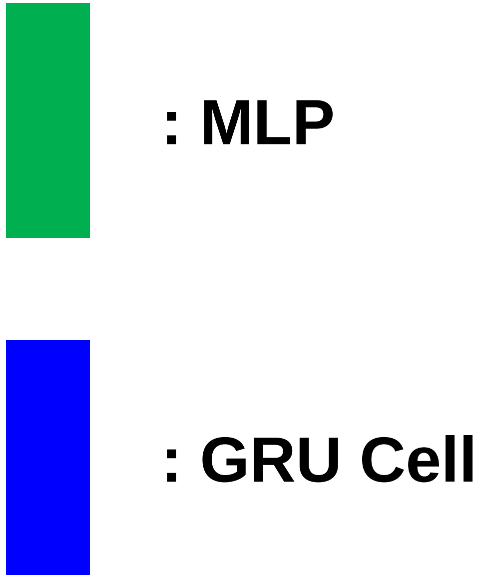


KAIST

RNN to predict logP

Model 2)

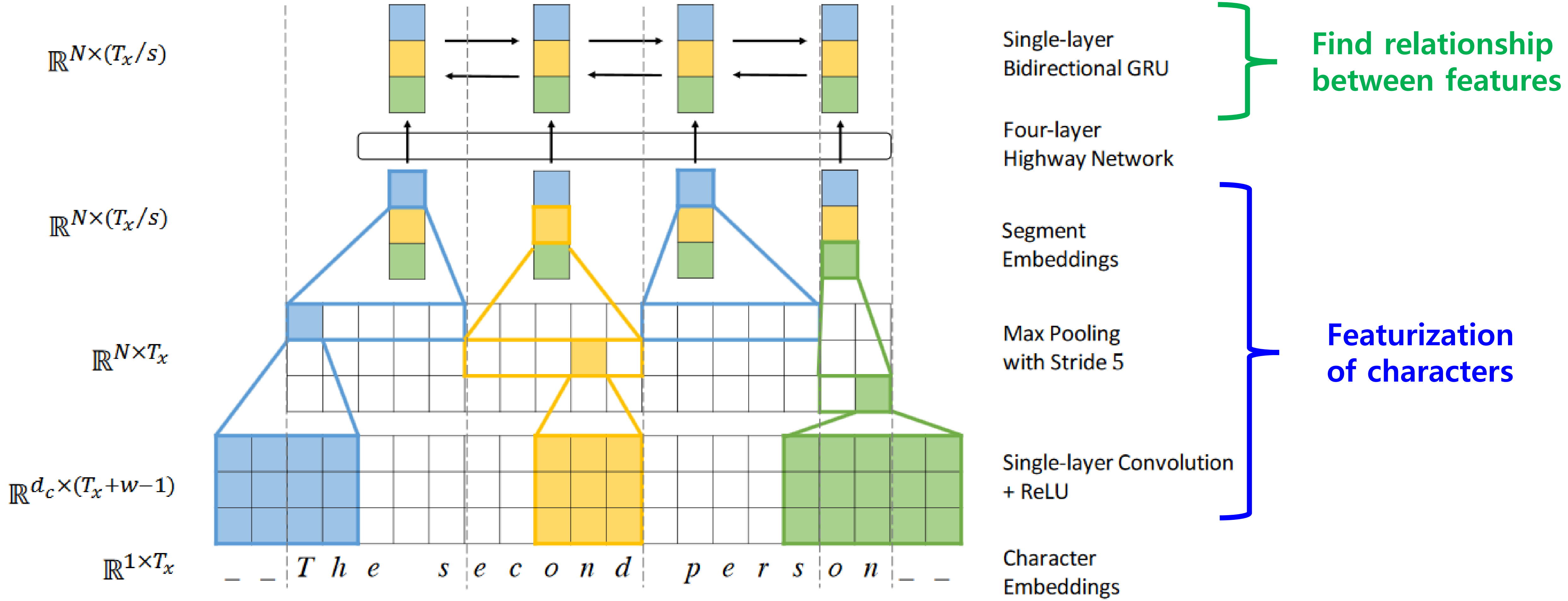




In this case, you don’t need a zero-padding to inputs.

Modern advanced RNN

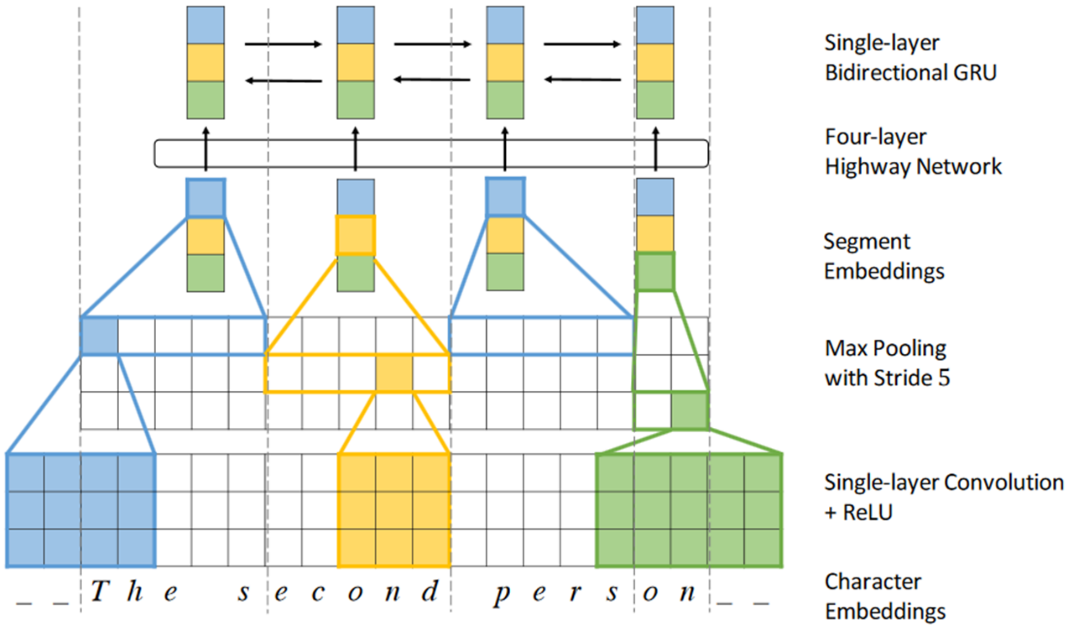
Character level NMT

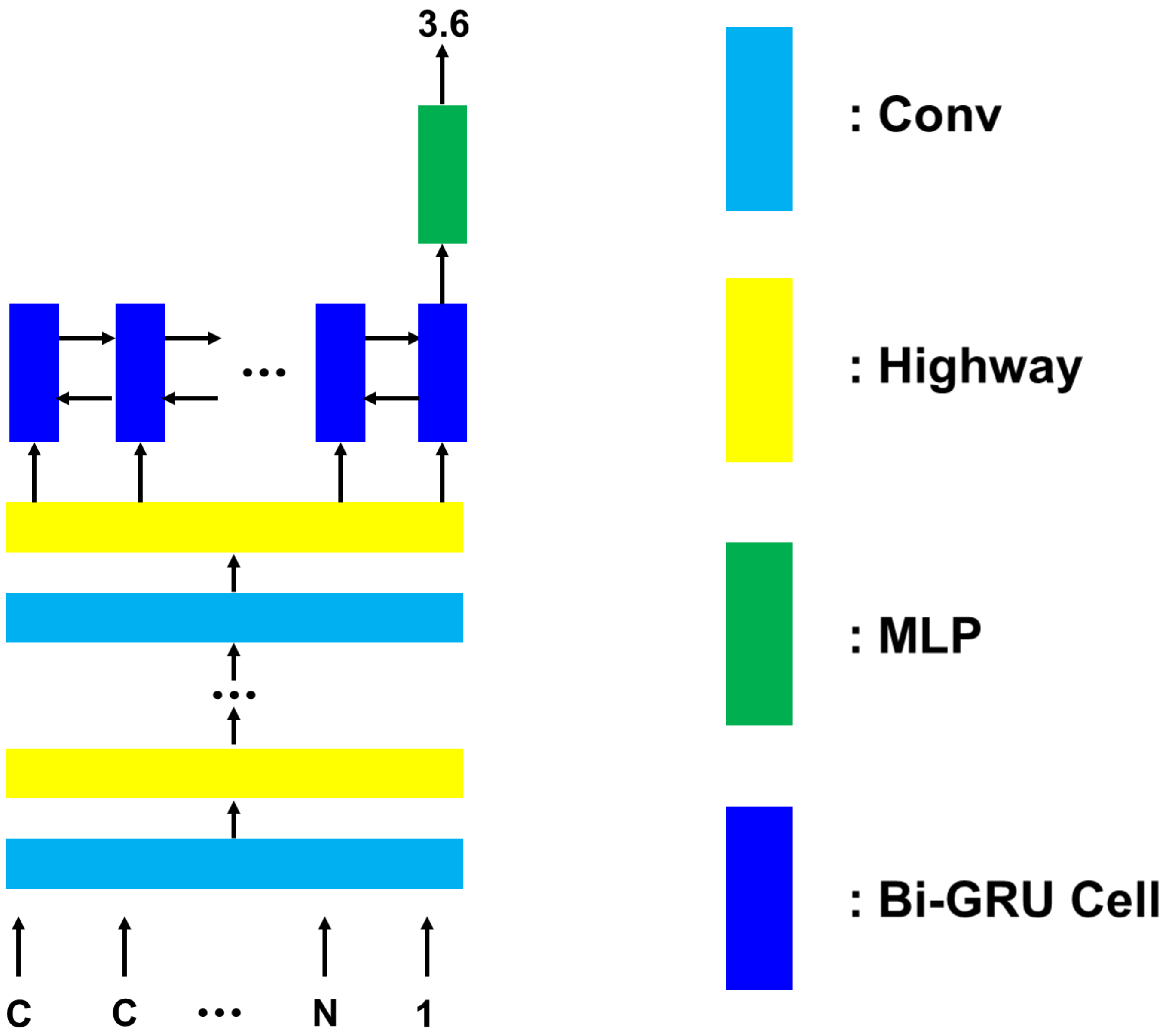


Lee, Jason, Kyunghyun Cho, and Thomas Hofmann. "Fully character-level neural machine translation without explicit segmentation." *arXiv preprint arXiv:1610.03017* (2016).

RNN to predict logP

Model 3)





Highway network

$$y = \mathbf{g} \odot \mathrm{ReLU}(\mathbf{W}_1\mathbf{x} + \mathbf{b}_1) + (1 - \mathbf{g}) \odot \mathbf{x}$$

$$\mathbf{g} = \sigma(\mathbf{W}_2\mathbf{x} + \mathbf{b}_2)$$

RNN to predict logP

Comparison: (input data representation, model architecture)

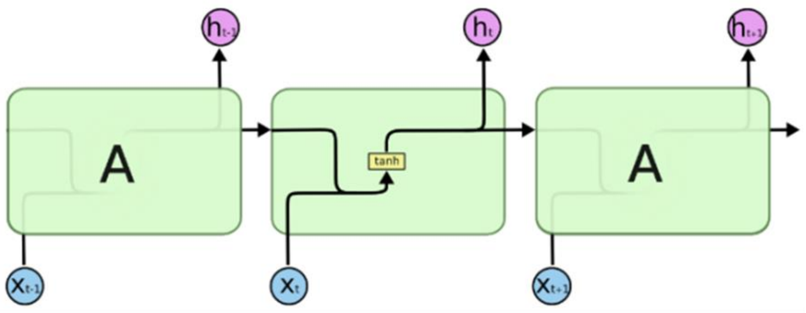
	FP - MLP	SMILES – CNN	Graph – GCN	SMILES – RNN1	SMILES – RNN2	SMILES – RNN3	…	My best (???)
MAE	0.31	0.15	0.088	0.13	0.05	0.072	…	0.01
Std.dev	0.42	0.20	0.137	0.18	0.08	0.11	…	-

Assignment #6

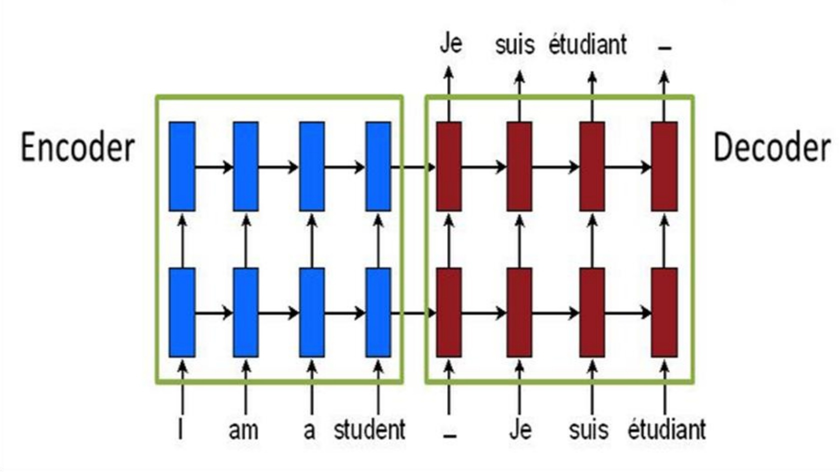
Build your own RNN for logP prediction

- In this class, TA showed the several RNN models for logP prediction
- In this week, we learned i) RNN, and ii) modern advanced RNNs.
- Therefore, **building your own RNN is an objective of this assignment.**
- **Report your results - MAE, std. dev, and truth-prediction plot.**
- **Your models do not need to be better than the TA's RNN model, improving the TA's model is not the goal of this assignment. However, please discuss why and how your neural network gives such results.**

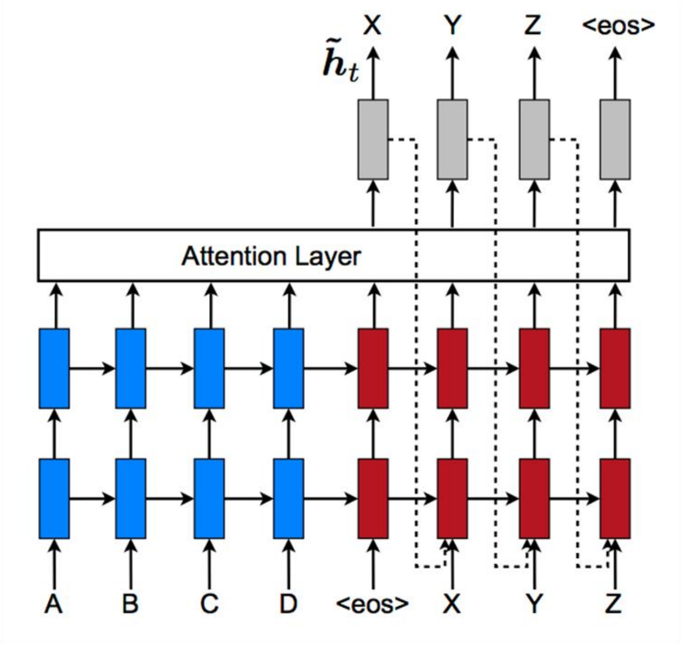
교수님이 수업에서 다루신 것



조교가 연습반에서 다룬 것



과제



Final term project