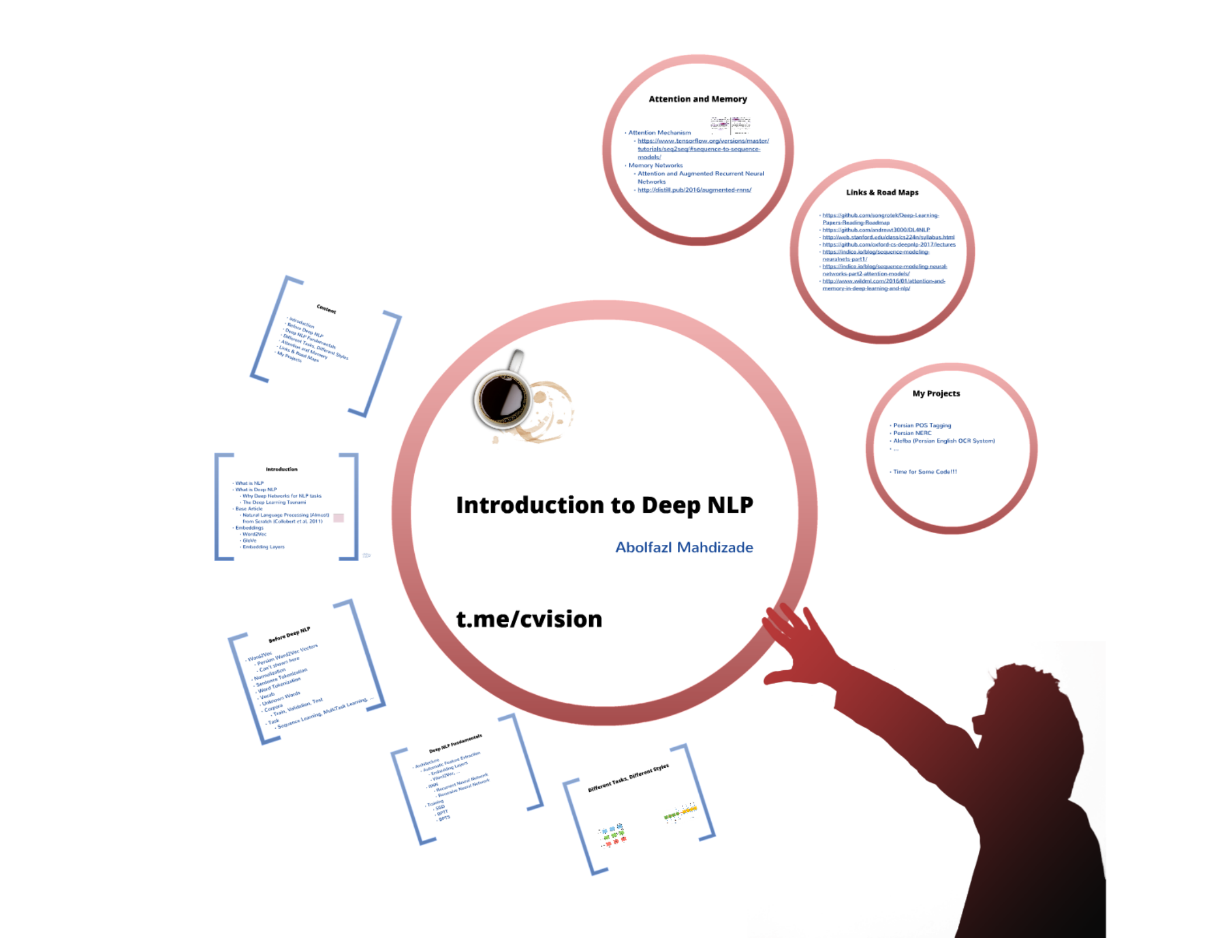
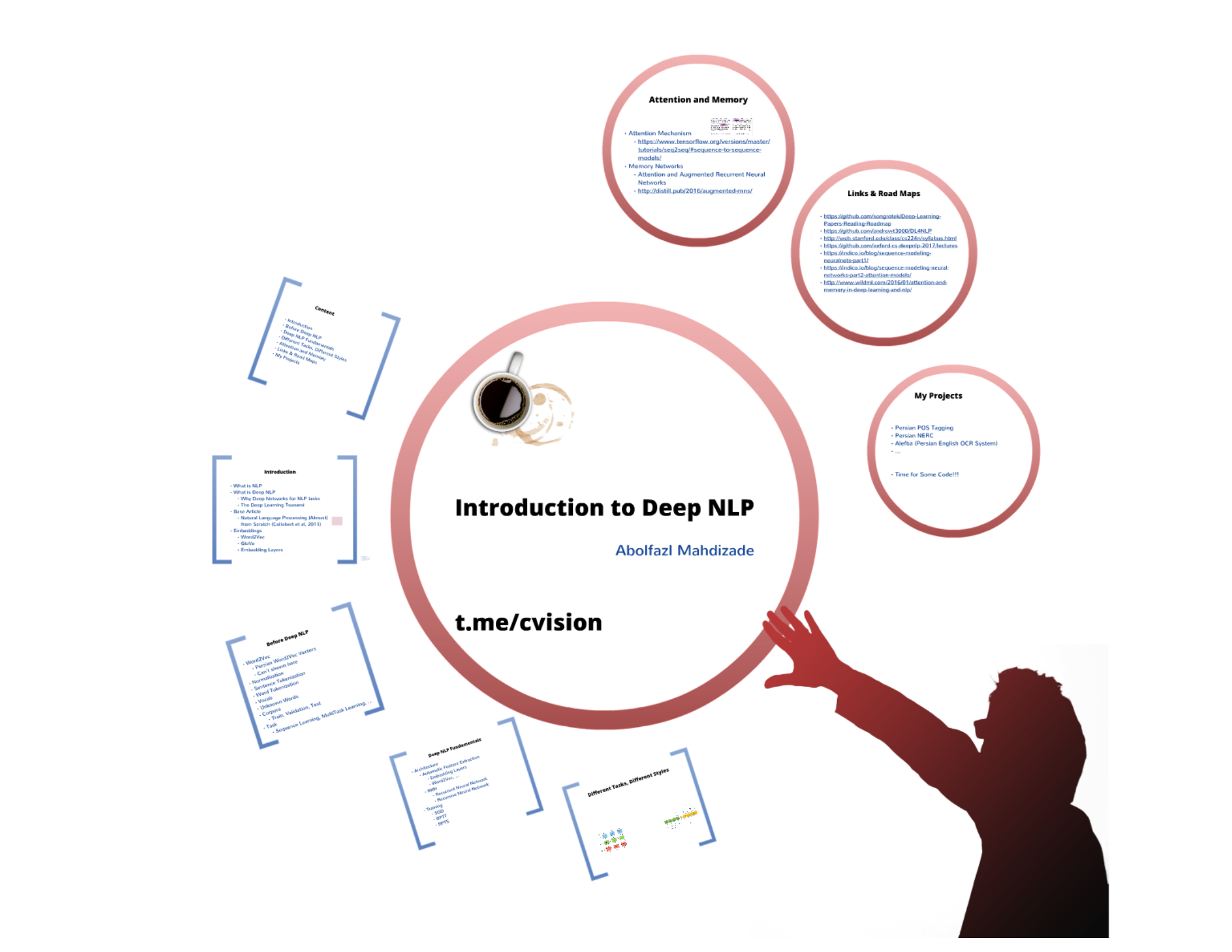

Attention and Memory
• Attention Mechanism
- https://www.tensorflow.org/versions/master/tutorials/seq2seq/#sequence-to-sequence-models/
• Memory Networks
- Attention and Augmented Recurrent Neural Networks
- http://distill.pub/2016/augmented-rnns/
Links & Road Maps
- https://github.com/songrotek/Deep-Learning-Papers-Reading-Roadmap
- https://github.com/andrewt3000/DL4NLP
- http://web.stanford.edu/class/cs224n/syllabus.html
- https://github.com/oxford-cs-deepnlp-2017/lectures
- https://indico.io/blog/sequence-modeling-neuralnets-part1/
- https://indico.io/blog/sequence-modeling-neural-networks-part2-attention-models/
- http://www.wildml.com/2016/01/attention-and-memory-in-deep-learning-and-nlp/
My Projects
- Persian POS Tagging
- Persian NERC
- Alefba (Persian English OCR System)
- ...
- Time for Some Code!!!
Content
- Introduction
- Before Deep NLP
- Deep NLP Fundamentals
- Different Tasks, Different Styles
- Attention and Memory
- Links & Road Maps
- My Projects
Introduction
- What is NLP
- What is Deep NLP
- Why Deep Networks for NLP tasks
- The Deep Learning Tsunami
- Base Article
- Natural Language Processing (Almost) from Scratch (Collobert et al. 2011)
- Embeddings
- Word2Vec
- GloVe
- Embedding Layers
Introduction to Deep NLP
Abolfazl Mahdizade
t.me/cvision
Before Deep NLP
- Word2Vec
- Persian Word2Vec Vectors
- Can't shown here
- Normalization
- Sentence Tokenization
- Word Tokenization
- Vocab
- Unknown Words
- Corpora
- Train, Validation, Test
- Task
- Sequence Learning, MultiTask Learning, ...
Deep NLP Fundamentals
- Architecture
- Automatic Feature Extraction
- Embedding Layers
- Word2Vec, ...
- RNN
- Recurrent Neural Network
- Recursive Neural Network
- Training
- SGD
- BPTT
- BPTS
Different Tasks, Different Styles


Attention and Memory
• Attention Mechanism
• https://www.tensorflow.org/versions/master/tutorials/seq2seq/#sequence-to-sequence-models/
• Memory Networks
• Attention and Augmented Recurrent Neural Networks
• http://distill.pub/2016/augmented-rnns/
Links & Road Maps
• https://github.com/songrotek/Deep-Learning-Papers-Reading-Roadmap
• https://github.com/andrewt3000/DL4NLP
• http://web.stanford.edu/class/cs224n/syllabus.html
• https://github.com/oxford-cs-deepnlp-2017/lectures
• https://indico.io/blog/sequence-modeling-neuralnets-part1/
• https://indico.io/blog/sequence-modeling-neural-networks-part2-attention-models/
• http://www.wildml.com/2016/01/attention-and-memory-in-deep-learning-and-nlp/
My Projects
• Persian POS Tagging
• Persian NERC
• Alefba (Persian English OCR System)
• ...
• Time for Some Code!!!
Content
• Introduction
• Before Deep NLP
• Deep NLP Fundamentals
• Different Tasks, Different Styles
• Attention and Memory
• Links & Road Maps
• My Projects
Introduction
• What is NLP
• What is Deep NLP
• Why Deep Networks for NLP tasks
• The Deep Learning Tsunami
• Base Article
• Natural Language Processing (Almost) from Scratch (Collobert et al, 2011)
• Embeddings
• Word2Vec
• GloVe
• Embedding Layers
Introduction to Deep NLP
Abolfazl Mahdizade
t.me/cvision
Before Deep NLP
• Word2Vec
• Persian Word2Vec Vectors
• Can't shown here
• Normalization
• Sentence Tokenization
• Word Tokenization
• Vocab
• Unknown Words
• Corpora
• Train, Validation, Test
• Task
• Sequence Learning, MultiTask Learning, ...
Deep NLP Fundamentals
• Architecture
• Automatic Feature Extraction
• Embedding Layers
• Word2Vec, ...
• RNN
• Recurrent Neural Network
• Recursive Neural Network
• Training
• SGD
• BPTT
• BPTS
Different Tasks, Different Styles

Introduction to Deep NLP

Abolfazl Mahdizade

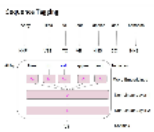
t.me/cvision

Content

- Introduction
- Before Deep NLP
- Deep NLP Fundamentals
- Different Tasks, Different Styles
- Attention and Memory
- Links & Road Maps
- My Projects

Introduction

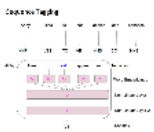
- What is NLP
- What is Deep NLP
  - Why Deep Networks for NLP tasks
  - The Deep Learning Tsunami
- Base Article
  - Natural Language Processing (Almost) from Scratch (Collobert et al, 2011)
- Embeddings
  - Word2Vec
  - GloVe
  - Embedding Layers



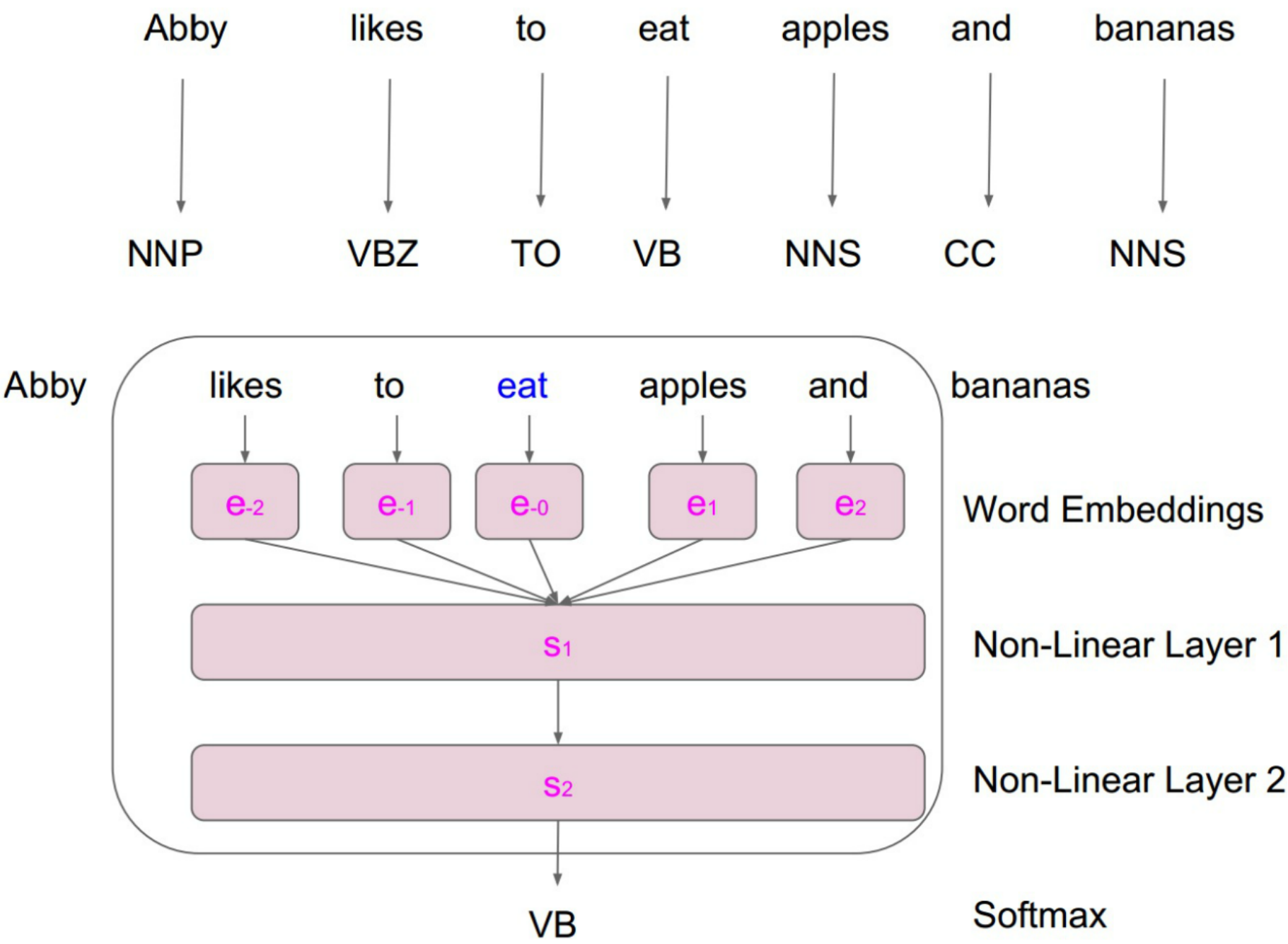
- The Deep Learning Tsunami
  - Christopher D. Manning:
  - Deep Learning waves have lapped at the shores of computational linguistics for several years now, but 2015 seems like the year when the full force of the tsunami hit the major Natural Language Processing (NLP) conferences.

Introduction

- What is NLP
- What is Deep NLP
  - Why Deep Networks for NLP tasks
  - The Deep Learning Tsunami
- Base Article
  - Natural Language Processing (Almost) from Scratch (Collobert et al, 2011)
- Embeddings
  - Word2Vec
  - GloVe
  - Embedding Layers

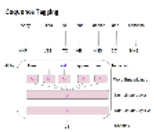


Sequence Tagging



Introduction

- What is NLP
- What is Deep NLP
  - Why Deep Networks for NLP tasks
  - The Deep Learning Tsunami
- Base Article
  - Natural Language Processing (Almost) from Scratch (Collobert et al, 2011)
- Embeddings
  - Word2Vec
  - GloVe
  - Embedding Layers

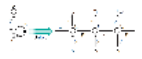


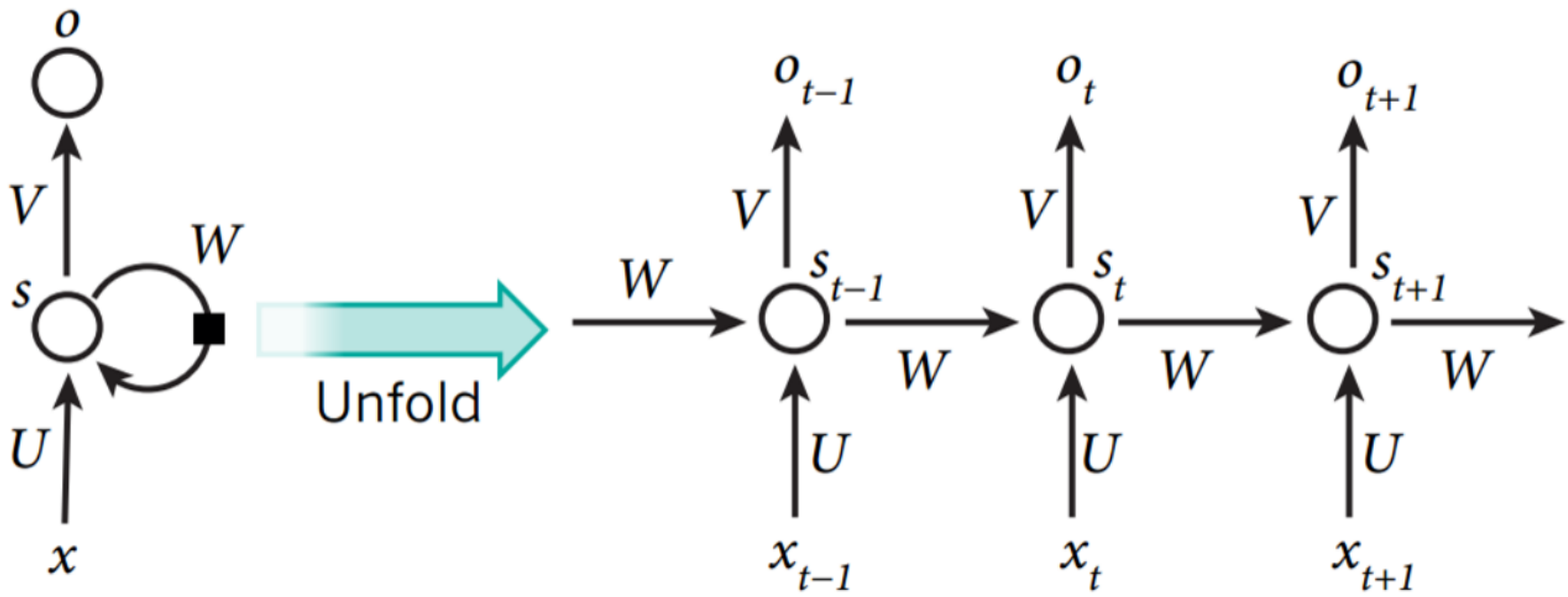
Before Deep NLP

- Word2Vec
  - Persian Word2Vec Vectors
  - Can't shown here
- Normalization
- Sentence Tokenization
- Word Tokenization
- Vocab
- Unknown Words
- Corpora
  - Train, Validation, Test
- Task
  - Sequence Learning, MultiTask Learning, ...

Deep NLP Fundamentals

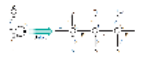
- Architecture
  - Automatic Feature Extraction
    - Embedding Layers
    - Word2Vec, ...
  - RNN
    - Recurrent Neural Network
    - Recursive Neural Network
- Training
  - SGD
  - BPTT
  - BPTS



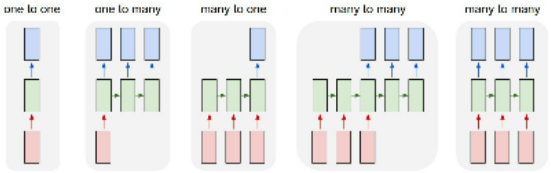

o
V
W
s
U
x
Unfold
W
o t−1
V
s t−1
U
x t−1
W
o t
V
s t
U
x t
W
o t+1
V
s t+1
U
x t+1
W

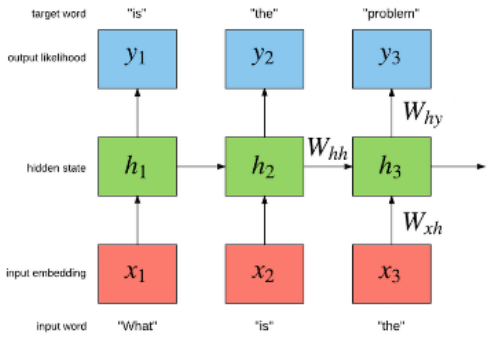
Deep NLP Fundamentals

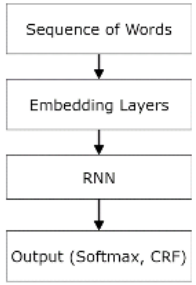
- Architecture
  - Automatic Feature Extraction
    - Embedding Layers
    - Word2Vec, ...
  - RNN
    - Recurrent Neural Network
    - Recursive Neural Network
- Training
  - SGD
  - BPTT
  - BPTS

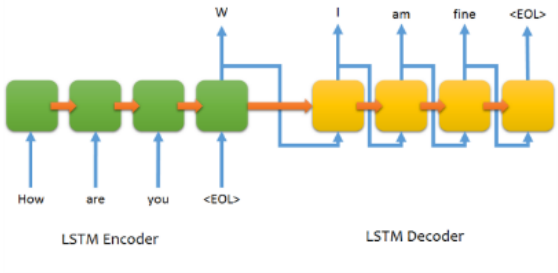


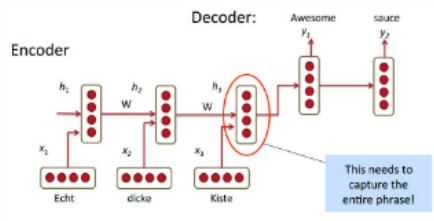
Different Tasks, Different Styles

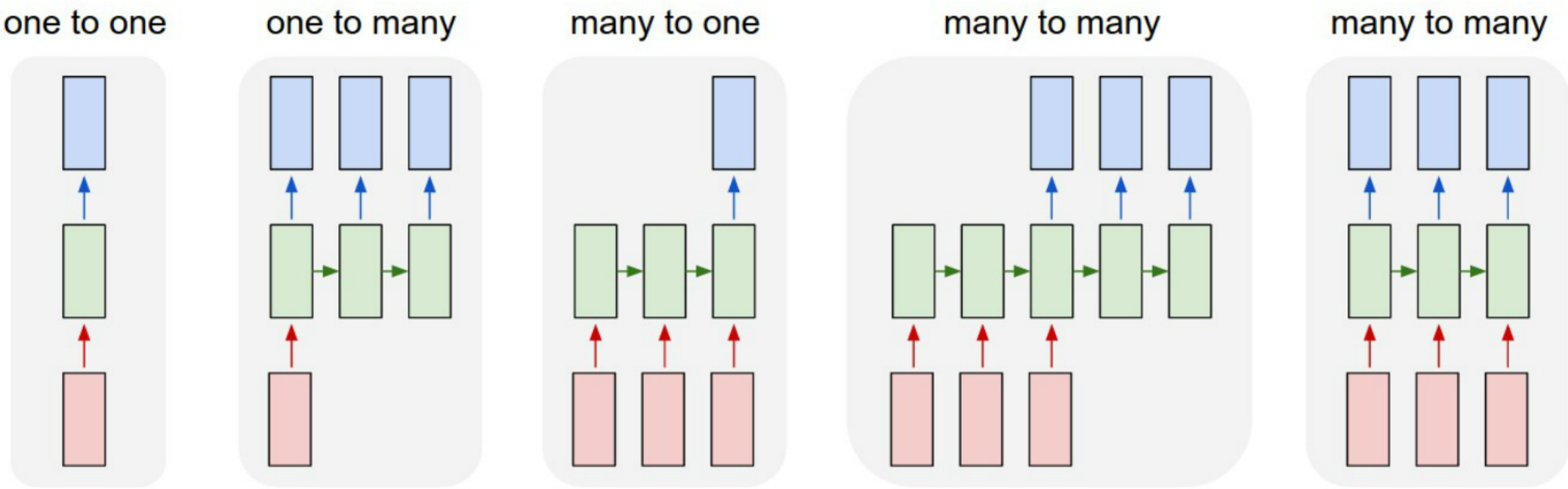


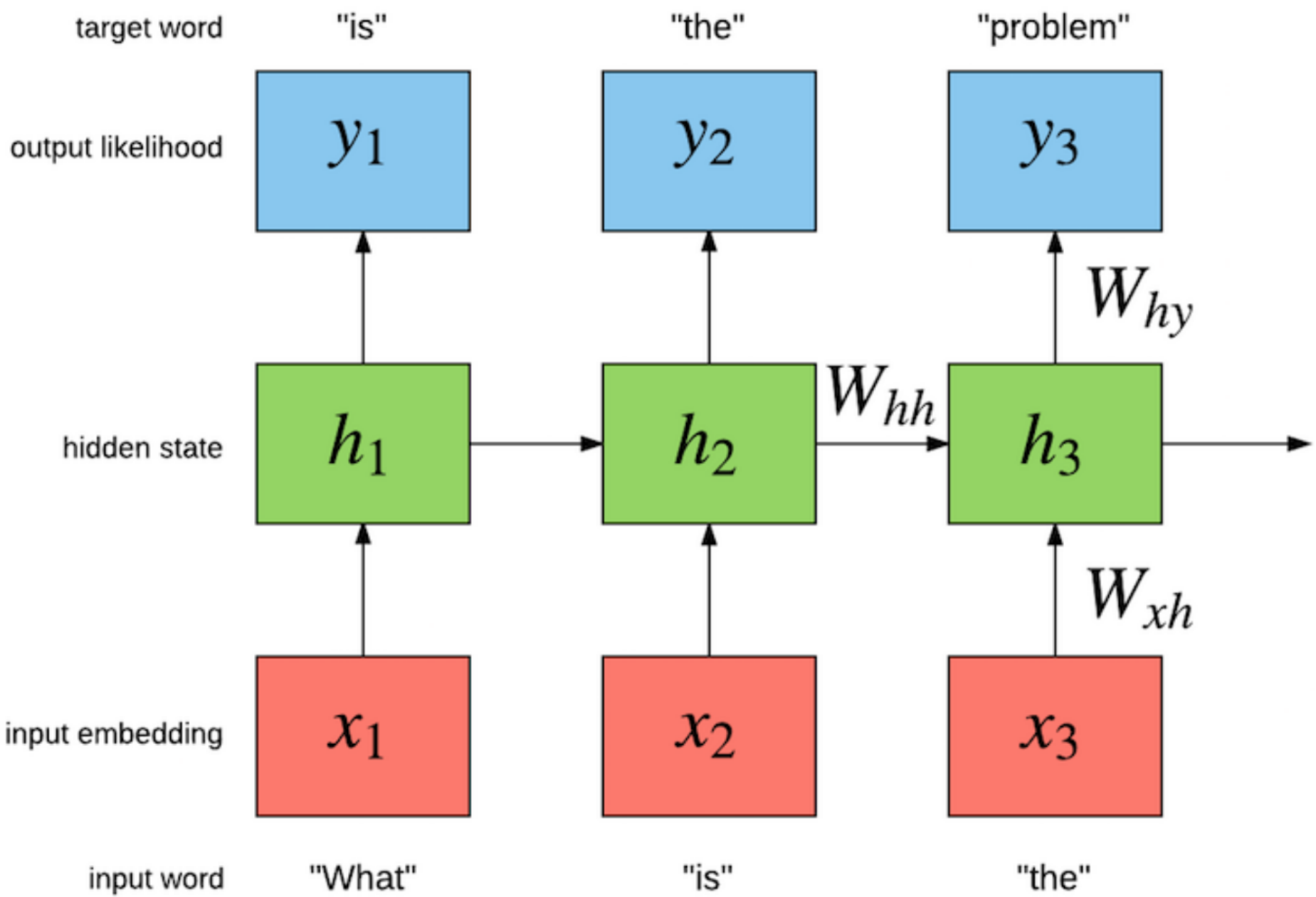


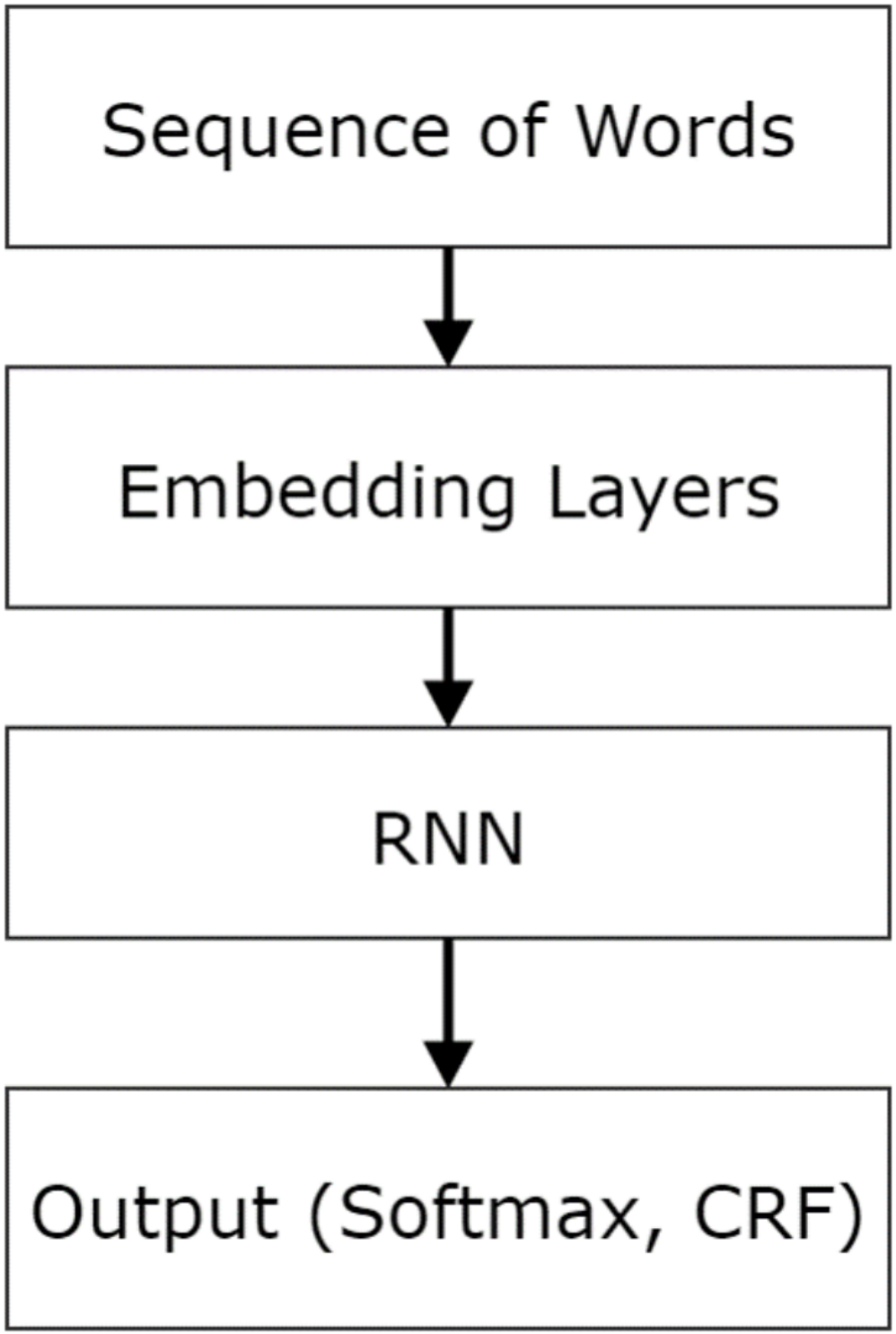


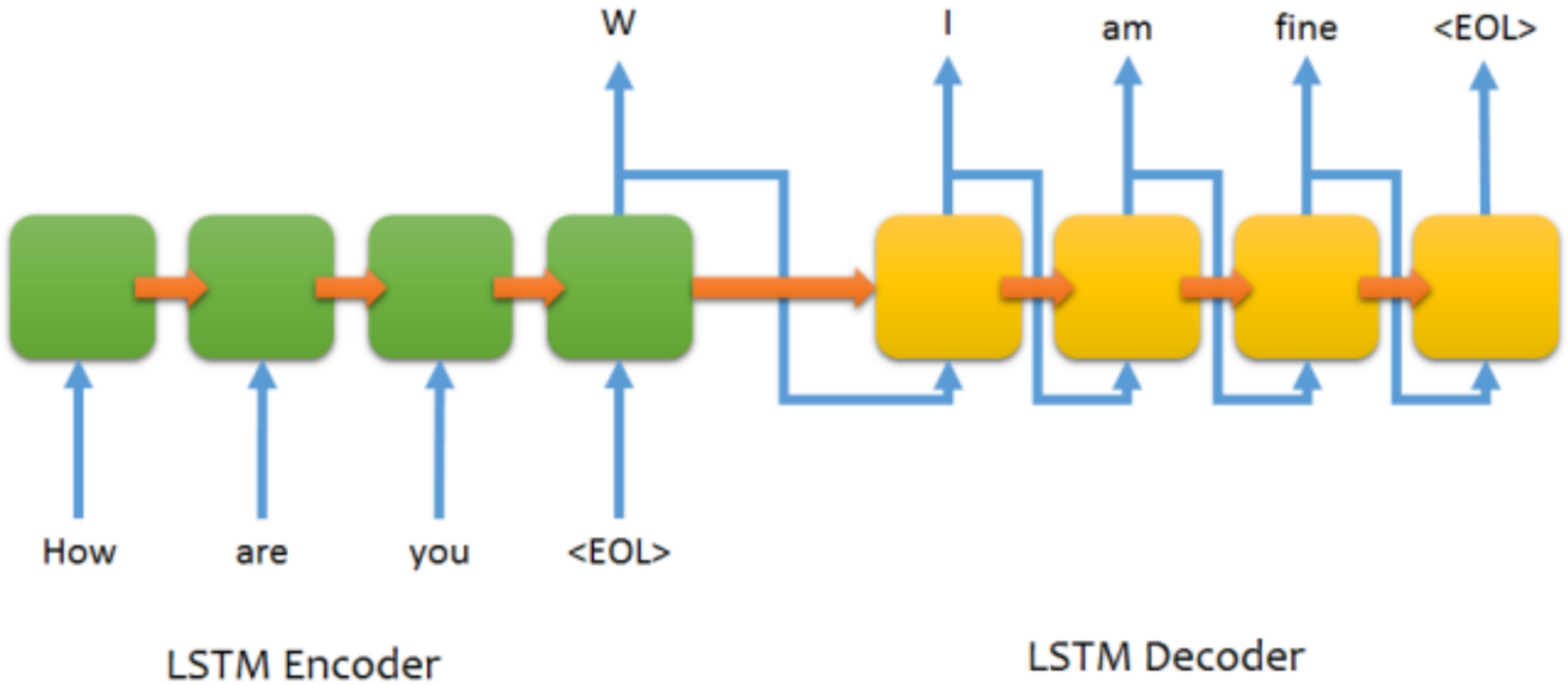


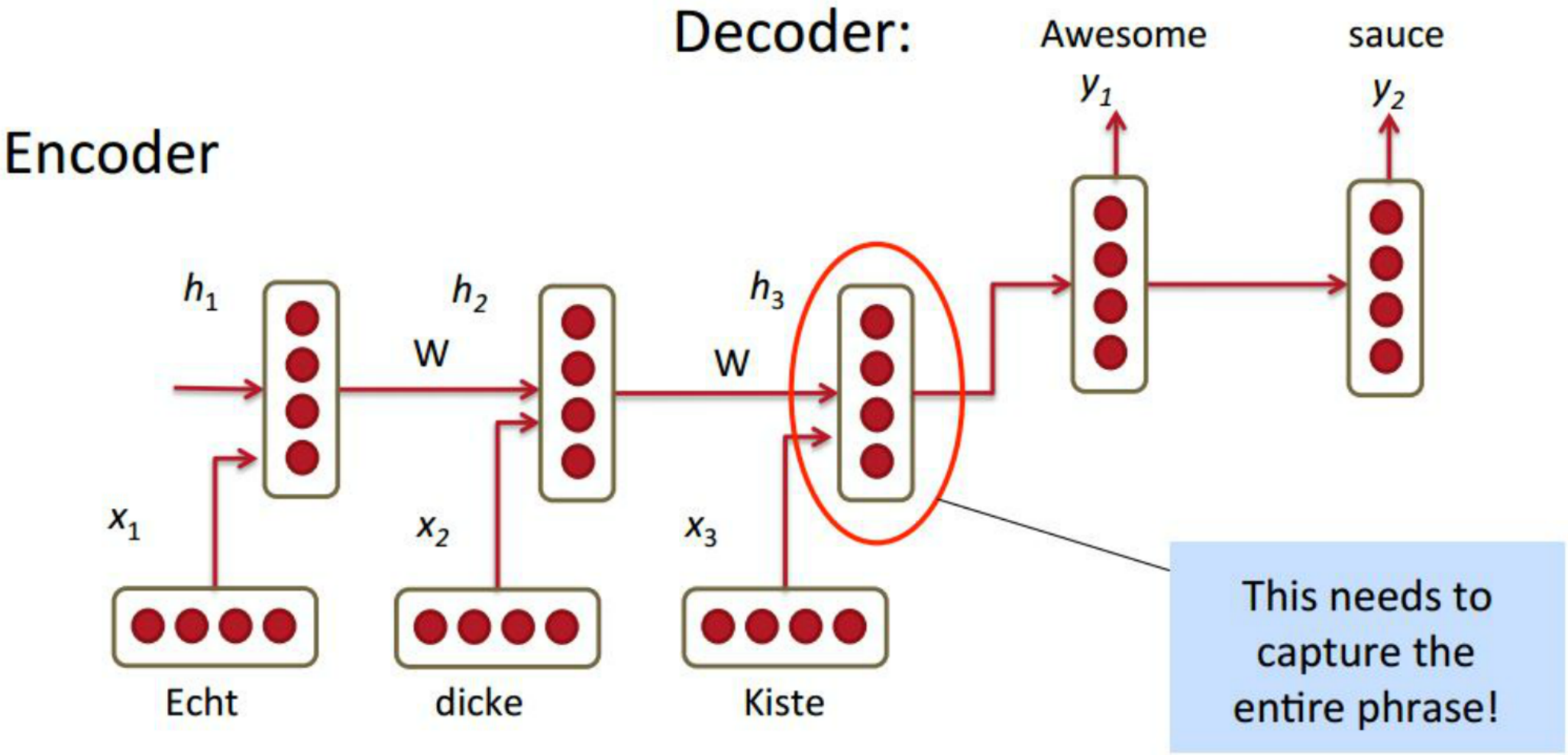



one to one
one to many
many to one
many to many
many to many

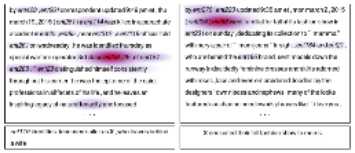

target word
"is"
"the"
"problem"
output likelihood
y_1
y_2
y_3
W_{hy}
W_{hh}
hidden state
h_1
h_2
h_3
W_{xh}
input embedding
x_1
x_2
x_3
input word
"What"
"is"
"the"


Sequence of Words
Embedding Layers
RNN
Output (Softmax, CRF)


W
I
am
fine
<EOL>
How
are
you
<EOL>
LSTM Encoder
LSTM Decoder


Decoder:
Awesome
sauce
y_1
y_2
Encoder
h_1
h_2
h_3
W
W
x_1
x_2
x_3
Echt
dicke
Kiste
This needs to capture the entire phrase!

Attention and Memory



- Attention Mechanism
  - https://www.tensorflow.org/versions/master/tutorials/seq2seq/#sequence-to-sequence-models/
- Memory Networks
  - Attention and Augmented Recurrent Neural Networks
  - http://distill.pub/2016/augmented-rnns/

by *ent423* , *ent261* correspondent updated 9:49 pm et , thu march 19 , 2015 (*ent261*) a *ent114* was killed in a parachute accident in *ent45* , *ent85* , near *ent312* , a *ent119* official told *ent261* on wednesday . he was identified thursday as special warfare operator 3rd class *ent23* , 29 , of *ent187* , *ent265* . `` *ent23* distinguished himself consistently throughout his career . he was the epitome of the quiet professional in all facets of his life , and he leaves an inspiring legacy of natural tenacity and focused

. . .

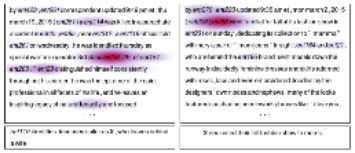
ent119 identifies deceased sailor as **X** , who leaves behind a wife

by *ent270* , *ent223* updated 9:35 am et , mon march 2 , 2015 (*ent223*) *ent63* went familial for fall at its fashion show in *ent231* on sunday , dedicating its collection to `` mamma '' with nary a pair of `` mom jeans '' in sight . *ent164* and *ent21* , who are behind the *ent196* brand , sent models down the runway in decidedly feminine dresses and skirts adorned with roses , lace and even embroidered doodles by the designers ' own nieces and nephews . many of the looks featured saccharine needlework phrases like `` i love you ,

. . .

X dedicated their fall fashion show to moms

Attention and Memory



- Attention Mechanism
  - https://www.tensorflow.org/versions/master/tutorials/seq2seq/#sequence-to-sequence-models/
- Memory Networks
  - Attention and Augmented Recurrent Neural Networks
  - http://distill.pub/2016/augmented-rnns/

Links & Road Maps

- https://github.com/songrotek/Deep-Learning-Papers-Reading-Roadmap
- https://github.com/andrewt3000/DL4NLP
- http://web.stanford.edu/class/cs224n/syllabus.html
- https://github.com/oxford-cs-deepnlp-2017/lectures
- https://indico.io/blog/sequence-modeling-neuralnets-part1/
- https://indico.io/blog/sequence-modeling-neural-networks-part2-attention-models/
- http://www.wildml.com/2016/01/attention-and-memory-in-deep-learning-and-nlp/

My Projects

- Persian POS Tagging
- Persian NERC
- Alefba (Persian English OCR System)
- ...

- Time for Some Code!!!