ADVANCED RNNS

ATTENTION

martin @ reddragon.ai
sam @ reddragon.ai

8 November 2017

`WiFi : SG-Guest`

Problems with Installation? **ASK!**

RED DRAGON AI

PLAN OF ACTION

TODAY

- Confirm : 27-Nov is Deepavali make-up
- Attention : AIAYN, captioning
- CNNs for text
- Personal Project work

PLAN OF ACTION

MONDAY

- GANs
- Personal Project work

PLAN OF ACTION

NEXT WEDNESDAY

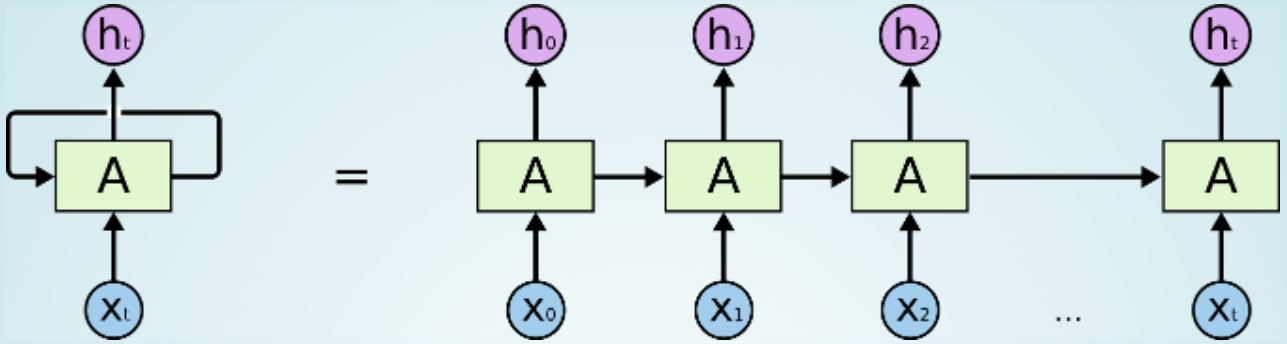
- Mobile Deep Learning
- Personal Project work

PROCESSING SEQUENCES

VARIABLE-LENGTH INPUT DOESN'T "FIT"

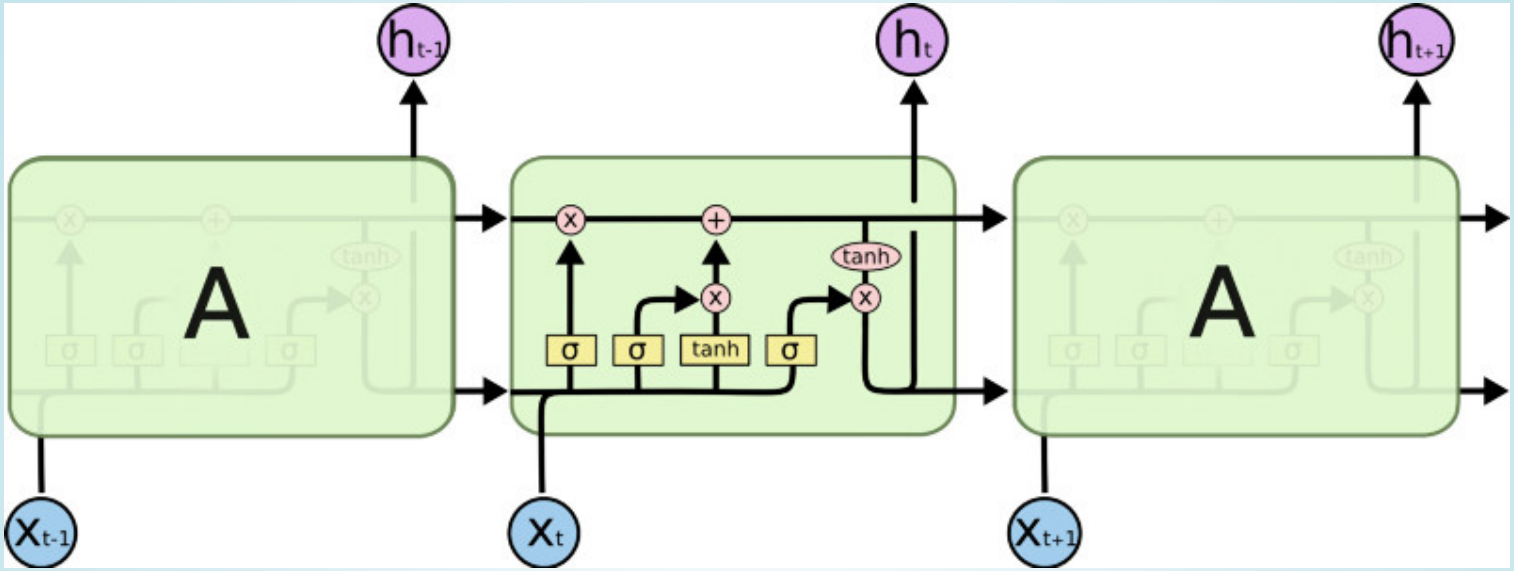
- Run network for each timestep
  - ... with the same parameters
- But 'pass along' internal state
- This state is 'hidden depth'
  - ... and should learn features that are useful
  - ... because everything is differentiable

BASIC RNN



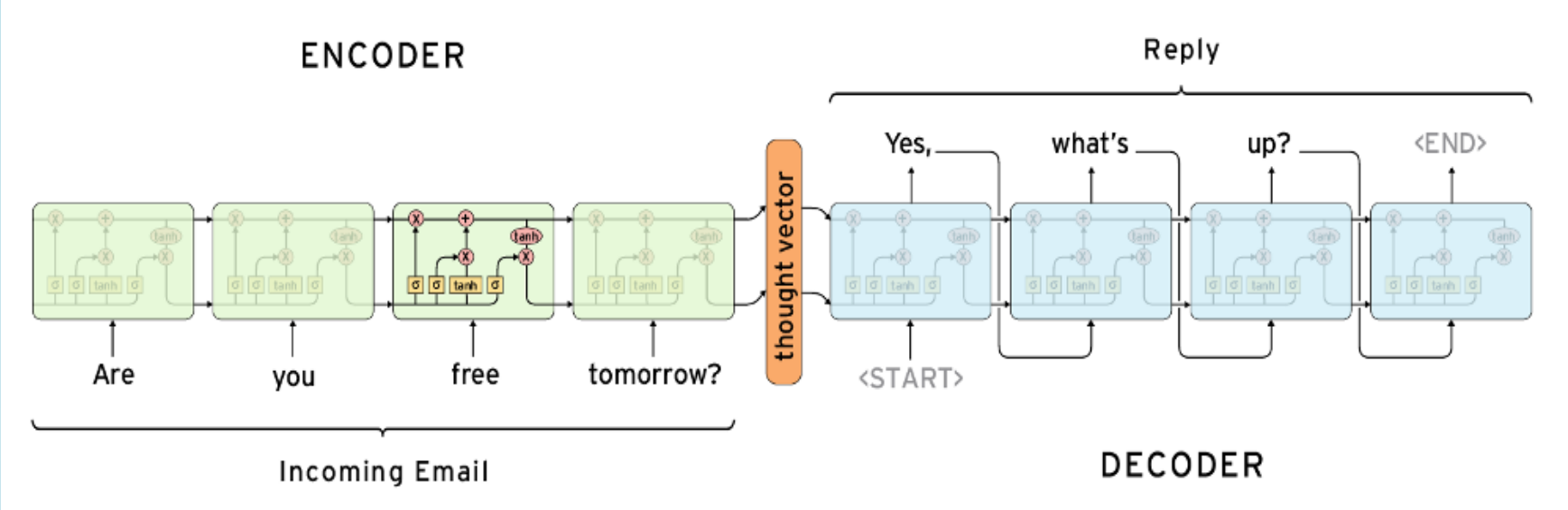
RNN chain : RNN overview

BASIC LSTM



LSTM chain : Understanding LSTMs

BASIC SEQ2SEQ

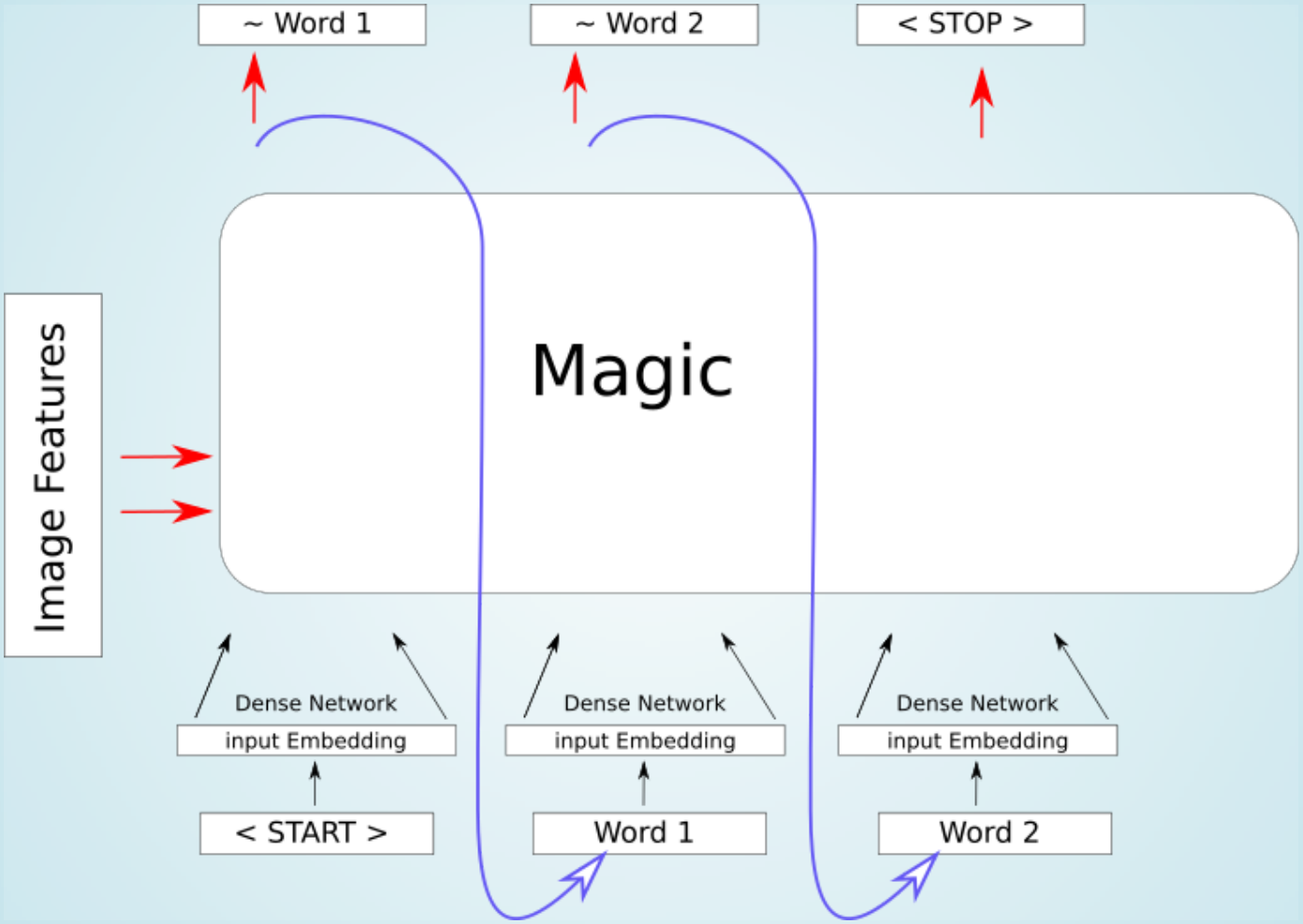


Practical seq2seq

SEQUENCES FROM NETWORKS

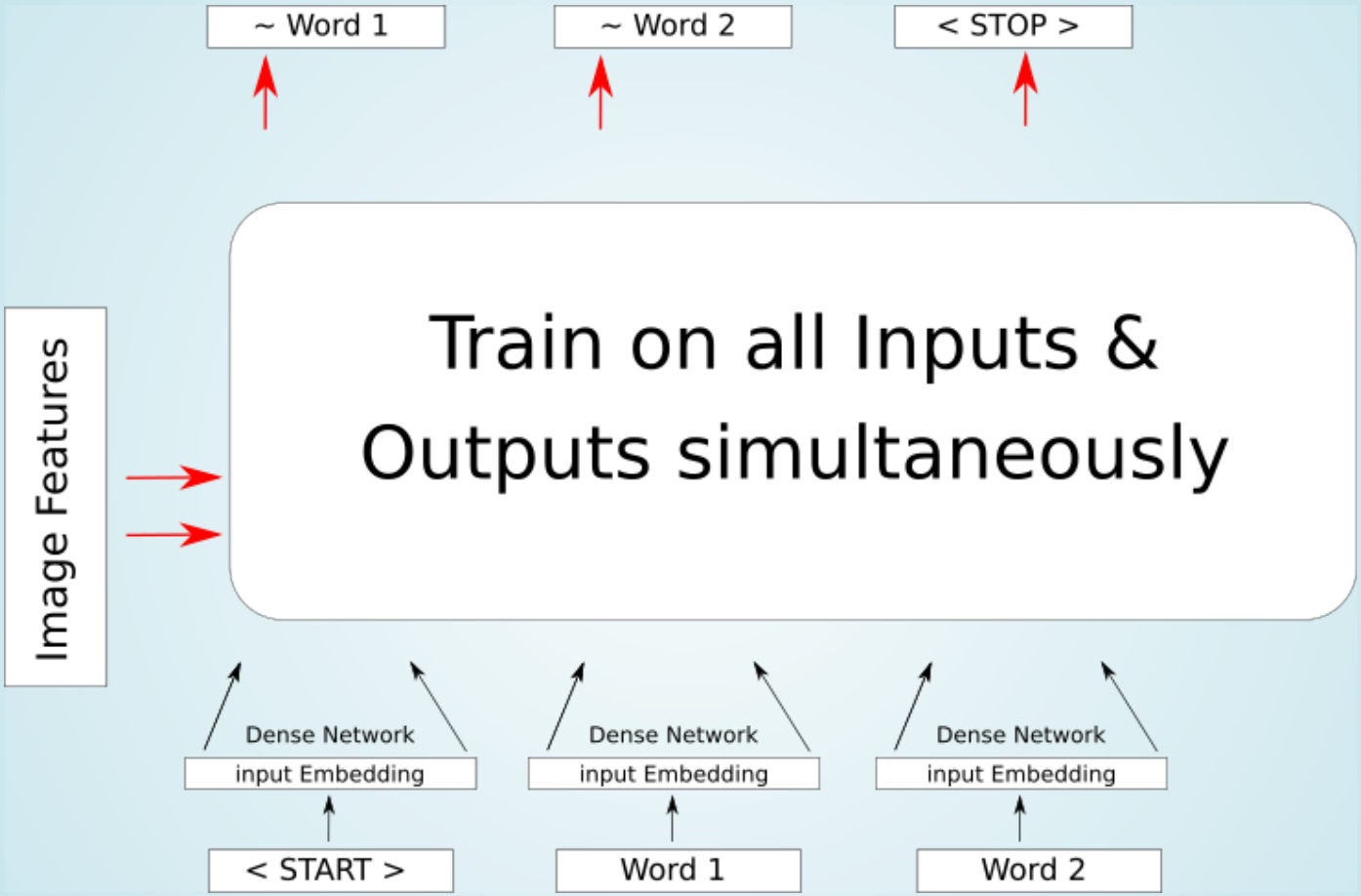
- Word-by-word (Test)
- Teacher forcing (Training)
- Embedding choices

GENERATING SEQUENCES



Basic Layout : Test Time

TRAINING TIME



"Teacher Forcing"

EMBEDDING CHOICES

- Word Vector
- One-Hot embedding
- Use each word's numeric index

WORD VECTORS

- Fixed dimension, independent of vocab size
- Stop words may be 'murky'
- Action words need definitions
- Often used as input stage

ONE-HOT

- Vocab ~7k ⇒ `vector.len == 7000`
- Very high number of 0/1 inputs
- Often used as output stage
- `idx = ArgMax( Softmax() )`

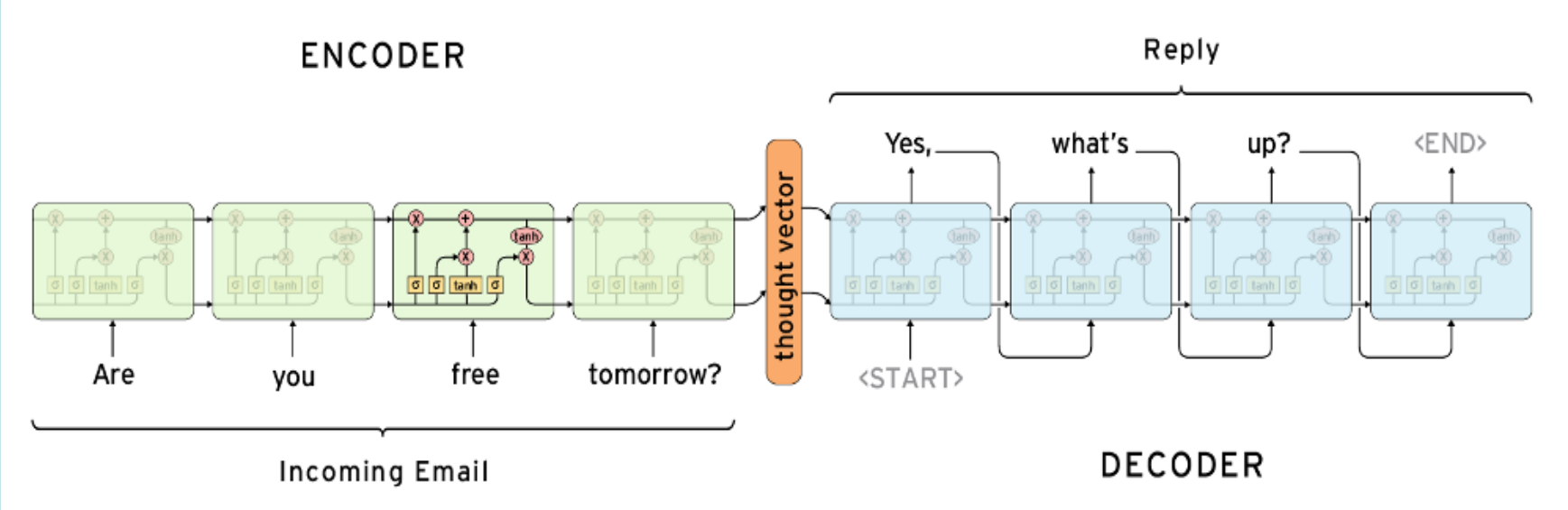
BINARY INDEX

- Low dimensionality
  - 14 binary digits for 7k vocabulary
- Difficult to believe it works
- Add resilience using ECC

COMBOS

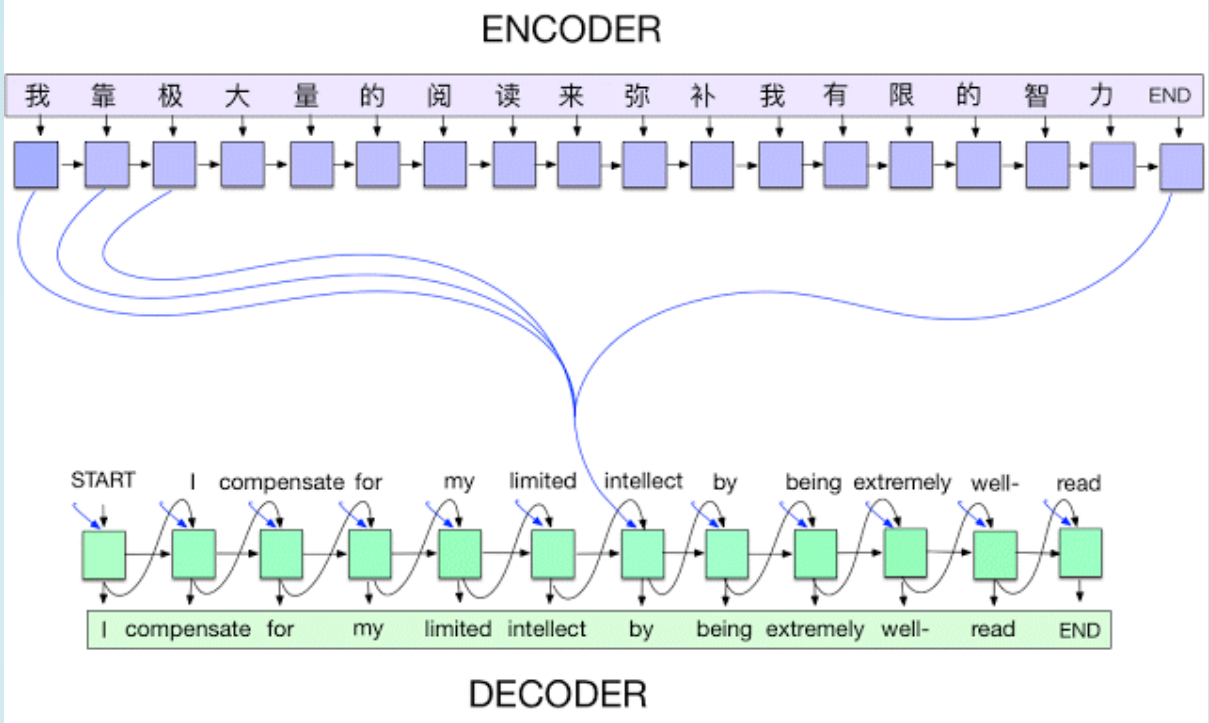
- Action+Stop words : 141-d
- Word Embedding : 50-d
- Concatenate them
- Use as input stage

BASIC SEQ2SEQ (AGAIN)



Practical seq2seq

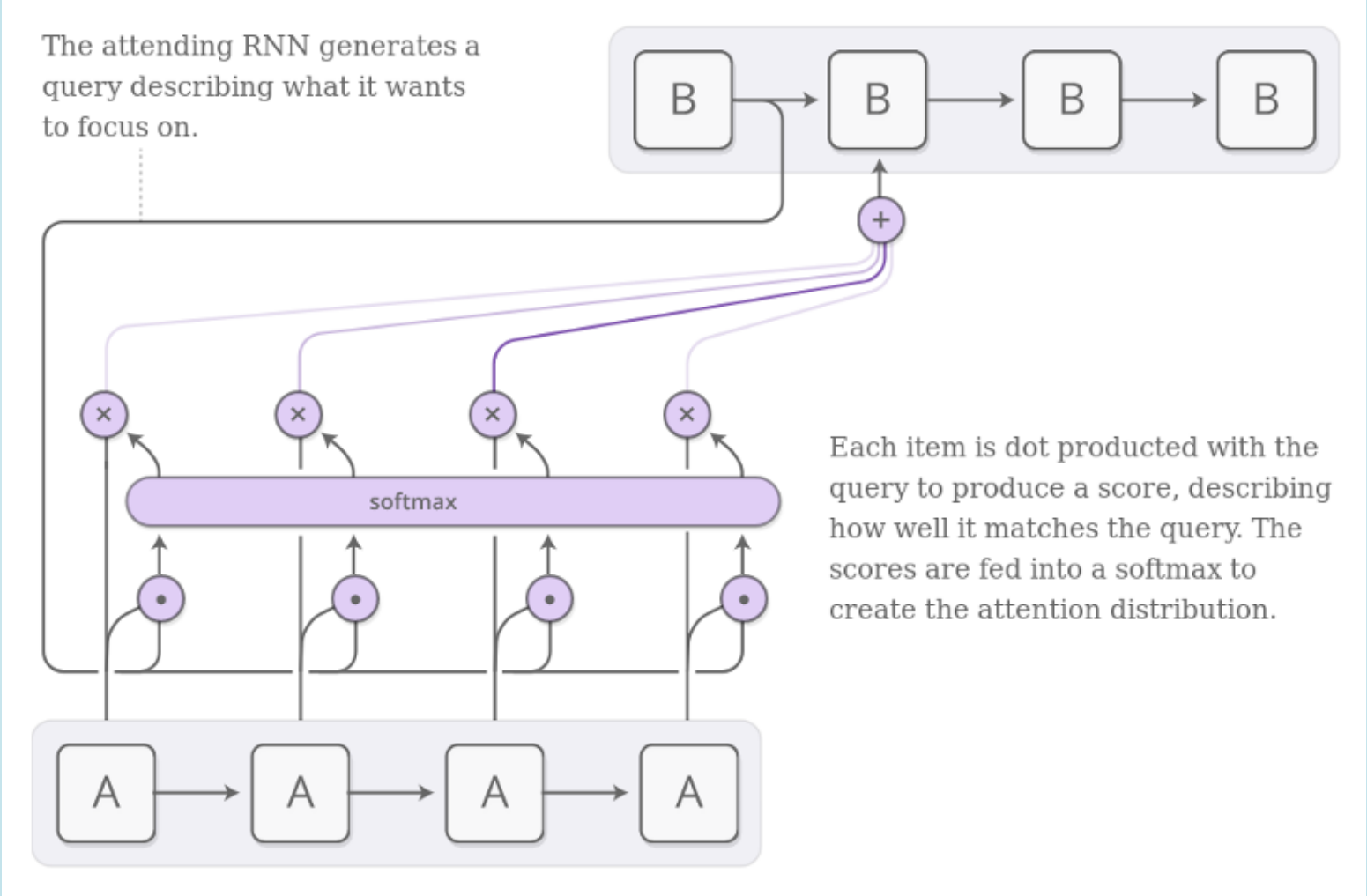
ATTENTION!



tf-seq2seq (Google)

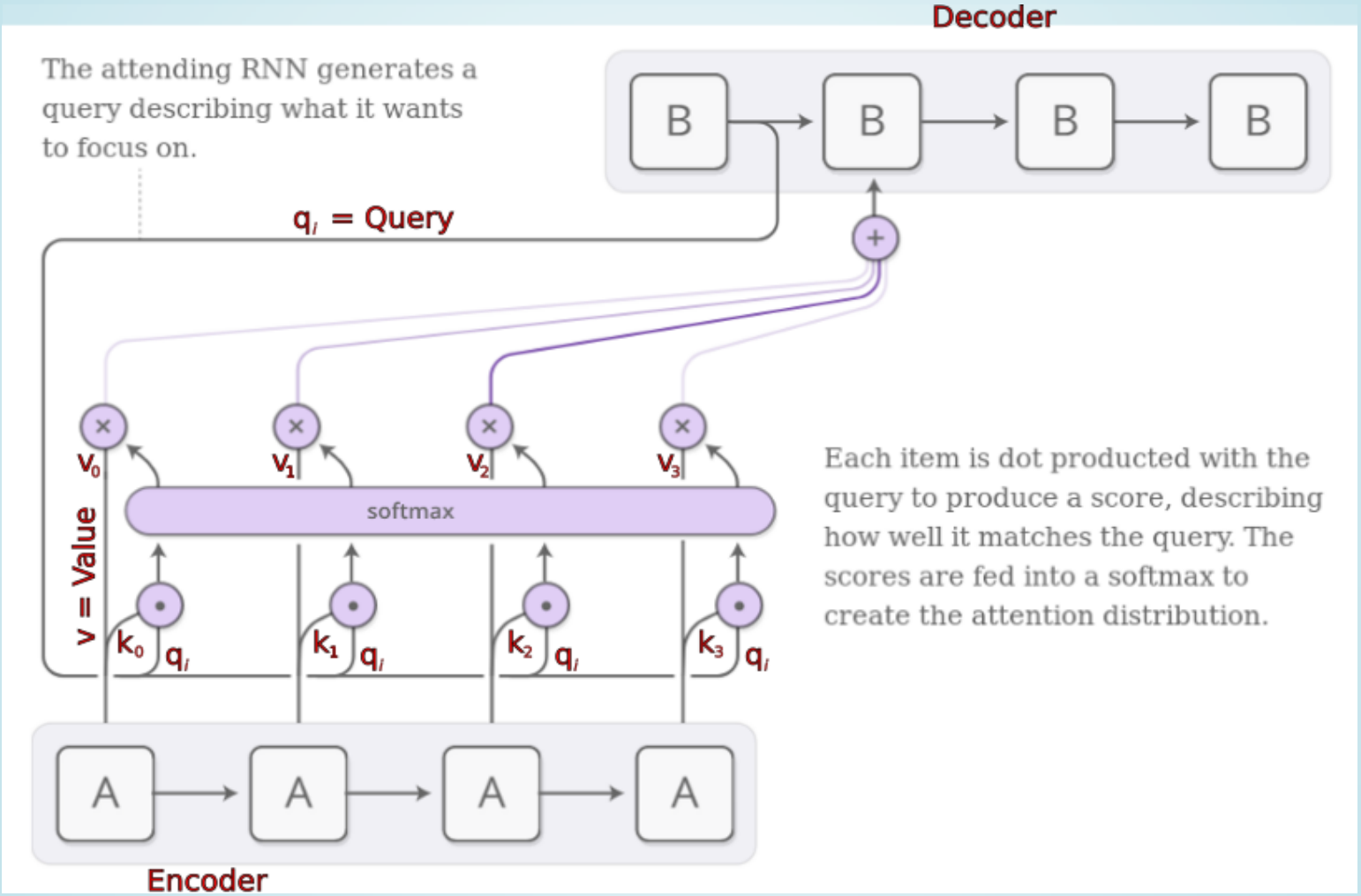
RED DRAGON AI

ATTENTION ORIGINAL



Augmented RNNs (distill)

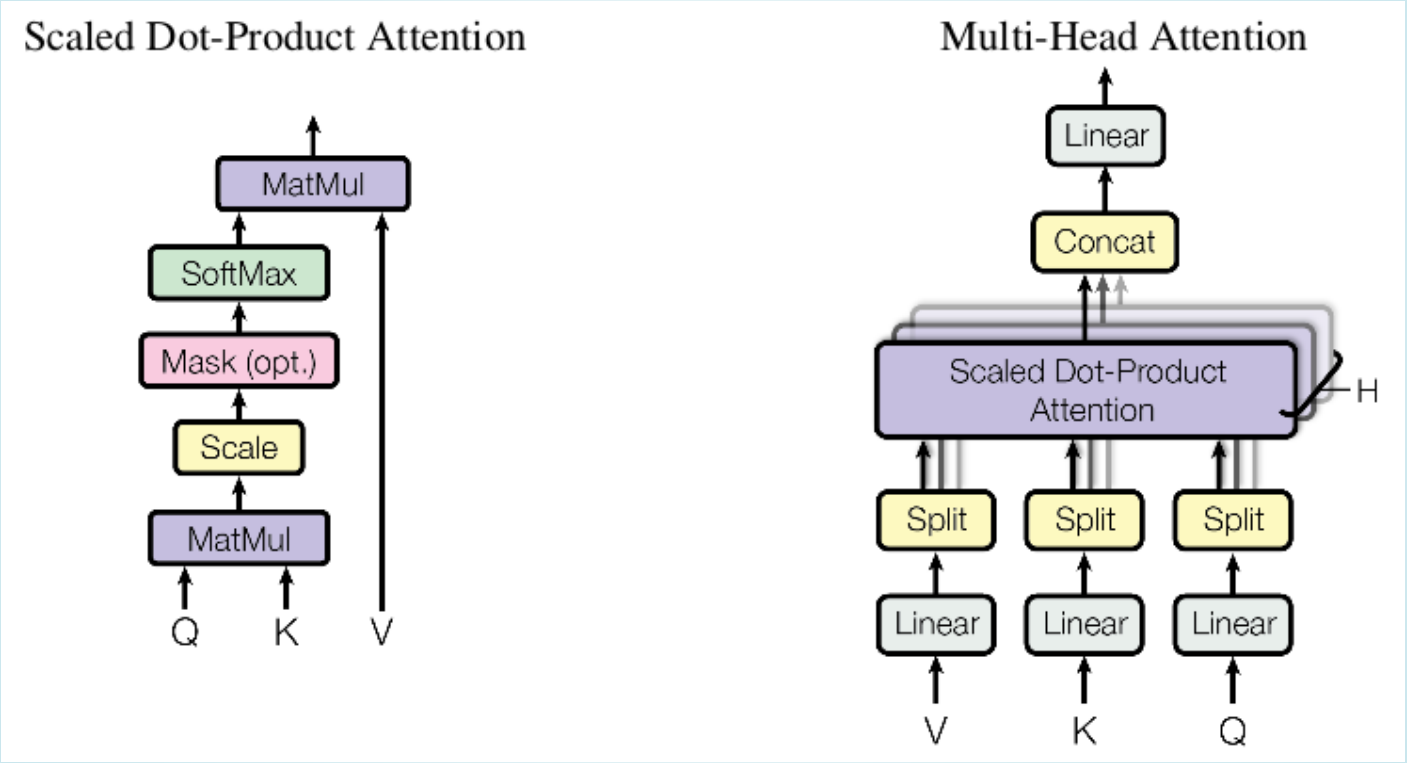
ATTENTION KEY-VALUE



Basic idea for "Attention is all you need" = AIAYN

RED DRAGON AI

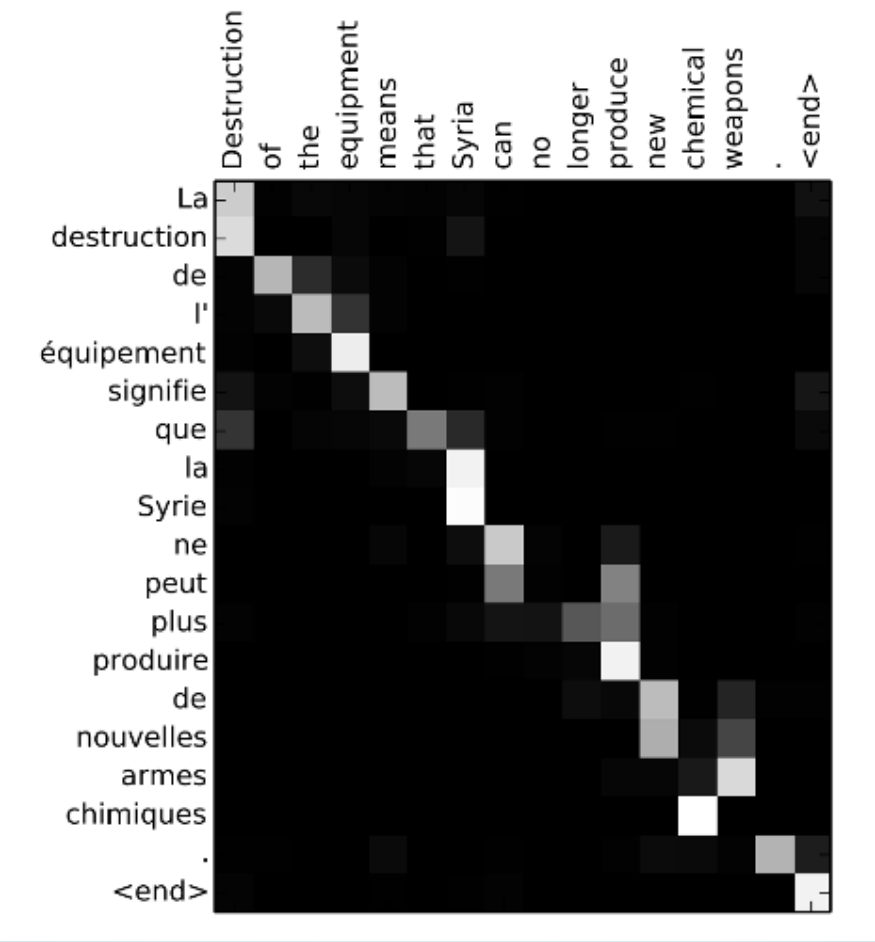
ATTENTION-IS-ALL-YOU-NEED DETAIL



AIAYN paper

RED DRAGON AI

ATTENTION BONUS



French-to-English

RED DRAGON AI

APPLICATION : CAPTIONING

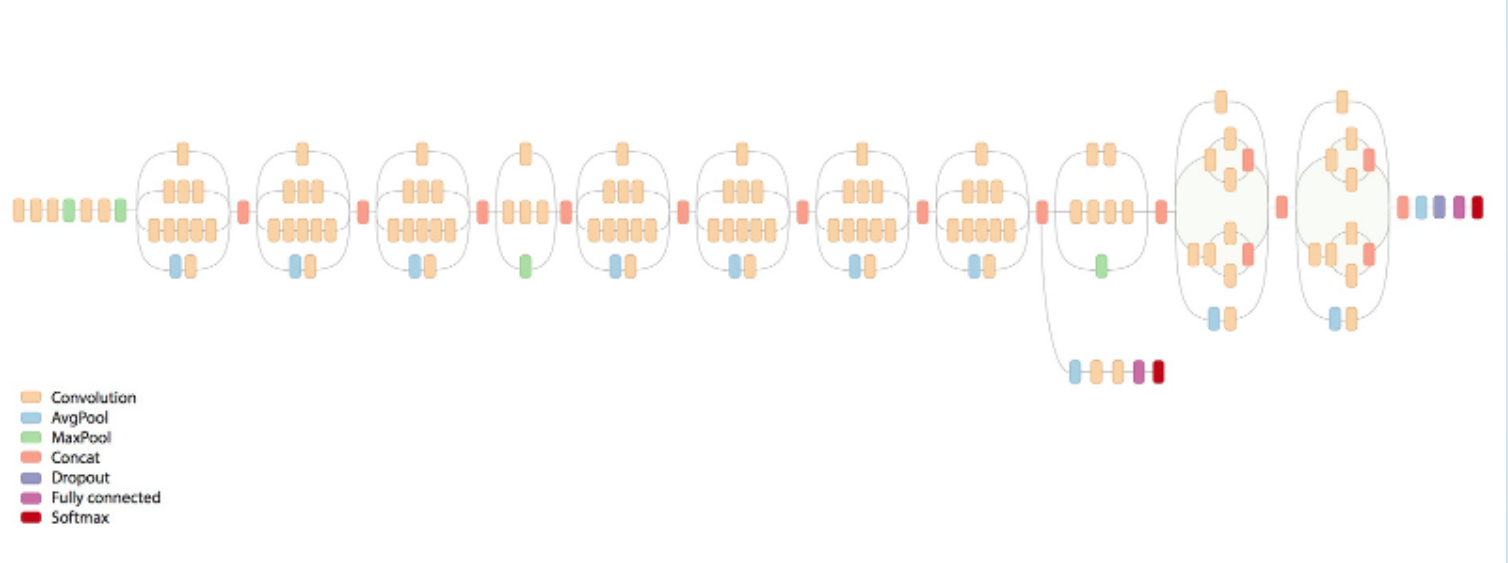
- large brown dog running away from the sprinkler in the grass .
- a brown dog chases the water from a sprinkler on a lawn .
- a brown dog running on a lawn near a garden hose
- a brown dog plays with the hose .
- a dog is playing with a hose .

DATA SET : FLICKR30K

- Summary statistics :
  - 31,783 images
  - 158,915 human-created captions
- Attribution-style licensing :
  - P. Young, A. Lai, M. Hodosh, and J. Hockenmaier. From image description to visual denotations: New similarity metrics for semantic inference over event descriptions

FLICKR30K : FEAT(IMAGE)

- Featurize all the images using InceptionV3
- Does not include any positional information



`github.com / mdda / deep-learning-workshop`
`/notebooks/2-CNN/7-Captioning/`
`1-folder-images-to-features.ipynb`

RED DRAGON AI

FLICKR30K : FEAT(TEXT)

- Want to make sure captions are *learnable*
  - Only use captions with "common-enough" words
  - All words must be in 5 different images
  - All words must be in GloVe 100k (50d) embedding
- Ensure 'stop' words are at start of dictionary

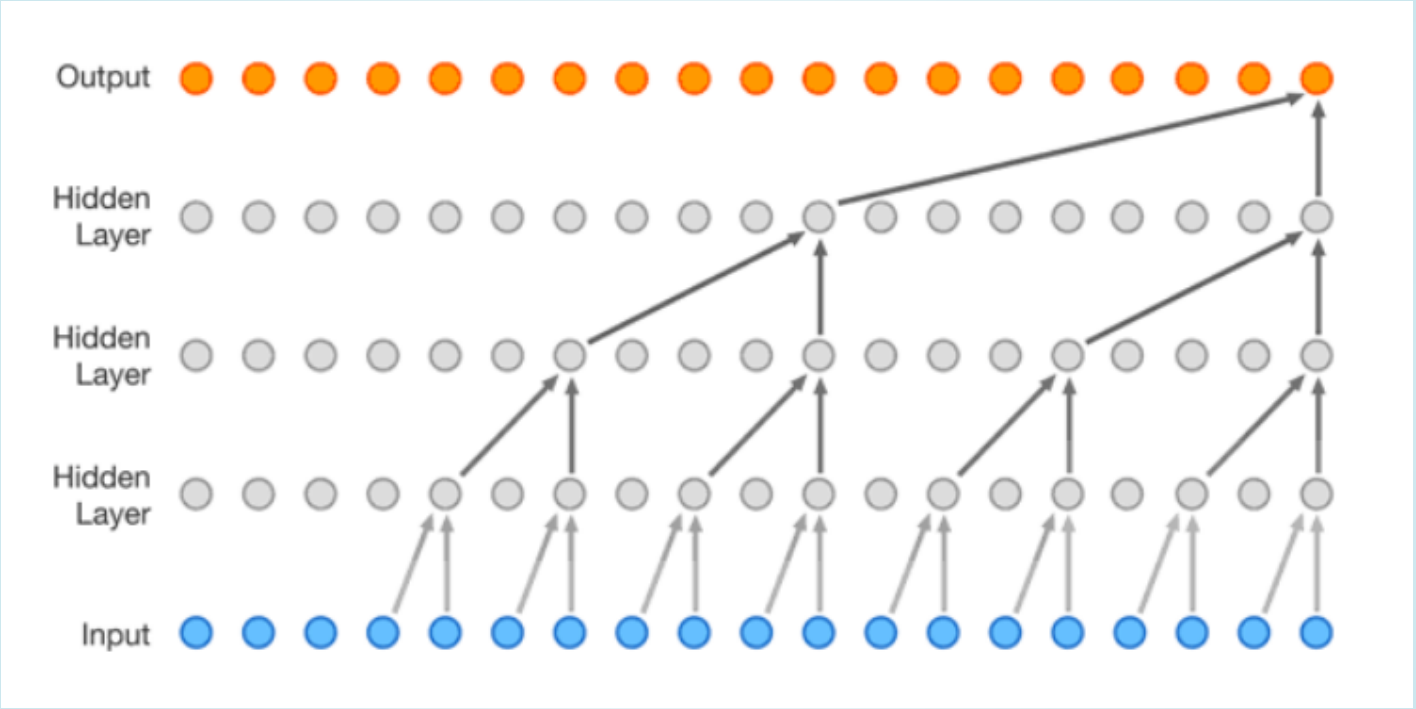
`github.com / mdda / deep-learning-workshop`
`/notebooks/2-CNN/7-Captioning/`
`2-Flickr30k-captions-to-corpus.ipynb`

EXTRA MACHINERY

- Dilated CNNs
- Gated Linear Units
- Fishing Nets

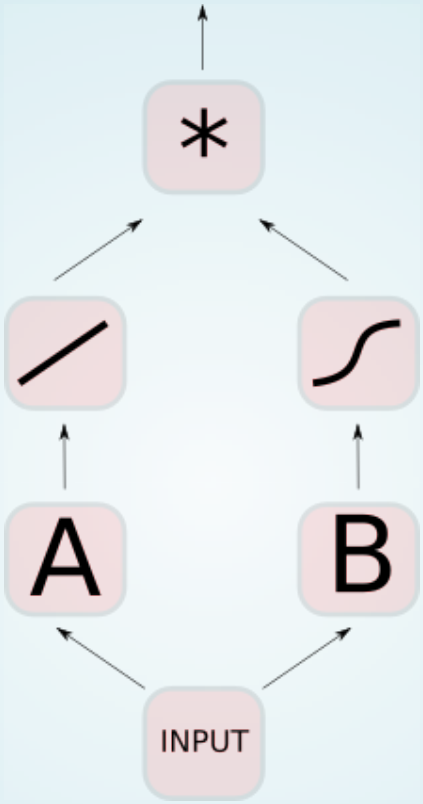
DILATED CNNS

DeepMind : WaveNet




```
Conv1D(padding='causal', dilation_rate=4)
```

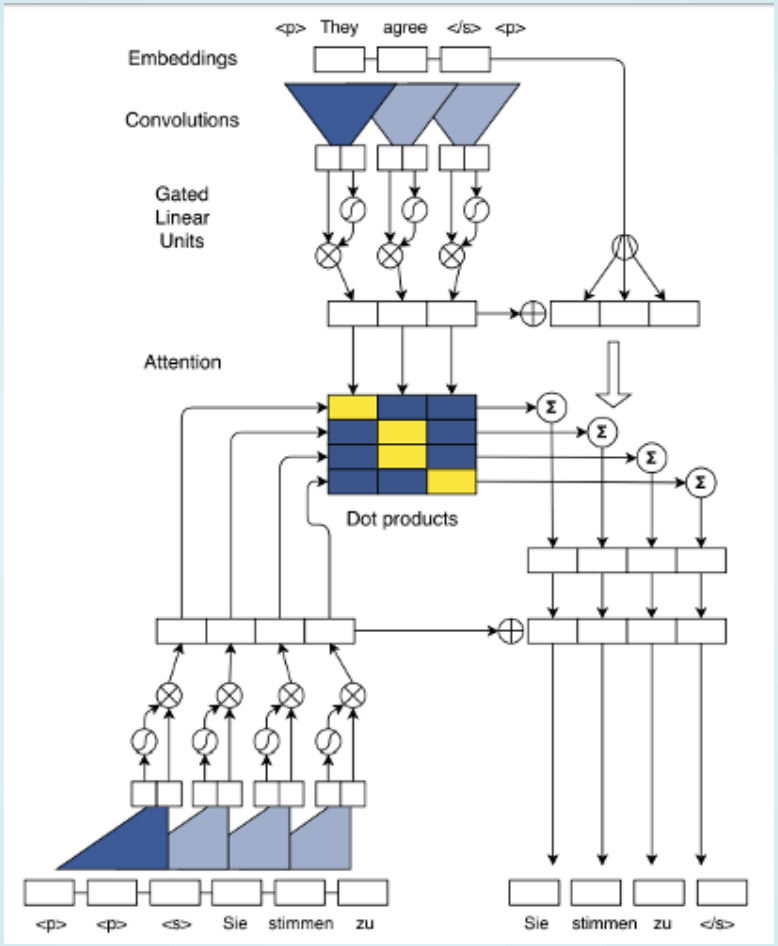
GATED LINEAR UNITS



Use one path to 'gate' another path

RED DRAGON AI

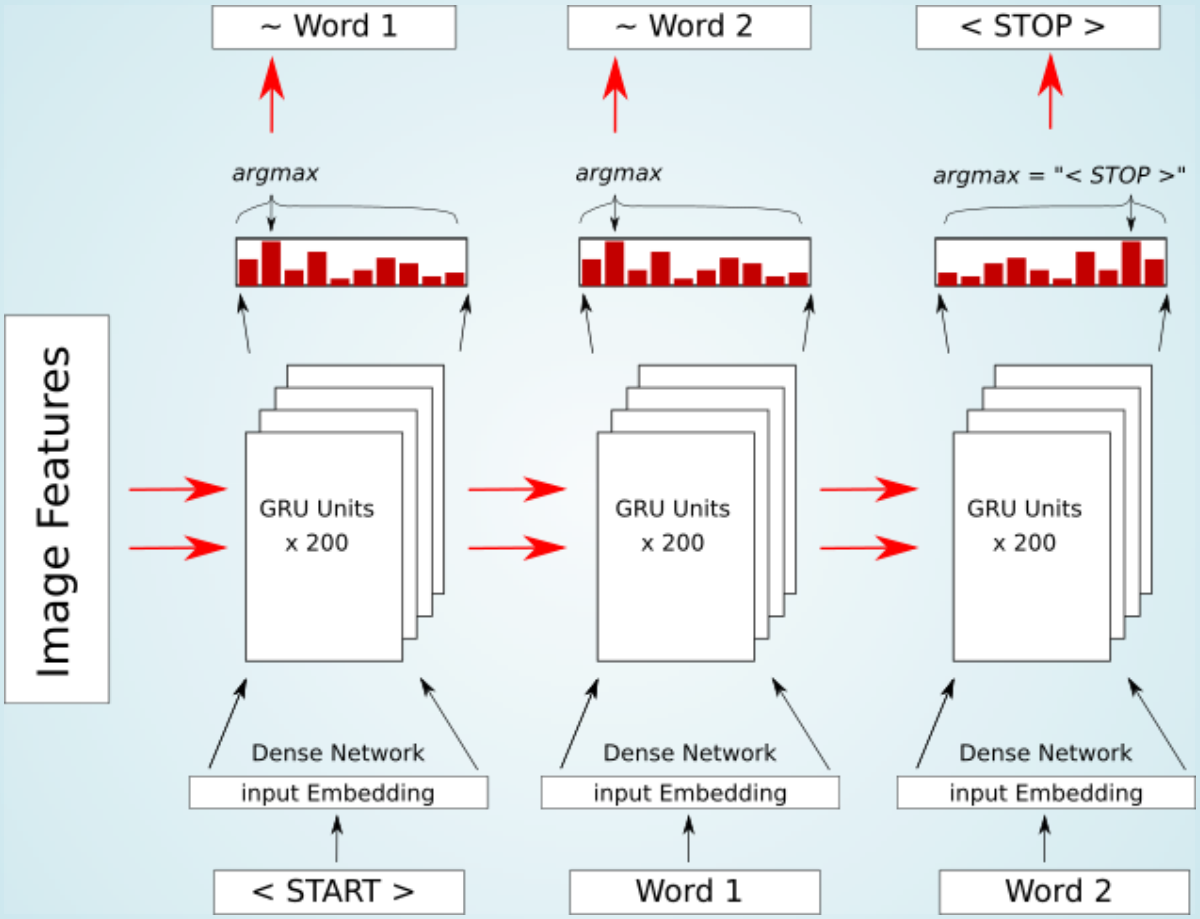
CNNS FOR SEQ2SEQ



Convolutional Seq-to-Seq Learning (Facebook)

RED DRAGON AI

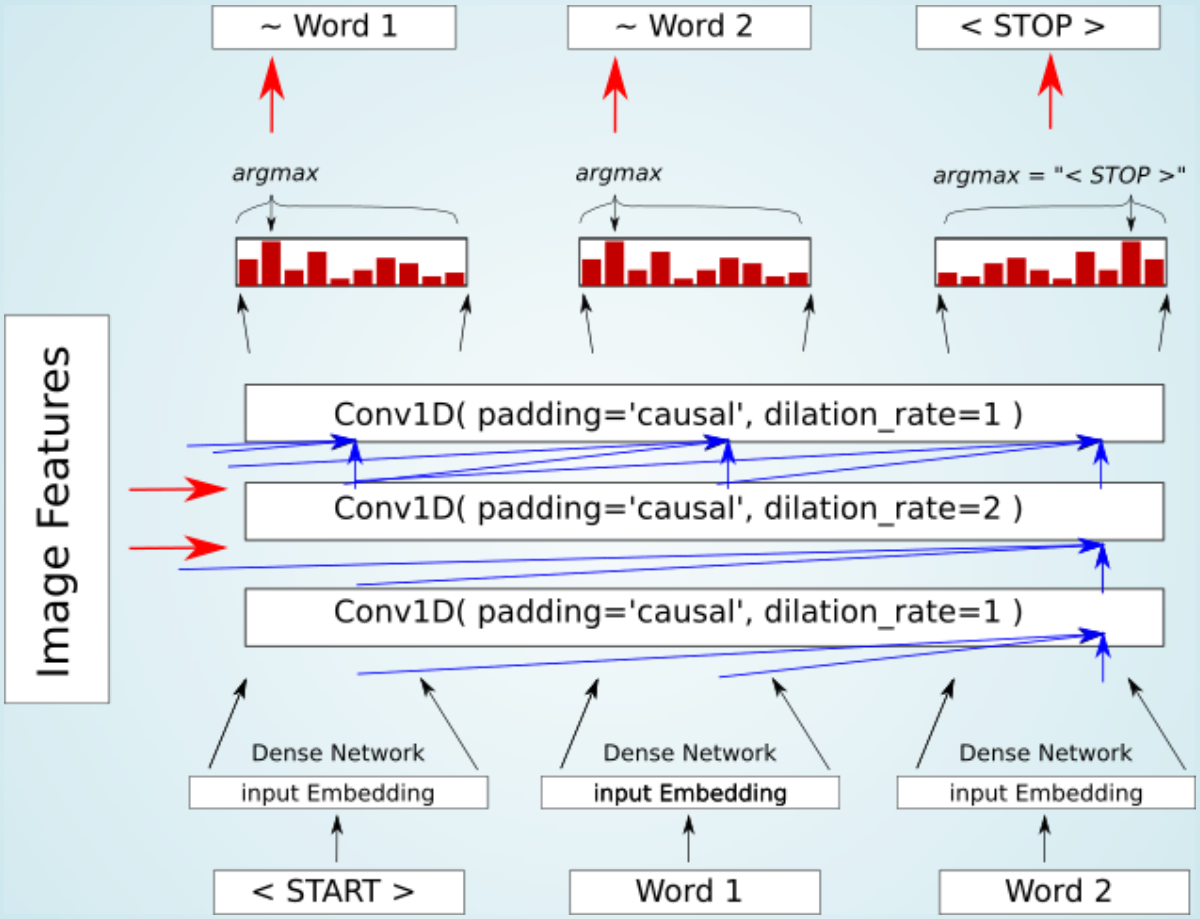
NETWORK PICTURE 1



'Standard' GRU set-up

RED DRAGON AI

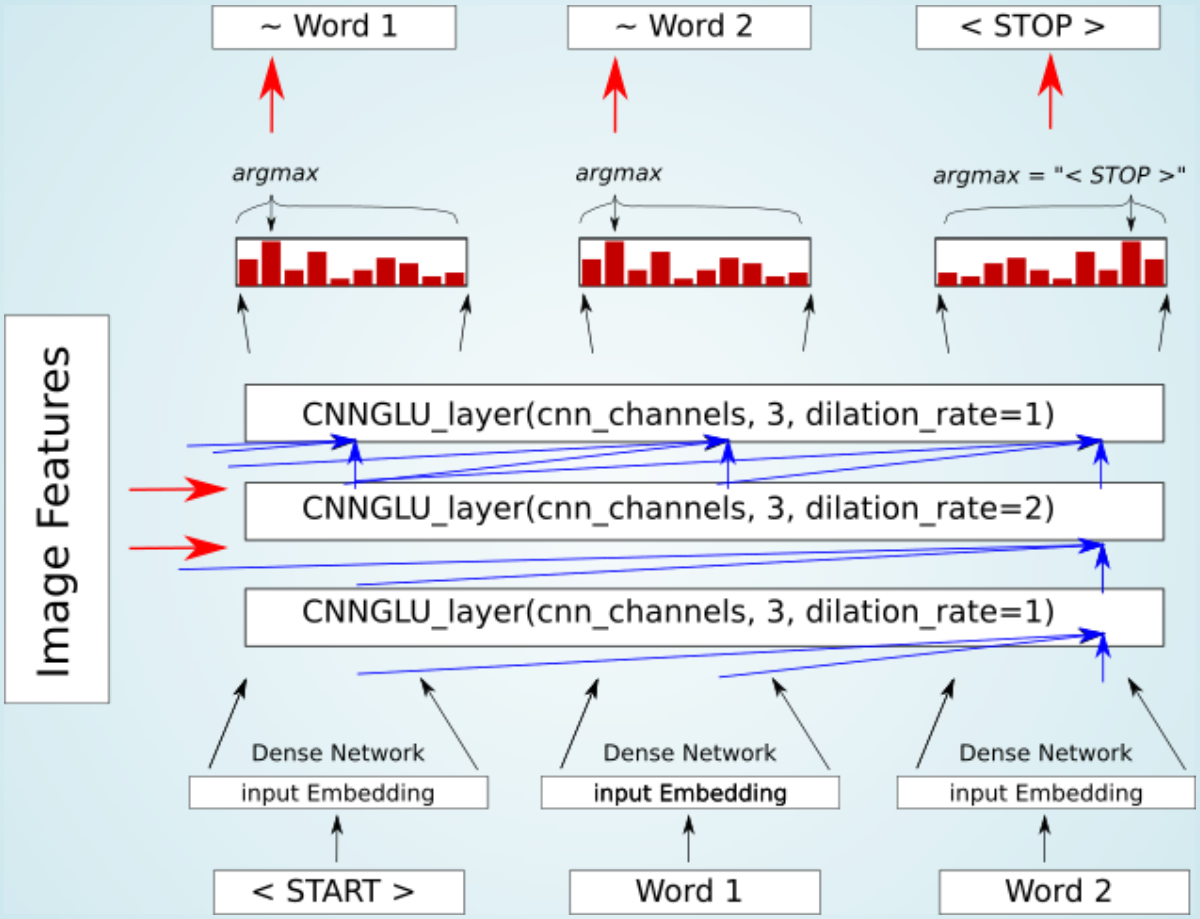
NETWORK PICTURE 2



Dilated CNN set-up (many variants)

RED DRAGON AI

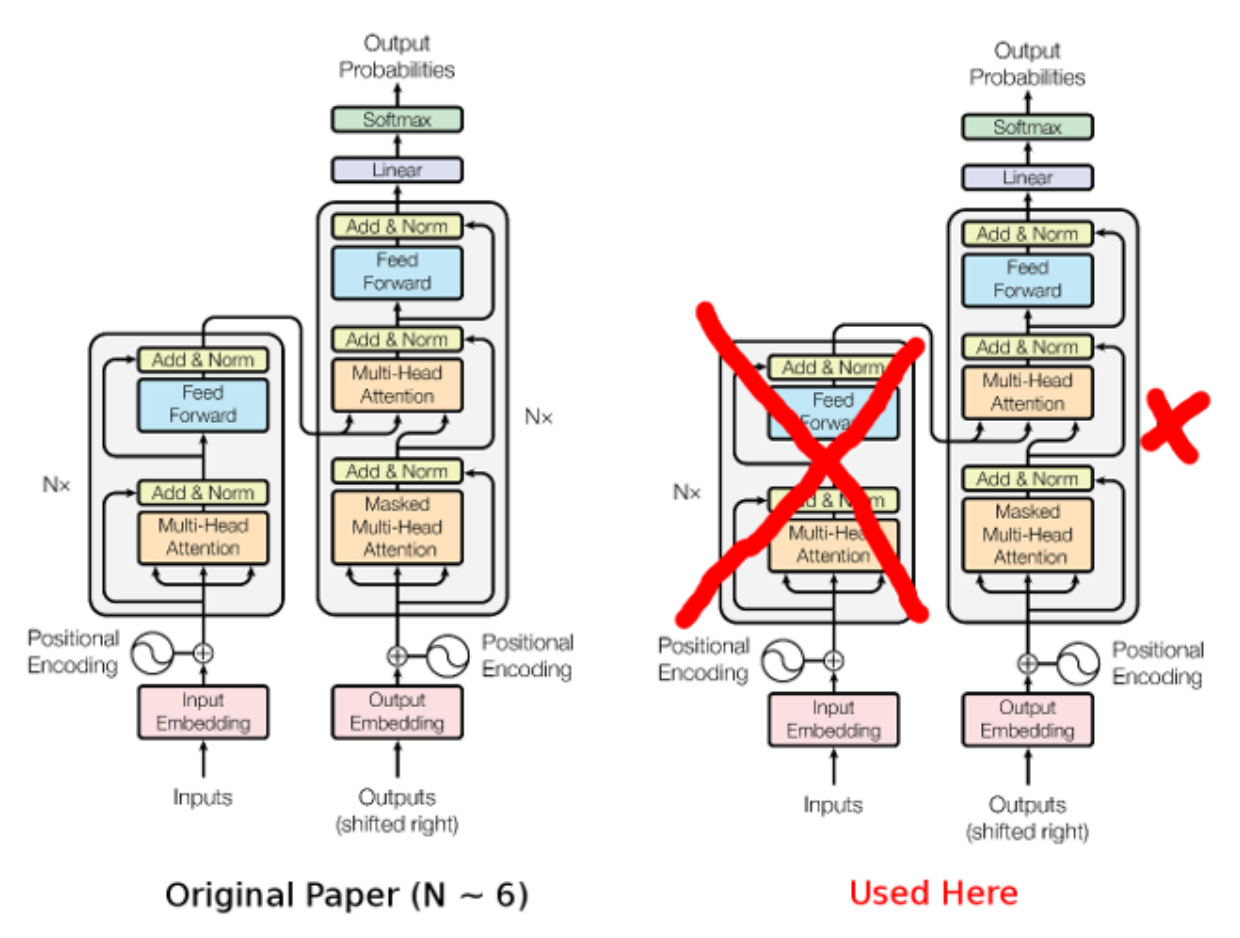
NETWORK PICTURE 3



Facebook CNN set-up (radically simplified)

RED DRAGON AI

NETWORK PICTURE 4



Google tensor2tensor

RED DRAGON AI

WALK-THROUGH

github.com / mdda / deep-learning-workshop
/notebooks/2-CNN/7-Captioning/
4-run-captioning.ipynb

IMAGE → CAPTION: UNTRAINED

- cables burning gracefully pin shine spoons arrange marshy solar board briefs claps tickets survey disinterested tractor looked movies guns rows engine technical town plaza fat captain paddlers historic motorcyclist soccer scales arabian
- does crown items bug pause ink what kayakers ohio lettering bikes battle squeezing person clad

TYPICAL TRAINING

- Input is 141d one-hot + 50d embedding
- Output is ~7,000 softmax one-hot
- Internal width ~200 units
- No special learning rate adjustments
- 50 epochs take ~ 3.5 hrs

IMAGE → CAPTION : GRUS

- a black dog running on a park .
- two big dogs play ball across the grass .
- the dog is being blocked by three other men each of it to something .
- a dog chases a ball while a man in a vest holding the hand .
- a man and a dog are chasing with a frisbee in the grass .

RESULTS : DILATED CNN

- the brown dog is standing on a yard .
- one dog bites another baseball player has found behind in the background .
- a dog running in a field leaps onto a field .
- two brown dogs are playing with a ball at a park .
- a brown dog runs his white dog while he is running along in winter grass .

RESULTS : GATED-LINEAR-UNITS

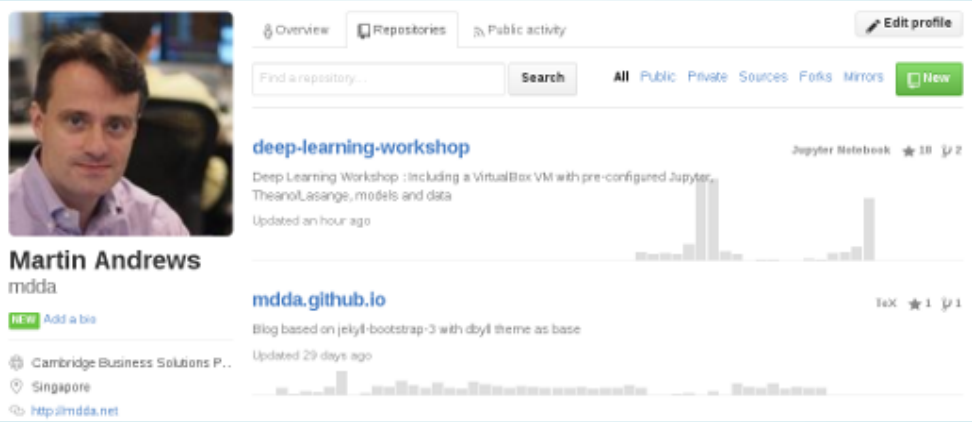
- a gray dog is running on a grass field .
- a dog jumping off over a bush .
- a dog on a leash is near a fountain .
- a brown dog is running through the muddy rain .
- a one dog with a brown jacket is playing in an enclosed setting .

RESULTS : AIAYN

- two dogs play in the grass .
- two dogs race by the two dogs fight to a grassy yard .
- the brown dogs lead beside two fire .
- two colored dog on a dogs to a metal tunnel .
- one dog chases after a brown dog on the park .

WRAP-UP

- Attention is very interesting
- Lots of innovation in NLP
- Having a GPU is VERY helpful



* Please add a star... *

RED DRAGON AI

- QUESTIONS -

MARTIN.ANDREWS @
REDDRAGON.AI

My blog : http://mdda.net/

GitHub : mdda