Recurrent Neural Networks

Oct.28.2017

Vijay Daultani
Rakuten Institute of Technology
Rakuten, Inc.

Rakuten

Education & Work Experience

Figure 9. Accenture Logo (2011)

(2010 : Ass. Software Engineer)

Figure 10. NEC Logo (2008)

(2014 : Researcher)

Figure 11. Rakuten Logo (2017)

(2017 : Researcher)

Figure 12. IIT Delhi Logo (2012)

(2012: Masters)

Figure 13. UC Berkeley Logo (2014)

(2015 : Visiting Scholar)

Table of Contents

Introduction

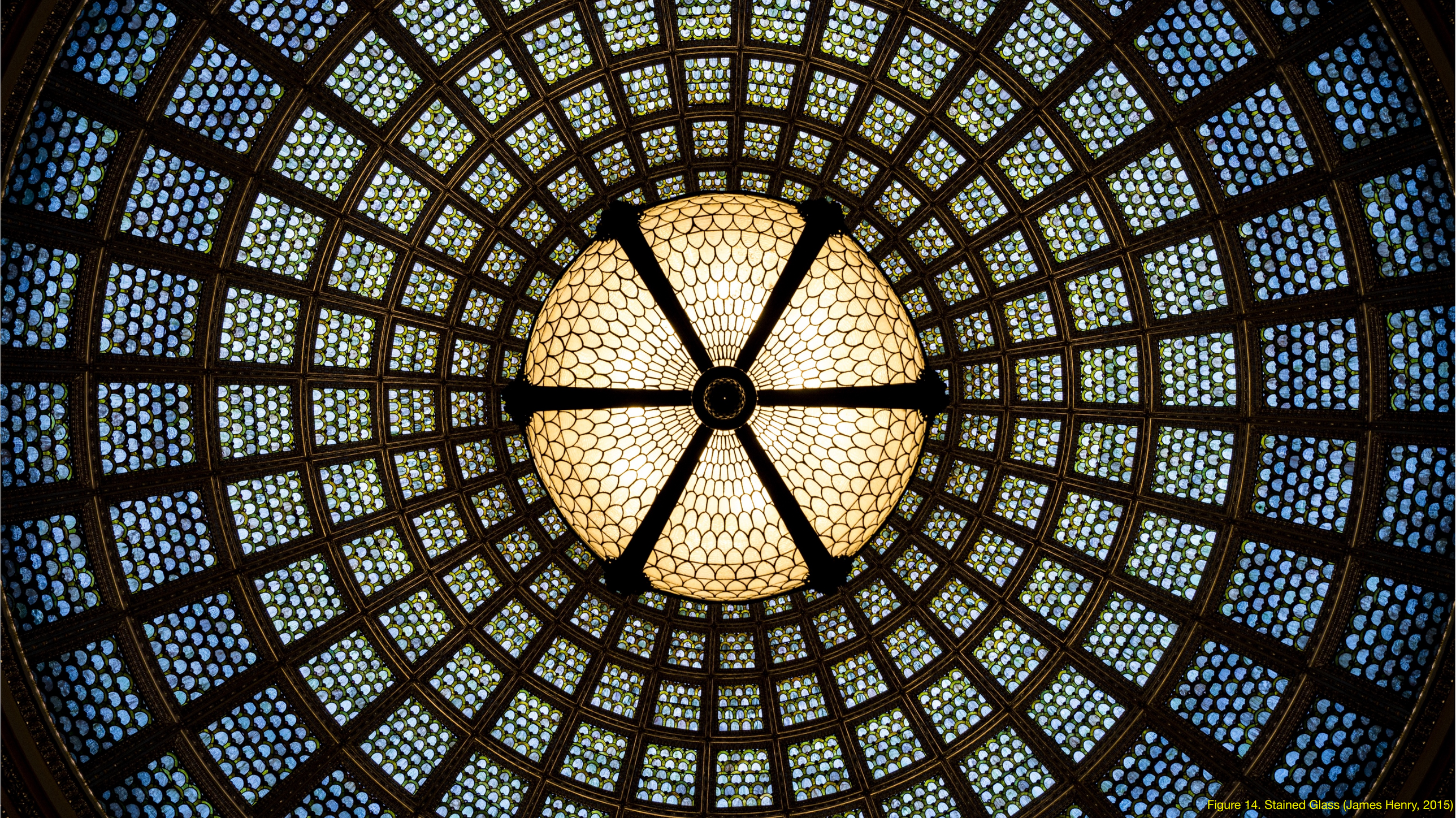


Figure 14. Stained Glass (James Henry, 2015)

Motivation for RNN

Why RNN?

Why Sequentiality explicitly?

Alan Turing

Figure 15. Alan Turing (1928)

Why not Markov Models?

Andrey Markov

Figure 16. Andrey Markov (1880)

Are RNN too expressive?

Edurado Daniel Sontag

Figure 17. Edurado Daniel Sontag (Greuel Martin, 2009)

RNN

Sequences

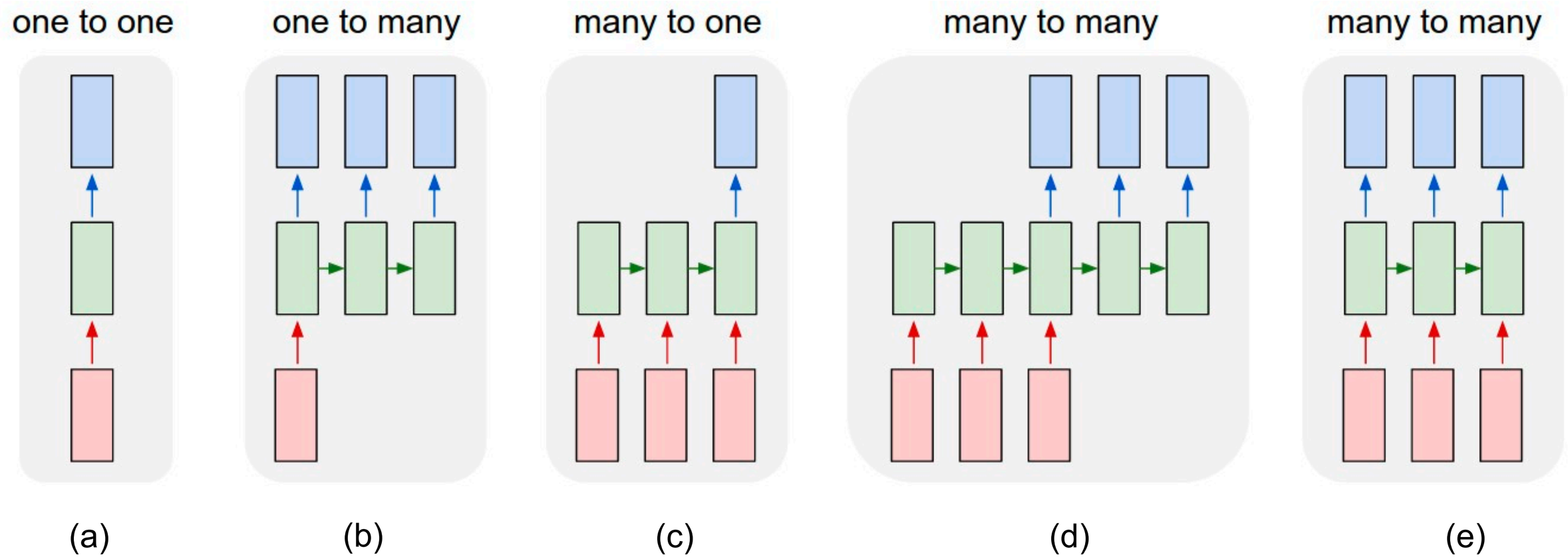


Figure 18. The Unreasonable Effectiveness of Recurrent Neural Networks (Andrej Karpathy, 2015)

RNN Architecture

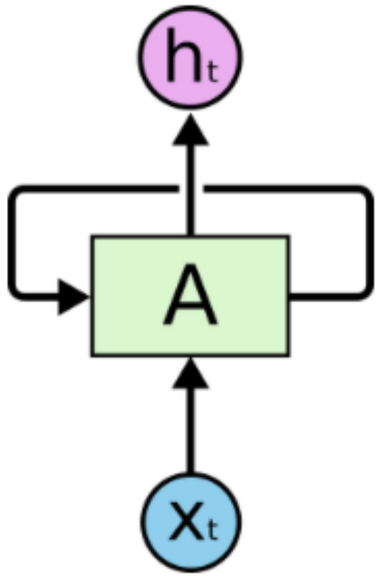


Figure 19. Understanding LSTM Networks (Christopher Olah, 2015)

RNN Architecture Unfolded

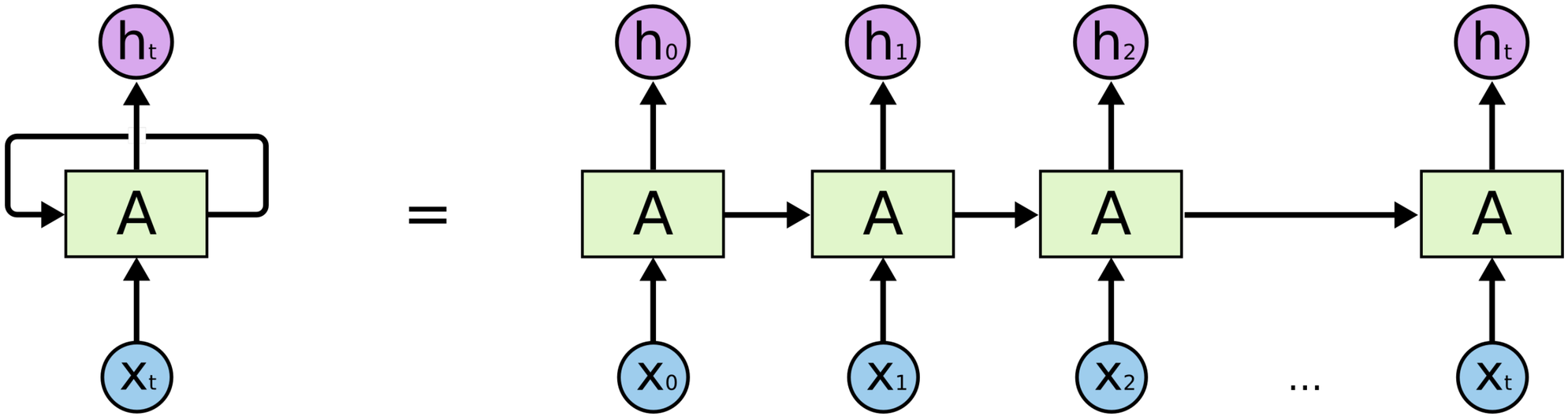


Figure 20. Understanding LSTM Networks (Christopher Olah, 2015)

Use Case

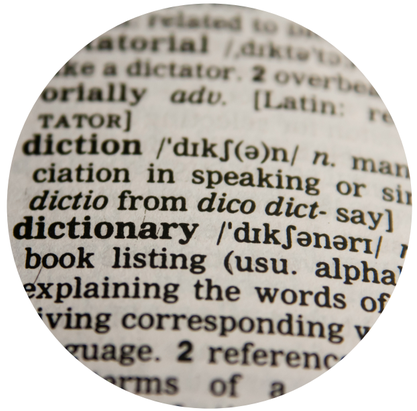
Neural Language Modelling

Figure 21. Alphabet boggle (Faby Green, 2015)

“black and white dog jumps over bar”

Image Captioning

Figure 22. Deep Visual-Semantic Alignments for Generating Image Descriptions (Andrej Karpathy, 2015)



Language Translation

Figure 23. Dictionary Words Grammar (PDPics, 2014)

Speech Recognition

Figure 24. Aluminium Audio Battery Broadcast (Vanleuven0, 2014)

Algorithms

Figure 25. Lunch Vegetables Healthy Meal (2014)

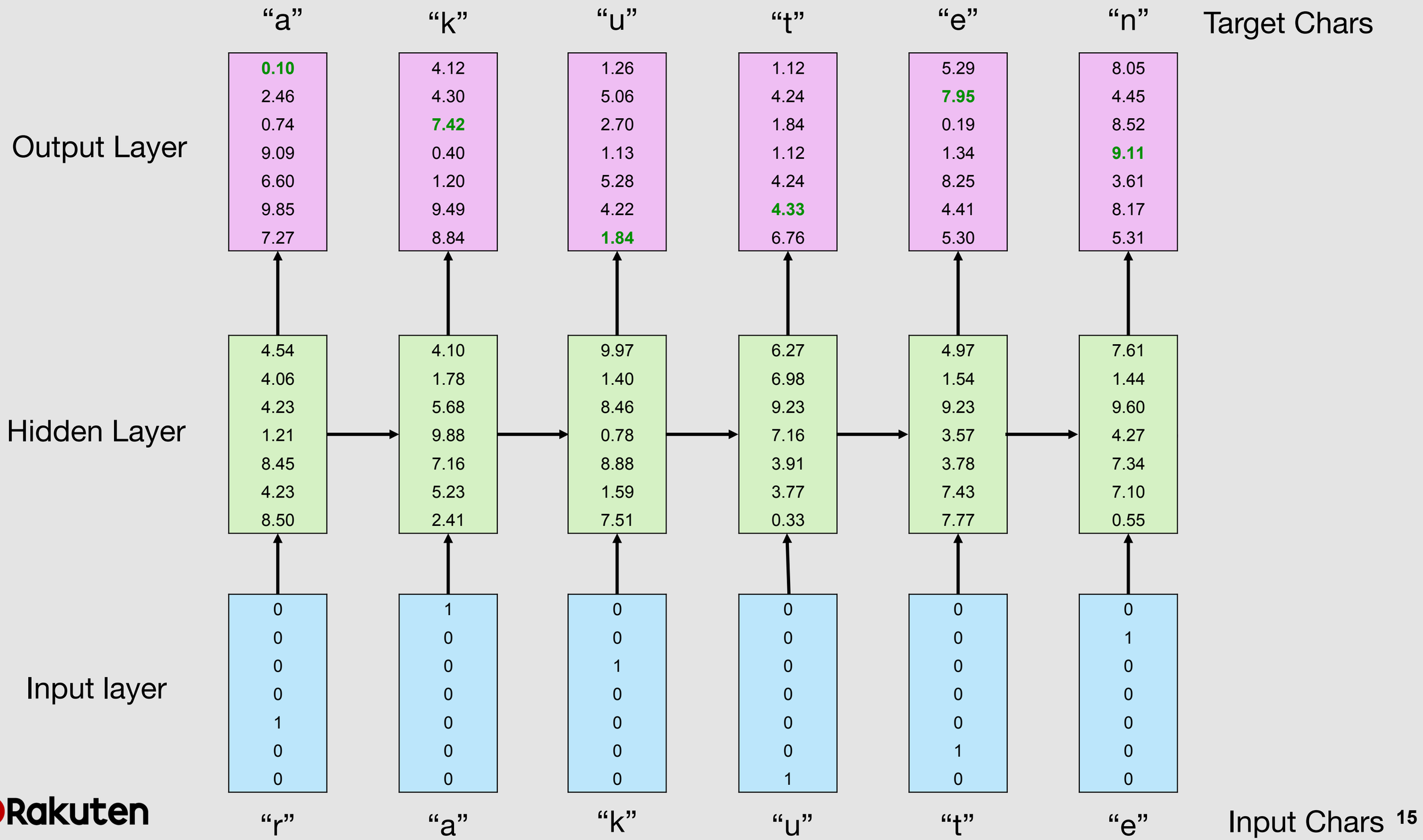
Questions & Answers

Figure 26. Question Mark Why Problem Solution (Tero Vesalainen, 2017)

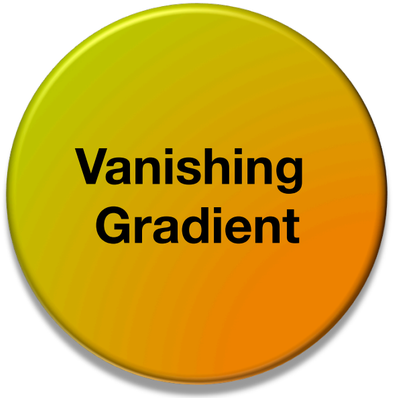
Neural Language Modelling

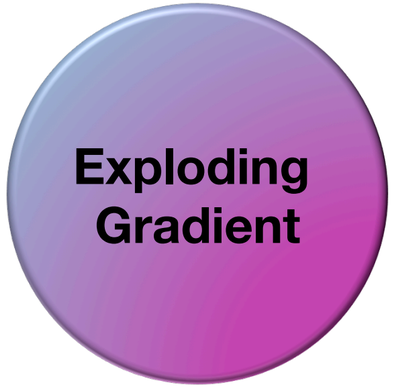
Example : Neural Language Modelling

a, e, k, n, r, t, u


“a”
“k”
“u”
“t”
“e”
“n”
Target Chars
Output Layer
0.10 2.46 0.74 9.09 6.60 9.85 7.27
4.12 4.30 7.42 0.40 1.20 9.49 8.84
1.26 5.06 2.70 1.13 5.28 4.22 1.84
1.12 4.24 1.84 1.12 4.24 4.33 6.76
5.29 7.95 0.19 1.34 8.25 4.41 5.30
8.05 4.45 8.52 9.11 3.61 8.17 5.31
Hidden Layer
4.54 4.06 4.23 1.21 8.45 4.23 8.50
4.10 1.78 5.68 9.88 7.16 5.23 2.41
9.97 1.40 8.46 0.78 8.88 1.59 7.51
6.27 6.98 9.23 7.16 3.91 3.77 0.33
4.97 1.54 9.23 3.57 3.78 7.43 7.77
7.61 1.44 9.60 4.27 7.34 7.10 0.55
Input layer
0 0 0 0 1 0 0
1 0 0 0 0 0 0
0 0 1 0 0 0 0
0 0 0 0 0 1 0
0 1 0 0 0 0 0
“r”
“a”
“k”
“u”
“t”
“e”
Input Chars

Limitations

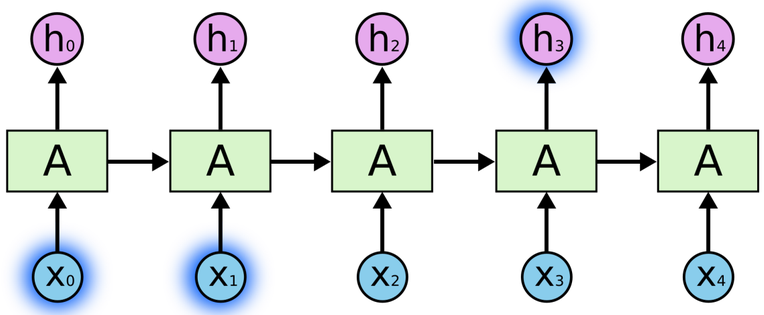




Long term
Dependencies

Long Term Dependencies

“The cloud is in the *sky*”



“ I grew up in Japan… I speak fluent *Japanese*”

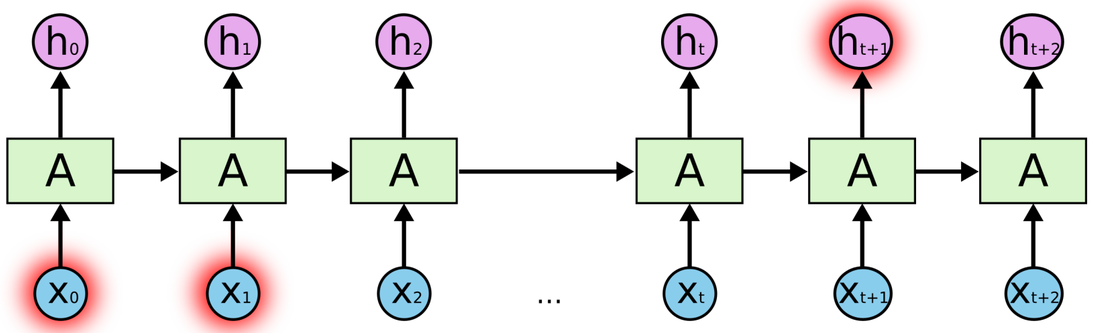


Figure 27. Understanding LSTM Networks (Christopher Olah, 2015)

LSTM

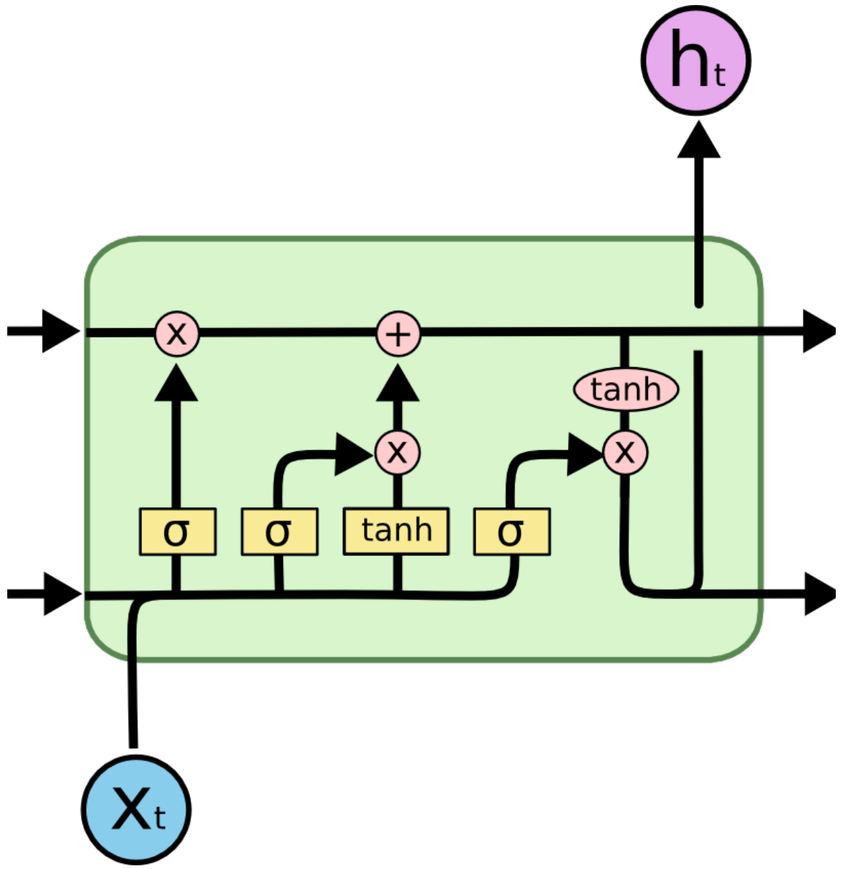
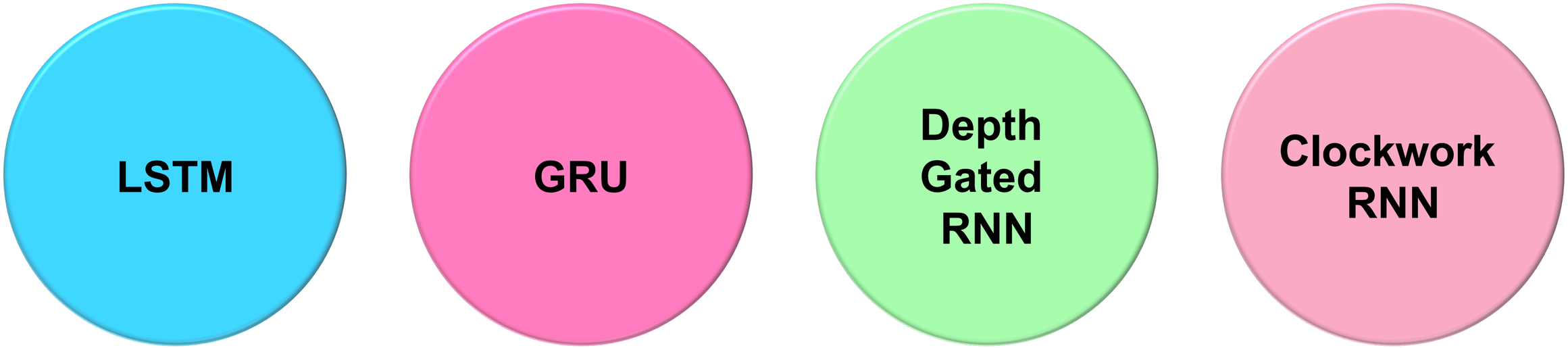
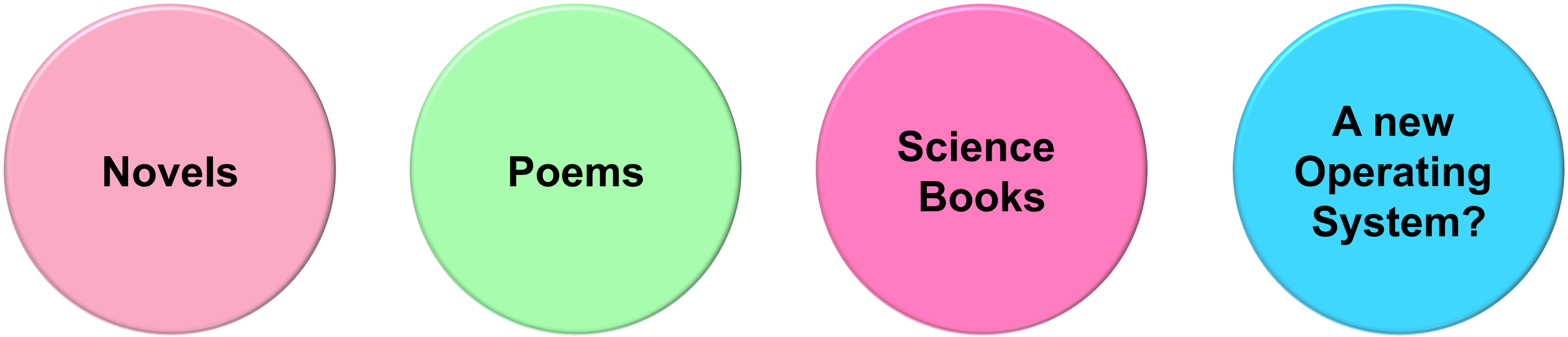


Figure 28. Understanding LSTM Networks (Christopher Olah, 2015)

Variations of RNN



Use Case : Neural Language Modelling



Example : Neural Language Modelling (Linux Source Code)

```
static void il3945_add_ibssi(struct ieee80211_hw *hw, void *data,
                             struct ieee80211_ht_callback *car_prio)
{
    struct lqp_information *icid;

    lps = &rtlpriv->linked_ring_list;

    iv = del_timer(&rtlpriv->woodlet_timer);
    if (single_rss == IL_SCOOP)
        rssi = time;
    else in->rs_start = rtlpriv->stats.rssi_lim.Lumb_M;
        struct mesh_configerr *malft = &min_id;

        if (p1 == 8 && max > clause) {
            if (level == MSG_DEVICES && (pHalData->config == (MCU_BASIC_PSPISR) &&
                (pdaddr < 0)))
                rtl_phy_acx(pd);
        }
        break;
    }
}
```

144 characters away

Distros

Do we need to have another distro of Linux?

Certainly Not !!

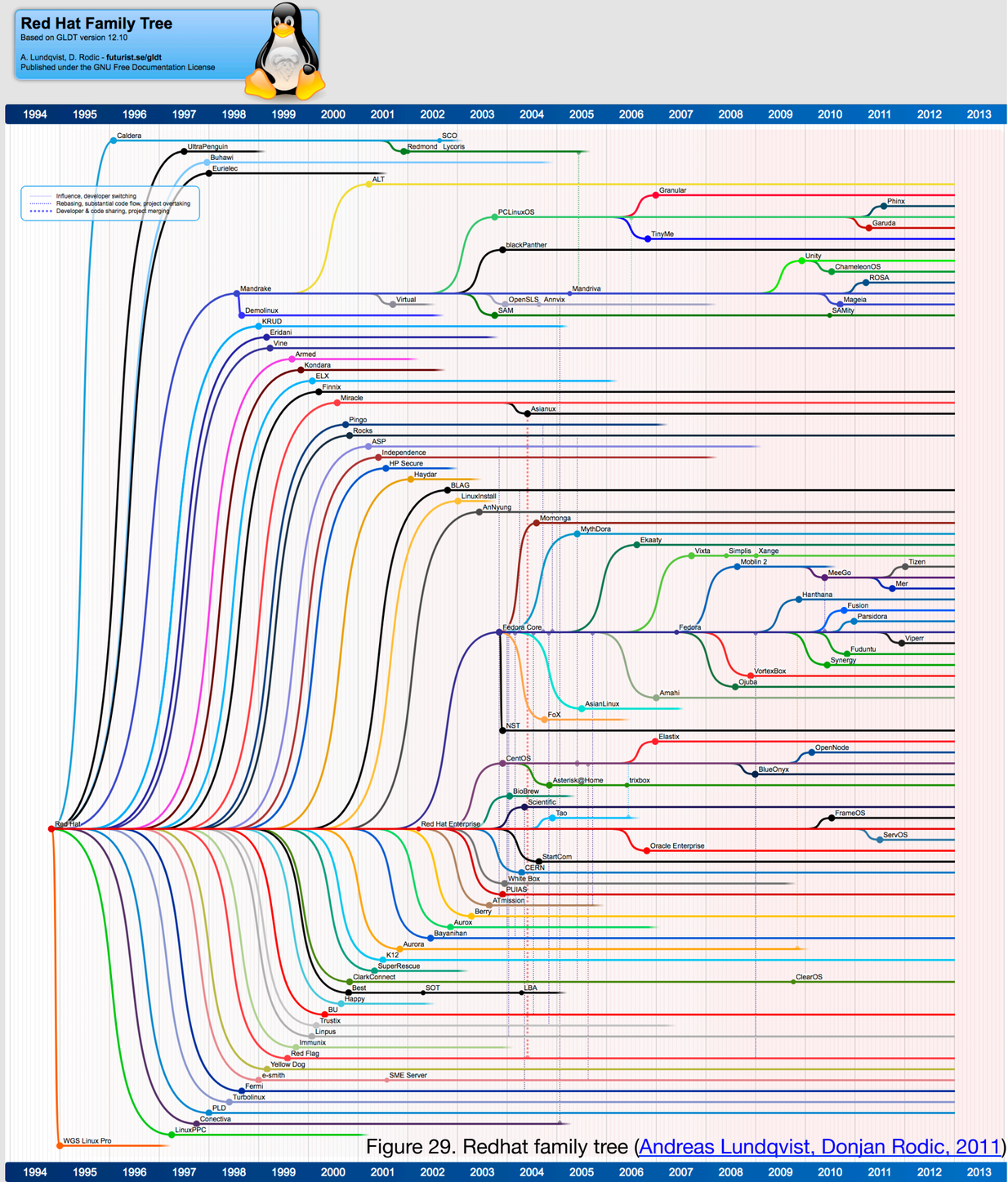


Figure 29. Redhat family tree (Andreas Lundqvist, Donjan Rodic, 2011)

Rakuten

Encoder - Decoder

Language Translation

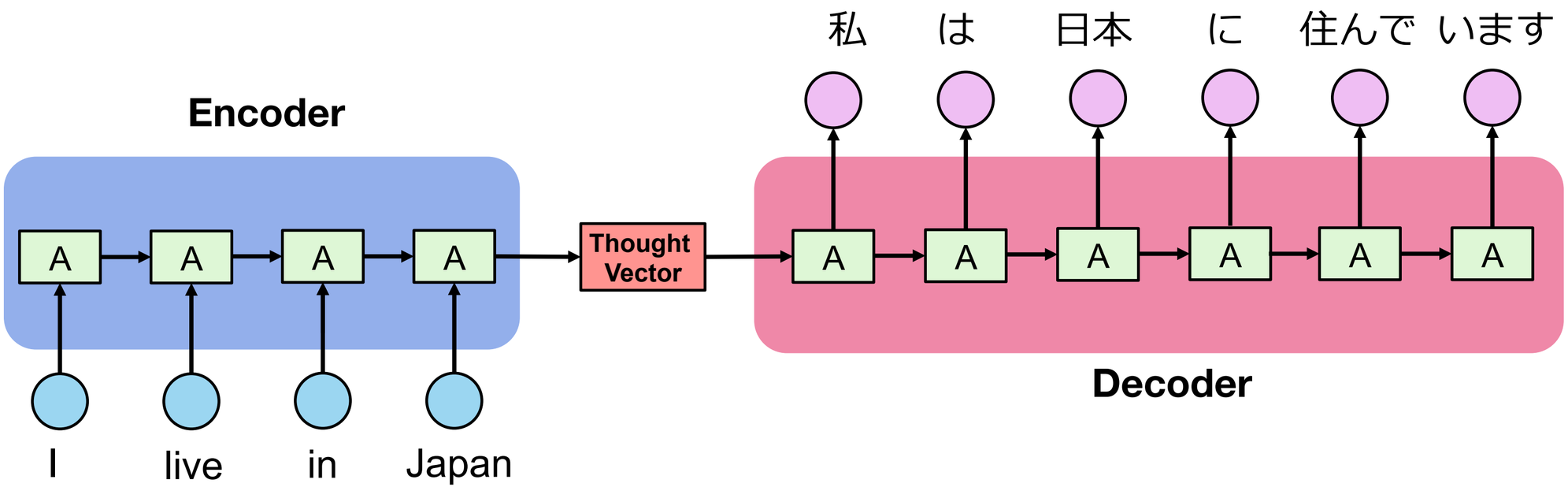
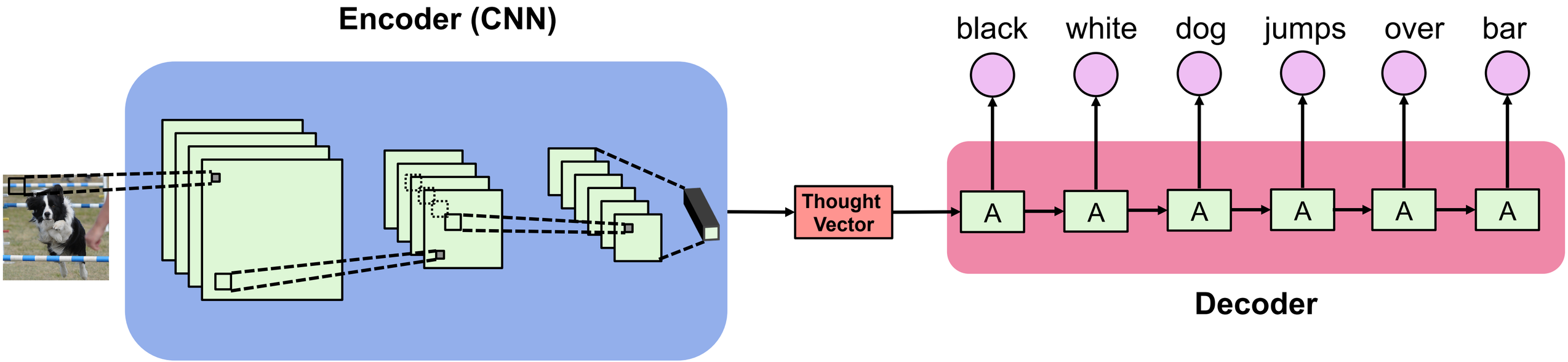


Image Captioning



Neural Turing Machine

NTM Architecture

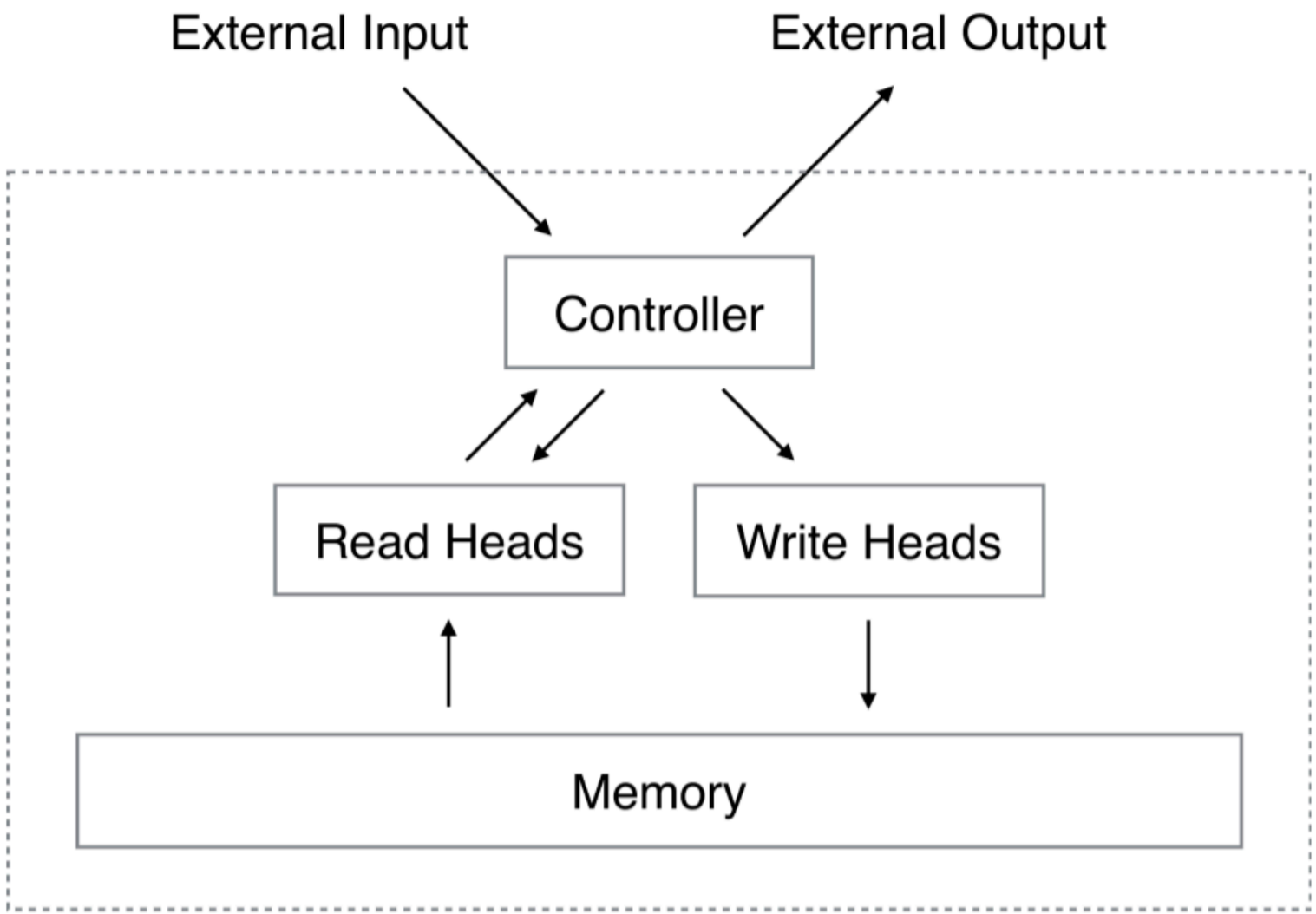


Figure 30. A. Graves, G.Wayne, I. Danihelka. (2014). Neural Turing Machines.

NTM Experiment – Copy Task Algorithm

initialise: move head to start location
while input delimiter not seen **do**
    receive input vector
    write input to head location
    increment head location by 1
end while
return head to start location
while true **do**
    read output vector from head location
    emit output
    increment head location by 1
end while

Figure 31. A. Graves, G.Wayne, I. Danihelka. (2014). Neural Turing Machines.

NTM Experiment – Copy Task

Network was trained on sequences of up to length 20

Target sequence length : 10, 20, 30, 50, 120,

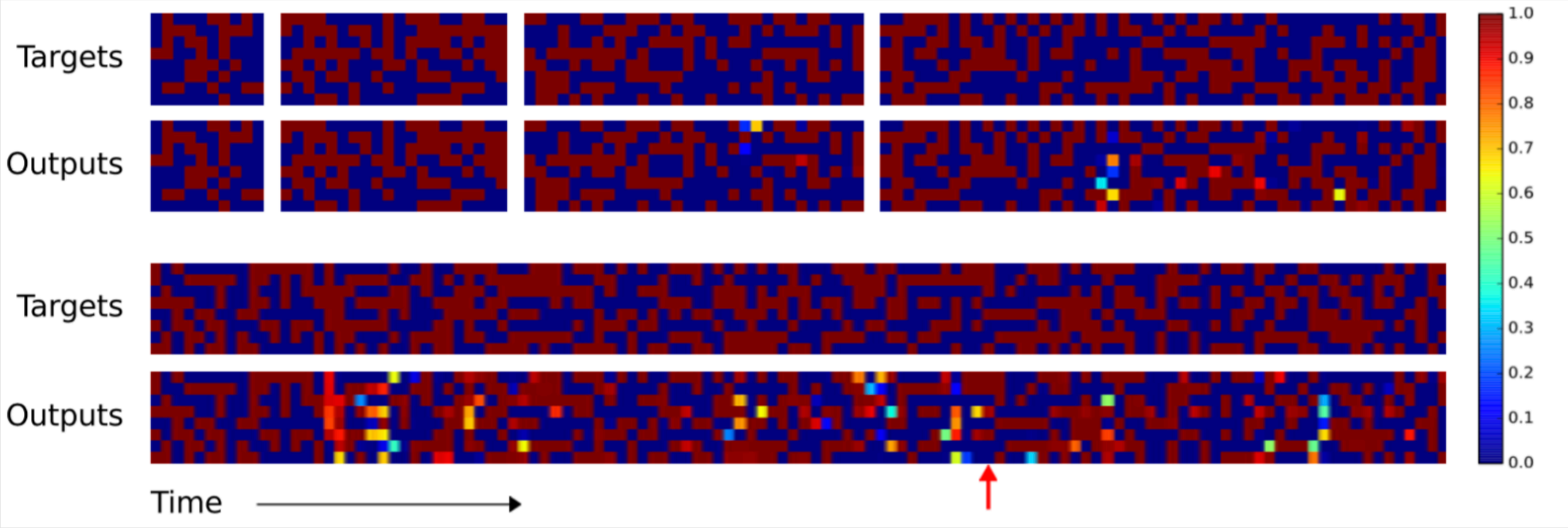


Figure 32. A. Graves, G.Wayne, I. Danihelka. (2014). Neural Turing Machines.

Memory Networks

Memory Networks Architecture

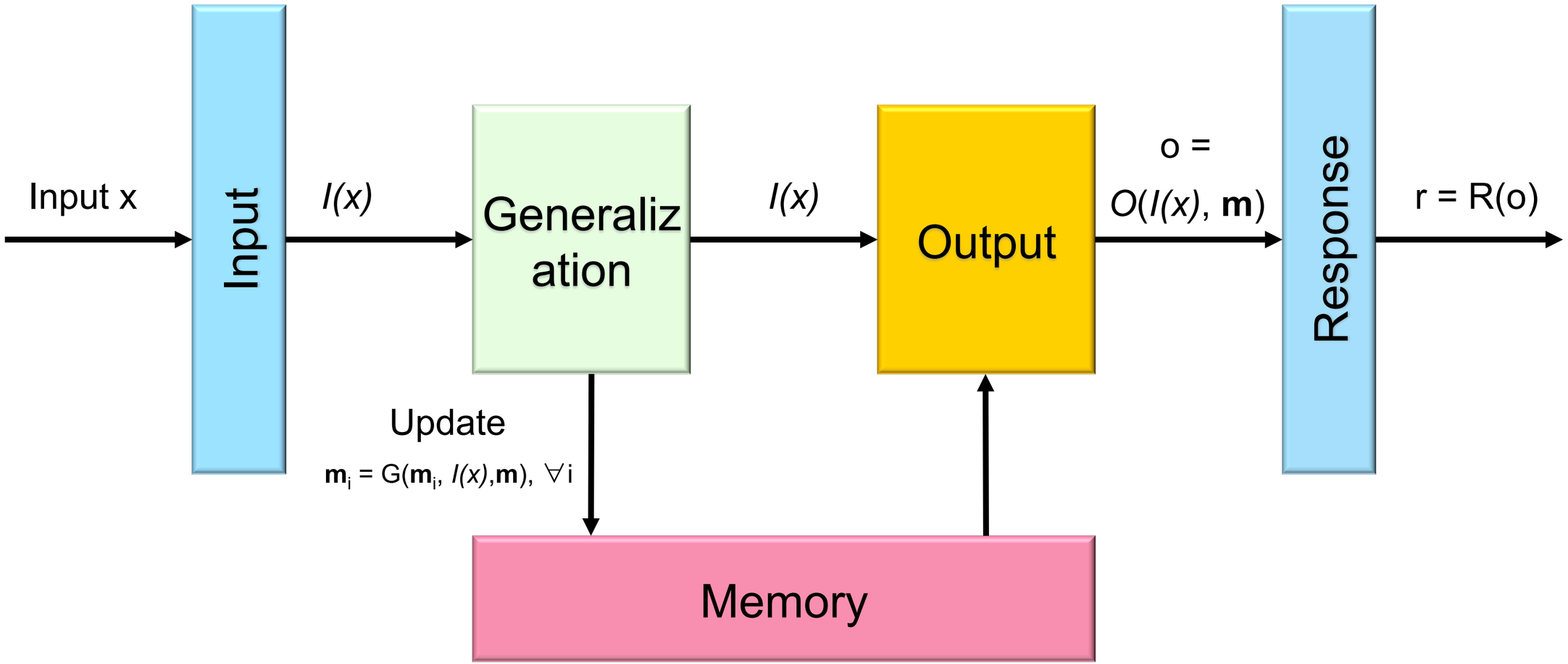


Figure 33. J.Weston, S.Chopra, A.Bordes (2014). Memory Networks.

Memory Networks - Experiment

Lord of the Rings in short :

Bilbo travelled to the cave. Gollum dropped the ring there. Bilbo took the ring.
Bilbo went back to the Shire. Bilbo left the ring there. Frodo got the ring.
Frodo journeyed to Mount-Doom. Frodo dropped the ring there. Sauron died.
Frodo went back to the Shire. Bilbo travelled to the Grey-havens. The End.

Q : Where is the ring?
A : Mount-Doom

Q : Where is Bilbo now?
A : Grey-havens

Q : Where is Frodo now?
A : Shire

Rakuten

Figure 34. J.Weston, S.Chopra, A.Bordes (2014). Memory Networks.

Applications

Applications

Robotics

Figure 35. BB-8 in action (Joseph Chan, 2017)

Computer Vision

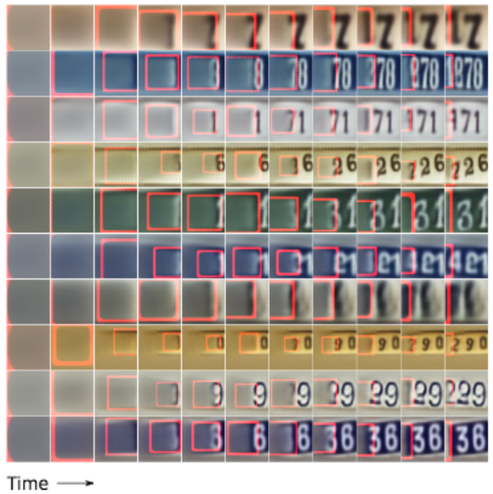


Figure 36. Gregor, Karol. Danihelka, Ivo. Graves, Alex. Rezende, Wierstra, Daan. (2015) DRAW: A Recurrent Neural Network For Image Generation

- Object Recognition
- Image Generation
- Video Analysis

Multimodal

Figure 37. Deep Visual-Semantic Alignments for Generating Image Descriptions (Andrej Karpathy, 2015)

- Image Captioning
- Video Captioning
- Visual Q & A

NLP

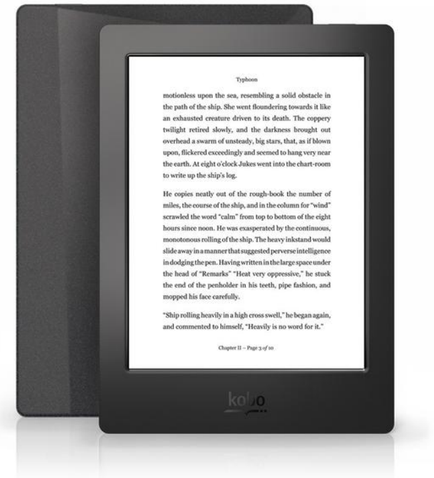


Figure 38. Rakuten Kobo (2017)

- Language Modeling
- Speech Recognition
- Machine Translation
- Conversation Model
- Text Q & A

Future

Future of RNN (Deep Learning)

- Unsupervised learning will become far more important
- Hybrid models like CNNs with RNNs will become more prevalent
- Combination of Deep Learning and Reinforcement learning are in their infancy and will become far more powerful.

References - I

- Doron (2003). *Jahangir Mahal* [Image file]. Retrieved from https://hi.wikipedia.org/wiki/%E0%A4%AE%E0%A4%A7%E0%A5%8D%E0%A4%AF_%E0%A4%AA%E0%A5%8D%E0%A4%B0%E0%A4%A6%E0%A5%87%E0%A4%B6_%E0%A4%95%E0%A4%BE_%E0%A4%AA%E0%A4%B0%E0%A5%8D%E0%A4%AF%E0%A4%9F%E0%A4%A8
- Bernard Gagnon (2013). *Maheshwar Fort - Jharokha* [Image file]. Retrieved from https://commons.wikimedia.org/wiki/File:Maheshwar_Fort_-_Jharokha_02.jpg
- "Lalbagh Palace" (Year). *Lalbagh Palace* [Image file]. Retrieved from http://www.iimidr.ac.in/about-us/explore-indore/Tfeayush. (2012).
- Nariman Point as seen from Marine Drive [Image file]. Retrieved from https://commons.wikimedia.org/wiki/File:Tfeayush1.jpg
- "Akshardham". (2016). *Akshardham* [Image file]. Retrieved from https://commons.wikimedia.org/wiki/File:BAPS_Akshardham_New_Delhi.jpg
- C.SHELARE. (2009). *Ajanta Caves* [Image file]. Retrieved from https://commons.wikimedia.org/wiki/File:AJANTA_CAVES_-_C.SHELARE_(4).jpg
- Free-Photos. (2015). *Golden Gate Bridge* [Image file]. Retrieved from https://pixabay.com/p-1031321/?no_redirect
- Taro Tokyo. (2010). *Tokyo Tower* [Image file]. Retrieved from https://commons.wikimedia.org/wiki/File:201010-TokyoTower-illuminated-fromWTC.jpg
- "Accenture Logo". (2011). *Logo of Accenture* [Image file]. Retrieved from https://en.wikipedia.org/wiki/Accenture#/media/File:Accenture.svg
- "NEC Logo". (2008). *NEC Logo* [Image file]. Retrieved from https://en.wikipedia.org/wiki/NEC
- "Rakuten Logo". (2017). *Rakuten Logo* [Image file]. Retrieved from https://en.wikipedia.org/wiki/Rakuten
- "IIT Delhi Logo". (2012). *Indian Institute of Technology Delhi Logo* [Image file]. Retrieved from https://en.wikipedia.org/wiki/Indian_Institute_of_Technology_Delhi
- "University of California Berkeley Logo". (2014). *University of California Berkeley Logo* [Image file]. Retrieved from https://en.wikipedia.org/wiki/University_of_California,_Berkeley
- James Henry. (2015). *Stained Glass* [Image file]. Retrieved from https://static.pexels.com/photos/161043/stained-glass-colorful-glass-stained-glass-window-161043.jpeg
- "Alan Turing". (1928). *Photo of Alan Turing at age 16* [Image file]. Retrieved from https://en.wikipedia.org/wiki/Alan_Turing

References - II

- BB-8 in action [Image file]. Retrieved from https://unsplash.com/photos/C8VWyZhcIIU
- "Andrey Markov". (1880). *Photo of mathematician Andrey Markov* [Image file]. Retrieved from https://en.wikipedia.org/wiki/Andrey_Markov
- Greuel Martin. (2009). *Workshop: Control Theory* [Image file]. Retrieved from https://en.wikipedia.org/wiki/Eduardo_D._Sontag
- Andrej Karpathy. (2015, May 21). *The Unreasonable Effectiveness of Recurrent Neural Networks* [Blog post]. Andrej Karpathy blog. Retrieved from http://karpathy.github.io/2015/05/21/rnn-effectiveness/
- Christopher Olah. (2015, August 27). *Understanding LSTM Networks* [Blog post]. colah's blog. Retrieved from http://colah.github.io/posts/2015-08-Understanding-LSTMs/
- Faby Green. (2015). *Alphabet boggle* [Image file]. Retrieved from https://pixabay.com/en/quotes-carpe-diem-word-diem-729173/
- Andrej Karpathy. (2015). *Deep Visual-Semantic Alignments for Generating Image Descriptions* [Image file]. Retrieved from http://cs.stanford.edu/people/karpathy/deepimagesent/
- PDPics. (2014). *Dictionary Words Grammar* [Image file]. Retrieved from https://pixabay.com/en/dictionary-words-grammar-abc-390055/
- Vanleuven0. (2014). *Aluminium Audio Battery Broadcast* [Image file]. Retrieved from https://pixabay.com/en/microphone-radio-broadcast-dj-772577/
- "Lunch Vegetables Healthy Meal". (2014). *Lunch Vegetables Healthy Meal* [Image file]. Retrieved from https://www.pexels.com/photo/lunch-vegetables-healthy-meal-8844/
- Tero Vesalainen. (2017). *Question Mark Why Problem Solution* [Image file]. Retrieved from https://pixabay.com/en/question-mark-why-problem-solution-2123967/
- Andreas Lundqvist, Donjan Rodic. (2011). *Redhat family tree // GNU/Linux Distribution Timeline 11.06* [Image file]. Retrieved from https://commons.wikimedia.org/wiki/File:Redhat_family_tree_11-06.png
- A. Graves, G.Wayne, I. Danihelka. (2014, Dec 10). *Neural Turing Machines*. Retrieved from https://arxiv.org/abs/1410.5401
- J.Weston, S.Chopra, A.Bordes (2014, Nov 29). *Memory Networks*. Retrieved from https://arxiv.org/abs/1410.3916Joseph Chan. (2017).


Rakuten