Exploring Efficacy of Embeddings on Relation Network for Natural Language Question Answering Task

Anonymous Author(s)
Affiliation
Address
email

1 Introduction

Deep learning has made it possible to do classification of objects in images and translation of languages, often with incredible accuracy. This is achieved due to the ability of neural networks to pick out important patterns that are inconceivable to the human eye, from large quantities of labeled data. However, just being able to learn patterns is not sufficient as it is not the only ability associated to intelligence; reasoning is another essential ability [1] that separates humans from machines. Hence, in recent years there is much work on reasoning related research, like visual reasoning [3, 4] where the machine is able to give an answer given an image and a visual question about the image, and text-based question answering [4] where the machine is able to answer a question based on the earlier sentences given to it.

In this project, we focus on the text-based question answering task using relation network (RN) [4] on the bAbI dataset [6]. RNs are networks that are designed based on relational reasoning, where its capacity to compute relations is baked into the architecture without having the neural network to learn it.

For any neural approach for natural language processing, word and sentence embeddings are indispensible. They allow us to represent words and sentences whose original forms are strings, as vectors which then we can feed it into a artificial neural network. There are various ways to embed words, and while unsupervised representations have been the more commonly used approach, using the assumption that you can tell a word by the company it keeps, there is an increased focus on supervised representations and also multi-task learning of representations. In our project, we explore how different types of embeddings, in particular, how the traditional unsupervised representations compare up to representations obtained from multi-task learning.

Our experiments involve using two different representations, the first one being the approach used by the original RN paper [4], to embed the context and questions into sentence embeddings using LSTMs, and the other uses the universal sentence encoder (USE) [2] to embed the context and questions into sentence embeddings. In the RN paper, sentence embeddings are called objects which the RN is uses to learn the relation between them. We then compare the performance of different embeddings on RN for bAbI question answering task.

2 Task

2.1 bAbI

The bAbI dataset is a pure text-based question answering (QA) dataset that contains a total of 20 tasks. Each task corresponds to a particular type of reasoning, such as deduction, induction and counting. Every question is associated with a set of supporting facts, which provides the context for

Submitted to Lu Wei for Machine Learning Project. Aiming to extend the work further and publish it as a paper.

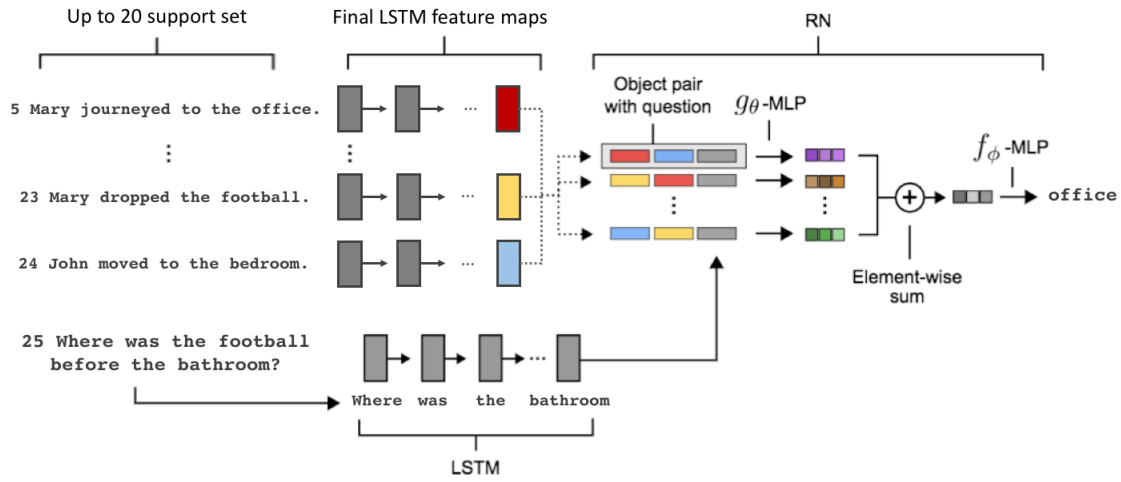

Figure 1: Text based QA architecture. Contexts and questions are processed with an LSTM to produce a set of context and question embedding. Objects, colored yellow, red, and blue, are constructed using LSTMs or USE. The RN considers relations across all pairs of objects, conditioned on the question embedding, and integrates all these relations to answer the question. Our alternative approach substities the LSTM with the USE to produce the embeddings.

the question being asked. An example "Sandra picked up the football" and "Sandra went to the office" support the question "Where is the football?", which we humans can arrive at an easy at the answer "office". A task is considered to be successfully passed if it attains an accuracy of 95% or higher.

2.2 Two Supporting Fact Task

Task 2 of the bAbI tasks requires chaining of two or three supporting facts to answer the question. As such, to answer the question "Where is the football?", it has to possibly be able to link information from the sentences "Mary moved to the bathroom", "Mary picked up the football there" and "Mary went back to the garden" for it to conclude that the football is at the garden. This tasks makes it challenging for the neural network as it requires some form of memory for it to be able to link previously acquired knowledge to answer the question.

(To talk about how other networks fare for task2, like those with memory perform better and those without memory) Thus Memory Networks [7] which uses long-term memory that can be read and written to

End-to-End Memory Networks [5]

3 Model

1. Overview of the original RN model, comment on the strength and weaknesses
2. Modifications to the RN model that will help improve the accuracy of the task. Motivations for the modifications.
3. (Optional) A paragraph on USE?
4. How long we take to train our model and the train/test accuracy, loss values etc. Use original RN paper as a guideline of what numbers to show.

3.1 Relation Network

The RN is a neural network module which is designed to do relational reasoning. It is a composite function whose form is given by:

$$RN(O) = f_\phi \left(\sum_{i,j} g_\theta(o_i, o_j) \right) \tag{1}$$

where the input is a set of "objects" $O = \{o_1, \ldots, o_n\}$, $o_i \in \mathbb{R}^m$ is the i^{th} object, and f_ϕ and g_θ are function with learnable parameters ϕ and θ respectively. f_ϕ and g_θ are mulit-layer perceptrons (MLP), where g_θ learns the relation between two given objects and f_ϕ maps the relations to one of the many possible answers.

RNs take in objects as its input and do not explicitly operate on natural language

In their simplest form RNs operate on objects, and hence do not explicitly operate on images or natural language. A central contribution of this work is to demonstrate the flexibility with which relatively unstructured inputs, such as CNN or LSTM embeddings, can be considered as a set of objects for an RN. Although the RN expects object representations as input, the semantics of what an object is need not be specified. Our results below demonstrate that the learning process induces upstream processing, comprised of conventional neural network modules, to produce a set of useful "objects" from distributed representations.

In the original RN model by [4] for the bAbI task,

In the model used by [4], up to a maximum of 20 sentences in the support set was processed through a 32 unit LSTM to produce an object. The g_θ is a four-layer MLP which contains 256 units per layer, and takes as input all possible pairings of of the sentences in the support set, concatenated with the

3.2 Embeddings

Short paragraph on embeddings,bert, Elmo, glove, etc.

3.2.1 LSTM

hashing on the word level, the lstm to obtain an embedding on the sentence level.

3.2.2 Universal Sentence Encoder

paper, tensorflow blog

4 Results

Example of results: Our model succeeded on 18/20 tasks. Notably, it succeeded on the basic induction task (2.1% total error), which proved difficult for the Sparse DNC (54%), DNC (55.1%), and EntNet (52.1%). Also, our model did not catastrophically fail in any of the tasks: for the 2 tasks that it failed (the "two supporting facts", and "three supporting facts" tasks), it missed the 95% threshold by 3.1% and 11.5%, respectively. We also note that the model we evaluated was chosen based on overall performance on a withheld validation set, using a single seed. That is, we did not run multiple replicas with the best hyperparameter settings (as was done in other models, such as the Sparse DNC, which demonstrated performance fluctuations with a standard deviation of more than ± 3 tasks passed for the best choice of hyperparameters). 5.5

5 Discussion and Conclusions

References

[1] L. Bottou. From Machine Learning to Machine Reasoning. *Arxiv preprint arXiv11021808*, page 15, 2011. ISSN 0885-6125. doi: 10.1007/s10994-013-5335-x. URL `http://arxiv.org/abs/1102.1808`.

[2] D. Cer, Y. Yang, S.-y. Kong, N. Hua, N. Limtiaco, R. S. John, N. Constant, M. Guajardo-Cespedes, S. Yuan, C. Tar, Y.-H. Sung, B. Strope, and R. Kurzweil. Universal Sentence Encoder. 2018. ISSN 1042-1629. doi: 10.1007/s11423-014-9358-1. URL http://arxiv.org/abs/1803.11175.

[3] J. Johnson, B. Hariharan, L. V. D. Maaten, J. Hoffman, L. Fei-Fei, C. L. Zitnick, and R. Girshick. Inferring and Executing Programs for Visual Reasoning. *Proceedings of the IEEE International Conference on Computer Vision*, 2017-Octob:3008–3017, 2017. ISSN 15505499. doi: 10.1109/ICCV.2017.325.

[4] A. Santoro, D. Raposo, D. G. T. Barrett, M. Malinowski, R. Pascanu, P. Battaglia, and T. Lillicrap. A simple neural network module for relational reasoning. pages 1–16, 2017. ISSN 21607516. doi: 10.1109/WACV.2017.108. URL http://arxiv.org/abs/1706.01427.

[5] S. Sukhbaatar, A. Szlam, J. Weston, and R. Fergus. End-To-End Memory Networks. pages 1–11, 2015. ISSN 10495258. doi: v5. URL http://arxiv.org/abs/1503.08895.

[6] J. Weston, A. Bordes, S. Chopra, A. M. Rush, B. van Merriënboer, A. Joulin, and T. Mikolov. Towards AI-Complete Question Answering: A Set of Prerequisite Toy Tasks. 2015. ISSN 03787753. doi: 10.1016/j.jpowsour.2014.09.131. URL http://arxiv.org/abs/1502.05698.

[7] J. Weston, S. Chopra, and A. Bordes. Memory Networks. pages 1–15, 2015. ISSN 1098-7576. doi: v0. URL https://arxiv.org/pdf/1410.3916.pdf.

6 Submission of papers to NIPS 2018

NIPS requires electronic submissions. The electronic submission site is

`https://cmt.research.microsoft.com/NIPS2018/`

Please read the instructions below carefully and follow them faithfully.

6.1 Style

Papers to be submitted to NIPS 2018 must be prepared according to the instructions presented here. Papers may only be up to eight pages long, including figures. Additional pages *containing only acknowledgments and/or cited references* are allowed. Papers that exceed eight pages of content (ignoring references) will not be reviewed, or in any other way considered for presentation at the conference.

The margins in 2018 are the same as since 2007, which allow for $\sim 15\%$ more words in the paper compared to earlier years.

Authors are required to use the NIPS LaTeX style files obtainable at the NIPS website as indicated below. Please make sure you use the current files and not previous versions. Tweaking the style files may be grounds for rejection.

6.2 Retrieval of style files

The style files for NIPS and other conference information are available on the World Wide Web at

`http://www.nips.cc/`

The file `nips_2018.pdf` contains these instructions and illustrates the various formatting requirements your NIPS paper must satisfy.

The only supported style file for NIPS 2018 is `nips_2018.sty`, rewritten for LaTeX 2_ε. **Previous style files for LaTeX 2.09, Microsoft Word, and RTF are no longer supported!**

The LaTeX style file contains three optional arguments: `final`, which creates a camera-ready copy, `preprint`, which creates a preprint for submission to, e.g., arXiv, and `nonatbib`, which will not load the `natbib` package for you in case of package clash.

New preprint option for 2018 If you wish to post a preprint of your work online, e.g., on arXiv, using the NIPS style, please use the `preprint` option. This will create a nonanonymized version of your work with the text "Preprint. Work in progress." in the footer. This version may be distributed as you see fit. Please **do not** use the `final` option, which should **only** be used for papers accepted to NIPS.

At submission time, please omit the `final` and `preprint` options. This will anonymize your submission and add line numbers to aid review. Please do *not* refer to these line numbers in your paper as they will be removed during generation of camera-ready copies.

The file `nips_2018.tex` may be used as a "shell" for writing your paper. All you have to do is replace the author, title, abstract, and text of the paper with your own.

The formatting instructions contained in these style files are summarized in Sections 7, 8, and 9 below.

7 General formatting instructions

The text must be confined within a rectangle 5.5 inches (33 picas) wide and 9 inches (54 picas) long. The left margin is 1.5 inch (9 picas). Use 10 point type with a vertical spacing (leading) of 11 points. Times New Roman is the preferred typeface throughout, and will be selected for you by default. Paragraphs are separated by ½ line space (5.5 points), with no indentation.

The paper title should be 17 point, initial caps/lower case, bold, centered between two horizontal rules. The top rule should be 4 points thick and the bottom rule should be 1 point thick. Allow ¼ inch space above and below the title to rules. All pages should start at 1 inch (6 picas) from the top of the page.

For the final version, authors' names are set in boldface, and each name is centered above the corresponding address. The lead author's name is to be listed first (left-most), and the co-authors' names (if different address) are set to follow. If there is only one co-author, list both author and co-author side by side.

Please pay special attention to the instructions in Section 9 regarding figures, tables, acknowledgments, and references.

8 Headings: first level

All headings should be lower case (except for first word and proper nouns), flush left, and bold.

First-level headings should be in 12-point type.

8.1 Headings: second level

Second-level headings should be in 10-point type.

8.1.1 Headings: third level

Third-level headings should be in 10-point type.

Paragraphs There is also a `\paragraph` command available, which sets the heading in bold, flush left, and inline with the text, with the heading followed by 1 em of space.

9 Citations, figures, tables, references

These instructions apply to everyone.

9.1 Citations within the text

The `natbib` package will be loaded for you by default. Citations may be author/year or numeric, as long as you maintain internal consistency. As to the format of the references themselves, any style is acceptable as long as it is used consistently.

The documentation for `natbib` may be found at

`http://mirrors.ctan.org/macros/latex/contrib/natbib/natnotes.pdf`

Of note is the command `\citet`, which produces citations appropriate for use in inline text. For example,

```
\citet{hasselmo} investigated\dots
```

produces

Hasselmo, et al. (1995) investigated…

If you wish to load the `natbib` package with options, you may add the following before loading the `nips_2018` package:

```
\PassOptionsToPackage{options}{natbib}
```

If `natbib` clashes with another package you load, you can add the optional argument `nonatbib` when loading the style file:

```
\usepackage[nonatbib]{nips_2018}
```

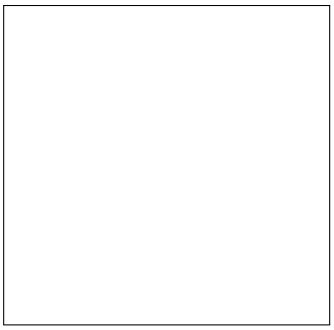

Figure 2: Sample figure caption.

As submission is double blind, refer to your own published work in the third person. That is, use "In the previous work of Jones et al. [4]," not "In our previous work [4]." If you cite your other papers that are not widely available (e.g., a journal paper under review), use anonymous author names in the citation, e.g., an author of the form "A. Anonymous."

9.2 Footnotes

Footnotes should be used sparingly. If you do require a footnote, indicate footnotes with a number[1] in the text. Place the footnotes at the bottom of the page on which they appear. Precede the footnote with a horizontal rule of 2 inches (12 picas).

Note that footnotes are properly typeset *after* punctuation marks.[2]

9.3 Figures

All artwork must be neat, clean, and legible. Lines should be dark enough for purposes of reproduction. The figure number and caption always appear after the figure. Place one line space before the figure caption and one line space after the figure. The figure caption should be lower case (except for first word and proper nouns); figures are numbered consecutively.

You may use color figures. However, it is best for the figure captions and the paper body to be legible if the paper is printed in either black/white or in color.

9.4 Tables

All tables must be centered, neat, clean and legible. The table number and title always appear before the table. See Table 1.

Place one line space before the table title, one line space after the table title, and one line space after the table. The table title must be lower case (except for first word and proper nouns); tables are numbered consecutively.

Note that publication-quality tables *do not contain vertical rules.* We strongly suggest the use of the `booktabs` package, which allows for typesetting high-quality, professional tables:

`https://www.ctan.org/pkg/booktabs`

This package was used to typeset Table 1.

10 Final instructions

Do not change any aspects of the formatting parameters in the style files. In particular, do not modify the width or length of the rectangle the text should fit into, and do not change font sizes (except perhaps in the **References** section; see below). Please note that pages should be numbered.

[1] Sample of the first footnote.

[2] As in this example.

Table 1: Sample table title

	Part	
Name	Description	Size (μm)
Dendrite	Input terminal	$\sim$100
Axon	Output terminal	$\sim$10
Soma	Cell body	up to 10^6

11 Preparing PDF files

Please prepare submission files with paper size "US Letter," and not, for example, "A4."

Fonts were the main cause of problems in the past years. Your PDF file must only contain Type 1 or Embedded TrueType fonts. Here are a few instructions to achieve this.

- You should directly generate PDF files using `pdflatex`.
- You can check which fonts a PDF files uses. In Acrobat Reader, select the menu Files>Document Properties>Fonts and select Show All Fonts. You can also use the program `pdffonts` which comes with `xpdf` and is available out-of-the-box on most Linux machines.
- The IEEE has recommendations for generating PDF files whose fonts are also acceptable for NIPS. Please see `http://www.emfield.org/icuwb2010/downloads/IEEE-PDF-SpecV32.pdf`
- `xfig` "patterned" shapes are implemented with bitmap fonts. Use "solid" shapes instead.
- The `\bbold` package almost always uses bitmap fonts. You should use the equivalent AMS Fonts:

  ```
  \usepackage{amsfonts}
  ```

  followed by, e.g., `\mathbb{R}`, `\mathbb{N}`, or `\mathbb{C}` for $\mathbb{R}$, $\mathbb{N}$ or $\mathbb{C}$. You can also use the following workaround for reals, natural and complex:

  ```
  \newcommand{\RR}{I\!\!R} %real numbers
  \newcommand{\Nat}{I\!\!N} %natural numbers
  \newcommand{\CC}{I\!\!\!\!C} %complex numbers
  ```

  Note that `amsfonts` is automatically loaded by the `amssymb` package.

If your file contains type 3 fonts or non embedded TrueType fonts, we will ask you to fix it.

11.1 Margins in LaTeX

Most of the margin problems come from figures positioned by hand using `\special` or other commands. We suggest using the command `\includegraphics` from the `graphicx` package. Always specify the figure width as a multiple of the line width as in the example below:

```
\usepackage[pdftex]{graphicx} ...
\includegraphics[width=0.8\linewidth]{myfile.pdf}
```

See Section 4.4 in the graphics bundle documentation (`http://mirrors.ctan.org/macros/latex/required/graphics/grfguide.pdf`)

A number of width problems arise when LaTeX cannot properly hyphenate a line. Please give LaTeX hyphenation hints using the `\-` command when necessary.

Acknowledgments

Use unnumbered third level headings for the acknowledgments. All acknowledgments go at the end of the paper. Do not include acknowledgments in the anonymized submission, only in the final paper.

References

References follow the acknowledgments. Use unnumbered first-level heading for the references. Any choice of citation style is acceptable as long as you are consistent. It is permissible to reduce the font size to `small` (9 point) when listing the references. **Remember that you can use more than eight pages as long as the additional pages contain *only* cited references.**

[1] Alexander, J.A. & Mozer, M.C. (1995) Template-based algorithms for connectionist rule extraction. In G. Tesauro, D.S. Touretzky and T.K. Leen (eds.), *Advances in Neural Information Processing Systems 7*, pp. 609–616. Cambridge, MA: MIT Press.

[2] Bower, J.M. & Beeman, D. (1995) *The Book of GENESIS: Exploring Realistic Neural Models with the GEneral NEural SImulation System.* New York: TELOS/Springer–Verlag.

[3] Hasselmo, M.E., Schnell, E. & Barkai, E. (1995) Dynamics of learning and recall at excitatory recurrent synapses and cholinergic modulation in rat hippocampal region CA3. *Journal of Neuroscience* **15**(7):5249-5262.