Deep Learning — Assignment 4

Anirudhan J. Rajagopalan
Department of Computer Science
New York University
New York, NY.
ajr619@nyu.edu

1 nnGraph

1.1 1. Warmup

The code for nngraph_warmup.lua can be found at `https://git.io/vwQco`

1.1.1 2. Grucell diagram

The gru cell was drawn using the following steps.

1. Code the cell in torch similar to the code in main.lua
2. Plot the code using graph.dot function passing the filename argument
3. Open the .svg file in browser and remove the unwanted nodes.

The cell diagram generated is included in 1.

2 Language Modeling

2.1 Generating sequences

The query_sentences.lua can be found at `https://git.io/vwQEc`.

The query_sentences.lua does the following

1. Loads the core network of the model.
2. Builds the vocabulary map (max 10,000) and the inverse vocabulary map.
3. Fetches the number of words to generate and the initial seed words (minimum 2).
4. Does a forward pass on the core_network for each and every word to generate the index for next word.
5. The index is generated by using a multinomial distribution over the probabilities generated by the logsoftmax layer (layer 44 in core_network)
6. Concatenates and returns the new sentence.

Steps to run the model: *th query_sentences.lua*

2.2 Improvements to the model

29th Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, Spain.

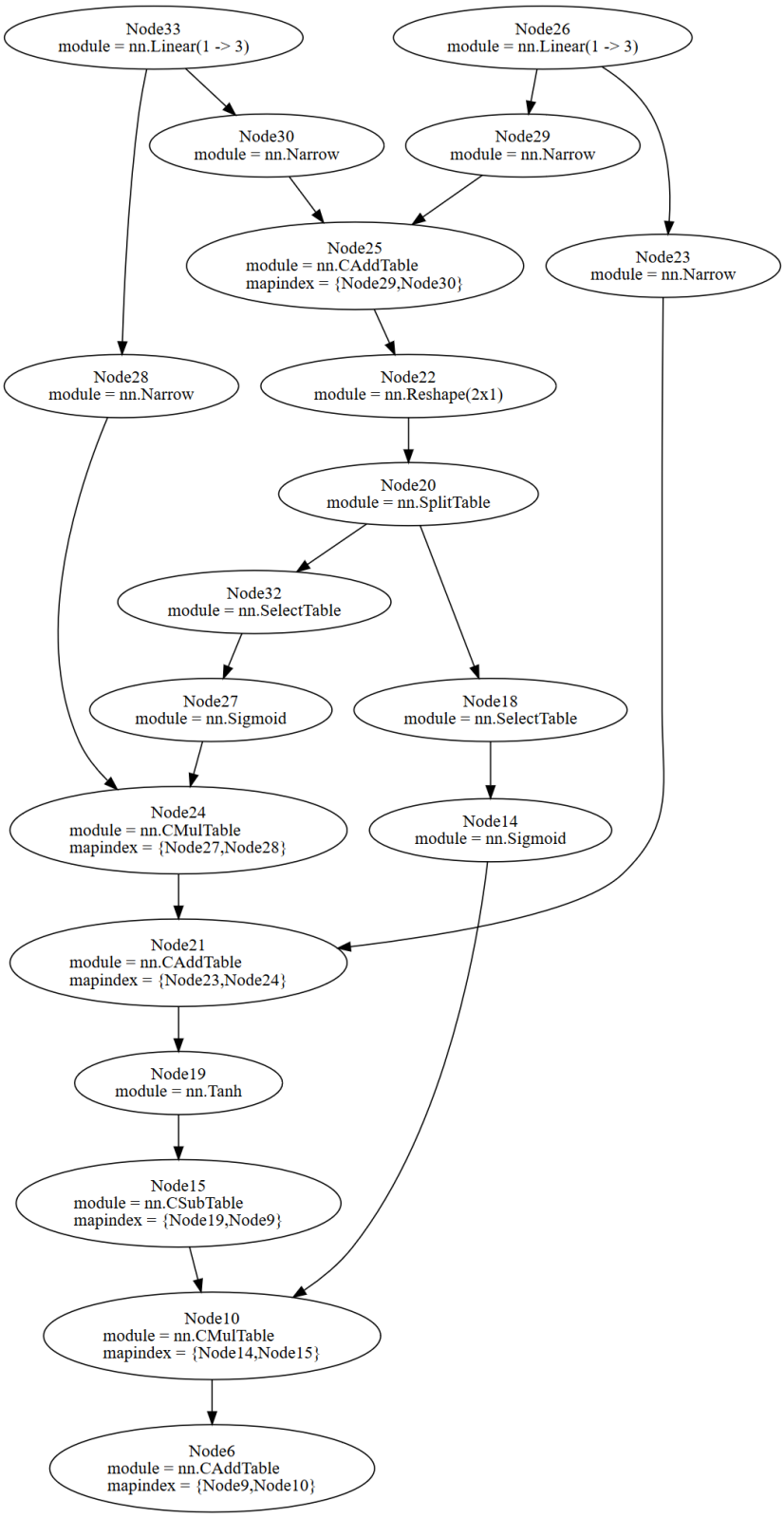

Figure 1: GRUCell given in slide 32 of talk by Armand Joulin