Tingyu Thomas Lin

Juan M. Navarro

CS 263B – Natural Language Processing

Final Project Report

Due: Monday, June 8th, 2009

**A Perceptron-Based Approach for Image-Word Association**

Our project has 2 initial purposes:

1. Associate an image to a concrete noun
2. Understand relative positioning of objects

Two modules were designed to handle each purpose:

1. Image learning module
2. Relation learning module

Image learning module

1. The input/output that we used to train the image module consists of 7 word/image pairs:
2. ghost
3. blizzard
4. tree
5. house
6. car
7. window
8. box
9. We tried two different architectures for our system:
10. First approach consisted of a multi-layer perceptron:
11. Architecture
    1. 3 layers: input, hidden and output layers
    2. Logistic sigmoid function applied at the outputs of hidden and output layers
    3. Learning is based on the standard feed-forward backwards propagating neural network
12. Three approaches were taken to implement the learning algorithm:
13. Online training, where the weights are updated after each training image, was our first approach.
14. Rprop
15. Rprop+
16. Second approach consisted of a reverberating Elman network
    1. Architecture
    2. Issues
17. Current status
    1. Online learning: converges in 1737 epochs for the 7 image set with a learning rate of 0.8
    2. Rprop learning: converges in 131 epochs for the 7 image set

Relations learning module

1. The input/output that we used to train the relations module consists of 196 <noun><relative position><noun> phrases:
2. Four relations were implemented: left, right, above, below
3. Example: house right car, tree above box
4. 16 phrases removed from training set and into testing set
5. Architecture: similar to image learning module
6. Current status

Three major scalability problems were identified:

1. Increasing training test size results in noisy output
2. Online learning method is obscenely slow for large training sets
3. Batch learning method get stuck at local minima

Current status:

1. Learning algorithms
2. Catastrophic forgetting

Description of software packages used:

1. Explain why Matlab was used

Four approaches were considered to prevent catastrophic forgetting:

1. Rehearsing original set of images
2. Generating pseudo-patterns and training on a single neural net
3. Coupling two neural networks
4. Evolution-based approach

Appendix

References