

IQBAL

Ahsan

ahsiq139

Post project paper

Transition-Based Dependency Parsing with Rich Non-Local Features

1497 words

Table of contents

Part 1	2
Part 2	Erreur ! Signet non défini.
Part 3	Erreur ! Signet non défini.

Part 1

The project selected by my group was to study a transition-based dependency parser and see how it can be improved by adding new features. To prepare myself for the project I started by reading all I could find regarding arc-standard algorithm and arc-eager algorithm to have a more proper understanding of how they work and their differences. Then I focused on the paper which the project was based on Transition-Based Dependency Parsing with Rich Non-Local Features by Zhang and Nivre. After that I had a better understanding of what our aim was and how we would do it, but I still had troubles with some of the new features introduced in the article, so I did other research to grasp their meaning and how to implement them.

At the beginning, the group agreed what level of analyse we were aiming for and with that in mind wrote the project plan. We then estimated the amount of work needed and divided the work knowing that it was not definitive and could change after. Since Anass and I had the less experience in NLP and machine learning more generally we agreed with Joar and Martin that they will be mainly in charge of implementing the baseline so we could have more time to works on the features. Therefore, I was in pair with Anass, and we had to implement 2 features. Then, when all the features were implemented, I was in charge to analyse them on a Chinese dataset.

After implementing the baseline, we implemented the features. With Anass we were in charge of the Distance and the Valency. We then ran tests to determine if there was an improvement of the Unlabelled Attachment Score (UAS). Unfortunately, this showed that our implementation of the distance was not working, and so Martin proposed to look it again to see if he could do something. He tried to employ other definition of distance, but we still could not make this feature useful. The results showed that the implementation of all features except distance gave a better improvement of the UAS of the parser than all features and that valency resulted in the largest improvement of the UAS for both English and Chinese data. The article of Zhang and Nivre implemented 5 features, but since we decided to have an arc-standard parser instead of an arc-eager (like in the article.) we had to alter the definition of some features and wipe out the ones we found incompatible with our system. By doing so, we ended with 4 features adapted from the ones in the article. We had Distance and Valency as stated before to which we added Third order and unigram model. When we first conducted our test with all the features, it took more than 4 hours for one epoch, so we had to go over our code to find a way to reduce this. Hopefully by changing some "for" loops, we managed to do it. Implementing the features did not take very long, but we lost a lot of time on our training loop, and by trying to make distance work by implementing different definition of distance we could found. Because of that we decided to stick to the English data and Chinese one and do a deeper analysis of it rather than trying other languages and features.

After the project presentation we received positive feedback on how our presentation was easy to follow, had some good examples and that the discussion on each feature was interesting. Nonetheless, by making the presentation easy to understand some people stated that we could have gone deeper on some concepts that we introduced instead of just showing some examples.

Part 2

The main purpose of our project was to implement a syntactic parser and improve it by adding new features. A syntactic parser takes a sentence as an input and gives each word in the sentence a part-of-speech tag in order to describe the word function and then it gives the relation between words in the sentence. There is also the baseline which was crucial for our project. Indeed, it is thanks to the baseline that we were able to evaluate the performance of our parser and compare it to when we add new features. We first measured the UAS of the baseline alone and then we added each feature one by one to see how each feature affected the score. After that we had a good overview of the influence of each feature which led us to try the baseline with all the features and the baseline with all features except distance.

Before attending this course I only had one course about machine learning and neural network. Therefore, it was hard for me to follow up with the labs and complete them in time. Because of that my understanding of the concepts used in the project was low but hopefully with the help of my group and my personal work I was able to overcome that. I believe that during the project I gained confidence in my understanding of the lab material which led me to have an accurate view on how to do a project in NLP, what are the different steps, how to test them and finally how to analyse them.

The course syllabus also states that at the end of the course a student should be able to “identify, assess and make use of NLP research literature.” I think that this is the point on which I have improved the most. Before the project I had barely looked at any NLP literature, I principally used the course video. But to overcome my shortcomings I had to read a lot of articles to better comprehend how the NLP techniques used in our project works and then how we can modify them in order to implement them in our arc-standard algorithm. The other point on which I most improved is “evaluate NLP components and systems with appropriate methods.” Given that we had to adapt the notion used in the article I had no other choice than manipulate the definitions and evaluate them to know if they had any interest for us.

Part 3

As I remarked it before, at the beginning it was very hard for me to be able to understand NLP articles. It is one of the reasons why we selected one of the first standard project, it was more accessible. In fact, our project was predominantly based on the article of Zhang and Nivre even if we all looked at other resources to help us implement the features. To overcome my difficulties, I had to read multiple times the article and focus on one aspect each time. When I was stuck on one, I tried to understand each specific word to better define what was blocking me and so I could ask more precise question to my group in order to get help. Sometimes I tried to find the same topic in a different article in hope to observe example which might help me to have a better grasp. For the future I will try to follow this method by first identifying the words and notion causing me trouble, then try to find other example in other resources in order to have a different approach of the subject and finally discuss what I have learned and understood of the topic with my group.

In the previous part I also mentioned that we had to evaluate the article to be able to work with it and modify some of its aspect. The first significant change we undertook was to go from an arc-eager parser to an arc-standard one. This led to adapt all the features to our baseline and choose how we were going to do it. Moreover, we used different treebanks, our was the Universal Dependencies Treebank whereas in the article it was the Penn Treebank. Lastly, we evaluated our parser with each feature individually, with all of them and all feature except distance while on the article they added them incrementally. All these differences are important to have in mind when we compare our results with the paper's one. Nonetheless, I regret some choice we made, and I think we could have improved our analyses. For example, when testing on Chinese dataset we found that distance was a useful feature, which was opposed to our first idea that distance was not a feature working with arc-standard parser. Instead of trying to find new definition I think we should have tested the distance on other languages from other family trees. I did not have this sort of reasoning before doing this project. During the lab I mostly tried to follow the instruction and was not able to have a critical opinion about works related to NLP or machine learning. In the future I will do my best to keep this in mind and have a global view of the project instead of sticking with one idea.