CMU Portugal
Undergraduate Internship Program - Summer 2017

António Lopes

antonio.vilarinho.lopes@tecnico.ulisboa.pt

Instituto Superior Técnico, Universidade de Lisboa

This report summarizes the activities carried out during my internship at Carnegie Mellon University, which took place from June 21st to September 1st, 2017. I was hosted and advised by Professor Graham Neubig and sponsored by Professor David Matos. During my stay, I had the opportunity to work at the Language Technologies Institute (LTI).

The initial proposal was to work in Natural Language Generation (NLG) for cross-lingual systems. However, due to the short nature of this internship and also Professor Neubig expertise on Machine Translation (MT), we decided to work on methods for document-level Machine Translation that can be applied to NLG. Therefore, we proposed setting a baseline for document-level MT and improve it with sophisticated methods used in machine learning.

During the internship, I implemented both the baseline and one of the sophisticated methods discussed with Professor Graham, working closely with one of his Master's students, Paul Michel, who helped me with some implementation details and also provided some working code as I used Professor Graham's deep learning framework and I was not proficient using it. Even though the internship has ended, I'm currently working on other methods for document-level MT discussed with Professor Graham, which will likely result in a paper.

Furthermore, I had the opportunity to work closely with Professor Graham and his Master/PhD students, we had weekly meetings where everyone would discuss their weekly progress as well as relevant papers. In addition, I also had the opportunity to be part of a weekly MT reading group with Professor Graham and other Professors and students. These meetings were one of the most valuable aspects of my stay as I had the opportunity to discuss important aspects of MT and Machine Learning with renowned scholars.

During my time at CMU, the LTI hosted the Jelinek Summer Workshop on Speech and Language Technology (JSALT) and I was able to attend the plenary lectures with renowned invited speakers, such as Brian MacWhinney, Luke Zettlemoyer, among others. As JSALT was hosted by LTI, a considerable amount of the scholars at the workshop were also at the MT meetings, which led to one of the highlights of my stay: Philipp Koehn, a renowned scholar from John Hopkins University who works on MT and was a visitor at the workshop, presenting one of his recent papers in one of the meetings.

Prior to my return, I had opportunity to attend a couple of classes with Apple's research director Professor Ruslan Salakhutdinov and, my host, Professor Graham Neubig, as classes at CMU started the week I returned to Portugal. Regarding Pittsburgh, I enjoyed my stay in the city and during my free time I visited the Carnegie Museum of Natural History, where I had a well spent time.

To conclude, my stay at CMU was an incredible and enriching opportunity, which was made possible by CMU-Portugal. Therefore, I would like to thank the CMU-Portugal Program, CMU, and the Portuguese Science and Technology Foundation for providing me this experience. Finally, a special thank you to Professor Graham and Professor David for hosting and sponsoring me.