

# Automated Characters Recognition and Family Relationship Extraction

Alisha Bajracharya (070bct504@ioe.edu.np)<sup>1</sup>, Saurav Shrestha<sup>1</sup>, Sharmila Upadhyaya<sup>1</sup>, Shrawan Bk<sup>1</sup>, Subarna Shakya<sup>1</sup>

<sup>1</sup> Department of Electronics and Computer Engineering, Pulchowk Campus, IOE, Tribhuvan University, Nepal.

## Abstract

“Automated characters recognition and family relationship extraction” is an application of Natural Language Processing to identify characters from the story and determine the family relationship among them. This application is the use of specialized computer programs to identify entities, classify them and extract characters from them and determine relationship between them.

This paper follows basic steps of NLP i.e. Tokenization, POS tagging, sentence parsing followed by the pronoun resolution implementing various algorithms and finally extracting entities and relations among them. Heretofore, we have successfully resolved pronoun from simple sentences by resolving Noun Phrase using the recursive algorithm for tree generation and hence extracting relation between the Noun Phrase (NP). Basic approach towards this project is to do Tokenization and POS tagging first. Then, sentence which is recursive composition of Noun phrase, verb phrase and prepositional phrase is parsed and recursive tree is generated. Then tree is traversed to determine the noun phrase which is replaced by the entity object of that particular noun phrase. Pronoun resolution is the essence of NLP and is of different type. Here, Co reference resolution has been used. After resolving the entire pronoun, then finally relationship is extracted from the story by comparing the relation ID of each Entity.

Given the simple story, entities are being extracted and relationship is also determined. Understanding the approach of NLP and implementing them to showcase its use is the main theme of this project which is being done with as accurate result as possible. This paper can act as a base for story summarization, grasping insight of story and analysis of characters of story as well.

**Keywords:** *Noun / Verb / Prepositional Phrase, Pronoun resolution, Natural language processing*