

Entity Relation and Attribute Extraction

Burhan Can Akkuş
Burak Çetin
Department of Computer Engineering
Bogazici University
Istanbul Turkey
can.akkus@boun.edu.tr
burak.cetin@boun.edu.tr

May 7, 2021

Abstract

Entity Relation and Attribute Extraction project is designed to scan a text and extract the entities in the text and the relations between these entities. These relations can be various things ranging from one character's opinions towards another to having certain ontological connections such as being married to each other. We aim to create an automatized tool to summarize news and stories in a similar fashion to Character Maps of novels.

1 Introduction

Entity Relation and Attribute Extraction(ERAE) is an NLP project that analyzes a given text (preferably a news story or a short novel or a play) and creates a Character Map for this text. In the Character Map we aim to include the characters and entities of the story, their relations with each other, their attributes, opinions and beliefs. We also aim to visualize these findings in an easy to digest manner. Final goal of the ERAE is to implement the above mentioned pipeline to process Turkish texts.

2 Related Work

2.1 Stanford CoreNLP

Stanford CoreNLP [1] is the State-of-the-Art natural language processing library. It includes various modules. We use the ones relevant to our project, namely tokenize, ssplit, pos, lemma, ner, parse and coref modules. The first four are the standart modules. NER is Named Entity Recognition which is the first step for our project. It lists named entities such as people and location names. The coref module is coreference resolution. We use this module to replace the pronouns with their respective proper nouns.

2.2 BBC's Storyline Ontology

BBC's Storyline Ontology [2] is an ontology of structured narratives like news. It partitions the story to its main components. We use this components to track the changes in characters' opinions through the text.

2.3 Ontological Representations of Narratives: a Case Study on Stories and Actions

Rossana Damiano and Antonio Lieto's paper [3] on the ontology of characters is another work that is related to our project. It allows users to explore a digital archive by following the narrative relations among the resources contained in it. This work is especially important for us because it lets our project to make inferences using information not present in the given text such as cultural references or geographical references.

3 Approach

3.1 Coreference Resolution

Our first step is to do a Coreference Resolution on the given text to replace pronouns with their proper nouns. For this task we used Stanford CoreNLP's coreference module. Each reference chain is resolved to its head noun. This step ignores Named Entities however this is not an issue since we will extract these Named Entities in the next step.

```
Original Text:  
Barack Obama was born in Hawaii. He is the president. Obama was elected in 2008.  
De-coreferenced Text:  
Barack Obama was born in Hawaii. Barack Obama is the president. Barack Obama was elected in 2008.
```

Figure 3.1: Example use of Coreference Resolution.

3.2 Named Entity Recognition

We use Stanford CoreNLP's Named Entity Recognition module for entity extraction. This module extracts Named Entities and their category ie. Person, Location, Date etc.. In Figure 3.2 you can see a NER extract of a news story about Turkey's Withdrawal from Istanbul Convention. [4]

```

entities found
  detected entity:  President    TITLE
  detected entity:  Recep Tayyip Erdogan    PERSON
  detected entity:  midnight    TIME
  detected entity:  last Friday DATE
  detected entity:  Turkey    COUNTRY
  detected entity:  violence    CAUSE_OF_DEATH
  detected entity:  Turkish NATIONALITY
  detected entity:  Council of Europe    ORGANIZATION
  detected entity:  domestic violence    CRIMINAL_CHARGE
  detected entity:  34    NUMBER
  detected entity:  European    NATIONALITY
  detected entity:  2014    DATE
  detected entity:  Istanbul    CITY
  detected entity:  one    NUMBER
  detected entity:  Erdogan PERSON
  detected entity:  Turkey    COUNTRY
  detected entity:  he    PERSON

```

Figure 3.2: Example output of Named Entity Recognition.

3.3 Relation Extraction

This step is not implemented yet.

3.4 Timeline of Relations

This step is not implemented yet

4 Status

For our initial test subject, we decided to use children's fairy tales because they are simple in structure, contain a small number of characters and the relations between these characters are simple. Little Red Riding Hood [5] is our choice of fairy tale as our first subject.

The first step is to use coreference resolution to replace pronouns with their respective proper nouns. You can see the results of this step in figure 4.2, 4.3 and 4.4. In this step each reference chain is collapsed to its head. We highlighted these heads in different colors.

Our next step is the Named Entity Recognition step. You can see the results of this step in figure 4.1. NER recognizes only the entities whose names are explicitly given. For this reason it misses the main characters the mother and the woodsman. We plan to solve this problem by using the data acquired in the Coreference Resolution part in the future.

entities found

detected entity:	Once	DATE
detected entity:	morning	TIME
detected entity:	Little Red Riding Hood	MISC
detected entity:	summer	DATE
detected entity:	day	DURATION
detected entity:	Wolf	PERSON
detected entity:	Wolf	PERSON
detected entity:	Wolf	PERSON
detected entity:	Wolf	PERSON
detected entity:	Granny	PERSON
detected entity:	Granny	PERSON
detected entity:	few minutes later	DATE
detected entity:	Wolf	PERSON
detected entity:	her	PERSON
detected entity:	he	PERSON
detected entity:	Wolf	PERSON
detected entity:	one	NUMBER
detected entity:	he	PERSON
detected entity:	he	PERSON
detected entity:	her	PERSON
detected entity:	Little Red Riding Hood	MISC
detected entity:	he	PERSON
detected entity:	he	PERSON

Figure 4.1: Little Red Riding Hood's entities after Coreference Resolution.

Little Red Riding Hood

Once upon a time, there was a **little girl who lived in a village near the forest.**

Whenever a **little girl who lived in a village near the forest** went out, the little girl wore a red riding cloak, so everyone in the village called a **little girl who lived in a village near the forest** **Little Red Riding Hood**.

One morning, **Little Red Riding Hood** asked a **little girl who lived in a village near the forest** **mother** if a **little girl who lived in a village near the forest** could go to visit a **little girl who lived in a village near the forest** **grandmother** as **Little Red Riding Hood** had been awhile since they'd seen each other.

"That's a good idea," a **little girl who lived in a village near the forest** **mother** said.

So they packed a nice basket for **Little Red Riding Hood** to take to a **little girl who lived in a village near the forest** **grandmother**.

When the basket was ready, the little girl put on a **little girl who lived in a village near the forest** red cloak and kissed a **little girl who lived in a village near the forest** **mother** goodbye.

"Remember, go straight to Grandma's house," a **little girl who lived in a village near the forest** **mother** cautioned.

"Don't dawdle along the way and please don't talk to strangers! The woods are dangerous.

"Don't worry, mommy," said **Little Red Riding Hood**, "I'll be careful."

But when **Little Red Riding Hood** noticed some lovely flowers in the woods, a **little girl who lived in a village near the forest** forgot a **little girl who lived in a village near the forest** promise to a **little girl who lived in a village near the forest** **mother**.

a **little girl who lived in a village near the forest** picked a few, watched the butterflies flit about for awhile, listened to the frogs croaking and then picked a few more.

Little Red Riding Hood was enjoying the warm summer day so much, that a **little girl who lived in a village near the forest** didn't notice a dark shadow approaching out of the forest behind a **little girl who lived in a village near the forest**.

.

.

Figure 4.2: Little Red Riding Hood's entities highlighted after Coreference Resolution.

Suddenly, the **wolf** appeared beside a **little girl who lived in a village near the forest**.

"What are you doing out here, little girl?" the **wolf** asked in a voice as friendly as **Wolf** could muster.

"**Wolf**'m on **Wolf** way to see **Wolf** Grandma who lives through the forest, near the brook," **Little Red Riding Hood** replied.

Then a **little girl who lived in a village near the forest** realized how late a **little girl who lived in a village near the forest** was and quickly excused a **little girl who lived in a village near the forest**, rushing down the path to a **little girl who lived in a village near the forest** Grandma's house.

The **wolf**, in the meantime, took a shortcut.

.

.

The **wolf**, a little out of breath from running, arrived at Grandma's(?) and knocked lightly at the door.

"Oh thank goodness dear! Come in, come in! I(?) was worried sick that something had happened to you(?) in the forest," said **my Grandma who lives through the forest, near the brook** thinking that the knock was a **little girl who lived in a village near the forest** granddaughter.

The **wolf** let himself(?) in.

Poor Granny did not have time to say another word, before the **wolf** gobbled a **little girl who lived in a village near the forest** up! The **wolf** let out a satisfied burp, and then poked through Granny's wardrobe to find a nightgown that **Wolf** liked.

Wolf added a frilly sleeping cap, and for good measure, dabbed some of Granny's perfume behind **Wolf** pointy ears.

A few minutes later, Red Riding Hood knocked on the door.

The **wolf** jumped into bed and pulled the covers over **Wolf** nose.

"Who is Red Riding Hood?" **Wolf** called in a cackly voice.

"Red Riding Hood's me, **Little Red Riding Hood**."

"Oh how lovely! Do come in, me dear," croaked the **wolf**.

When **Little Red Riding Hood** entered the little cottage, a **little girl who lived in a village near the forest** could scarcely recognize a **little girl who lived in a village near the forest** Grandmother.

Figure 4.3: Little Red Riding Hood's entities highlighted after Coreference Resolution.

"Grandmother! a **little girl who lived in a village near the forest** voice sounds so odd.

Is something the matter?" a **little girl who lived in a village near the forest** asked.

"Oh, a **little girl who lived in a village near the forest** just have touch of a cold," squeaked the **wolf** adding a cough at the end to prove the point.

"But **Grandmother!** What big ears you have," said **Little Red Riding Hood** as a **little girl who lived in a village near the forest** edged closer to the bed.

"The better to hear you with, my dear," replied the **wolf**.

"But **Grandmother!** What big eyes you have," said **Little Red Riding Hood**.

"The better to see you with, my dear," replied the **wolf**.

"But **Grandmother!** What big teeth you have," said **Little Red Riding Hood** a **little girl who lived in a village near the forest** voice quivering slightly.

"The better to eat you with, my dear," roared the **wolf** and he(?) leapt out of the bed and began to chase the little girl.

Almost too late, **Little Red Riding Hood** realized that the person in the bed was not a **little girl who lived in a village near the forest** **Grandmother**, but a hungry **wolf**.

She ran across the room and through the door, shouting, "Help! **wolf!**" as loudly as a **little girl who lived in a village near the forest** could.

A woodsman who was chopping logs nearby heard a **little girl who lived in a village near the forest** cry and ran towards the cottage as fast as he could.

He grabbed the **wolf** and made he(?) spit out the poor **Grandmother** who was a bit frazzled by the whole experience, but still in one piece.

"Oh **my Grandma who lives through the forest, near the brook**, I was so scared!" sobbed **Little Red Riding Hood**, "I'll never speak to strangers or dawdle in the forest again.

"There, there, child.

You've learned an important lesson.

Thank goodness You shouted loud enough for this kind woodsman to hear You!" The woodsman knocked out the **wolf** and carried he deep into the forest where he wouldn't bother people any longer.

Little Red Riding Hood and a **little girl who lived in a village near the forest** **Grandmother** had a nice lunch and a long chat.

Figure 4.4: Little Red Riding Hood's entities highlighted after Coreference Resolution.

5 Open Questions

5.1 Ontological Relations Map

We hope to create a system which can fit characters of a story into an ontology of characters that maps and makes inferences on inter-character relations. A really basic example of this can be seen in figure 5.1.

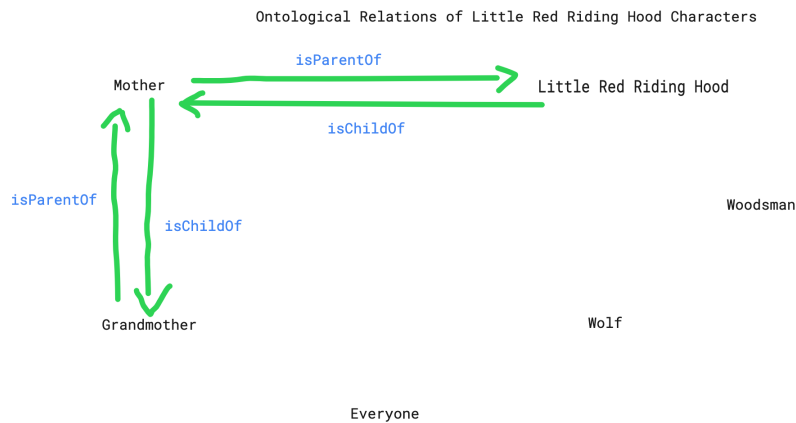


Figure 5.1: Ontological Relations map of Little Red Riding Hood

5.2 Event-Attribute Map

We aim to extract characters of a story and make inferences on their actions and reactions throughout the story arc. We want to explore what kind of relations we can discover by investigating the drama in fairy tales. We also want to extract sentimental analysis of each characters' actions and speeches regarding other characters or events. You can see an example of this, done by hand in Figure 5.2

	Little Red Riding Hood	Mother	Grandma	Wolf	Woodsman
Events	Wore a red Cloak				
Asks(Permission)(Little Red Riding Hood,Mother,to visit Grandma)	Opinion(Mother,respects authority) Opinion(Grandmother, likes)				
Warns(Mother, Little Red Riding Hood)	Cautious Opinion(Woods, dangerous)				
Reassures(Little Red Riding Hood,Mother)	Opinion(Mother, cares for)				
Forgets(her Promise to Mother)	Easily Distracted				
OutSmarts(Wolf,Little Red Riding Hood)	Naive				
OutSmarts(Wolf,Grandma)	Naive				
Eats(Grandma)	Naive				
Disguises(Wolf)	WeakerThan(Wolf)				
Eats(Grandma)	StrongerThan(Grandma)				
OutSmarts(Wolf, Little Red Riding Hood)	Cunning				
Asks(Little Red Riding Hood, Grandma)	Naive				
Asks(Little Red Riding Hood, Grandma)	Opinion(Grandma(Wolf),sounds odd)				
Asks(Little Red Riding Hood, Grandma)	Opinion(Grandma(Wolf),has big ears)				
Asks(Little Red Riding Hood, Grandma)	Opinion(Grandma(Wolf),has big teeth)				
Chases(Wolf, Little Red Riding Hood)	Opinion(Grandma(Wolf), wants to eat better,Little Red Riding Hood)				
Escapes(Little Red Riding Hood, Wolf)	BadIntentionsTowards(Little Red Riding Hood)				
Helps(Woodsman, Little Red Riding Hood)	FasterThan(Wolf)				
Beats(Woodsman, Wolf)	SlowerThan(Little Red Riding Hood)				
Saves(Woodsman, Little Red Riding Hood and Grandma)	Opinion(Little Red Riding Hood, likes)				
	WeakerThan(Woodsman)				
	StrongerThan(Wolf)				
	Opinion(Little Red Riding Hood and Grandma, likes)				

Figure 5.2: Event-Attribute map of Little Red Riding Hood

5.3 Fairy Tale Ontology

Story arcs of Fairy tales are simplistic structures. There is a significant overlap between different tales across multiple cultures. We want to explore what kind of elements are common in children's stories. We want to further develop this model to include short stories such as news or blog posts.

5.4 Story Through Multiple Perspectives

Stories usually have one or more main characters and a few side characters. Each of these characters goes through a different set of events. We want to explore each version of the story experienced by different characters and draw some conclusions regarding the functions of the characters in the story based on this information.

Bibliography

- [1] Mihai Surdeanu John Bauer Jenny Finkel Steven J. Bethard Manning, Christopher D. and David McClosky. The stanford corenlp natural language processing toolkit in proceedings of the 52nd annual meeting of the association for computational linguistics: System demonstrations, 2014.
- [2] Paul Wilton Jarred McGinnis, Jeremy Tarling. Storyline ontology, 2013.
- [3] Antonio Lieto Rossana Damiano. Ontological representations of narratives: a case study on stories and actions, 2013.
- [4] Esra Yalcinalp. Turkey erdogan: Women rise up over withdrawal from istanbul convention, 2021.
- [5] Leanne Guenther. Little red riding hood.