UNIVERSITY OF TEXAS ARLINGTON

IDIR

FrameAnnotator
A frame-semantic annotation tool

Sarbajit Roy
M.S. Thesis, IDIR Lab

Supervising Professor
Dr. Chengkai Li

Committee Members
Dr. Ramez Elmasri
Dr. Deokgun Park

Master's Thesis   **April 10, 2019**

I would really like to dedicate my M.S. Thesis to my father.
This one is for you Baba!

Shyamal Roy
(1st May 1959 – 6th March 2016)

Outline

- **The importance of Fact-Checking**
- **Motivation**
- **Modeling Factual Claims with Frames**
- **FrameNet**
- **Challenges**
- **FrameAnnotator**
  - Definition
  - Understanding with an example
  - Live Demonstration
- **Questions?**
- **References**
- **Acknowledgement**

The importance of Fact-checking

There is a struggle with......
- Unprecedented amount of falsehoods,
- Hyperboles and half-truths

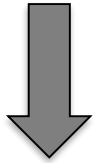Which do harm to wealth, democracy, health, and national security

https://ktwop.com/2017/11/03/will-recognition-of-fake-news-be-followed-by-fake-science/

In fighting against false information, the number of active fact-checking organizations has grown
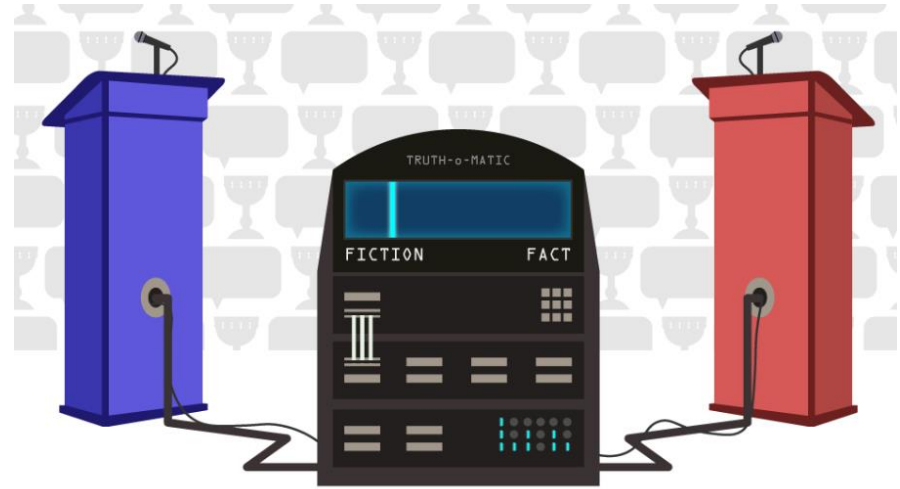
Motivation

- Human based Fact-checking is hard :
  - Intellectually demanding,
  - Laborious,
  - Time-consuming

An opportunity for automated fact-checking systems

https://nordicapis.com/automated-fact-checking-the-holy-grail-of-political-communication/

Modeling Factual Claims with Frames

Increasing demands
for fact-checking ⟶ growing interest in automating
various fact-checking steps

Representing claims in a structured and semantic way will help to capture various aspects :

- The domain and topic,
- The expression of facts,
- The entities involved and their relationships, quantities and many more

This claim modeling capability is useful for a variety of fact-checking steps

Modeling Factual Claims with Frames

The rank based on

Ranking entity

The rank

The U.S. only ranks 25th worldwide on defense spending as a percentage of GDP

The rank comparison

Predicate
Meaning : stand, position, order

Predicate Sense: **rank**
Ranking entity: **The U.S.**
The rank: **25th**
The rank comparison: **worldwide**
The rank based on: **on defense spending.....**

Modeling Factual Claims with Frames

Factual-claim specific frames

- Collected fact-checked claims from PolitiFact,
- Examined a subset of these claims one by one and grouped similar ones,
- Created new frames, if it doesn't already exist in FrameNet collection,
- 20 frames (13 newly created) and 900 labelled factual claims so far

Few examples of factual-claim specific frames

- Oppose and support consistency,
- Vote,
- Correlation,
- Occupy rank

FrameNet

- **FrameNet** is a project housed at the International Computer Science Institute in Berkeley, California
  - A rich knowledge base that contains information about words by providing their description and associated frames,
  - A conceptual structure describing an event, relation, or object and the participants in it

    - The FrameNet lexical database contains over 1,200 semantic frames,
    - 13,000 lexical units, and
    - 202,000 example sentences

  - FrameNet has been used in applications like question answering, paraphrasing, information extraction, machine translation, and many more.

FrameNet

What is a Frame?

"Stand, top, rank"

↑

All having a common-sense of background information

- The holistic background knowledge that unites these words

- A frame contains a textual frame definition, associated frame elements, lexical units, example sentences, and frame-to-frame relations

FrameNet

Frame Elements

- Provide additional information to the semantic structure of a sentence
- Types of Frame Elements:
  - core : Essential to the meaning of the frame
  - non-core : Generally descriptive (such as time, place, manner, etc.)

Lexical Units

- Lexical units (LU) are lemmas , with their POS , that evoke a specific frame

  Lexical units that suggest the Occupy_rank frame include the words "rank", "stand", and "top"

Frame-to-frame relations

- Relationships between different frames

Some relations are subframe, inheritance, Perspectivized_in and so on

Understanding an example with Occupy_rank frame

Under Gov. Mitt Romney , Massachusetts ranked 47th in job creation .

LU

Under Gov. Mitt Romney , Massachusetts `ranked` 47th in job creation .

LU

`Under Gov. Mitt Romney`, Massachusetts `ranked` 47th in job creation .

Comparison_set	Time	Item	Rank	Dimension

Modeling Factual Claims with Frames
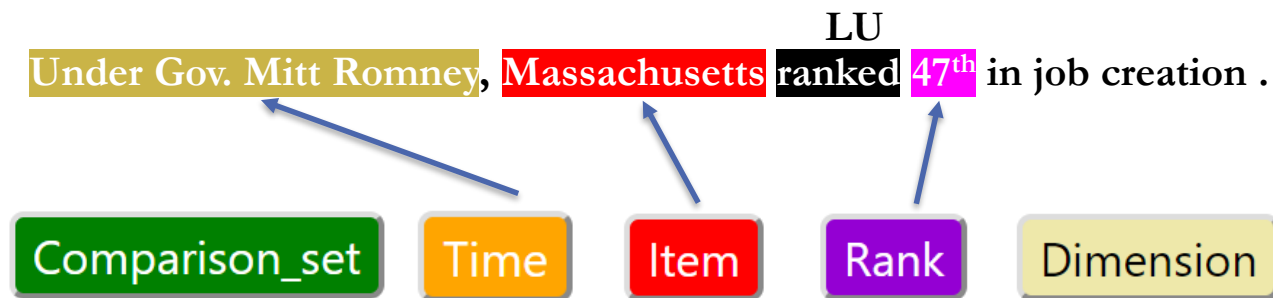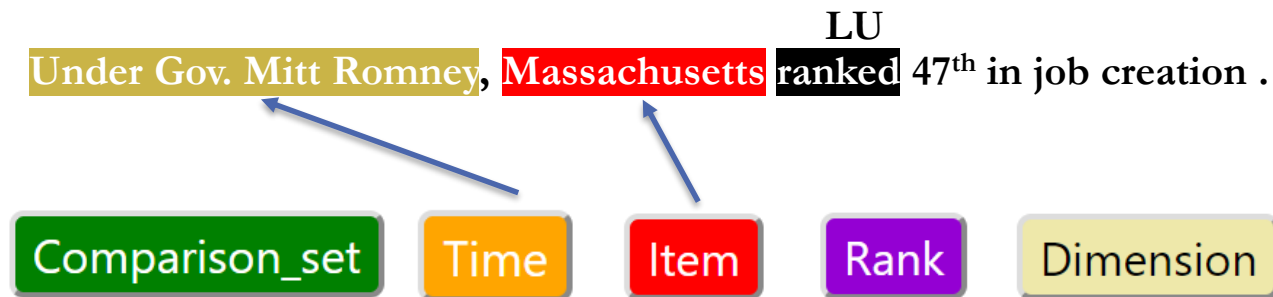
Modeling Factual Claims with Frames

LU

Under Gov. Mitt Romney, **Massachusetts** ranked **47th** in job creation .

Comparison_set	Time	Item	Rank	Dimension
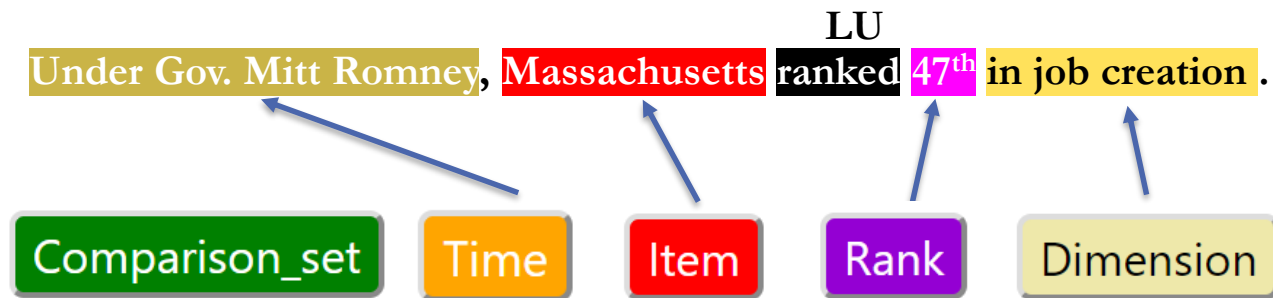
Under Gov. Mitt Romney , **Massachusetts** ranked **47th** in job creation .

Core :
Dimension – Along which the ranking is defined,
Item –In which occupies the rank
Rank - In which the item occupies

Non-Core :
Comparison_set - Comparison among the entities
Time – Time over which the item occupies the rank

Modeling Factual Claims with Frames

Understanding an example with Vote frame

Mitch MacConnell voted two times against the violence Against Women Act .

LU

Mitch MacConnell **voted** two times against the violence Against Women Act .

LU

Mitch MacConnell voted two times against the violence Against Women Act .

Agent   Frequency   Position   Side   Time   Issue

Challenges

- The current state-of-the-art frame semantic parsers suffer from lack of a large labeled dataset,
- Manual labeling of data is time consuming, and
- There are no open-source annotation tools available

To overcome all those challenges we developed the idea of
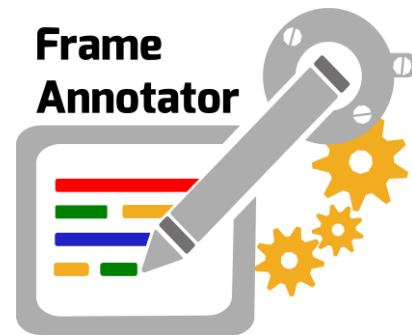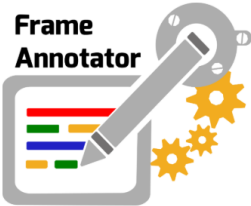
FrameAnnotator

A web-based public frame semantic annotation tool

FrameAnnotator

It allows users to create frame-semantic datasets and codify sentences

```xml
<sentence corpID="100" docID="20007" sentNo="0" paragNo="1" aPos="0" ID="4000071">
  <text>Bernie Sanders voted against the Brady Bill -- background checks and waiting periods .</text>
  <annotationSet cDate="06/04/2019 19:34:41 CDT Sat" status="UNANN" ID="9000071">
    <layer rank="1" name="PENN">
      <label end="5" start="0" name="NNP"/>
      <label end="13" start="7" name="NNP"/>
      <label end="19" start="15" name="VBD"/>
      <label end="27" start="21" name="IN"/>
      <label end="31" start="29" name="DT"/>
      <label end="37" start="33" name="NNP"/>
      <label end="42" start="39" name="NNP"/>
      <label end="44" start="44" name="HYPH"/>
      <label end="46" start="46" name="HYPH"/>
      <label end="57" start="48" name="NN"/>
      <label end="64" start="59" name="NNS"/>
      <label end="68" start="66" name="CC"/>
      <label end="76" start="70" name="VBG"/>
      <label end="84" start="78" name="NNS"/>
      <label end="86" start="86" name="."/>
    </layer>
    <layer rank="1" name="NER"/>
    <layer rank="1" name="WSL"/>
  </annotationSet>
  <annotationSet cDate="06/04/2019 19:34:41 CDT Sat" luID="90000" luName="vote.v" frameID="9000" frameName="Vote" status="MANUAL" ID="6000013">
    <layer rank="1" name="Target">
      <label cBy="IDIR" end="19" start="15" name="Target"/>
    </layer>
    <layer rank="1" name="FE">
      <label cBy="IDIR" feID="50000" bgColor="FF0000" fgColor="FFFFFF" end="13" start="0" name="Agent"/>
      <label cBy="IDIR" feID="50004" bgColor="A52A2A" fgColor="FFFFFF" end="27" start="21" name="Position"/>
      <label cBy="IDIR" feID="50002" bgColor="008000" fgColor="FFFFFF" end="83" start="29" name="Issue"/>
    </layer>
    <layer rank="1" name="GT"/>
    <layer rank="1" name="PT"/>
    <layer rank="1" name="Other"/>
    <layer rank="1" name="Sent"/>
    <layer rank="1" name="Verb"/>
  </annotationSet>
</sentence>
```

FrameAnnotator

Understanding with an example - Frame Name : Vote

Bernie Sanders voted against the Brady Bill -- background checks and waiting periods .

Bernie Sanders voted against the Brady Bill -- background checks and waiting periods .

Lexical Units

Bernie Sanders voted against the Brady Bill -- background checks and waiting periods .

Agent   Action   Issue   Side   Position   Frequency   Time

Bernie Sanders voted against the Brady Bill -- background checks and waiting periods.

FrameAnnotator

Understanding with an example - Frame Name : **Vote**

Demonstration

https://idir.uta.edu/frameannotator/

Questions ?

References

- Hassan, Naeemul, et al. "Toward automated fact-checking: Detecting check-worthy factual   claims by ClaimBuster." Proceedings of the 23rd ACM SIGKDD International Conference   on Knowledge Discovery and Data Mining. ACM, 2017.
- Baker, Collin F., Charles J. Fillmore, and John B. Lowe. "The berkeley framenet project." Proceedings of the   17th international conference on Computational linguistics-Volume 1. Association for Computational   Linguistics, 1998.
- Wu, You, et al. "Toward computational fact-checking." Proceedings of the VLDB Endowment 7.7 (2014):   589-600.
- Fillmore, Charles J., Christopher R. Johnson, and Miriam RL Petruck. "Background to framenet." International journal of lexicography 16.3 (2003): 235-250.
- en.wikipedia.org/wiki/FrameNet
- framenet.icsi.berkeley.edu/fndrupal/
- aclweb.org/anthology/N16-1019
- www.techopedia.com/why-is-data-annotation-important-in-some-machine-learning-projects/7/33331
- appen.com/why-human-annotated-data-key/
- web.stanford.edu/~jurafsky/slp3/slides/22_SRL.pdf
- web.stanford.edu/~jurafsky/slp3/18.pdf

Acknowledgement

Thank you!