„Al. I. Cuza" University, Iasi, Romania

Faculty of Computer Science

UAIC Participation at RTE5

Adrian Iftene
Mihai Alex Moruz
{adiftene,mmoruz}@info.uaic.ro

Overview

- 3-way RTE5 System
  - Newly added components
  - Positive and Negative rules
  - Results
- Pilot task
  - Application of QA techniques
  - Results
- Conclusions
- Further work

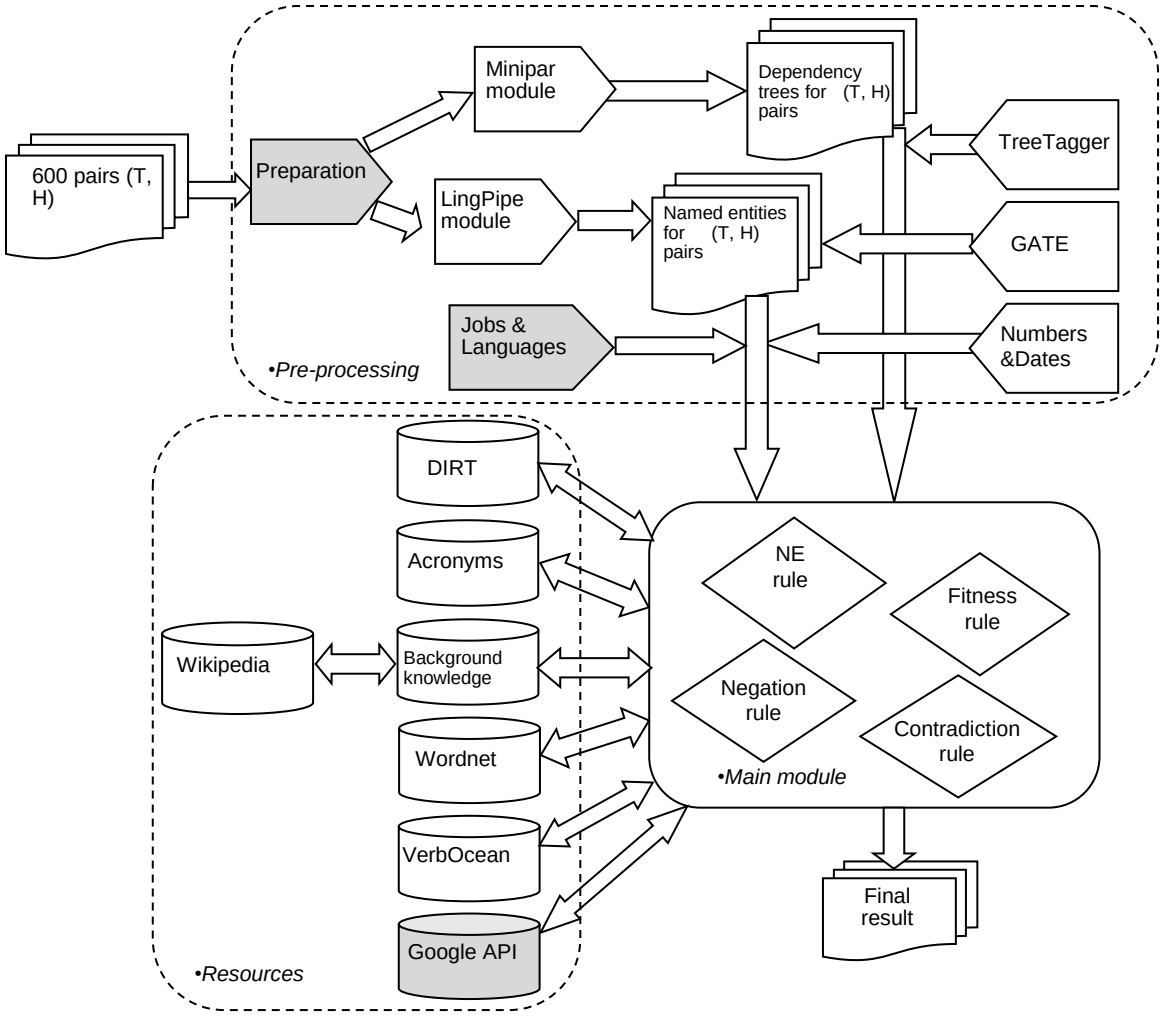
RTE Competition

The **two-way RTE task** (2005-2009) is to decide whether:

- **T entails H - ENTAILMENT**
- **T does not entail H - NO ENTAILMENT**

- The **three-way RTE task** (2007-2009) is to decide whether:
  - **T entails H - ENTAILMENT**
  - **T contradicts H - CONTRADICTION**
  - The **truth of H cannot be determined on the basis of T – UNKNOWN**

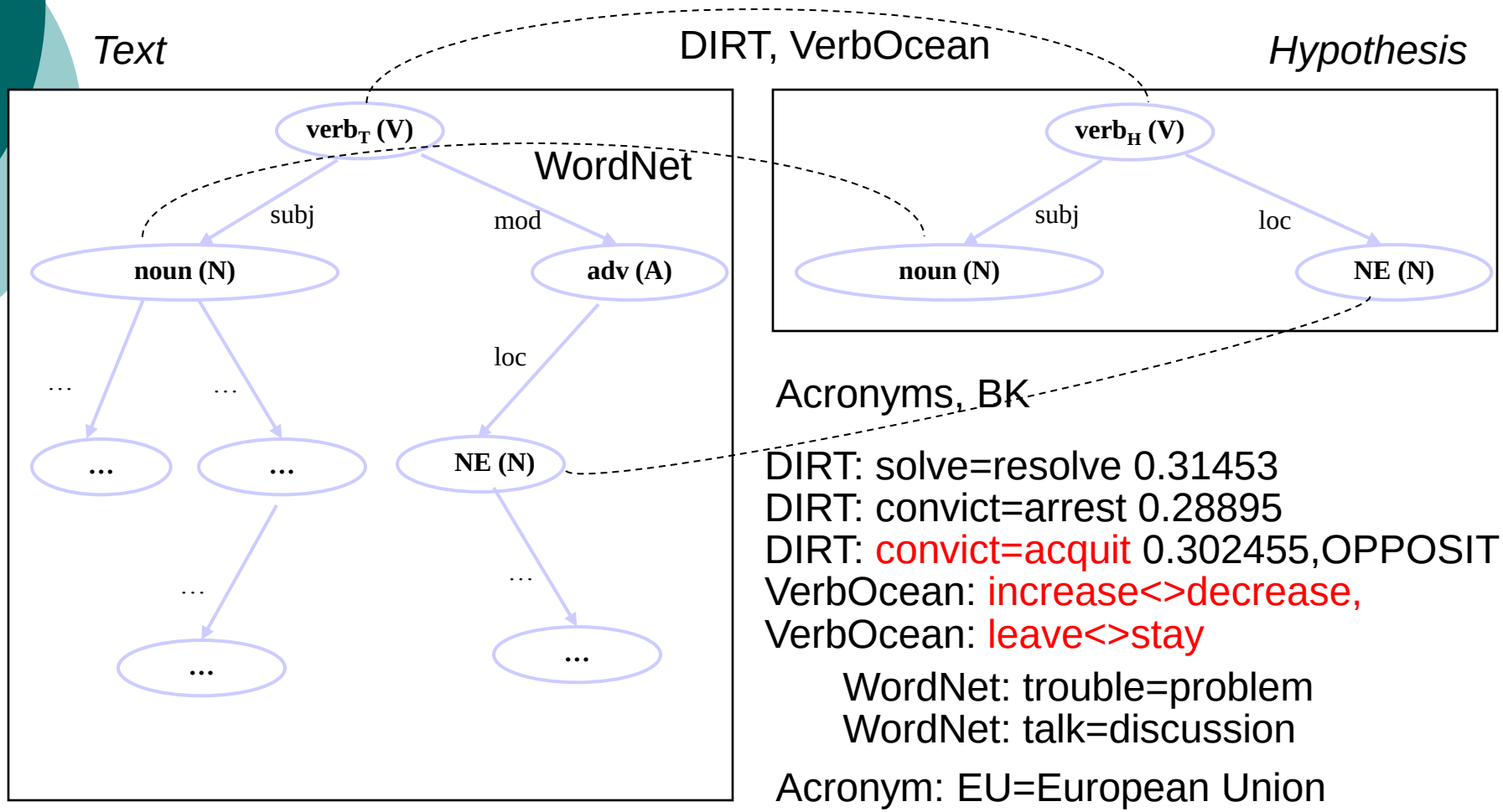
System Presentation



Preparation and New Modules

- We replace “hasn’t” with “has not”, “isn’t” with “is not”, “couldn’t” with “could not” and pad with spaces some punctuation
- In the case of Named Entities of type JOB and LANGUAGE, we additionally used GATE, which contains finer-grained classes of entities
- In order to cope with misspelled words (particularly Named Entities) we used the Google API

Example


Text
DIRT, VerbOcean
Hypothesis
verb$_T$ (V)
WordNet
subj
mod
noun (N)
adv (A)
loc
NE (N)
verb$_H$ (V)
subj
loc
noun (N)
NE (N)
Acronyms, BK
DIRT: solve=resolve 0.31453
DIRT: convict=arrest 0.28895
DIRT: convict=acquit 0.302455,OPPOSITE
VerbOcean: increase<>decrease,
VerbOcean: leave<>stay
WordNet: trouble=problem
WordNet: talk=discussion
Acronym: EU=European Union
BK: Buenos Aires [in] Argentine
BK: 16 [is] sixteen

Rules

- For every type of possible answer we will present the rules that lead to it
- Possible cases are:
  - Entailment cases
  - No entailment cases
    - Contradiction cases
    - Unknown cases

Entailment Cases

- Every type of mapping: *direct* (lexical) or *indirect* (using knowledge bases)
- Verb similarity is computed using DIRT
  - *passed away ≈ has died*
- For named entity we use an acronym database and background knowledge
  - *United States ≈ US*
  - *Basel in Switzerland ≈ European City*

Entailment Cases (cont.)

- For nouns and adjectives we use WordNet and some of the relations from eXtended WordNet to look up synonyms
- For every transformation with DIRT or WordNet, we will consider local fitness to be the similarity value indicated by these resources
- Stop words from the hypothesis artificially increase the value of global fitness and are ignored

Entailment Cases for Numbers

- When numbers from T or H are separated by "and" or ",", add them
  - T: *10 people were killed and more than 30 died* ≈ H: *killing more than 40 people*
- *Positive rules for Numbers (context rules)*: quantification words: at least, more than, less than, over, under, etc.
  - T: *at least 80 percent* ≈ H: *more than 70 percent*

Contradiction Cases

- *F*or every verb subtree, we check for words such as “not”, “never”, “may”, “might”, “cannot”, etc. and modify the negation value of the verb
  - T: *New Line Cinema has announced that movie director Peter Jackson will **never** be allowed to **work** …*
  - H: *New Line **wants to work** with Peter Jackson.*
- Verbs in the long infinitive receive special treatment

Contradiction Cases (cont.)

- When before the long infinitive we have “refuse”, “deny”, “ignore”, “plan”, “intend”, “proposal”, “able”, etc.
- *Antonymy relation*: use [*opposite-of*] relation from VerbOcean and antonymy relation from WordNet
- We consider a combination of synonyms from WordNet and antonyms from WordNet or opposites from VerbOcean

Unknown Cases

- *W*hen verbs are modified by words such as “may”, “can”, “should”, “could”, “must”, “might”, “infrequent”, “rather”, “probably”, etc.
  - T: *…could eventually be taken over …* and H: “*… is taken over…*”
- Related to verbs in the infinitive, we will consider as Unknown those cases which are not included in the contradiction cases
- If we cannot map a NE from H, either directly or by using the acronym database and background knowledge, the result for the current pair is *Unknown*

Unknown Cases (cont.)

- We make an exception from the named entity rule when the type of named entity is *first name*
  - T: *The man accused of killing **Ms. Zapata**, …*
  - H: ***Angie Zapata** has been killed with a fire extinguisher.*
- In this case we only insert a penalty in the global fitness.

Unknown Cases (cont…)

- If any of the numbers in the text or the hypothesis has an attached unit of measure, it is always kept
  - T: *At least 14 people have been killed in a suicide bomb attack in southern Sri Lanka, police say. The telecoms minister was among about 35* ***people*** *injured in the blast at the town of Akuressa…*
  - *H:35* ***government officials*** *were injured by a suicide bomber in Akuressa*

Results in RTE5

Answer Type	In Gold	Correct offered by our system	Total offered by our system	Precision	Recall	F-measure
Entailment	300	260	379	68.60%	**86.67%**	**76.58%**
Contradiction	90	22	44	50.00%	24.44%	32.84%
Unknown	210	128	177	**72.32%**	60.95%	66.15%
Total	**600**	**410**	**600**			**68.33%**

Answer Type	In Gold	Correct offered by our system	Total offered by our system	Precision	Recall	F-measure
Yes	300	260	379	68.60%	86.67%	76.58%
No	300	181	221	**81.90%**	60.33%	69.48%
Total	**600**	**441**	**600**			**73.50%**

Ablation Tests

System Description	2-way (73.5 %)		3-way (68.33 %)	
	P (%)	C (%)	P (%)	C (%)
Without DIRT	73.33	0.17	68.00	0.33
Without WordNet	72.50	1.00	67.00	1.33
Without Acronyms	73.33	0.17	68.17	0.17
Without BK	72.33	1.17	66.83	1.50
Without NE rule	67.33	**6.17**	63.33	**5.00**
Without the Negation rule	73.50	0.00	66.83	1.50
Without the Contradiction rule	71.50	**2.00**	69.67	- 1.34
Without additional processing steps	69.33	**4.17**	64.33	**4.00**
Total		**14.85**		**12.49**

Ablation Tests Overview

System description	RTE-3 (69.13 %)			RTE-4 (72.1 %)			RTE-5 (73.5 %)		
	P	C	WR	P	C	WR	P	C	WR
Without DIRT	68.76	0.37	0.54	71.40	**0.7**	**0.97**	73.33	0.17	0.23
Without WordNet	68.00	1.13	1.63	69.10	**3.0**	**4.16**	72.5	1.00	1.36
Without Acronyms	68.38	**0.75**	**1.08**	71.80	0.3	0.42	73.33	0.17	0.23
Without BK	67.75	1.38	2.00	70.40	**1.7**	**2.36**	72.33	1.17	1.59
Without the NE rule	57.58	**11.55**	**16.71**	66.90	5.2	7.21	67.33	**6.17**	**8.39**
Without the Negation rule	67.63	1.50	2.17	68.70	**3.4**	**4.72**	73.5	0.00	0.00
Without the Contradiction rule	-	-	-	68.10	**4.0**	**5.55**	71.5	**2.00**	**2.72**
No additional processing steps	-	-	-	-	-	-	69.33	**4.17**	**5.67**
Total		16.68	24.13		18.3	25.39		14.85	20.20

Pilot Task

- Extraction of text from a series of newspaper articles that yielded positive entailment for a given set of hypotheses
  - the texts are not modified in any way as compared to the original source
  - a large numbers of candidate pairs, as for every one of the nine topics there are about ten hypotheses

Pilot Task (cont.)

- In order to reduce the search space, we have made use of a technique used for our question answering systems
- Using Lucene, we have indexed the articles from each topic at the sentence level
- We have built queries for all the hypotheses
- The snippets with the highest chance of yielding positive entailment are clustered around the top scoring snippets

Pilot Task (cont.)

- In order to determine the entailment value of the candidate pairs (approx. 1700 in all), we have applied a lightweight version of our entailment system

Result	Precision	Recall	F-measure
Micro-average	51.12%	22.88%	**31.61%**
Macro-average Topic	53.03%	24.08%	33.12%
Macro-average Hypothesis	46.55%	26.42%	33.71%

Conclusions

- Main idea of our TE system is to map every node from hypothesis to a node from text, either *lexically* or *semantically*
- The rules regarding Named Entity processing were more elaborate
- Preprocessing for our RTE-5 is more elaborate
- RTE-5 also introduced a pilot task, to which we applied QA techniques to reduce the solution space

Further Work

- Using semantic roles
  - "LOSAIL, Qatar (AFP) - Torrential rain caused the season-opening Qatar MotoGP to be cancelled on Sunday, …"
  - "Valentino Rossi won the season-opening Qatar MotoGP."
- "the season-opening Qatar MotoGP did not finish" ≈ "the season-opening Qatar MotoGP was cancelled"
- In order to win a race, the race must finish, the winner must finish that race and the winner needs to be first when he finishes

Further Work (cont.)

- Enhancing entities with ontological knowledge
  - "He has long been linked to some of the world's most notorious conflicts, allegedly supplying arms to former Liberian dictator Charles Taylor and Libyan leader Colonel Gaddafi."
  - "Gaddafi is the Liberian dictator."
- In ontological knowledge, we find that a person can only have one occupation
- We attempt to unify the property sets in the text with those in the hypothesis

Acknowledgments

- Students:
  - Alexandra Balahur-Dobrescu, Daniel Matei
- NLP group of Iasi:
  - Supervisor: Prof. Dan Cristea
  - Maria Husarciuc, Ionut Pistol, Marius Raschip, Diana Trandabat
- Support from SIR-RESDEC project

THANK YOU!