EPE 2017: The Trento–Gothenburg Opinion Extraction System

UNIVERSITY OF
GOTHENBURG

CHALMERS

Richard Johansson
richard.johansson@gu.se

EPE shared task
September 20, 2017

the third EPE downstream task

- the third task is **extraction of opinion expressions**
- we use the MPQA annotation model [Wiebe et al., 2005]

  "The report is full of absurdities", Xirao-Nima said.

- the downstream application is the Trento–Gothenburg system [Johansson and Moschitti, 2013]

types of expressions annotated in MPQA

- direct-subjective expressions (DSEs):

  Paolo likes Pisa
- expressive-subjective elements (ESEs):

  Pisa is a wonderful city
- objective speech events (OSEs):

  Paolo says that Pisa is widely appreciated

polarity annotation

- direct-subjective expressions (DSEs):

  Paolo likes Pisa [positive]
- expressive-subjective elements (DSEs):

  Pisa is a wonderful city [positive]
- objective speech events (OSEs):

  Paolo says that Pisa is widely appreciated

opinion holders

- explicitly mentioned:

  Paolo likes Pisa
- writer:

  Pisa is a wonderful city
- implicit:

  Pisa is widely appreciated

definition of the task

- **extract expressions** and label them (DSE, ESE, OSE)
- determine the **polarity** of DSEs and ESEs
- find the **holders** of all expressions, including writer and implicit

scoring the participating systems

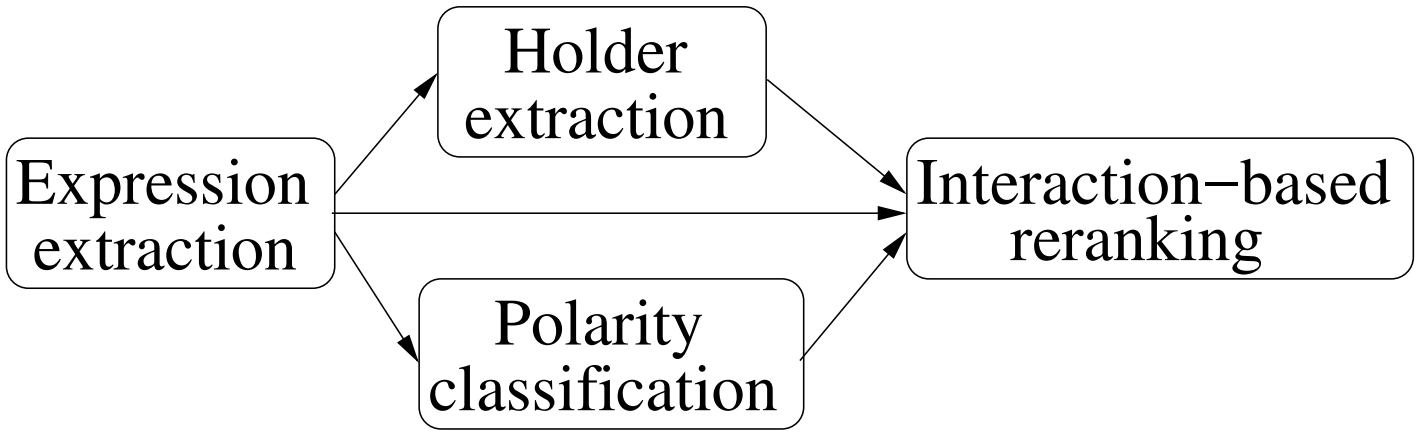
- precision, recall, and F-score for all three subtasks
- we use a lenient scoring approach:

  gold standard: **The report is full of absurdities**

  system output: **The report is full of absurdities**

  gives P = 1.0, R = 0.58

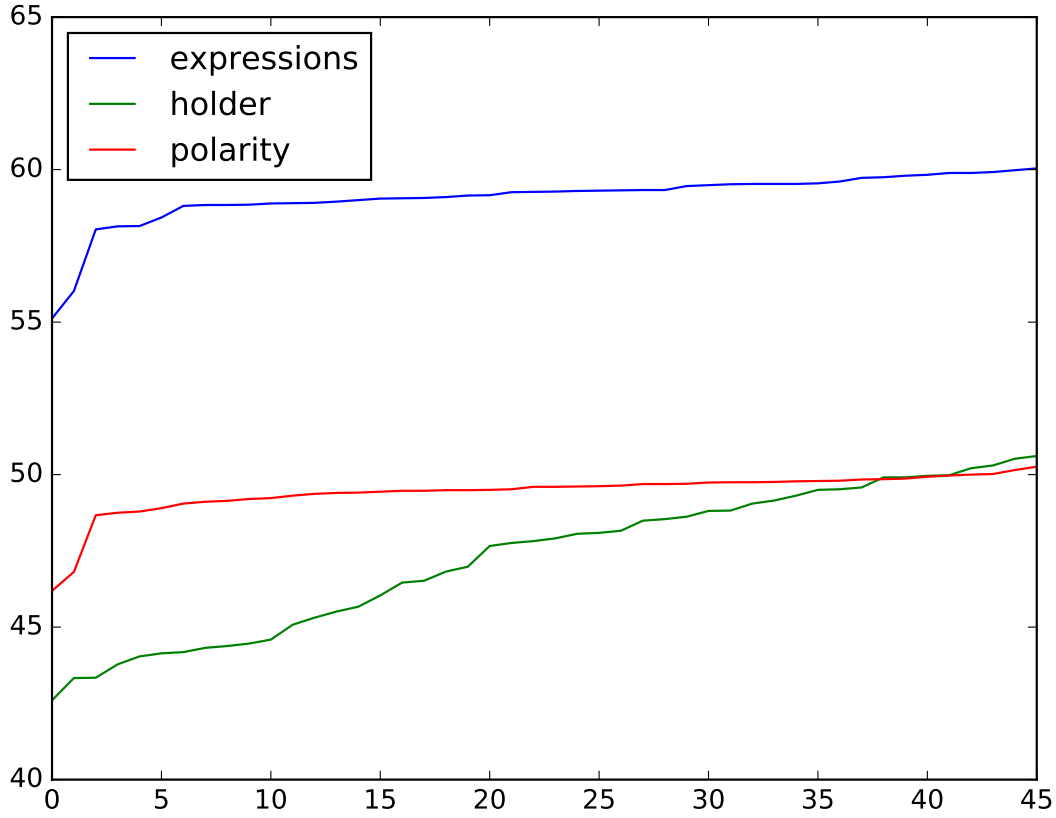
overview of the system by Johansson and Moschitti (2013)



how does the linguistic analysis affect the modules?

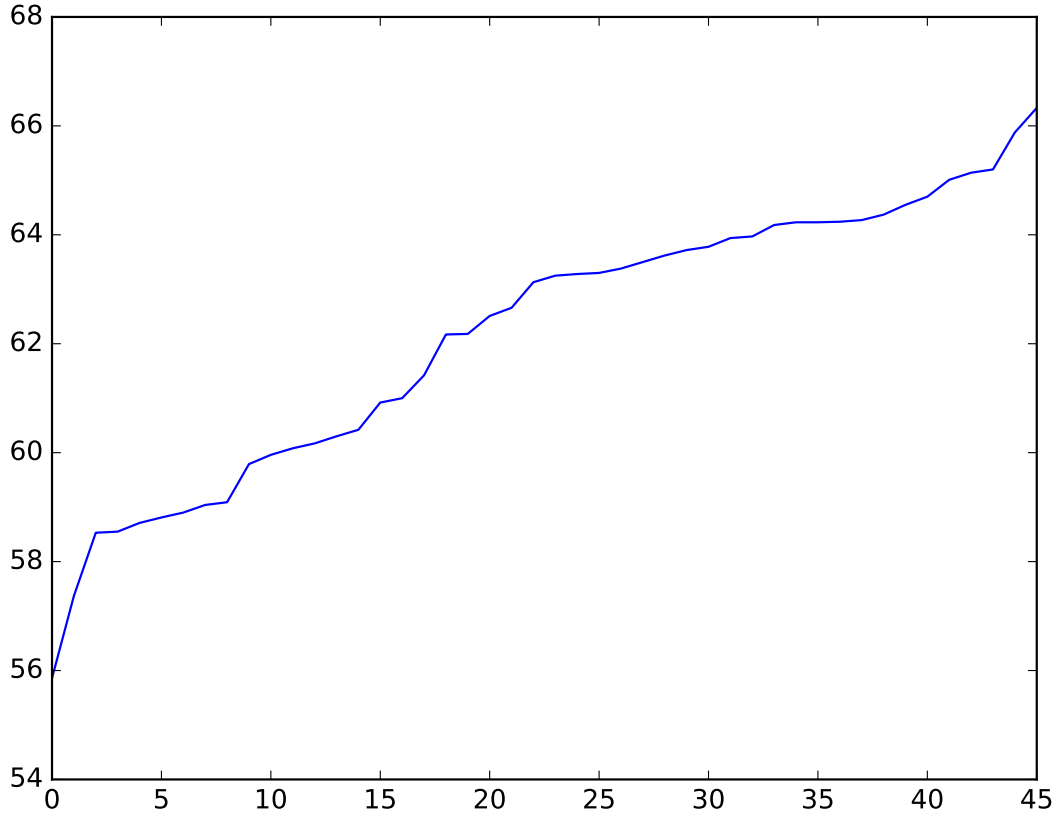
- expression extraction: **tags**, **lemmas**
- polarity classification: **tags**, **lemmas**
- holder extraction: **tags**, **lemmas**, **dependencies**
- reranking: **tags**, **lemmas**, **dependencies**

results: high-level trends



- holder extraction results show much more variation than the other two subtasks

detailed results: “in vitro” holder extraction



- we evaluated the holder extraction module using gold-standard opinion expressions
- this is the scenario we used in the final overall ranking

honorable mentions

Szeged	1	66.3	post-processed CoNLL-08
Stanford–Paris	6	65.2	UD v1 enhanced
Paris–Stanford	3	64.3	UD v1 enhanced

conclusions: some tentative observations

- how much does the choice of dependency style matter?
  - hard to say: most systems are UD-based, but much variation inside this group
  - not many datapoints for other dependency styles
- how well do parsers producing "semantic" representations perform?
  - not very well! mean F-score 58.8, vs 62.9 for the "syntactic" representations
  - but the features in the downstream system were never designed for this type of representation

references I

- Johansson, R. and Moschitti, A. (2013). Relational features in fine-grained opinion analysis. *Computational Linguistics*, 39(3):473–509.
- Wiebe, J., Wilson, T., and Cardie, C. (2005). Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation*, 39(2-3):165–210.