Speaker Attribution in Cabinet Protocols

Josef Ruppenhofer, Caroline Sporleder, & Fabian Shirokov

May 19, 2010

Part I

Introduction

Cultural heritage data

- Many efforts to make cultural heritage data more accessible by digitizing them and making them publically searchable

Cultural heritage data

- Many efforts to make cultural heritage data more accessible by digitizing them and making them publically searchable
- Support for more sophisticated search requires enriching the data with additional information

Cultural heritage data

- Many efforts to make cultural heritage data more accessible by digitizing them and making them publically searchable
- Support for more sophisticated search requires enriching the data with additional information
- One kind of enrichment is attributing speech events in cabinet protocols to their speakers.

Cultural heritage data

- Many efforts to make cultural heritage data more accessible by digitizing them and making them publically searchable
- Support for more sophisticated search requires enriching the data with additional information
- One kind of enrichment is attributing speech events in cabinet protocols to their speakers.
- Attribution information allows historians to search systematically for statements made by a particular politician.

Cultural heritage data

- Many efforts to make cultural heritage data more accessible by digitizing them and making them publically searchable
- Support for more sophisticated search requires enriching the data with additional information
- One kind of enrichment is attributing speech events in cabinet protocols to their speakers.
- Attribution information allows historians to search systematically for statements made by a particular politician.
  - Statements frequently reflect opinions of their speakers
  - They also provide information about which facts were known by a particular person at a given time.

German Cabinet Protocols: Example

(1) Der Bundeskanzler **erklärt**, daß er dem Kabinett zur Saarfrage alles **gesagt** habe, was er wisse.
'The chancellor states that he has told the cabinet everyting about the Saar question that he knows.'

(2) Seitdem **SEI** nichts geschehen und es werde auch nichts geschehen.
'Since then nothing had happened and nothing would happen.'

- minutes, not transcripts
- almost all sentences in the minutes report utterances by the meeting participants
- only a few sentences contain background or meta information

Part II

Related work

Related work on speaker attribution and point of view

- Bergler's (1992) thesis studies reported speech in newspaper articles
- Krestel et al (2008) work on finding sources of reported speech but only do this for explicitly marked reported speech
- Wiebe (1990) provides an implemented system for tracking psychological point of view in narratives

Related work on sentiment analysis

- Finding sources of opinions is one sub-task in automatic sentiment analysis

Related work on sentiment analysis

- Finding sources of opinions is one sub-task in automatic sentiment analysis
- In some contexts (e.g. reviews) there is only one relevant source

Related work on sentiment analysis

- Finding sources of opinions is one sub-task in automatic sentiment analysis
- In some contexts (e.g. reviews) there is only one relevant source
- Sources are found only for opinionated sentences

Related work on sentiment analysis

- Finding sources of opinions is one sub-task in automatic sentiment analysis
- In some contexts (e.g. reviews) there is only one relevant source
- Sources are found only for opinionated sentences
- Typically, sources are sought within the same sentence (Bethard 2004, Choi et al. 2005, Kim and Hovy 2006)

Related work on sentiment analysis

- Finding sources of opinions is one sub-task in automatic sentiment analysis
- In some contexts (e.g. reviews) there is only one relevant source
- Sources are found only for opinionated sentences
- Typically, sources are sought within the same sentence (Bethard 2004, Choi et al. 2005, Kim and Hovy 2006)
- But Seki et al. 2009 do use information from prior sentences

Part III

Data and Annotation

Data

- minutes of the weekly meetings of the German cabinet between 1949 and 1960[1]
- obtained from German federal archive (Bundesarchiv)
- total collection of 58,310 sentences
- randomly extracted
  - a development set (566 (687) sentences)
  - a test set (323 (400) sentences)

[1] First female cabinet member only at end of 1961.

Annotation

Example

(3) <sentence id="149" hasSpeaker="281,5"> <person id="281"> Der Bundesinnenminister </person>**schließt** sich der **Auffassung** <person id="5"> des Bundeskanzlers </person> **an**, wird den Entwurf noch zurückhalten und verschiebt die von ihm vorgesehenen Besprechungen. </sentence>
'The Secretary of the Interior concurs with the opinion of the Chancellor, is going to hold back the proposal for a while, and postpones the talks he had planned.'

Annotation II

- Record for every sentence the set of speakers for all actual present or past speech events and private states (Wiebe et al. 2005) expressed in the sentence

Annotation II

- Record for every sentence the set of speakers for all actual present or past speech events and private states (Wiebe et al. 2005) expressed in the sentence
- Future or hypothetical speech events are left unannotated (cf. *insubstantial* category of Wiebe et al. 2005)

(5) Es besteht **Übereinstimmung**, daß dieses der Öffentlichkeit nicht **bekanntzugeben** ist.
'There is consensus that it will not be made known to the public.'

Annotation III

- Speakers are resolved to IDs in a biographical database (total of 1932 possible speakers)
- Assign value 'Unknown' when (1) speaker not in database ; (2) speaker cannot be identified; or (3) sentence is background or meta info by minute taker
- Inter-annotator F-score of 0.87 and 0.88 on strict and loose measures, respectively

Annotation IV

- Sentences may have more than one speaker associated
- The embedding of speakers is not captured

	Total	Avg. per S
private states/speech	493	1.6
insubstantial events	84	0.3
speakers	405	1.4
unknown speakers	58	0.2

Table: Statistics on test data

Linguistic background

We exploit the following tendencies in our data:

- New speakers appear as the subject of a reporting verb
- Contents of reported speech typically in subjunctive mood
- Reported speech is marked by subjunctive mood even when there is no reporting clause
- Whenever a potential speaker appears as subject of a sentence, he is typically an actual speaker (at some depth of embedding)

Linguistic background: example

- Staatssekretär Hartmann **bemerkt** ergänzend, daß über die in dieser Vorlage angeschnittenen Fragen soeben eine Chefbesprechung stattgefunden habe, die zu keiner **Einigung** geführt habe.
- Überdies **wolle** der Verkehrsminister das Ermäßigungsprogramm umarbeiten und auf Kinder bis zu 25 Jahren ausdehnen.
- Der Bundesminister für Verkehr **erklärt** hierzu, daß er diese **Absicht** nicht mehr habe.
- Der Bundesminister für Familienfragen **betont** demgegenüber, daß man sich in der genannten Chefbesprechung **einig geworden** sei.
- Man solle vorläufig an der Vorlage festhalten und sie möglicherweise später verbessern.

Linguistic background: example

Observes	**agreement**	Undersecretary of state Hartmann **observes** in addition that, concerning the issues broached in this proposal, a principals' meeting had just taken place, which had not produced an **agreement**.
Observes	**wanted**	On top of that, the transportation secretary **wanted** to revise the discount program and extend it to children up to 25 years.
Explains	**intention**	The transportation secretary **explains** that he no longer has this **intention**.
Stresses	**agreement**	The Secretary for Family Affairs **stresses**, by contrast, that there had been an **agreement** in the aforementioned principals' meeting.
Stresses	**should**	One **should** hold fast to the proposal and improve it later, if possible.

Part IV

Experiments

Measures

- Precision, Recall, F-score
  - **Loose** precision counts a sentence as correctly labeled if at least one of the recognized speakers is correct.
  - **Strict** precision requires all recognized speakers to be correct.
  - **Loose** recall: a sentence counts as correctly labeled if at least one of the speakers in it was found by our system.
  - **Strict** recall: a sentence counts as correctly labeled if all speakers in it have been found.
- Development set
- Test set

Baseline algorithm

- if there is evidence for speaker continuity (subjunctive verb forms, pronoun *Er* 'he')
  - if there is a prior sentence with known speaker
    - ★ assign that speaker
  - else
    - ★ set speaker to unknown
- else
  - if current sentence mentions potential speakers
    - ★ choose first mentioned potential speaker as speaker
  - else
    - ★ assign unknown

Baseline performance

	Development		Test	
	Loose	Strict	Loose	Strict
Prec.	77%	77%	83%	83%
Recall	44%	36%	35%	35%
F-score	56%	49%	49%	49%

Table: Performance of baseline algorithm

- too many unknown speakers
- only one speaker per sentence
- first mentioned potential speaker need not be a speaker
- too few known subjunctive forms; too many instances that are not in main clause

Subject-based algorithm

Our first algorithm following on the baseline is **subject-based** in that it addresses the problem that the first mention of a person in a sentence is not necessarily the subject by using the output of the Stanford parser (Klein & Manning 2003). The new algorithm works as follows:

1. If the current sentence s_i has a main clause subject go to step 2. Otherwise assign the person mentioned first in s_i as its speaker.
2. If the subject(s) occurring in s_i refer to persons from the biographical database, assign them as speakers. Otherwise, go to 3.
3. If s_i contains references to potential speakers, assign the first one as the subject. Otherwise, assign as speaker of s_i the speaker of s_{i-1}

Performance of subject-based algorithm

	Development		Test	
	Loose	Strict	Loose	Strict
	Baseline			
Prec.	77%	77%	83%	83%
Recall	44%	36%	35%	35%
F-score	56%	49%	49%	49%
	Subject-based			
Prec.	81%	79%	80%	79%
Recall	65%	56%	70%	70%
F-score	72%	65%	75%	74%

Table: Performance of subject-based algorithm

Syntax-based algorithm

1. If current sentence s_i has a subjunctive mood main verb, assign speaker of s_{i-1}. Go on to 2
2. If s_i has a subject referring to potential speakers, add them to the set of speakers. If not, add the first-mentioned person in s_i to the set of speakers. Go on to 3.
3. If no speaker has been assigned so far, assign the speakers of s_{i-1}.
4. If the head verb is passive, assign the virtual speaker representing the cabinet as a whole.

Performance of syntax-based algorithm

	Development		Test	
	Loose	Strict	Loose	Strict
	Baseline			
Prec.	77%	77%	83%	83%
Recall	44%	36%	35%	35%
F-score	56%	49%	49%	49%
	Subject-based			
Prec.	81%	79%	80%	79%
Recall	65%	56%	70%	70%
F-score	72%	65%	75%	74%
	Syntax-based			
Prec	86%	69%	86%	72%
Recall	87%	79%	88%	88%
F-score	87%	74%	87%	79%

Table: Performance of syntax-based algorithm

Conclusion

- We presented a rule-based system for speaker attribution in cabinet protocols
- We improved over our baseline by exploiting linguistic cues
- Not yet taken into consideration
  - embedding of speech events
  - speech events denoted by nouns
- Extensions
  - use of semantic role labeler
  - use our rule-based system to label initial training data for a second stage supervised classifier, which can then exploit a larger set of linguistic cues to deal with the more difficult cases as well.
  - use topic identification: not all speakers are equally likely to speak on any given topic

Part V

Extra material

Speaker continuity cues in English

- Sir Eric Geddes **said** that it was proposed so to throw the net as to get more men than we require.
- The A.S.C on the lines of communication contained a large proportion of the older men.
- In the combatant services there were many older men who were *pivotal* N.C.O.'s and who *must* be retained.
- He *therefore* did not **see** why it should be necessary to discriminate against the A.S.C.

Speaker continuity in North American news

- sample of 10 Associated Press newswire stories from 2003 totalling 4122 words; 122 expressions of speech events and private states.
- the only type of speaker continuity that occurs is of the type exemplified by (6), where direct speech is continued

  (6) "The domestic leisure market is growing rapidly and now represents over 60 percent of all passengers," Qantas Chief Executive Officer Geoff Dixon **said** Monday. "Jetstar will concentrate on growing this market with value fares while opening up new destinations."

- no cases where indirect speech is continued past a reported speech-sentence marked by a reporting verb.
- This confirms Bergler's (1995) finding that so-called free indirect speech is virtually absent from North American newspaper writing.

Rule optimization

- Optimize: Inventory and order of rules
  - Given a set of ordered rules R
    1. calculate F-score of R
    2. for every rule r in R, try to substitute it at every position in the order of R and calculate the F-score
    3. if any substitution produces a better F-score than the current best result, adopt the resulting ordered rule set as new best rule set B
    4. perform manual error analysis and propose new rules, create new rule inventory R_{man}
    5. for every rule r in R_{man}, try to substitute it at every position in the order of R and calculate the F-score
    6. if any substitution produces a better result than the current best result, adopt new rule set as new best rule set B
    7. go back to 1 with current best B as new R