Discourse Structure in Twitter Conversations

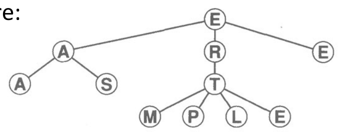
Manfred Stede, Univ Potsdam
ESSLLI 2016

Overview

- Twitter conversations ??
- From speech acts via dialog acts to tweet acts
- Coherence relations and „rhetorical structure“
- Deriving rhetorical structure automatically
- Rhetorical structure in Twitter conversations

Twitter Conversations

- **reply-to**-function creates conversations on Twitter
- ~20-25% of tweets are replies
- ~40% of tweets = part of conversations
- tree structure:



Analysis: What do people do ?

WTF? I have green energy and am supposed to finance nuclear and coal? What nonsense. WHAT NONSENSE!

wtf? Why and in which way?

Oh, and I have coal and nuclear energy and have to finance green power thanks to [new law] just so you can have it cheaper? WTF!

…

Agree. But it's also a fact that …

right, I'm destroying the climate. Put a windmill in front of your house and we'll talk. …

Speech acts

- John Austin: *How to do things with words* (1962)

- Performative utterances do not merely *describe* the world but *change* it
(and determining their truth value is pointless)

- I name this ship the „Queen Elizabeth.“
- I give and bequeath my watch to my brother.
- I bet you sixpence it will rain tomorrow.
- I apologize.

Speech acts

- **John Austin**: *How to do things with words* (1962)

- Levels of analyzing an utterance:
  - **locution**: performing the action of uttering
  - **illocution**: the action/intention of the speaker
  - **perlocution**: the effect of the utterance on the addressee

Speech Acts

- **John Searle**: *Speech acts* (1969); *A taxonomy of illocutionary acts* (1975)

- Assertives = speech acts that commit a speaker to believing the expressed proposition
- Directives = speech acts that are to cause the hearer to take a particular action, e.g. requests, commands and advice
- Commissives = speech acts that commit a speaker to doing some future action, e.g. promises and oaths
- Expressives = speech acts that express the speaker's attitudes and emotions towards the proposition, e.g. congratulations, excuses and thanks
- Declarations = speech acts that change the social sphere, e.g. baptisms or pronouncing someone husband and wife

Speech acts, empirically

Huge literature on speech act theory: their definition; linguistic realization; indirect speech acts, ...

In practice,
the vast majority of illocutions encountered in text are **assertives**

Example: User-generated hotel reviews

I stayed at this Hilton for the third time.
As usual, the staff was extremely attentive and friendly.
The concierge is called Pierre; he always wears a bowtie.
He's particularly sweet, even early in the morning.
I guess he's getting very good coffee at his place.
...

An inventory of illocutions for subjective text

- Report: The waiter brought us a new cocktail.
- Report_author: I payed for it right away.
- Ident: I was afraid to see him again.
- Evaluation: The cocktail turned out to be lousy.
- Estimate: It probably was made in a hurry.
- Commitment: I'll never have it again!
- Directive: And you should avoid it, too.

M. Stede, A. Peldszus: The role of illocutionary status in the usage conditions of causal connectives and in coherence relations. *Journal of Pragmatics* 44(2), 2012

Dialog act

- ...captures the functional relevance of an utterance in context
- Example from *Verbmobil* appointment scheduling (Alexandersson et al. 1997)
  - Can we meet in the second half of may? SUGGEST-DATE
  - Well, that's not so good, REJECT
    because I'll be on holiday. GIVE-REASON
    How about early June, such as the 3rd? SUGGEST-DATE
  - Hm, I don't know, HESITATE
    all right, I guess I can do that. ACCEPT-DATE

Dialog act

- Some proposals of DA taxonomies
  - DAMSL (Allen & Core, 1997)
  - DIT++ (Bunt 2006) http://dit.utv.nl

- Automatic DA recognition
  - (Stolcke et al. 00) on minutes of meetings
  - nice overview:
    P. Kral: Dialogue act recognition approaches. Computing and Informatics 29:227-250, 2010

Why study this on Twitter?

- In the conversations, do people
  - talk to each other or past each other?
  - exchange information?
  - exchange opinions?
  - exchange arguments?
  - display their emotions?
  - follow „standard“ dialog protocols?
  - …
- Knowing this is relevant, inter alia, for building good Twitter bots
  - **User**: QUESTION
    Bot: ANSWER
  - **User**: OPINION
    Bot: AGREE | DISAGREE
  - …

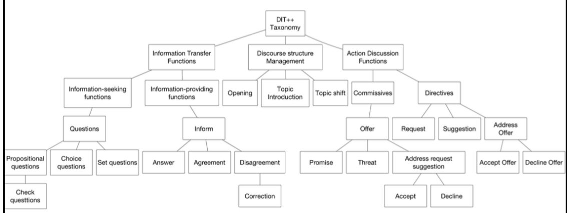
From dialog acts to „tweet acts“

- Tweets collected on the topic of renewable energy
  - keyword spotting
  - re-crawl missing tweets to get trees as complete as possible
- 1566 tweets in 172 conversations

E. Zarisheva, T. Scheffler: Dialog act annotation for Twitter conversations. Proc. of SIGDIAL, Prague, 2015

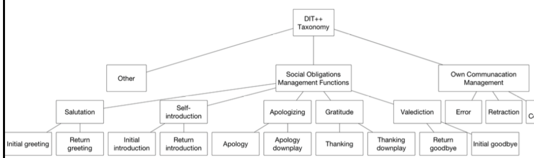
Manfred Stede ESSLLI 2016

DA taxonomy (part 1)



Manfred Stede ESSLLI 2016

DA taxonomy (part 2)



Manfred Stede ESSLLI 2016

Annotation

- Total: 51 DA labels
- Annotators have to choose one label per tweet segment
- *[True, unfortunately.]*$_{\text{Agreement}}$ *[But what about the realization of high solar activity in the 70s and 80s?]*$_{\text{SetQuestion}}$

Manfred Stede ESSLLI 2016

Annotation

- minimally-trained undergraduate students
- two steps
  - segmentation: Fleiss multi-pi 0.88
  - DA labelling: Fleiss multi-pi 0.56
    - notice: disagreements on subcategories are punished like disagreements on main categories
    - (with 10 DAs: multi-pi 0.76)
    - most disagreement is among types of Information-providing
- Finally, all annotations merged into a single gold-standard: 1213 tweets / 2936 segments

Manfred Stede ESSLLI 2016

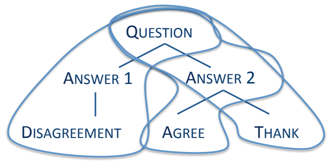
Tweet-internal structure

Number of segments per tweet	Tweets
1 segment	89 times
2 segments	671 times
3 segments	320 times
4 segments	114 times
5 segments	17 times
6 segments	2 times

[@TheBug0815 @Luegendetektor @McGeiz]$_0$ [Exactly, we don't need a base load, it's only a capitalist construct]$_{\text{Agreement}}$ – [Wind/PV are sufficient?]$_{\text{PropQuestion}}$ [Lol]$_{\text{Disagreement}}$

Automatic recognition

- Assume gold segments
- Split conversation trees into single strands



Tatjana Scheffler and Elina Zarisheva: Dialog act recognition for Twitter conversations. Proc. of the LREC Workshop on Normalisation and Analysis of Social Media Texts, 2016

Features

feature combinations
UD
UD + top50
UD + top100
UD + embeddings
UD + top100 + embeddings

- user defined (UD):
  - segment length
  - author
  - position of segment in tweet
  - presence of question marks, links, hashtags, etc.
- top 50/100 words for the dialog act (by TF-IDF)
- word embeddings (pre-calculated, 64 dimensions)

Experiment: Full DA set

Hidden Markov Model
  Multinomial distribution (Discrete values)
  Gaussian distribution (M-dimensional vectors)
Conditional Random Fields

majority baseline (INFORM) f = 0.09

	MHMM			GHMM			CRF		
	f	acc.	π	f	acc.	π	f	acc.	π
UD	0.22	0.22	0.33	0.18	0.16	0.25	**0.28**	**0.32**	**0.44**
UD + L50	0.05	0.03	0.42	0.20	0.19	0.49	**0.31**	**0.37**	**0.62**
UD + L100	0.04	0.02	0.42	0.20	0.19	0.50	**0.31**	**0.37**	**0.62**
UD + WE				0.18	0.16	0.45	**0.31**	**0.36**	**0.61**
ALL				0.21	0.22	0.50	**0.31**	**0.37**	**0.62**

Results: Reduced/Minimal DA set

F_1	Baseline	GHMM	CRF
Full (50 DAs)	0.09	0.21	**0.31**
Reduced (12 DAs)	0.16	0.36	**0.51**
Minimal (8 DAs)	0.34	0.51	**0.72**

previous work:
- (Zhang et al., 2011): F_1 = 0.695 for 5 classes
- (Arguello and Shaffer, 2015) MOOC forum posts: average precision ~0.65; ours, 0.70
- (Vosoughi and Roy, 2016): F_1 = 0.70 for 6 classes

Overview

- Twitter conversations ??
- From speech acts via dialog acts to tweet acts
- Coherence relations and „rhetorical structure"
- Deriving rhetorical structure automatically
- Rhetorical structure in Twitter conversations

Coherence

John took a train to Istanbul. He has family there.

John took a train to Istanbul. He likes spinach.
(Hobbs 76)

Coherence := Coreference + Coherence relations

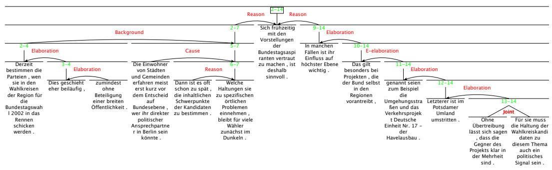
Coherence relations

- John took a train to Istanbul. He first visited the Hagia Sophia. (temporal-sequence)
- John took a train to Istanbul. His sister went to Rome. (contrast)
- John took a train to Istanbul. It was a comfortable Eurocity. (elaboration)

Claim

- In text, coherence arises because every clause is linked via a coherence relation to its left context.
- OK, but...
  - how exactly do you determine the minimal units?
  - what inventory of relations do you assume?
  - what can the current clause be attached to?

Rhetorical Structure Theory (Mann/Thompson 88)



Overview

- Twitter conversations ??
- From speech acts via dialog acts to tweet acts
- Coherence relations and „rhetorical structure“
- Deriving rhetorical structure automatically
- Rhetorical structure in Twitter conversations

Two notions of discourse parsing

- **„Shallow“**
  - Rooted in the *Penn Discourse Treebank* corpus
  - Assign arguments to connectives
  - *[Other reviewers said it's a great place,]$_{Arg1}$ but [my impression was otherwise.]$_{Arg2}$*
  - Identify other relations and arguments between sentences
- **„Full structure“**
  - Build a complete structure for the text
  - *Rhetorical Structure Theory* (Mann/Thompson 88)
  - *Segmented Discourse Representation Theory* (Asher/ Lascarides 03)

RST Discourse Treebank

- 385 articles from Wall Street Journal
- Overlap with the Penn Treebank

- 78 relations, many of them arising from nuclearity reversals
  - Evaluation (nucleus: evaluated / sat: evaluating)
  - Evaluation (nucleus: evaluating / sat: evaluated)

L. Carlson et al.: Building a discourse-tagged corpus in the framework of Rhetorical Structure Theory. Proc. of SIGDIAL, 2001

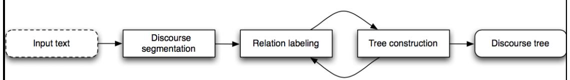
HILDA: RST-parsing via SVM classification

H. Hernault et al.L Building a Discourse Parser Using Support Vector Machine Classification. Dialogue & Discourse 1(3), 2010

HILDA: RST-parsing via SVM classification

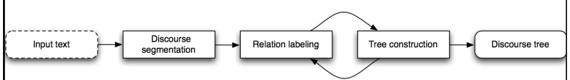
- Training & Test: RST Discourse Treebank
- Reduce 78 relations to 18 „families"
- *Attribution, Background, Cause, Comparison, Condition, Contrast, Elaboration, Enablement, Evaluation, Explanation, Joint, Manner-Means, Summary, Temporal, Topic-Change, Topic-Comment, Same-unit, Textual-organization*
- Any n-ary trees with n>2 are converted to binary trees

Idea



- Complexity: linear with respect to length of input text
- Both segmentation and relation labelling run as supervised classification tasks, using support vector machines
- (Here, we skip the segmentation step (F-score 0.95)

Idea



- STRUCT classifier: What is the probability of segments i, i+1 being connected?
- LABEL classifier: What is the most likely relation to hold between segments i, i+1?

STRUCT classifier

REL(list(elabNS(Seg1,Seg2), evalSN(Seg3,condNS(Seg4,Seg5))),Seg6)

list(elabNS(Seg1,Seg2), evalSN(Seg3,condNS(Seg4,Seg5)))

evalSN(Seg3,condNS(Seg4,Seg5)

elabNS(Seg1,Seg2) condNS(Seg4,Seg5)

Seg1 Seg2 Seg3 Seg4 Seg5 Seg6

REL classifier

elabNS(list(elabNS(Seg1,Seg2), evalSN(Seg3,condNS(Seg4,Seg5))),Seg6)

list(elabNS(Seg1,Seg2), evalSN(Seg3,condNS(Seg4,Seg5)

evalSN(Seg3,condNS(Seg4,Seg5)

elabNS(Seg1,Seg2) condNS(Seg4,Seg5)

Seg1 Seg2 Seg3 Seg4 Seg5 Seg6

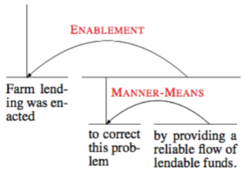
Features for relation labeling (1)

Table 1: Features encoding textual organization

Feature name	Scope
Belong to same sentence	F
Belong to same paragraph	F
Number of paragraph boundaries	S
Number of sentence boundaries	S
Length in tokens	S
Length in EDUs	S
Distance to beginning of sentence in tokens	S
Size of span over sentence in EDUs	S
Size of span over sentence in tokens	S
Size of both spans over sentence in tokens	F
Distance to beginning of sentence in EDUs	S
Distance to beginning of text in tokens	S
Distance to end of sentence in tokens	S

Features for relation labeling (2)

- Cue words: modeled as 3-grams at beginning and end of spans
- For Manner-Means: (to,correct,this) (of,lendable,funds)
- 12.000 3-grams
- 2 x 3 x 384 POS tags



Features for relation labeling (3)

Feature name	Scope
Distance to root of the syntax tree	S
Distance to common ancestor in the syntax tree	S
Delta of distances to common ancestor	F
Dominating node's lexical head in span	S
Common ancestor's POS tag	F
Common ancestor's lexical head	F
Dominating node's POS tag	F
Dominating node's lexical head	F
Dominated node's POS tag	F
Dominated node's lexical head	F
Dominated node's sibling's POS tag	F
Dominated node's sibling's lexical head	F
Relative position of lexical head in sentence	S

Results: Relation labeling

- **STRUCT**
  - Trained on 52.683 instances (1/3 positive)
  - Tested on 8.558 instances
  - Feature space dimensionality: 136.987
  - Accuracy with polyn. Kernel: **85.0**
- **LABEL**
  - Trained on 17.742 instances
  - Tested on 2.887 instances
  - Accuracy of multi-class SVM: **66.8**

Performance on individual relations

SVM Class	Precision	Recall	F-score
ATTRIBUTION[N][S]	93.6	96.2	94.9
ATTRIBUTION[S][N]	95.7	93.7	94.7
BACKGROUND[N][S]	47.8	41.5	44.4
BACKGROUND[S][N]	38.7	20.7	27.0
CAUSE[N][S]	33.3	2.1	3.9
COMPARISON[N][S]	50.0	5.9	10.5
CONDITION[N][S]	100.0	47.8	64.7
CONDITION[S][N]	85.7	72.0	78.3
CONTRAST[N][N]	31.1	21.9	25.7
CONTRAST[N][S]	50.0	20.8	29.4
CONTRAST[S][N]	51.1	39.7	44.7
ELABORATION[N][S]	58.1	94.5	72.0
ENABLEMENT[N][S]	61.9	59.1	60.5
ENABLEMENT[S][N]	50.0	50.0	50.0

Overview

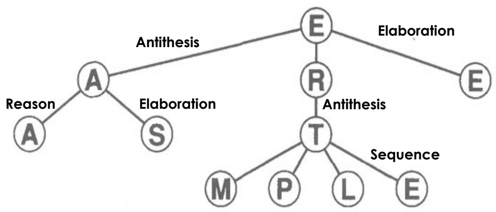
- Twitter conversations ??
- From speech acts via dialog acts to tweet acts
- Coherence relations and „rhetorical structure“
- Deriving rhetorical structure automatically
- Rhetorical structure in Twitter conversations

Inter-Tweet-Relations

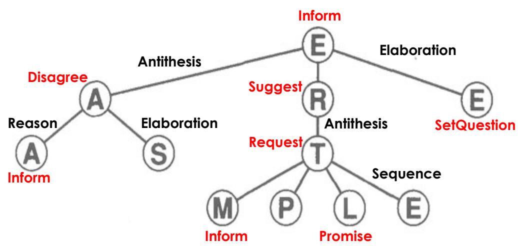
- **UdoSieverding**
  #Offshore-Ausbau: Warum schweigen Dauer-#EEG-Kritiker @Der_BDI @iw_koeln @insm @DICEHHU @RolandTichy @bdew_ev @igbce? http://t.co/WfZsirxMiC
- mapro67
  @UdoSieverding weil sie alle Interessen der großen Stromkonzerne vertreten @Der_BDI @iw_koeln @insm @DICEHHU @RolandTichy @bdew_ev @igbce

- **UdoSieverding**
  #Offshore expansion: Why are the big critics of the renewable energy law so quiet?
- mapro67
  @UdoSieverding because they represent the interests of the big energy companies

Ongoing work



To be clarified: DA ≈ CoherenceRel



Add: Tweet-internal structure

- *WTF? I have green energy and am supposed to finance nuclear and coal? What nonsense. WHAT NONSENSE!*

- *Agree. But it's also a fact that …*