Automatically Evaluating Text Coherence Using Discourse Relations

Discourse Coherence Theory and Modeling

Noushin Fadaei
Department of Computational Linguistics
Saarland University

July 1, 2013

Overview

- ❑ Motivation
- ❑ Background
  - – Different discourse frameworks
  - – State of the art development
  - – How to utilize PDTB in this task?
- ❑ The Refined Approach
  - – Discourse Role
  - – Discourse Role Transitions and Lexical Chains
  - – Discourse Role Matrix (Exercise)
- ❑ Evaluation of Results
- ❑ Conclusion
- ❑ Summary (Exercise)

Motivation

- Distinguish a Coherent Text from an Incoherent Text Automatically

- Improves state of the art model: Synergistic with the previous approach suggested by Barzilay and Lapata(2005)

- Use in Summarization, Text Generation and Essay Scoring.

Background

Different discourse frameworks

If I had an alarm clock, I would have broken it every morning.

I could not wake up with my alarm clock, which has been broken since yesterday.

RST
Discourse Relations

Concessive
Conditional

Satellite + Nucleus

Elaboration
Evidence

Nucleus+ Satellite

Background

Different discourse frameworks

[The constitution does not expressly give the president such power.]$\mathbf{S_1}$ [However the president does have a duty not to violate the constitution.]$\mathbf{S_2}$ [The question is whether his only means of defense is the veto.]$\mathbf{S_3}$

Pattern: **Contrast-followed-by-Cause**

Coherent text exhibits measurable preferences for specific intra- and inter-discourse relation ordering.

Background

State of the art related work

Local entity transition model by Barzilay and Lapata(2005)

	Department	Trial	Microsoft	Evidence	Competitors	Markets	Products	Brands	Case	Netscape	Software	Tactics	Government	Suit	Earnings	
1	S	O	S	X	O	–	–	–	–	–	–	–	–	–	–	1
2	–	–	O	–	–	X	S	O	–	–	–	–	–	–	–	2
3	–	–	S	O	–	–	–	–	S	O	O	–	–	–	–	3
4	–	–	S	–	–	–	–	–	–	–	–	S	–	–	–	4
5	–	–	–	–	–	–	–	–	–	–	–	–	S	O	–	5
6	–	X	S	–	–	–	–	–	–	–	–	–	–	–	O	6

Background

Use of PDTB

You want to model the Discourse Relation Transition 

Identify Discourse Relations by some automatic annotators

Discourse **L**exicalized **T**ree **A**djoining **G**rammar (**D-LTAG**)

PDTB

- Implicit → Inferred by reader
- Explicit → e.g. Because

Background

Use of PDTB

D-LTAG tags relations with two levels of types:

Level-1

Comparison(Comp)

The use of compound subjects help lengthen simple sentences but simple sentences are often short.

Temporal(Temp)

They have not spoken to each other since the huge argument they had last year.

Contingency(Cont)

Since you said you'll not drop by, I left home at 8PM.

Expansion(Exp)

Approach: Discourse Role

[Japan normally depends heavily on the Highland Valley and **Cananea** mines as well as the Bougainville mine in Papua New Guinea.] S_1

[Recently, Japan has been buying copper elsewhere.] S_2

[[But as Highland Valley and **Cananea** begin operating,] $C_{3:1}$ [they are expected to resume their roles as Japan's suppliers.] $C_{3:2}$] S_3

[[According to Fred Demler, metals economist for Drexel Burnham Lambert, New York,] $C_{4:1}$["Highland Valley has already started operating] $C_{4:2}$[and **Cananea** is expected to do so soon."] $C_{4:3}$] S_4

1. Implicit Comparison < **S_1:Arg1 , S_2:Arg2** >
2. Explicit Comparison < **S_2:Arg1, S_3:Arg2** >
3. Explicit Temporal < **$C_{3:1}$:Arg1, $C_{3:2}$:Arg2**>
4. Implicit Expansion < **S_3:Arg1, S_4:Arg2** >
5. Explicit Expansion < **$C_{4:2}$:Arg1, $C_{4:3}$:Arg2** >

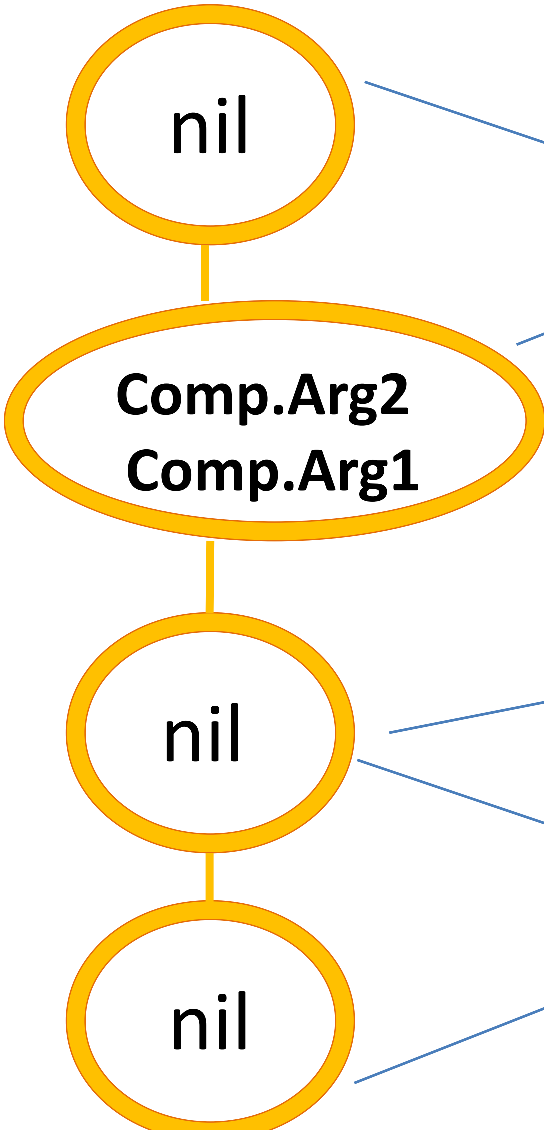
Discourse Role=Discourse Relation Type + Argument Span Position

Approach: Discourse Role Transitions

Sub-sequence of discourse roles for a term in multiple consecutive sentences.

1. Find the discourse roles of a word over the text (for each sentence)
2. Put the discourse roles of the word for a particular sentence in a node
3. Connect the nodes as their associated sentences are connected over the text
4. Now you have a **Lexical Chain** with simplified nodes and typed with relations
5. Any sub-sequence at any length extracted from this chain is a Discourse Role Transition

Approach: Discourse Role Transitions



Sub-sequences of this chain with length 2:

nil→**Comp.Arg2**
nil→**Comp.Arg1**

1x2=2

Comp.Arg2→nil
Comp.Arg1→nil

2x1=2

nil→nil

1x1=1

A total of **5** sub-sequences with the length 2

Approach: Discourse Role Matrix

S#	Copper	Cananea	Operate	Depend
S_1	nil	Comp.Arg1	nil	Comp.Arg1
S_2	Comp.Arg2 Comp.Arg1	nil	nil	nil
S_3	nil	Comp.Arg2 Temp.Arg1 Exp.Arg1	Comp.Arg2 Temp.Arg1 Exp.Arg1	nil
S_4	nil	Exp.Arg2	Exp.Arg2 Exp.Arg1	nil
	5	7	10	3

S#	Mike	Frank	story
S_1	Temp.Arg1 Temp.Arg2	nil	Temp.Arg1
S_2	nil	Comp.Arg2 Con.arg1	nil
S_3	Temp.Arg2 Temp.Arg1 Con.Arg2	Temp.Arg2 Temp.Arg1 Con.Arg2	nil
	5	8	2

15 sub-sequences with length2

How does the approach model Discourse Relation Transitions?

Imagine The matrices of the whole text:

- ❑ Identify all the possible sub-sequences of any term(column) with any length.

  Length 2: 25

  Length 3: …

  Length 4: …

- ❑ What is the probability of each of the sub-sequences? P(Comp.Arg2→Exp.Arg2)?

Approach: Distribution of Discourse Role

- By computing the probabilities of all the possible sub-sequences, we access the "Distribution of Discourse Roles" in the text.

- Hypothesis: This distribution is distinguishable between coherent and incoherent texts.

Further Refine the approach
Salience

Divide the Matrix into two parts

Salient

Term Frequency in the column is greater than a threshold

Non-Salient

A low frequency

Approach: Preference Ranking

Coherence of a text (or the level to which we comfortably read) is relative

Text 1
Text 2
Text 3

Input

Full Model

Text 2
Text 3
Text 1

Output

Experiment

- **Task:** Which of two texts is more coherent? A source text or its permutations?

- **Data collections:** Wall Street Journal articles and AP articles about earthquakes also articles about accidents(NTSB)

- **Parameters:** Term Frequency as threshold for salience(TF>=2) and length of sub-sequences as features(<=3)

Results

- This model outperforms the baseline(Local Entity Transition Model)

- Eliminating each of the information sources: Argument tags and salience play more important roles than relations.

- The combination of baseline and this approach outperforms both of the approaches: As it collects features from different aspects of coherence.

Test Set Ranking Accuracy

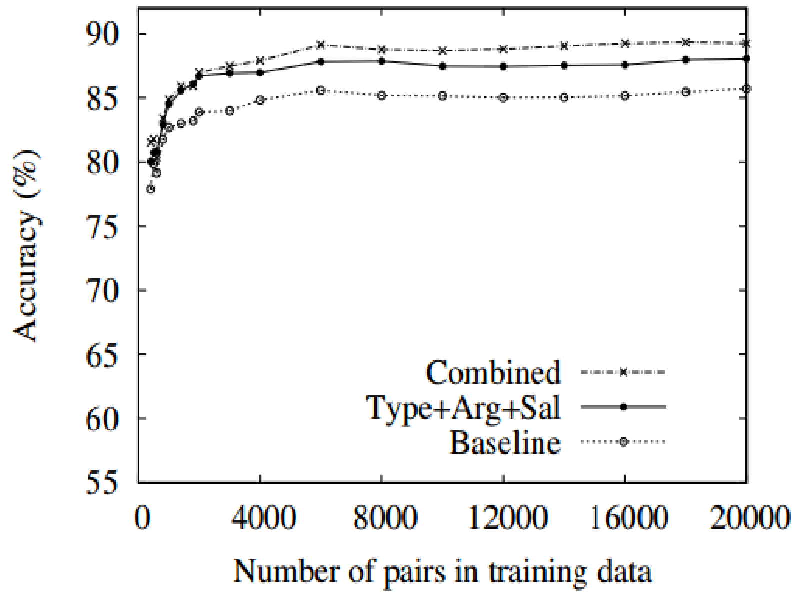
	WSJ	Earthquakes	accidents
Baseline	85.71	83.59	89.93
Type+Arg+Sal	**88.06**	**86.50**	**89.38**
Arg+Sal	88.28	85.89	87.06
Type+Sal	87.06	82.98	86.05
Type+Arg	85.98	82.67	87.87
Baseline & Type+Arg+Sal	89.25	89.72	91.64

Conclusion

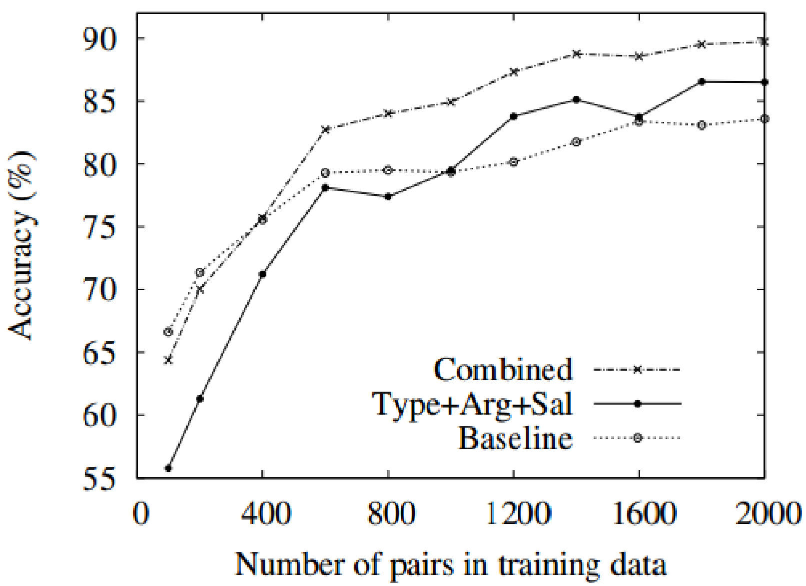
- The combined model performs better than each of its constituents .

- The full model outperforms the baseline in the mid-segment of the curve:
  - —It fails in some permutations.
  - —PDTB doesn’t work with implicit relations as it works with explicit ones.

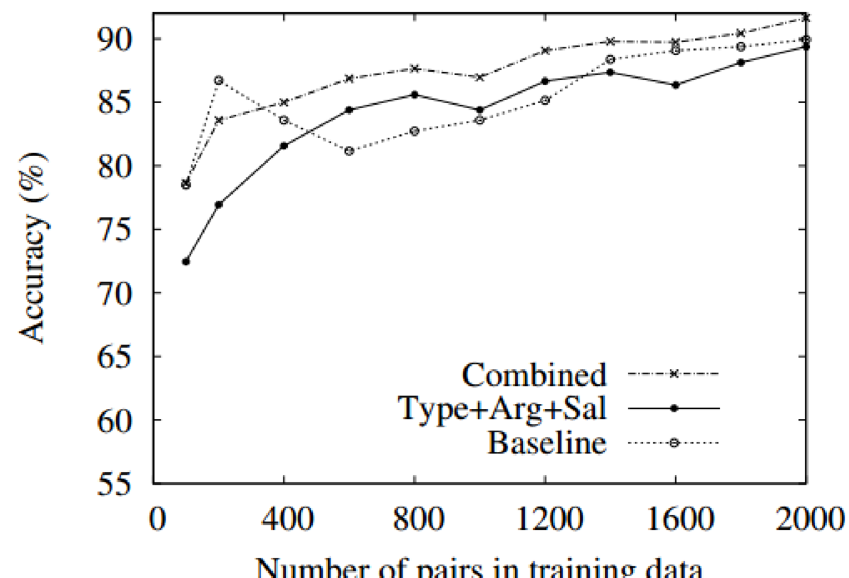
- Relation/length ratio determines why the accuracy is better in accidents



(a) WSJ



(b) Earthquakes



(c) Accidents

Summary

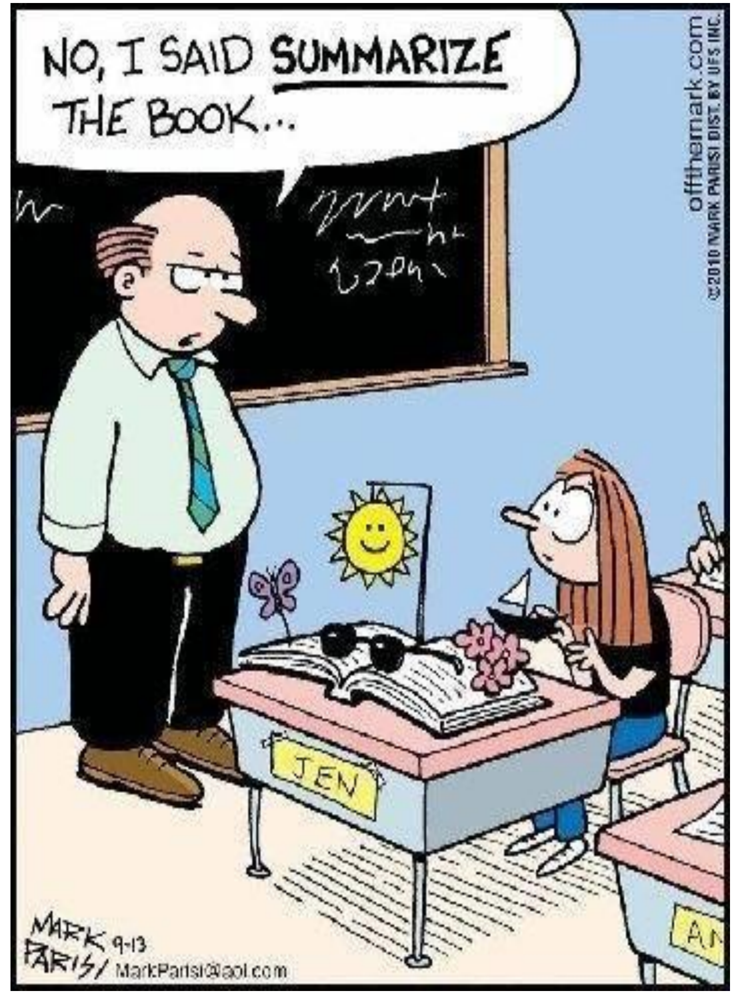
Distribution of Discourse role

Discourse Role Matrix

Salience

Baseline : Barzilay & Lapata(2005)

Preference Ranking



waits for a question

......

References

Ziheng Lin, Hwee Tou Ng and Min-Yen Kan. 2011. Automatically Evaluating Text Coherence Using Discourse Relations. *ACL:* 997-1006

Regina Barzilay and Mirella Lapata. 2008. Modeling local coherence: An entity-based approach. Computational Linguistics, 34(1), 1-34.

Thank you!