Natural Language Inference
NatLog

Marc Schulder

Saarland University

18. Juli 2011

Overview

Natural
Language
Inference

Marc Schulder

Entailment
Relations

Compositional
Entailment

NatLog
System
FraCaS Test Suite
RTE3 Test Suite

Recap: Entailment Relations

Symbol	Name	Example	Set Theoretic Definition
$x \equiv y$	equivalence	couch $\equiv$ sofa	$x = y$
$x \sqsubset y$	forward entailment	crow $\sqsubset$ bird	$x \subset y$
$x \sqsupset y$	reverse entailment	bird $\sqsupset$ crow	$x \supset y$
$x \wedge y$	negation	able $\wedge$ unable	$x \cap y = \emptyset \wedge x \cup y = U$
$x \mid y$	alternation	cat $\mid$ dog	$x \cap y = \emptyset \wedge x \cup y \neq U$
$x \smile y$	cover	animal $\smile$ non-ape	$x \cap y \neq \emptyset \wedge x \cup y = U$
$x \# y$	independence	hungry # hippo	(all other cases)

Natural
Language
Inference

Marc Schulder

Entailment
Relations

Compositional
Entailment

NatLog
System
FraCaS Test Suite
RTE3 Test Suite

Entailment joins

⋈	≡	⊏	⊐	∧	|	⌣	#
≡	≡	⊏	⊐	∧	|	⌣	#
⊏	⊏	⊏	≡⊏⊐|#	|	|	⊏∧|⌣#	⊏|#
⊐	⊐	≡⊏⊐⌣#	⊐	⌣	⊐∧|⌣#	⌣	⊐⌣#
∧	∧	⌣	|	≡	⊐	⊏	#
|	|	⊏∧|⌣#	|	⊏	≡⊏⊐|#	⊏	⊏|#
⌣	⌣	⌣	⊐∧|⌣#	⊐	⊐	≡⊏⊐⌣#	⊐⌣#
#	#	⊏⌣#	⊐|#	#	⊐|#	⊏⌣#	≡⊏⊐∧|⌣#

Recap: Compositional Entailment

Natural
Language
Inference

Marc Schulder

Entailment
Relations

**Compositional
Entailment**

NatLog
System
FraCaS Test Suite
RTE3 Test Suite

i	e_i	$x_i = e_i(x_{i-1})$	$\beta(e_i)$	$\beta(x_{i-1}, e_i)$	$\beta(x_0, x_i)$
		Stimpy is a cat.			
1	SUB(*cat, dog*)		|	|	|
		Stimpy is a dog.			
2	INS(*not*)		^	^	⊏
		Stimpy is not a dog.			
3	SUB(*dog, poodle*)		⊐	⊏	⊏
		Stimpy is not a poodle			

NatLog Pipeline

Natural
Language
Inference

Marc Schulder

Entailment
Relations

Compositional
Entailment

NatLog
System
FraCaS Test Suite
RTE3 Test Suite

1. Linguistic analysis
2. Alignment
3. Entailment classification
    1. Lexical entailment classification
    2. Entailment projection
    3. Entailment joining

Linguistic Analysis
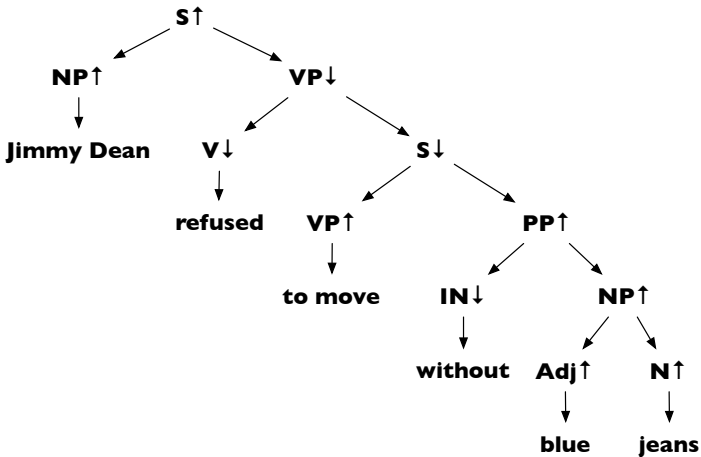
1. Tokenisation
2. POS tagging
3. Syntactic phrase-structure parsing
   ⇒ Stanford parser
4. Lemmatisation
   ⇒ Finite-State Analyser *(Minnen et al.)*
5. Monotonicity marking
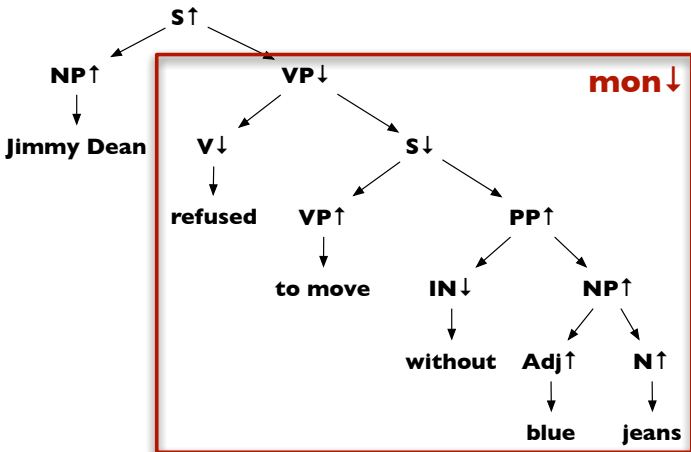
Tasks not performed

▶ Named entity recognition

▶ Coreference resolution

▶ Dependency parsing

▶ Semantic role labeling

Natural
Language
Inference

Marc Schulder

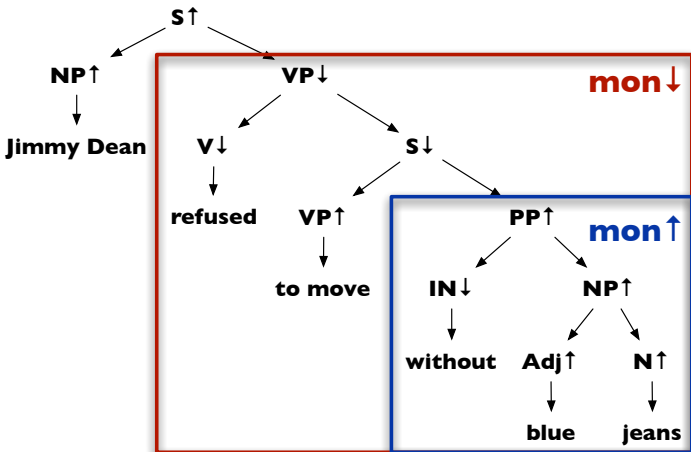Entailment
Relations

Compositional
Entailment

NatLog
System
FraCaS Test Suite
RTE3 Test Suite

Monotonicity marking

Natural
Language
Inference

Marc Schulder

Entailment
Relations

Compositional
Entailment

NatLog
System
FraCaS Test Suite
RTE3 Test Suite

Jimmy	Dean	refused				to move	without	blue	jeans
	SUB		DEL	INS	INS	SUB	EQ	DEL	SUB
James	Dean			did	n't	dance	without		pants

Sequence of atomic edits

1. Generate alignments
   - Atomic Edits: EQ, SUB, DEL, INS
   - ⇒ Stanford RTE aligner, string edit distance
2. Heuristic order of edits
   - DEL
   - SUB
   - All downward/non-monotone edits
   - INS

Lexical entailment classification

i	e_i	$x_i = e_i(x_{i-1})$	$\beta(e_i)$	$\beta(x_{i-1}, e_i)$	$\beta(x_0, x_i)$
		Stimpy is a cat.			
1	SUB(*cat, dog*)		|	|	|
		Stimpy is a dog.			
2	INS(*not*)		^	^	⊏
		Stimpy is not a dog.			
3	SUB(*dog, poodle*)		⊐	⊏	⊏
		Stimpy is not a poodle			

1 Linguistic analysis

2 Alignment

3 **Entailment classification**

    1 **Lexical entailment classification**
    2 Entailment projection
    3 Entailment joining

Lexical entailment classification

Jimmy Dean	refused				to move	without	blue	jeans
SUB		DEL	INS	INS	SUB	EQ	DEL	SUB
James Dean			did	n't	dance	without		pants

Predicting entailment relations

1 Find lexical features
- ▶ **WordNet** Synonymy, hypernymy, ...
- ▶ **String similarity** Levenshtein, ...
- ▶ **Relatedness** Semantic, derivational, distributional
- ▶ **Lexical category** Prepositions, light, quantifiers,...
- ▶ **Misc. features** Number mismatches, special cases

2 Look up in decision tree
⇒ Weka

i	e_i	$x_i = e_i(x_{i-1})$	$\beta(e_i)$	$\beta(x_{i-1}, e_i)$	$\beta(x_0, x_i)$
		Stimpy is a cat.			
1	SUB(*cat, dog*)		|	|	|
		Stimpy is a dog.			
2	INS(*not*)		^	^	⊏
		Stimpy is not a dog.			
3	SUB(*dog, poodle*)		⊐	⊏	⊏
		Stimpy is not a poodle			

1 Linguistic analysis

2 Alignment

3 **Entailment classification**

    1 Lexical entailment classification
    2 **Entailment projection**
    3 Entailment joining

Entailment joining

i	e_i	$x_i = e_i(x_{i-1})$	$\beta(e_i)$	$\beta(x_{i-1}, e_i)$	$\beta(x_0, x_i)$
		Stimpy is a cat.			
1	SUB(*cat, dog*)		|	|	|
		Stimpy is a dog.			
2	INS(*not*)		^	^	⊏
		Stimpy is not a dog.			
3	SUB(*dog, poodle*)		⊐	⊏	⊏
		Stimpy is not a poodle			

1 Linguistic analysis

2 Alignment

3 **Entailment classification**

    1 Lexical entailment classification

    2 Entailment projection

    3 **Entailment joining**

- ► Manually designed simple sentences
- ► Premise and hypothesis quite similar
- ► Designed to reflect various phenomena
- ► Answer types:
  - ► YES (entailment) (59%) ($\equiv \sqsubset$)
  - ► NO (contradiction) (28%) (^|)
  - ► UNK (compatibility) (10%) ($\sqsupset \smile \#$)
- ► NatLog ignores multi-premise problems
  - ► Remaining: 192 problems (55%)

p No delegate finished the report.
h Some delegate finished the report on time. **NO**

p Either Smith, Jones or Anderson signed the contract.
h Jones signed the contract. **UNK**

p John wrote a report, and Bill said Peter did too.
h Bill said Peter wrote a report. **YES**

FraCaS test suite

System	P (YES)	R (YES)	Acc
Baseline	55.7%	100.0%	55.7%
Bag of words	59.7%	87.2%	57.4%
NatLog	89.3%	65.7%	70.5%

Section	#	P (YES)	R (YES)	Acc
Quantifiers	44	95.2%	100.0%	97.7%
Plurals	24	90.0%	64.3%	75.0%
Anaphora	6	100.0%	60.0%	50.0%
Elipsis	25	100.0%	5.3%	24.0%
All sections	183	89.3%	65.7%	70.5%

- RTE3 = 3rd Recognizing Textual Entailment Challenge
- Real premises, hand-written hypotheses
- Premises a lot longer than hypotheses (30 vs 7 words)
- Binary answers:
  - YES (Entailment) (50%) ($\equiv \sqsubset$)
  - NO (Non-entailment) (50%) (^| $\sqsupset \smile \#$)
- Development & test set (800 problems each)
- Potential problems for NatLog:
  - Other types of inference required
  - Large edit distance

p As leaders gather in Argentina ahead of this weekends regional talks, Hugo Chávez, Venezuela's populist president is using an energy windfall to win friends and promote his vision of 21st-century socialism.

h Hugo Chávez acts as Venezuela's president.

► **YES**

p Mr. Fitzgerald revealed he was one of several top officials who told Mr. Libby in June 2003 that Valerie Plame, wife of the former ambassador Joseph Wilson, worked for the CIA.

h Joseph Wilson worked for CIA.

► **NO**

RTE3 test suite

Natural
Language
Inference

Marc Schulder

Entailment
Relations

Compositional
Entailment

NatLog
System
FraCaS Test Suite
RTE3 Test Suite

Section	Data	YES	P (YES)	R (YES)	Acc
NatLog	dev	22.5%	**73.9%**	32.4%	59.3%
	test	26.4%	**70.1%**	36.1%	59.4%
Stanford	dev	50.3%	68.7%	67.0%	67.3%
	test	50.0%	61.8%	60.2%	60.5%
Hybrid	dev	56.0%	69.2%	75.2%	**70.0%**
	test	54.5%	64.5%	68.5%	**64.5%**

- ▶ **Stanford system** predicts via *inference score*
- ▶ **Hybrid system** biasses Stanford towards NatLog decision

Discussion

NatLog

- ▶ Performs well for certain problems.
- ▶ Not yet fit for real world application.
- ▶ Low noise resistance.
- ▶ Strong bias towards independence relation.
- ▶ Potential as a component of a larger system