

# Discussion: Really Difficult Natural Language Inference problems

Johan Bos

University of Groningen

# The basic idea...



This school was founded in 1650.

| x1                      | e1 | t1 |
|-------------------------|----|----|
| school(x1)              |    |    |
| time(t1)                |    |    |
| YearOfCentury(t1, 1650) |    |    |
| t1 < now                |    |    |
| establish(e1)           |    |    |
| Time(e1, t1)            |    |    |
| Theme(e1, x1)           |    |    |



Diese Schule wurde 1650 gegründet.

| x1                      | e1 | t1 |
|-------------------------|----|----|
| school(x1)              |    |    |
| time(t1)                |    |    |
| t1 < now                |    |    |
| YearOfCentury(t1, 1650) |    |    |
| establish(e1)           |    |    |
| Time(e1, t1)            |    |    |
| Theme(e1, x1)           |    |    |

# Cross-lingual entailment

- Use parallel corpus of translations
- Translations ought to be meaning-preserving

**Parallel Meaning Bank (tomorrow)**

# Data-driven

- No farmers, bathrooms, donkeys, ...
- No scope ambiguities

# Recognizing Textual Entailment

- Input: two phrases/sentences/texts
- Output: one of three labels
  - Not informative (entailment)
  - Informative (neutral)
  - Very informative (contradiction)

# Observation

Observation:

you can also do RTE with a single sentence

- If we get more trade, then we get more commerce (G.W. Bush)
- Tom got himself new shoes.
- Every man shaved himself and John didn't.

# History of RTE

- Stanford NLI corpus

# History of RTE

- Stanford NLI corpus
- SICK (Trento)



# History of RTE

- Stanford NLI corpus
- SICK (Trento)
- Pascal RTE

# History of RTE

- Stanford NLI corpus
- SICK (Trento)
- Pascal RTE
- Monz & De Rijke (2001)

# History of RTE

- Stanford NLI corpus
- SICK (Trento)
- Pascal RTE
- Monz & De Rijke (2001)
- Fracas test suite (1996)

# History of RTE

- Stanford NLI corpus
- SICK (Trento)
- Pascal RTE
- Monz & De Rijke (2001)
- Fracas test suite (1996)
- Aristotle's syllogisms

# Predicting inference labels

- “today we are in a strange new world” (Vigdis)
- In near neural times the labels won't be hard to predict
- So we need to deliver more!  
[[we]] = semanticists/representationalists
- Explanation-based RTE? (more about this later!)

# Labelling Data

- Should be done by linguists!  
Preferably semanticists!!  
Not by ordinary language users!!!
- A dog is swimming in the pool.  
A dog is not swimming in the pool.

Semantic phenomena that I find hard

# Negation affixes

- I am unhappy --> I am not happy YES
- I am not happy --> I am unhappy NO



# Negation affixes

- I am unhappy --> I am not happy YES
- I am not happy --> I am unhappy NO
  
- I am unarmed --> I am not armed YES
- I am not armed --> I am unarmed YES

# False negation affixes

- This wood is inflammable --> This wood is not flammable NO
- This wood is not flammable --> this wood is inflammable NO
- That lecture was disappointing -->  
That lecture was not appointing. NO!

# Manner of Movement

- He is walking to school.
- He is going on foot to school. YES

# What do we learn from translations?

PMB: 14/0849

EN: She removed **the dishes** from the table.

NL: Ze ruimde de tafel af.



# What do we learn from translations?

PMB: 59/1946

EN: I do like ice cream.

DE: Ich mag **wirklich** Eiscreme.



PMB: 68/2811

EN: I do believe it's called a leek.

IT: Io credo **davvero** che si chiama porro.



# What do we learn from translations?

PMB: 59/1946

EN: I **do** like ice cream.

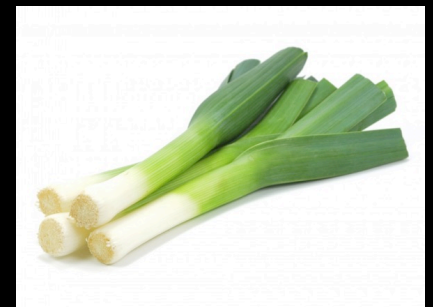
DE: Ich mag **wirklich** Eiscreme.



PMB: 68/2811

EN: I **do** believe it's called a leek.

IT: Io credo **davvero** che si chiama porro.



# Definiteness

- He walks to school.
- Er geht zu Fuß **zur** Schule.
- Tom also plays guitar.
- Tom suona anche **la** chitarra.

# Figurative Paraphrasing

- I am very hungry.
- Ik rammel van de honger.



# Noun-noun compounds

- We arrived at a car park.
- We arrived at a place where several cars were parked.
  
- The girls were playing with water pistols.
- The girls were playing with water.
  
- John got a bicycle helmet for his birthday.
- John has a bike.

# Implicit negation

- I hate birthday parties
- I don't like birthday parties

# Agent nouns

- The cook entered the building.
- Someone who cooks food entered.
- Someone who was cooking food entered.

# Plausible inferences

- John bought cat food.
- John bought food for his cat.
- Mary asked John to buy her some cat food.
- John bought food for Mary's cat.

# The elephant in the room: adjectives

fairly tall -/-> tall

very tall --> tall

extremely tall --> very tall

tall for a Dutchman

a tall Italian

tall for his age

# too

- He was **too** tired to do his homework.
  - He was tired. YES
  - He didn't do his homework. YES
- Stephan was **too** short to reach the top shelf.
  - Stephan was short. NO
  - Stephan could not reach the top shelf. YES

# near

- My sister lives near Yokohama.
- My sister lives in Yokohama. NO
- Tom lives near us.
- Tom lives at a place near where we live.

# Other vagueness introducers

- My cat weighs **about** ten pounds.
- Performance declined by **ca.** 2 %.
- There will be **approx.** 3000 invited guests.



# More numeral modifiers

- I saw **at least** 50 movies.
- I have known him for **more than** 10 years.
- The baby has been crying for **almost** 10 minutes.
- He gave me **no less than** 10 dollars.

# nearly (sentence modifier)

- She was nearly run over by a truck.
- I nearly made a mistake.
- Tom nearly got killed.
- We nearly missed the train.

# nearly (quantifier modification)

- Nearly all of us left.
- Almost everyone left.

# Universal quantifiers

- He flirts with every woman he meets.
- He runs to the station every morning.
- I buy newspapers every day.
- She plays tennis every day.
- The bus stopped in every village.

Back to RTE

# Explanation-based RTE

- A dog is swimming in the pool.
- A dog is not swimming in the pool.  
CONTRADICTION if there is exactly one dog.

# Explanation-based RTE

- John got a bicycle helmet for his birthday.
- John has a bike.

PLAUSIBLE

# Explanation-based RTE

- John bought cat food.
- John bought food for his dog.

IMPLAUSIBLE



# Explanation-based RTE

- I love Italian food.
  - Cristiano **Ronaldo** is a football player.
- DISCONNECTED

# Conclusion/Summary

- Look at corpora.
- Look at parallel corpora.
- Classical logic isn't enough for NL inferences.
- Negation is superinteresting.
- Adjectives require more work.
- We could revive RTE if we add more labels