Shenlock Holmer meets dr. Watsonson

A conversational game

Zuzana Nevěřilová

Natural Language Processing Centre, Faculty of Informatics
Masaryk University, Brno, Czech Republic

xpopelk@fi.muni.cz, xrambous@fi.muni.cz

RASLAN 2013, Karlova Studánka, 2013-12-06

The Game: Conversation

---
[1]Illustration by Paget

The Game: Inspiration

"It has always been my habit to hide none of my methods, either from my friend Watson or from any one who might take an intelligent interest in them."
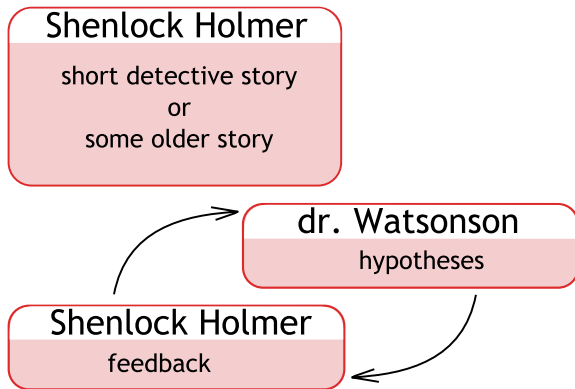
Sherlock Holmes
– The Reigate Puzzle

The Game Roles

Human: the detective (Shenlock Holmer)
Computer: the assistant (dr. Watsonson)

Game Loop

Game Screenshot

Game Elements

- points
  - new story
  - annotation
  - annotation agreement
- levels
- Watsonson's emotions

Shenlock Holmer meets dr. Watsonson

The Purpose

Definition (Textual Entailment)

A text t entails a hypothesis h if humans reading t will infer that h is most likely true. [Dagan et al., 2009]

The Purpose

Definition (Textual Entailment)

A text t entails a hypothesis h if humans reading t will infer that h is most likely true. [Dagan et al., 2009]

What do humans think about entailments $t \Rightarrow h$?

The Purpose
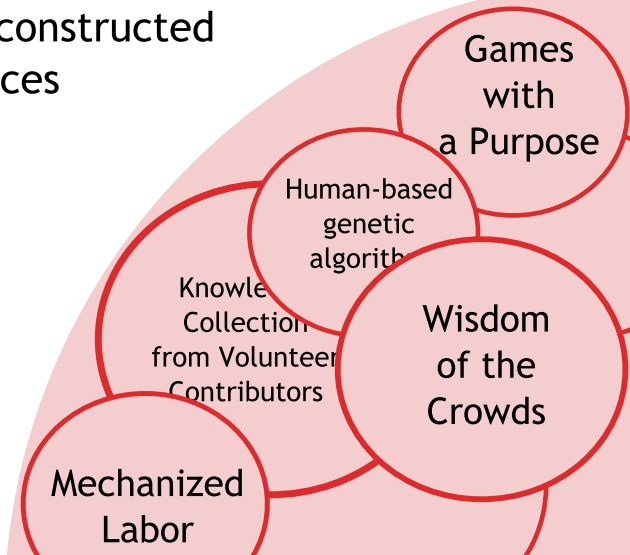
Definition (Textual Entailment)

A text t entails a hypothesis h if humans reading t will infer that h is most likely true. [Dagan et al., 2009]

What do humans think about entailments $t \Rightarrow h$?

manually annotated data

Distributed Human Computation Context

collaboratively constructed
language resources

Games
with
a Purpose

Human-based
genetic
algorithm

Knowle
Collection
from Volunteer
Contributors

Wisdom
of the
Crowds

Mechanized
Labor

Games with a Purpose: Related Work

- Common Sense Propositions [Ahn et al., 2006]
- Coreference Annotation [Chamberlain et al., 2009]
- Paraphrase Corpora Collection [Chklovski, 2005]
- Semantic Relations Collection [Vickrey et al., 2008]

The Game: Modules

- tokenization, disambiguation
- anaphora resolution and replacement by antecedent
- parsing
- phrase re-ordering
- wordnet synonym replacement
- wordnet hypernym replacement (w.r.t. sentence polarity)
- verb frame inference
- sentence generation
- sentence scoring

Interesting Modules: Verb Frame Inference

zapálit si: [c1, +person] [c4, -person]

$\Rightarrow$

být kuřák: [c1, +person]

hořet: [c1, -person]

VerbaLex + Sysel

Interesting Modules: Verb Frame Inference

zapálit si: [c1, +person] [c4, -person]

$\Rightarrow$

být kuřák: [c1, +person]

hořet: [c1, -person]

VerbaLex + Sysel

Interesting Modules: Verb Frame Inference

zapálit si: [c1, +person] [c4, -person]

$\Rightarrow$

být kuřák: [c1, +person]

hořet: [c1, -person]

VerbaLex + Sysel

Interesting Modules: Sentence Scoring

2 to 5-gram language model based on words

Na gauči se tlustá kočka válela.	0.521350
Na gauči se válela tlustá kočka.	2.139306
Na gauči válela tlustá kočka se.	0.574347
Válela se tlustá kočka na gauči.	2.748889
Tlustá kočka se válela na gauči.	4.706297

Interesting Modules: Sentence Generation

declension: noun phrases, modifier of noun phrases, adjective complements

conjugation: (in past tense) depends on the subject, multi-word verb phrases

Annotation Results

- 1635 sentences annotated
- 1298 annotated rather true
- 245 annotated rather false
- 319 annotated confused

- 14 115 positive votes
- 4 387 negative votes
- 4 458 confused votes

# Annotation Results: Sources of Negative Annotations

- WordNet: pes → obratlovec
- WordNet: pes → forma života
- WordNet: auto → skutek (auto–aut)
- WordNet: kolo → limonáda (kolo–kola)
- Saara
- syntactically incorrect sentences

Future Work... there are many ways to say the same thing

- semi-automatic generation of verb inference frame
- hedging: inference about subordinate clauses (Není pravda, že Martin žije v Praze. → Martin nežije v Praze.)
- inference about time (Včera byl Martin v Praze. 5.12. byl Martin v Praze.)
- inference about place (Včera byl Martin mimo republiku. Včera Martin nebyl v Praze.)
- "contradictions" (Petr si zapálil cigaretu. Petr byl nekuřák.)
- encyclopedic knowledge (Včera byl Martin v Praze. Včera byl Martin v hlavním městě.)

Future Work... players like fun

- annotation overdose
- reality paraphrasing (Raději bych si taky zapálil cigaretu, Holmere.)
- Majka Značková 18+
- badges?
- leaderboards?
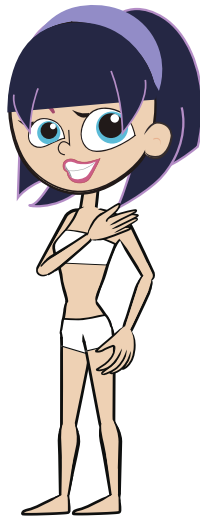
📄 Ahn, L. v., Kedia, M., and Blum, M. (2006).
Verbosity: a game for collecting common-sense facts.
In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 75–78, New York, NY, USA. ACM.

📄 Chamberlain, J., Kruschwitz, U., and Poesio, M. (2009).
Constructing an anaphorically annotated corpus with non-experts: Assessing the quality of collaborative annotations.

In *Proceedings of the 2009 Workshop on The People's Web Meets NLP: Collaboratively Constructed Semantic Resources*, People's Web '09, pages 57–62, Stroudsburg, PA, USA. Association for Computational Linguistics.

📄 Chklovski, T. (2005).
Collecting paraphrase corpora from volunteer contributors.
In *Proceedings of the 3rd International Conference on Knowledge Capture*, K-CAP '05, pages 115–120, New York, NY, USA. ACM.

📄 Dagan, I., Dolan, B., Magnini, B., and Roth, D. (2009).
Recognizing textual entailment: Rational, evaluation and approaches.
Natural Language Engineering, 15(Special Issue 04):i–xvii.

📄 Vickrey, D., Bronzan, A., Choi, W., Kumar, A., Turner-Maier, J., Wang, A., and Koller, D. (2008).
Online word games for semantic data collection.
In *EMNLP '08: Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 533–542, Morristown, NJ, USA. Association for Computational Linguistics.