

Computational Approaches to Creative Language: Summary

Caroline Sporleder

Computational Linguistics
Universität des Saarlandes

Sommersemester 2010

20.07.2010

Wrapping-Up

Creativity

The generation of ideas that are **novel** and **valuable**. Value is typically negotiated by social groups.

AI models of creativity

- combinatorial creativity
- exploration/transformation of conceptual space
 - exploratory creativity
 - transformational creativity

Detection & Processing

- idioms (type-based, token-based)
- metaphor
- neologisms
- irony, sarcasm
- humor

Generation

- metaphor
- humor
- poetry
- stories

Detection & Analysis: Techniques (1)

Idioms

Metaphor

Detection & Analysis: Techniques (1)

Idioms

- type-based
 - association measures (idiom components) (Villavicencio et al., 2007)
 - statistical measures of linguistic fixedness (Bannard, 2007)
- token-based
 - cohesion with context (Li & Sporleder, 2009)
 - vector space models (Katz & Giesbrecht, 2006)

Metaphor

Detection & Analysis: Techniques (1)

Idioms

- type-based
 - association measures (idiom components) (Villavicencio et al., 2007)
 - statistical measures of linguistic fixedness (Bannard, 2007)
- token-based
 - cohesion with context (Li & Sporleder, 2009)
 - vector space models (Katz & Giesbrecht, 2006)

Metaphor

- knowledge-based systems
 - met* (literal, metonymy, metaphor, anomaly) (Fass, 1991)
 - MIDAS (source target mapping; metaphor maps; new metaphors via analogy) (Martin, 1992)
 - slipnets (Veale & Hao, 2008)
- knowledge-lean systems for metaphor detection (similar to idiom detection) (Birke & Sarkar, 2006)

Detection & Analysis: Techniques (2)

Neologisms

Irony, Sarcasm

Humor

Neologisms

- Zeitgeist (Wikipedia / WordNet, plus analysis rules) (Veale & Butnario, 2010)
- statistical methods (Cook & Stevenson, 2010)

Irony, Sarcasm

Humor

Detection & Analysis: Techniques (2)

Neologisms

- Zeitgeist (Wikipedia / WordNet, plus analysis rules) (Veale & Butnario, 2010)
- statistical methods (Cook & Stevenson, 2010)

Irony, Sarcasm

- bag-of-words, surface cues (Burfoot & Baldwin, 2009; Carvalho et al., 2009; Kreuz & Caucci, 2007)
- pattern, n-gram based (Tsur et al., 2010)

Humor

Neologisms

- Zeitgeist (Wikipedia / WordNet, plus analysis rules) (Veale & Butnario, 2010)
- statistical methods (Cook & Stevenson, 2010)

Irony, Sarcasm

- bag-of-words, surface cues (Burfoot & Baldwin, 2009; Carvalho et al., 2009; Kreuz & Caucci, 2007)
- pattern, n-gram based (Tsur et al., 2010)

Humor

- classification based on phonological, stylistic, lexical properties (Mihalcea & Strapparava, 2005)

Generation: Techniques

Metaphor

Humor

Poetry

Stories

Generation: Techniques

Metaphor

- web harvesting plus WordNet (Sardonicus, simile generation) (Veale & Hao, 2007)

Humor

Poetry

Stories

Generation: Techniques

Metaphor

- web harvesting plus WordNet (Sardonicus, simile generation) (Veale & Hao, 2007)

Humor

- knowledge-based pun generation using templates, phonological & taxonomic knowledge (Manurung et al., 2008)

Poetry

Stories

Metaphor

- web harvesting plus WordNet (Sardonicus, simile generation) (Veale & Hao, 2007)

Humor

- knowledge-based pun generation using templates, phonological & taxonomic knowledge (Manurung et al., 2008)

Poetry

- Haiku generation using word association norms and structure templates (Netzer et al., 2009)

Stories

Metaphor

- web harvesting plus WordNet (Sardonicus, simile generation) (Veale & Hao, 2007)

Humor

- knowledge-based pun generation using templates, phonological & taxonomic knowledge (Manurung et al., 2008)

Poetry

- Haiku generation using word association norms and structure templates (Netzer et al., 2009)

Stories

- knowledge-based, e.g. using case-based reasoning (Gervás et al., 2004)
- knowledge-lean, schemas harvested from existing stories and re-combined (McIntyre & Lapata, 2010)

Techniques

- knowledge-based (KBs and semantic networks; metaphors, story generation)
- less knowledge-based (templates, taxonomies; poetry and pun generation)
- using internet resources (Wikipedia; neologisms, simile generation)
- data-driven
 - supervised ML using shallow features (detection of metaphors, irony, sarcasm, humor)
 - corpus-based, statistical, unsupervised (type- and token-based idiom detection, neologisms, story generation)

Types of Creativity involved

- combinatorial / exploratory (web corpus-based re-combination, implicit conceptual space)
 - Veale & Hao (2007) - simile generation
 - McIntyre & Lapata (2010) - story generation
- exploratory (explicitly defined conceptual space)
 - Gervás et al. (2004) - story generation
 - Manurung et al. (2008) - pun generation
 - ???Netzer et al. (2009) - Haiku generation
- transformational
 - ??? does this make sense for machines ???

What does a computer have to be able to do?

	detect?	understand?	generate?
idioms			
metaphors			
humor			
irony			
poetry			
stories			

What does a computer have to be able to do?

	detect?	understand?	generate?
idioms	definitely	definitely	possibly
metaphors			
humor			
irony			
poetry			
stories			

What does a computer have to be able to do?

	detect?	understand?	generate?
idioms	definitely	definitely	possibly
metaphors	definitely	definitely	possibly
humor			
irony			
poetry			
stories			

What does a computer have to be able to do?

	detect?	understand?	generate?
idioms	definitely	definitely	possibly
metaphors	definitely	definitely	possibly
humor	probably	possibly	possibly
irony			
poetry			
stories			

What does a computer have to be able to do?

	detect?	understand?	generate?
idioms	definitely	definitely	possibly
metaphors	definitely	definitely	possibly
humor	probably	possibly	possibly
irony	definitely	definitely	possibly
poetry			
stories			

What does a computer have to be able to do?

	detect?	understand?	generate?
idioms	definitely	definitely	possibly
metaphors	definitely	definitely	possibly
humor	probably	possibly	possibly
irony	definitely	definitely	possibly
poetry	possibly	???	possibly
stories			

What does a computer have to be able to do?

	detect?	understand?	generate?
idioms	definitely	definitely	possibly
metaphors	definitely	definitely	possibly
humor	probably	possibly	possibly
irony	definitely	definitely	possibly
poetry	possibly	???	possibly
stories	—	possibly	possibly

What does a computer have to be able to do?

	detect?	understand?	generate?
idioms	definitely	definitely	possibly
metaphors	definitely	definitely	possibly
humor	probably	possibly	possibly
irony	definitely	definitely	possibly
poetry	possibly	???	possibly
stories	—	possibly	possibly

Challenge: development of relatively knowledge-lean methods for these tasks!

- Ekaterina Shutova. “Models of Metaphor in NLP”. *ACL 2010*
- Beata Beigman Klebanov, Eyal Beigman. “A Game-Theoretic Model of Metaphorical Bargaining”. *ACL 2010*
- Neil McIntyre, Mirella Lapata. “Plot Induction and Evolutionary Search for Story Generation”. *ACL 2010*

Term Paper

Writing term papers

- answer a concrete research question (or a set of related questions)
- show evidence of independent thinking and reasoning
- independent bibliographic work
- meet academic standards (proper citations/references, well-structured, well-written etc.)

Don't plagiarise!

- can include some practical work but doesn't have to
- around 100 hours of work (2-3 weeks full-time)
- 15-20 pages long
- 10-15 references

Please discuss topic with me beforehand!