

# Computational Semantics: an Introduction

Katrin Erk

# Computational linguistics is about...

- Automatically processing human language
    - **Linguistics:**  
Using computational models to gain a better understanding of how language works
    - **Engineering:**  
Building language technology
- We look at natural language meaning from both perspectives.

# Computational semantics

- Automatically analyzing the meaning of natural language
- How can you describe the meaning of a word?
- How can you tell if someone has understood the meaning of a sentence?

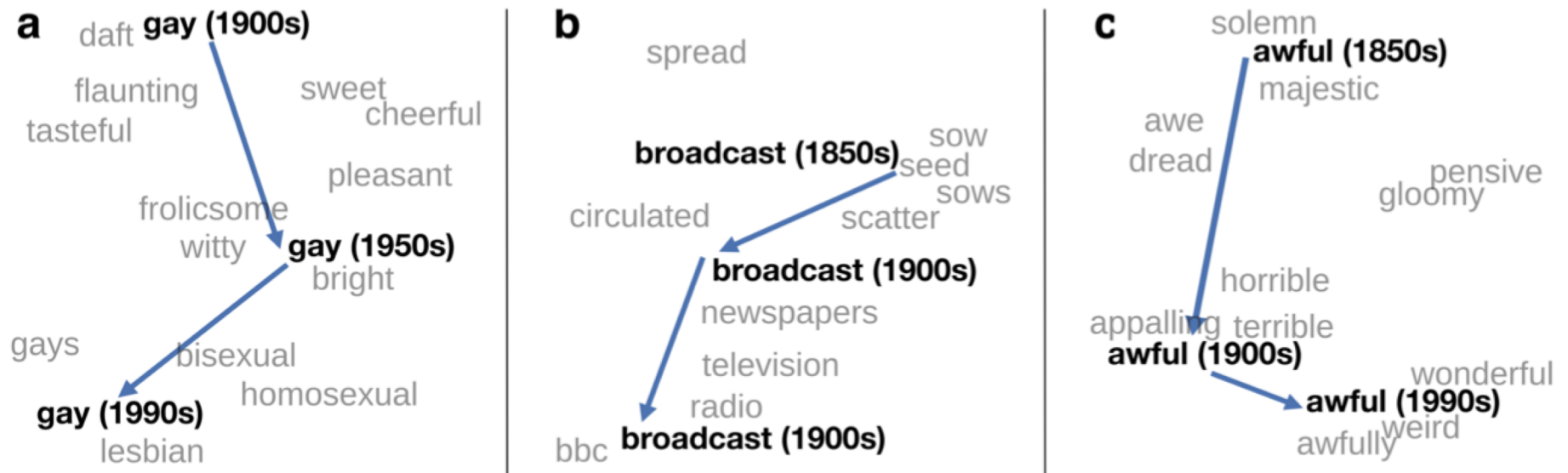
# Computational semantics

- Where can we use a system for automatic meaning analysis?

# Automatically building a thesaurus

- **Merriam-Webster:**
  - **Synonyms** [adversary](#), [antagonist](#), [foe](#), [hostile](#), [opponent](#)
  - **Related Words** [archenemy](#), [archfoe](#), [nemesis](#); [ill-wisher](#); [bane](#), [bête noire](#); [assailant](#), [attacker](#), [combatant](#), [invader](#); [competitor](#), [emulator](#), [rival](#)
  - **Near Antonyms** [buddy](#), [chum](#), [compadre](#), [crony](#)...
- **Automatically generated thesaurus:**
  - *adversary*: enemy, foe, ally, antagonist, opponent, rival, detractor, neighbor, supporter, competitor, partner, trading partner, accuser, terrorist, critic, Republican, advocate, skeptic, challenger

# Figuring out how word meanings changed over time



Hamilton, Leskovec, Jurafsky: Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change  
ACL 2016

# How many senses does this word have, and which are they?

## “Lie” clusters

### Mathematical sense - verb

- A skew polygon does not **lie** in a flat plan, but zigzags in three (or more) dimensions
- As an open string propagates through spacetime, its endpoints are required to **lie** on a D-brane..

### Untruth - verb

- Take for example the declaration "I will **lie** for personal benefit."
- Rob reveals to Tracy that everything was a **lie** and that he still hated her.

### Lie down - verb

- There Fenrir will **lie** until Ragnarok.
- They **lie** down to sleep deeply

### Geographical (island) - verb

- Some 3,579 islands **lie** adjacent to the peninsula.
- The islands **lie** on the Kerguelen Plateau in the Indian Ocean.

### Conceptual placement - verb

- According to Dewey, conversation, debate and dialogue **lie** at the heart of a democracy
- The origins of mathematical thought **lie** in the concepts of number, magnitude and form.

### Geographical (other) - verb

- Very small portions **lie** within the Pueblo County School District 70.
- The ruins of the town **lie** along the river Ziz in the Tafilalt oasis near the town of Rissani.

Coenen et al: Visualizing  
and Measuring the  
Geometry of BERT,  
arXiv  
<https://arxiv.org/abs/1906.02715>

# How many senses does this word have, and which are they?

## “Fair” clusters

### legal sense - adjective

- Using most or all of a work does not bar a finding of **fair use**.
- Examples of such doctrines are the **fair use** and fair dealing doctrine.

### gathering - noun

- In 2000, Hanover hosted the **world fair** expo 2000.
- Around 2 million people visit this **fun fair** every year

### mathematical sense - adjective

- For example, the entropy of a **fair coin** toss is 1 bit, and the entropy of tosses is its bits
- The gambler's fallacy can be illustrated by considering the repeated toss of a **fair coin**

### just and equitable - adjective

- Is all **fair** in biological warfare?
- In 1994 and again in 2002, they won the Hodgson Trophy for **fair play**.

### amount - adjective

- There is a **fair number** of bright stars, both single and double, in Lepus
- The rivalry has had its **fair share** of fights as well

Coenen et al: Visualizing  
and Measuring the  
Geometry of BERT,  
arXiv

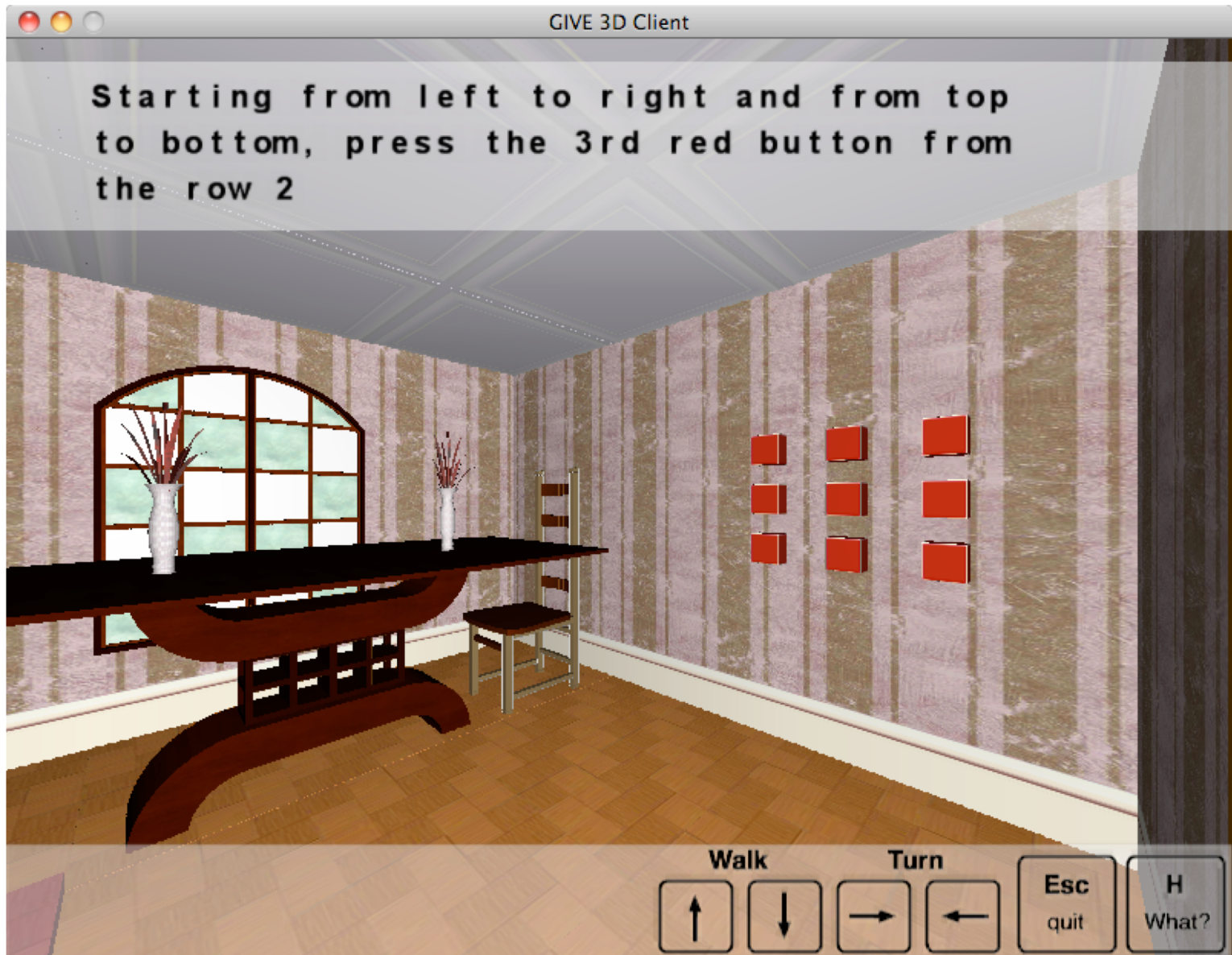
<https://arxiv.org/abs/1906.02715>



# Information Extraction

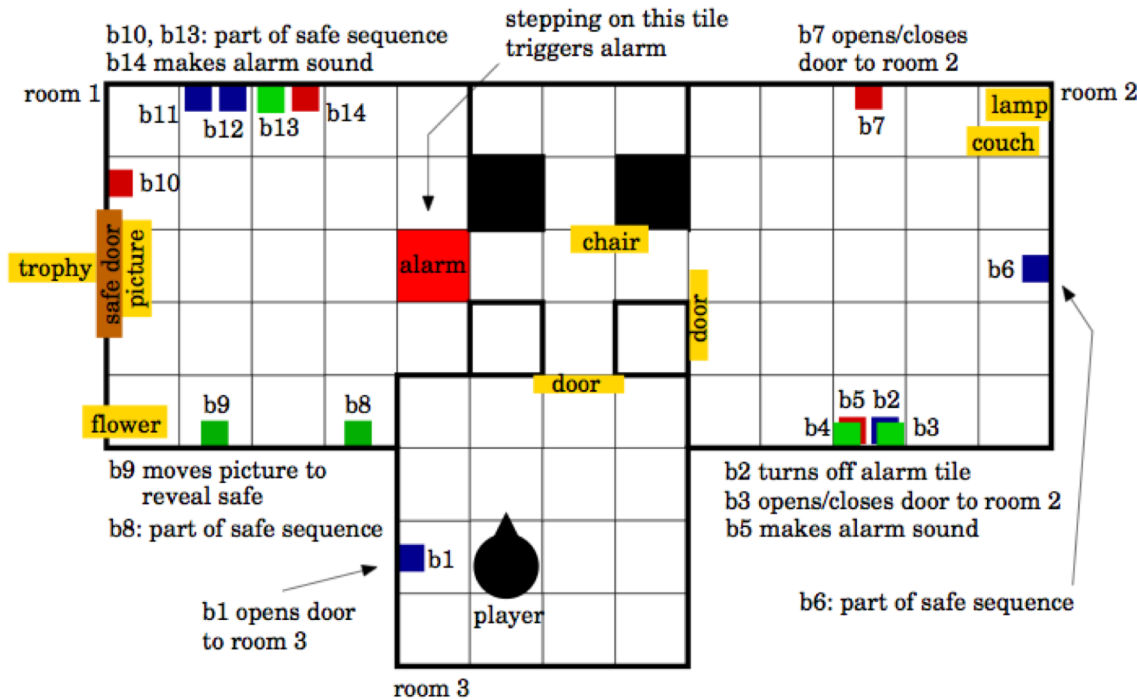
- Automatically extracting “structured data” from “unstructured text”<sup>\*</sup>
  - “[Citrix](#) announced [today](#) that it has completed its [acquisition](#) of [Sanbolic](#), an innovator and leader in workload-oriented storage virtualization technologies.[...] [Published Monday, January 12, 2015 9:54 AM](#) by David Marshall.
  - Merger(Citrix, Sanbolic, 2015-01-12)
- Need to figure out:
  - “acquisition” can mean merger, if it’s between companies
  - Citrix, Sanbolic are companies
  - link “today” to last line of text
  - that “Citrix”, “Sanbolic” are the two participants in the “acquisition”
- <sup>\*</sup>: What counts as “structured” depends on one’s perspective. Linguists usually laugh heartily at the idea that ordinary text is unstructured.

# Natural Language Generation



GIVE challenge, image from <http://www.give-challenge.org/research/>

# Natural language generation



Formal representation of the world



Natural language

- to win you have to retrieve the trophy from the safe in room 1
- use button b9 to move the picture (and get access to the safe)
- press buttons b8, b6, b13, b13, b10 (in this order) to open the safe; if a button is pressed in the wrong order, the whole sequence is reset
- if the alarm sounds, the game is over and you have lost



# Natural language generation

- We can define translations for pieces of the formal representation:
  - `badPass(P1, P2) =>`  
[translation(P1) makes a bad pass that was intercepted by [translation(P2)]]

# Another task: database queries

- “Which country is Athens in?”
- SQL query:  
`SELECT country FROM city_table WHERE city=“athens”`
- Here is the database city\_table:

<b>City</b>	<b>Country</b>	<b>Population</b>
athens	greece	1368
bangkok	thailand	1178
barcelona	spain	1280
berlin	east_germany	3481
birmingham	united_kingdom	1112

- More practically: dialog agents for flight planning/ train schedules

# Another task: database queries

- “Which country is Athens in?”
- SQL query:  
SELECT country FROM city\_table WHERE  
city=“athens”
- We can define translations for pieces of the sentence, and put them together:
  - “Which country is ... in” =>  
SELECT country FROM city\_table WHERE city=“...”
  - “Where is ... located” =>  
SELECT country FROM city\_table WHERE city=“...”

# Another task:

## Natural language inference

- After reading sentence T, would a person conclude that sentence H is most likely also true?

T: Phish disbands after a final concert in Vermont on Aug. 15

H: Rock band Phish holds final concert in Vermont.

T: Crude oil for April delivery traded at \$37.80 a barrel, down 28 cents

H: Crude oil prices rose to \$37.80 per barrel



# Natural language inference

- One way of doing this: with inference rules at the level of words and phrases

T: Phish disbands after a final concert in Vermont on Aug. 15

H: Rock band Phish holds final concert in Vermont.

– X disbands after [concert]  $\Rightarrow$  X holds [concert]

– Phish = Rock band Phish (in the context of “hold concert”)

# Natural language inference

- Approximate matching:
  - T: Crude oil for April delivery traded at \$37.80 a barrel, down 28 cents
  - H: Crude oil prices rose to \$37.80 per barrel
    - X trades at..., down = X prices fell
    - Rise  $\Leftrightarrow$  fall
    - A barrel = per barrel (in the context of \$XX)

# Natural language inference

T: Crude oil for April delivery traded at \$37.80 a barrel, down 28 cents

H: Crude oil prices rose to \$37.80 per barrel

- A more recent way of doing this:
- The computational representations of the sentences do not have to be readable by us.
- But they need to help the machine solve the problem.
- Iteratively adapt representations of the sentences to best work on the task

# Plan for the semester

- Part 1: Learning word meanings automatically from textual context
  - How to find & use pre-computed learned word meanings
  - How to do this yourself
    - Including a very short introduction to neural machine learning methods
  - What do these representations tell us about word contexts? What do they tell us about meaning representations in our heads?

# Plan for the semester

- Part 2: Representing sentence meaning using logic
  - Why logic for representing sentence meaning?
  - Short introduction to propositional first-order logic
  - Why can you figure out the meaning of a sentence you have never seen before, like, say, “My iguana is on fire”?
    - Putting together logic-based meaning representations piece by piece

# Plan for the semester

- Part 3: Structured meaning representations in applications
  - Logic-based representations have inspired tons of practical applications:
    - The “semantic web”
    - Knowledge graphs
    - Information extraction
  - We look at mechanisms and applications