RI MasterClass: Natural Language Processing Applications

Mehrnoosh Sadrzadeh

Julian Hough

Chris Madge

Julian Hough (Lecturer, QMUL)

- Interested in improving interaction with robots and computers through speech and gesture!

- Field: Natural Language Processing
  - Dialogue (conversation) systems and dialogue analysis
 - human-robot dialogue.

OUTLINE

1) What is NLP and where is it used?

2) Dialogue Systems

What is Natural Language Processing?

- **Natural Language Processing (NLP)** (sometimes called ***Computational Linguistics***) is the automatic processing of human language for some task or purpose.

What is Natural Language Processing?

- **<u>BIG PICTURE 1:</u>** We really want to build machines that **understand** human language in a human way, and **produce/generate** human language in a human way.

- Alan Turing (1950) originally posed the **Turing Test** as being key to solving artificial intelligence.

- Can you 'fool' someone they're talking to a human, when it's a machine? That has solved **Artificial Intelligence.**
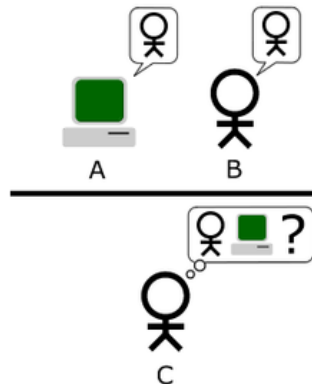
What is Natural Language Processing?

- **BIG PICTURE 2:** We want tools that allow us to do tasks more effectively.

- This technology might assist you with **organizing** huge amounts of text information, accessing parts of it, and extracting data from it.

- It can help you **create** your own text data: e.g. spelling and style correction.

- It can **help** those who need it: text-to-speech from screens for the blind; speech-to-text for those with manual problems.

Levels of analysis (small to large)

- Phonemes/sounds (Speech recognition, prosody)
- **Words (can be broken down into morphemes)**
- **Phrases**
- **Sentences/Turns**
- **Texts/Dialogues**

- At QM we focus approximately the level of the word upwards as an increment of analysis.

Why is NLP difficult/interesting? Because human language is…

- **Ambiguous (can mean several things at once)** (unlike programming languages)

- **Not always explicit and depends on context**. You leave out "code"- the listener/reader fills in the gaps!
  - **Context** includes real-world knowledge. Do words 'mean' anything without reference to real things/situations?

- **Rich** in its ability to express lots of things.

- **Creative**- you can always create a new word/phrase!

Applications: main areas

- Machine Translation (since the 1950s)

- Search (Google)

- **Managing BIG data**:
  - Analysing social media for advertising e.g. **sentiment analysis** for products.
  - Finance: buy/sell decisions based on social media texts. Health: Which hospitals are good?

- **Dialogue systems/Chatbots**:
  - Personal assistants (Amazon's Alexa, Apple's Siri).
  - Automating customer service.

Sentiment Analysis

POSITIVE about IPhone ☺

1. Id: Abc123 on 5-1-2008 "I bought an iPhone a few days ago. It is such a nice phone. The touch screen is really cool. The voice quality is clear too.

2. It is much better than my old Blackberry, which was a terrible phone and so difficult to type with its tiny keys. However, my mother was mad with me as I did not tell her before I bought the phone. She also thought the phone was too expensive, …"

Sentiment Analysis

POSITIVE about IPhone ☺

> 1. Id: Abc123 on 5-1-2008 "I bought an iPhone a few days ago. It is such a nice phone. The touch screen is really cool. The voice quality is clear too.

> 2. It is much better than my old Blackberry, which was a terrible phone and so difficult to type with its tiny keys. However, my mother was mad with me as I did not tell her before I bought the phone. She also thought the phone was too expensive, …"

POSITIVE about IPhone ☺   **NEGATIVE about Blackberry** ☹

Sentiment Analysis

- A typical NLP task
- You have a large amount of data available to you (a **corpus**). E.g. collection of tweets or comments.
- You need to build something to make the automatic decision:
  - **Positive ☺ vs Negative ☹**
    - *I'm really happy!*
    - *I'm having a terrible day*
    - *Oh man this is so great <3*
    - *I just can't believe it*
- How could we go about this?

Sentiment analysis 1: Dictionaries

- We could build dictionaries:
  - List of "positive" words
  - List of "negative" words

- Problem with ambiguity- is this positive or negative?:

  `i love @justinbieber #sarcasm`

- We might need a more data-driven approach…

Sentiment Analysis 2: Data-Driven Classification

- We could **learn** the dictionaries of 'positive' and 'negative' words from:
    - List of "positive" examples
    - List of "negative" examples

- Learn a classifier based on observed words … and combinations thereof

- We can use maths: **statistics** and **geometry**
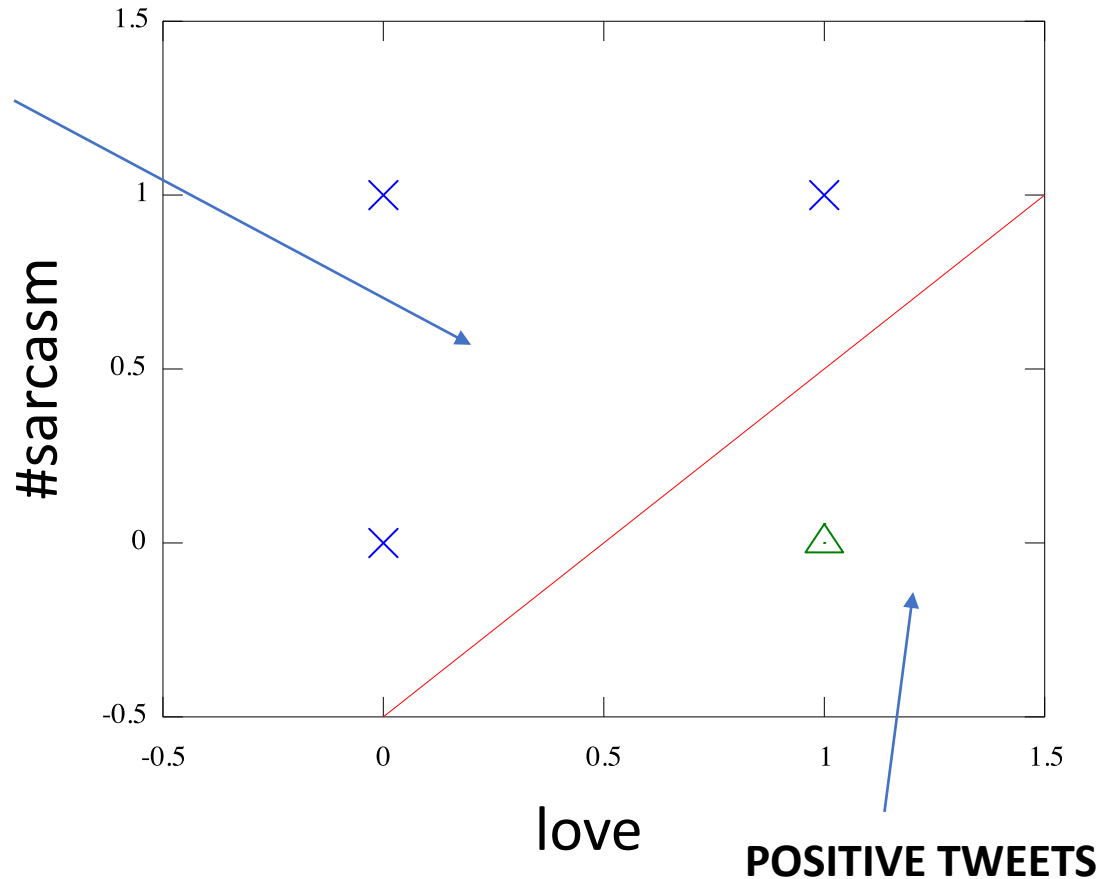
Sentiment Analysis 2: Data-Driven Classification

- Use Geometry for binary classification using **Machine Learning**
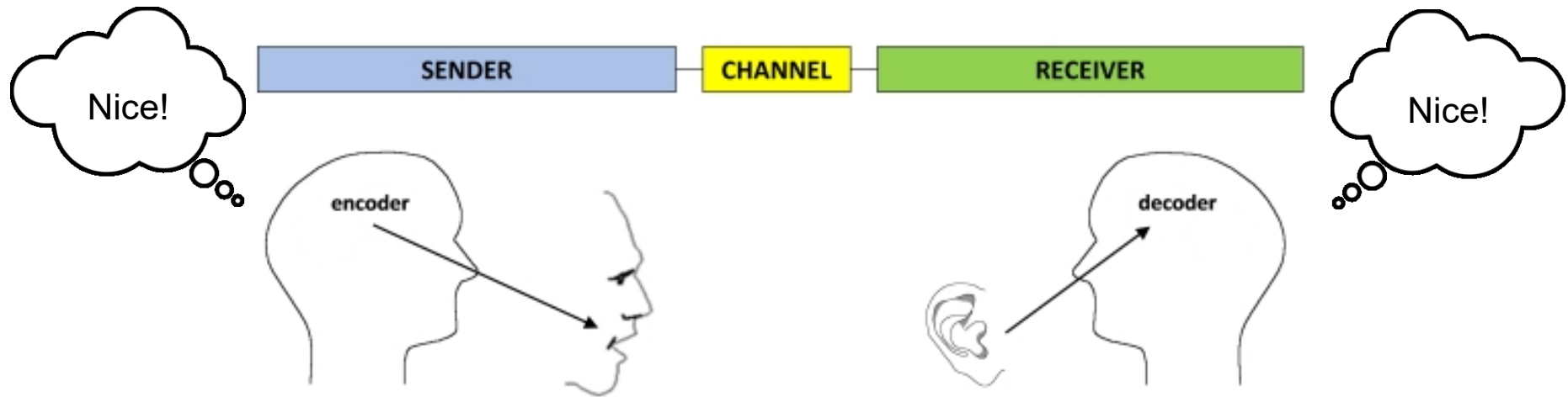
  - **i love @justinbieber #sarcasm**
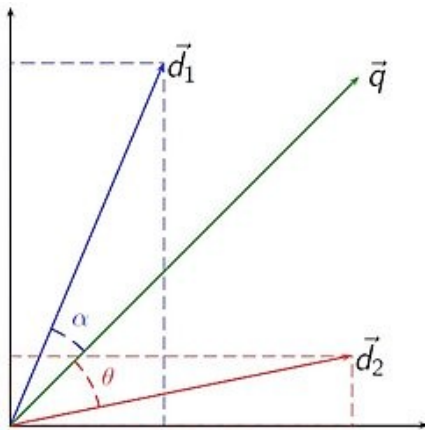
OUTLINE

1) What is NLP and where is it used?

2) Dialogue Systems

How do people communicate?

- First models similar to encoder/decoder model.

- Communication based on a common code.

How can people *mis*communicate?

- Just noise in signal? More recent theories about aligning internal representations via **communicative grounding** mechanisms.

- A. 'Put the apple over there'

  – B. 'Where did you mean?' (clarification)
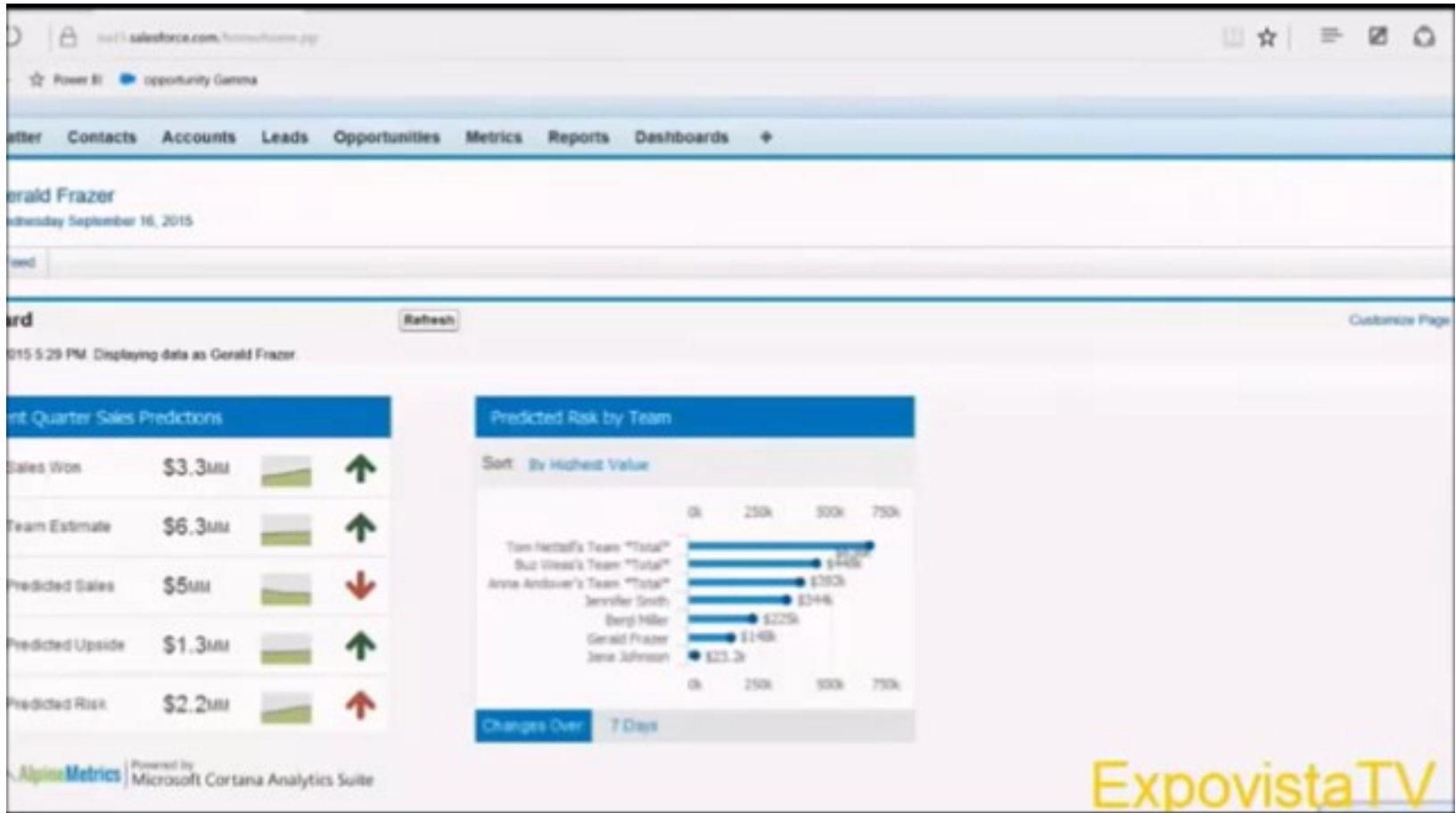
  – A. 'No, in the corner' (repair)
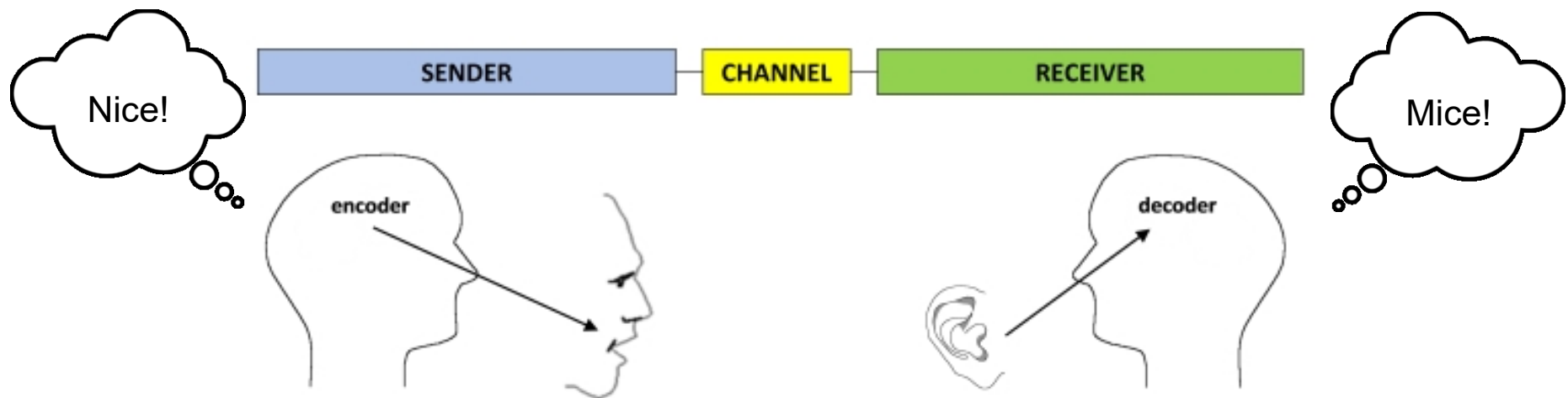
Robot (mis)communication

- What about chatbots and robots?

- Inherently different internal representations.

- Inherent uncertainty about common ground.

- Made harder if they don't give feedback through speech.
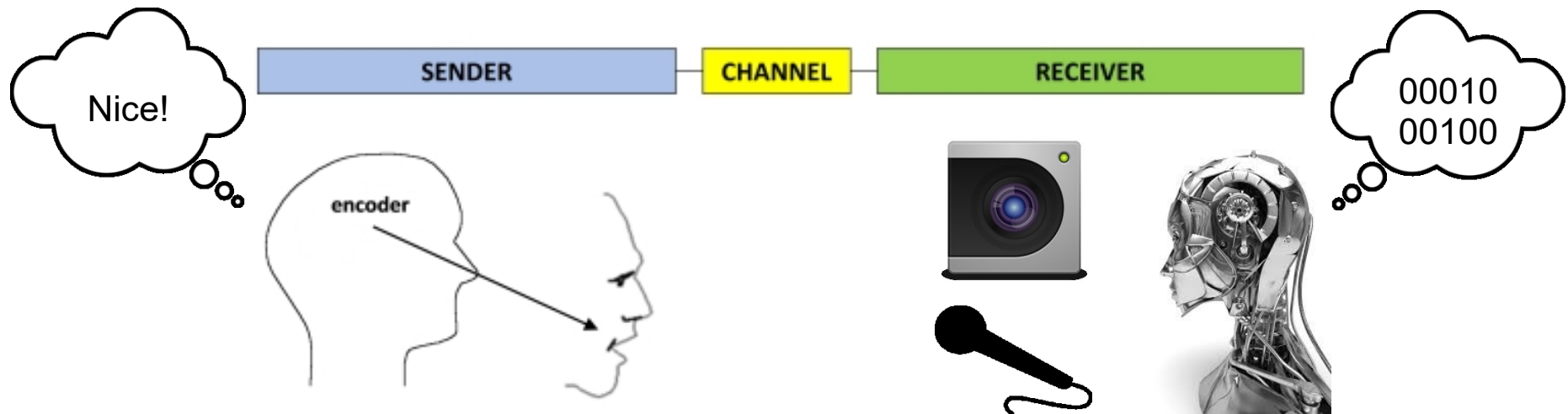
And hard for systems…

How do we build systems to speak with humans?

- Dialogue system designers struggle to deal with the rich range of human dialogue behaviour and what people **mean** in their utterances/texts.

- However, many useful systems use simple assumptions to get things working.

- https://www.cleverbot.com/

- You can even get things 'working' with robots (my own work with simple robots). I work on **avoiding delays** in robots.

Your dialogue task

- Google Dialogflow uses breaks things down to **intents** and **entities** and context variables.
- An intent is the recognized meaning of the user's intention e.g. I want a pizza -> *#orderfood*
- An entity is an individuated thing e.g. I want a pizza -> *entity:food=pizza*
- You will build a simple Google Dialogflow chatbot.
- http://tiny.cc/RIMasterClassNLP