# Facing NLP



German Rigau i Claramunt

http://adimen.si.ehu.es/~rigau

IXA group

Departamento de Lenguajes y Sistemas Informáticos

UPV/EHU

# Facing NLP

- From Cyc (adapted) (I)

  - Fred saw the plane flying over Zurich.

# Facing NLP

- From Cyc (adapted) (2)

  - Fred saw the train flying over Zurich.

# Facing NLP

- From Cyc (adapted) (3)

    - Fred saw the plane flying over Zurich.

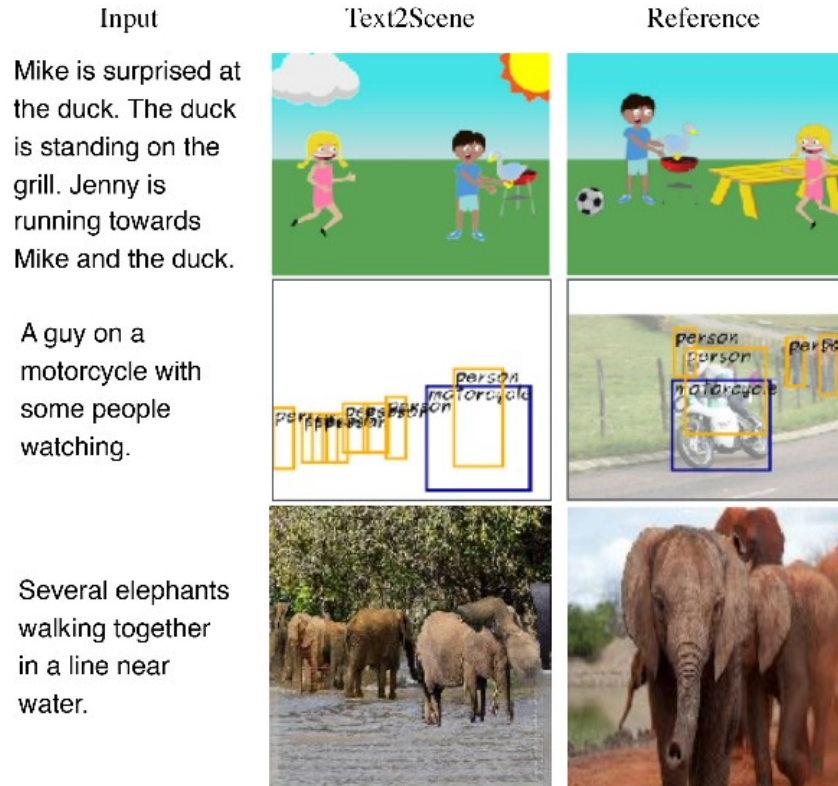    - Fred saw the train flying over Zurich.

# Text2Scene



Figure 1. Sample inputs (left) and outputs of our Text2Scene model (middle), along with *ground truth* reference scenes (right) for generation of abstract scenes (top), object layouts (middle), and synthetic image composites (bottom).

**Text2Scene: Generating Abstract Scenes from Textual Descriptions.(2019)**

Fuwen Tan, Song Feng, Vicente Ordonez

AI and NLP

# Facing NLP

- Don't think about a pink elephant!

# Setting

- From Winograd Schema Challenge (I):

  - The trophy would not fit in the brown suitcase because it was too big (small). What was too big (small)?
  - Answer 0: the trophy
  - Answer 1: the suitcase

# Setting

- From Winograd Schema Challenge (II):

  - The bee landed on the flower because **it** had pollen.
  - The bee landed on the flower because **it** wanted pollen.

# Setting

- Dificulty of NLP
- Levels of NLP processing
- Research areas related to NLP
- Setting
- Outline of the Seminar

# Difficulty of NLP

- Language is <u>dinamic!</u>

- More than 5000 languages!

- … and ~6000 millions of people!

- <u>Complexity</u>: several and complex levels of processing

- <u>Ambiguity</u>!

- Incomplete knowledge, fuzy, …

- Requires <u>World Knowledge</u>!

- Within a <u>social interaction</u> system!

Ontologies & large-scale KBs for NLP
# Levels of NLP processing (1)

- <u>Phonetic</u>: relating sounds with words

- <u>Morphologic</u>: building words: puño, empuñar, …

- <u>Syntactic</u>: building sentences with words and the role they play:

  - E.on will buy Endesa / Endesa will be acquired by por E.on

- <u>Semantic</u>: denoting meaning from words and sentences

  - *Zapatos de piel de señora*

  - Lady leather shoes

- <u>Pragmatic</u>: … in a context

  - Me dás hora? Tienes hora? … in the street / in the dentist

# **Levels of NLP processing (2)**

- <u>Discourse</u>:

  - *Él le dijo después que lo pusiera encima.*

- <u>World knowledge</u>: how to manage (and acquire)

  - Lucy in the sky with diamonds

  - Clever & Smart

  - GM drives to make Saturn a star again

  - *They are to see you better- said the wolf imitating the grandmother's voice.*

- <u>Generation</u>: how to generate correct text/sounds

  - 16/02/2007 => *dieciseis de febrero del dos mil siete*

# Levels of NLP processing (3)

Different types of ambiguity:

- Lexical ambiguity
- Sintactic ambiguity
- Semantic ambiguity
- Reference

# Levels of NLP processing (4)

Lexical ambiguity (examples):

- *Mi amigo Juan Mesa se mesa la barba al lado de la mesa.*
- *El cura recibió una cura completa.*

- From Financial Times

  - US officials has expected Basra <u>to fall</u> early
  - Music sales <u>will fall</u> by up to 15% this year
  - No missiles <u>have fallen</u> and …

# Levels of NLP processing (5)

Sense 10

fall -- (be captured; "The cities fell to the enemy")

=> yield -- (cease opposition; stop fighting)

Sense 2

descend, fall, go down, come down -- (move downward but not necessarily all the way; "The temperature is going down"; "The barometer is falling"; "Real estate prices are coming down")

=> travel, go, move, locomote -- (change location; …)

Sense 1

fall -- (descend in free fall under the influence of gravity; "The branch fell from the tree"; "The unfortunate hiker fell into a crevasse")

=> travel, go, move, locomote -- (change location; …)

# Levels of NLP processing (6)

Sintactic ambiguity (examples):

- *La vendedora de periódicos del barrio.*

- *El policia observó al sospechoso con unos prismáticos.*

Different meanings depending on parsing!

# **Levels of NLP processing (6)**

Semantic ambiguity (examples):

- *Para el cumpleaños les daré un pastel a los niños*
  - One for all? One to one?

Reference ambiguity (examples):

- *Él le dijo después que lo pusiera encima.*
  - Who? To whom? After what? What? Where?

**Levels of NLP processing (7)**

Semantic:

- John is sick. He has the flu.

Pragmatic:

- John cannot come. He has the flu.

# Levels of NLP processing (7)

Exercice:

- John was hungry.
- He opened the refrigerator.

# Levels of NLP processing (6)

Multidisciplinar research area:

- **Linguistics**: Study of language
- **Psycholinguistics**: how people communicate.
- **Computer Science**: computer models (algorithms) for NLP
- **Philosophy**: semantics, meaning, understanding
- **Logic**: formal reasoning mechanisms
- **Artificial Intelligence**: techniques, knowledge representation, commonsense, etc.
- **Statistics**: probabilistic models of language.
- **Machine Learning**: learning rules and models
- **Deep Learning**: learning large neural language models
- **Linguistic Engineering**: implementation of large and complex NLP systems

# Setting

- From NLP to NLU
- Large-scale Semantic Processing dealing with concepts (senses) rather than words
- Two complementary problems:

- Acquisition bottleneck
  - Autonomous large-scale knowledge acquisition systems
- Ambiguity
  - Highly accurate and robust semantic systems

# Setting

- This course focuses on:
  - the semantic components used NLP applications:
    - ontologies and
    - large-scale knowledge-bases.
  - automatic acquisition of lexical resources from textual corpora.
  - methods for reasoning about the implicitly/explicitly knowledge represented into the large-scale knowledge bases.

Ontologies & large-scale KBs for NLP
# Outline

- Introduction
- Words & Works
- Ontologies:
  - Mikrokosmos
  - SUMO ontology
- Large-scale Knowledge Bases:
  - WordNet & EuroWordNet
  - ThoughtTreasure, ConceptNet, MindNet, …
  - Framenet, VerbNet, PropBank, …
- Building Wordnets
- WordNet extensions:
  - eXtended WordNet, Meaning project, Omega, …
- Reasoning

# Facing NLP



German Rigau i Claramunt
http://adimen.si.ehu.es/~rigau
IXA group
Departamento de Lenguajes y Sistemas Informáticos
UPV/EHU