Using Semantic Models for Robust Natural Language Human-Robot Interaction
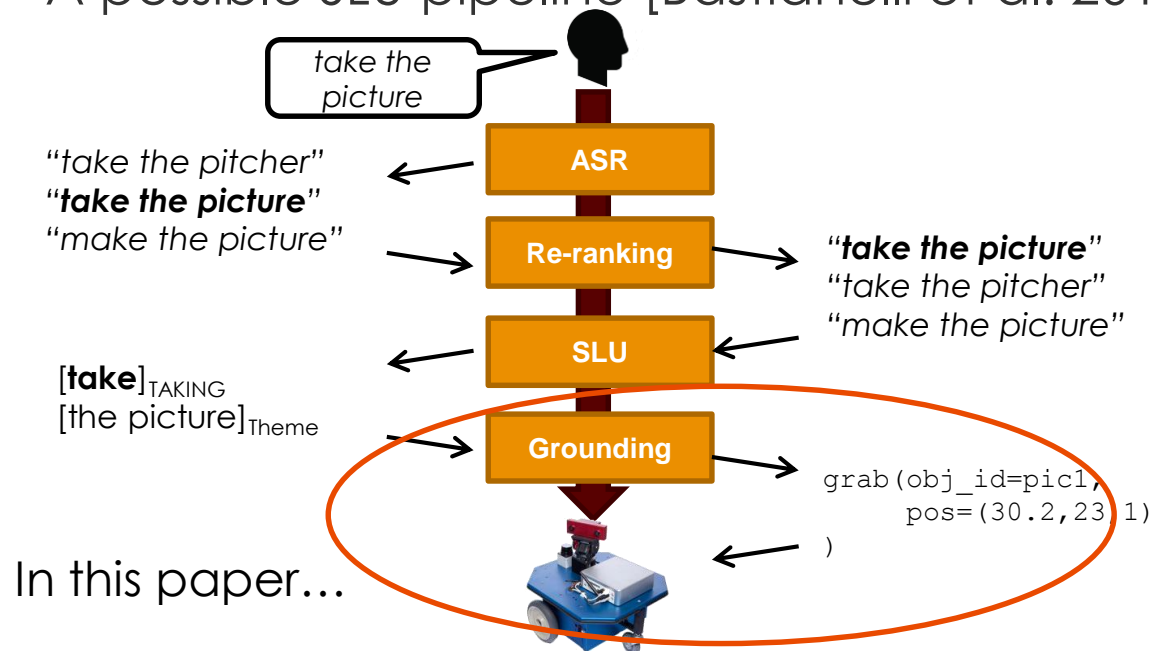
Emanuele Bastianelli[1], **Danilo Croce[2]**, Roberto Basili[2] and Daniele Nardi[3]

([1]) DICII, ([2]) DII. University of Rome Tor Vergata,
([3]) DIIAG, Sapienza University of Rome

AI*IA 2015, Ferrara 24th September 2015

Reference Scenario

- Natural Language Interfaces for Human Robot Interaction
  - Spoken Language Understanding and Robotics

- A possible SLU pipeline [Bastianelli et al. 2014]

take the picture

"*take the pitcher*"
"***take the picture***"
"*make the picture*"

ASR

Re-ranking

"***take the picture***"
"*take the pitcher*"
"*make the picture*"

SLU

[**take**]$_{TAKING}$
[the picture]$_{Theme}$

Grounding

```
grab(obj_id=pic1,
     pos=(30.2,23,1)
)
```

In this paper…

Grounding with Semantic Maps

■ Semantic Maps and Lexical References

```
objType(pic1,picture)
objPosition(pic1,5,7,180)
objName(pic1,'picture')
```


■ Language Grounding

How to deal with all possible variants?
How to preserve robus and accurate gounding w.r.t. linguistic variability?

take the picture

take the photograph

Objective

- A Lexicalized grounding-enabling function
  - Robust to mis-transcription
  - Robust to lexical variability (e.g. synonymy)
  - … or both…

*"take the **picture**"*

*"take the **pitcher**"*

*"take the **photograph**"*

*"take the **phonograph**"*

```
objType(pic1,picture)
objPosition(pic1,5,7,180)
objName(pic1,'picture')
```
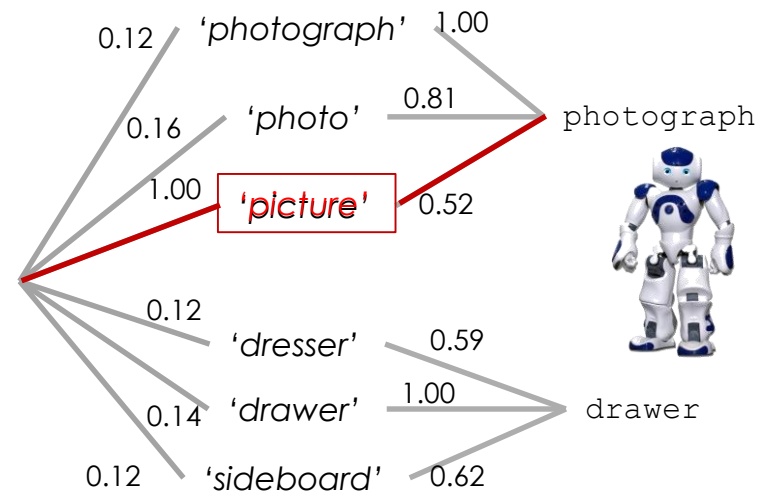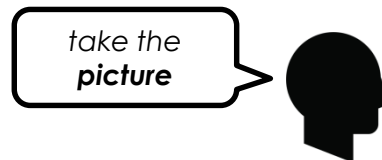
A lexicalized grounding enabling function

- Given a noun phrase w_c in a command, we evaluate a score between w_c and an entity e in the semantic map

$$g(w_c, e) = \max_{w \in S_\tau^e}(ph(w_c, w) * sim(w_e, w))$$

- $sim(\cdot, \cdot)$ is a semantic similarity function
- with $ph(\cdot, \cdot)$ phonetic similarity function
- specific selection policies are applied

take the **picture**

0.12   *'photograph'*   1.00

0.16   *'photo'*   0.81   photograph

1.00   *'picture'*   0.52

0.12   *'dresser'*   0.59

0.14   *'drawer'*   1.00   drawer

0.12   *'sideboard'*   0.62

Experimental Findings

- Different combinations of **phonetic distances** and **selection policies** have been evaluated

- The experimental evaluation considered a brand new corpus
  - `HuRIC`: a **Hu**man **R**obot **I**nteraction **C**orpus
  - `http://sag.art.uniroma2.it/demo-software/huric/`

- It work!
  - As the sentence complexity grows, our function performs better than the simple identity function