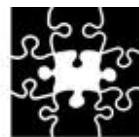


Visually Grounded, Task-oriented Dialogue

Elia Bruni



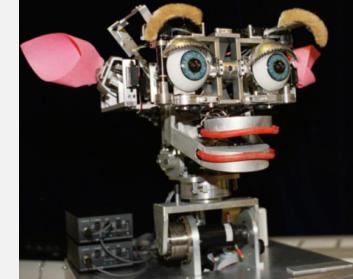
INSTITUTE FOR LOGIC,
LANGUAGE AND COMPUTATION



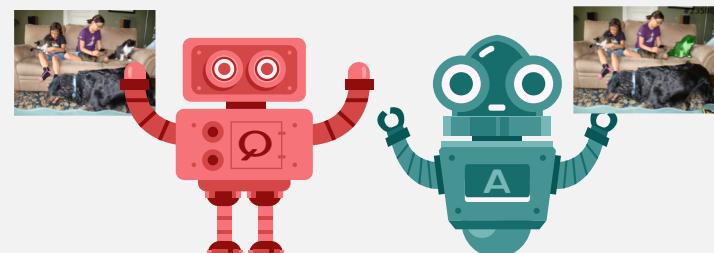
UNIVERSITY OF AMSTERDAM

Outline

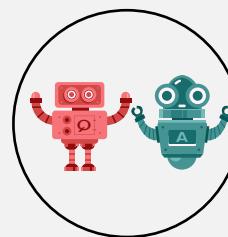
Language grounding



Visual dialogue



Appendix: Current and future work

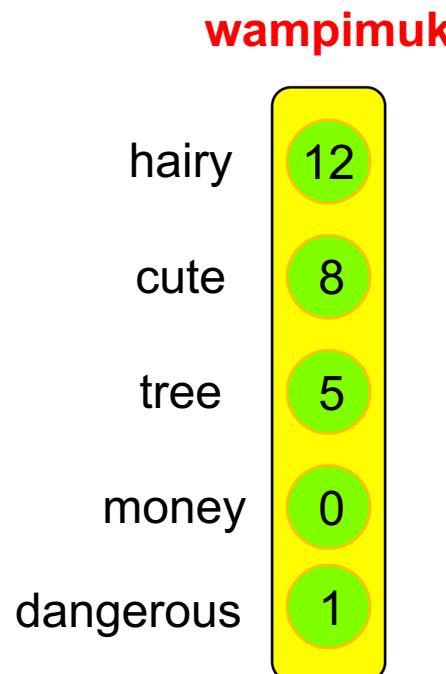


Distributional semantics

We found a cute, hairy **wampimuk** sleeping behind
the tree

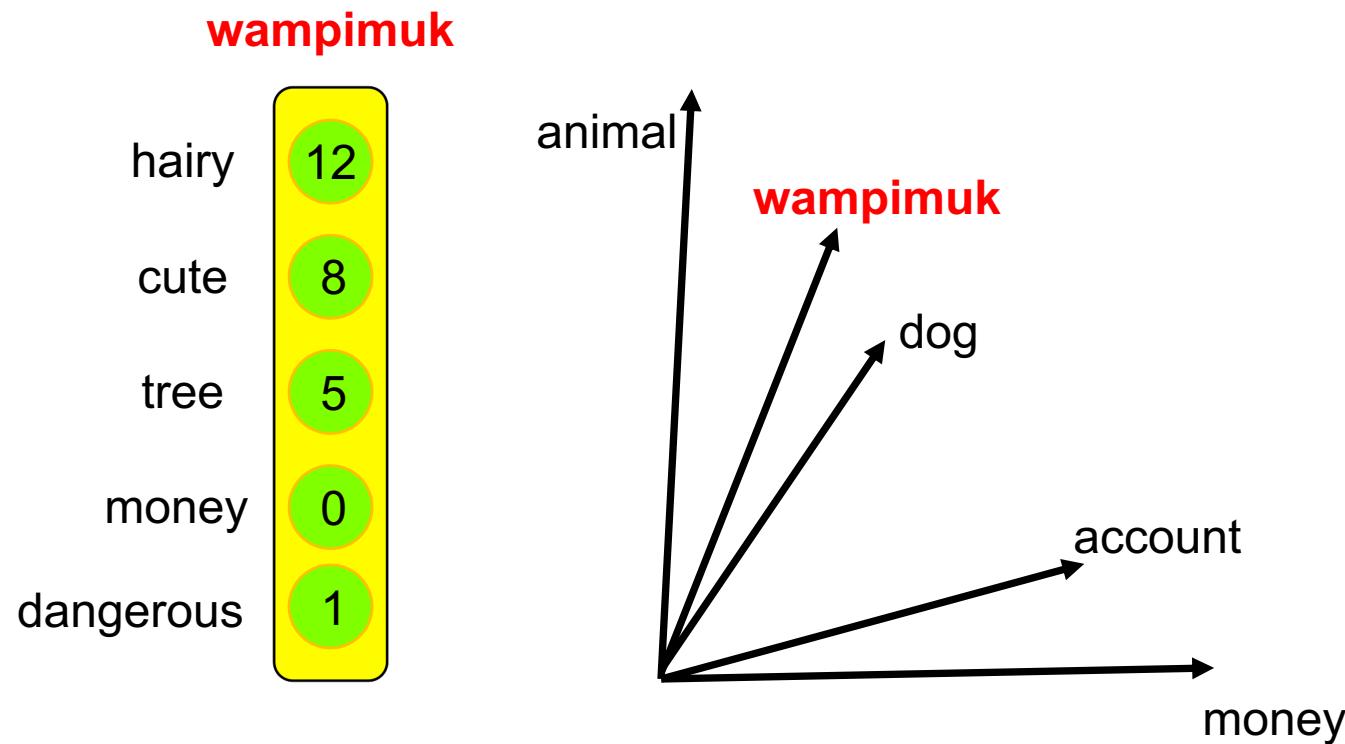
Distributional semantics

We found a cute, hairy **wampimuk** sleeping behind the tree



Distributional semantics

We found a cute, hairy **wampimuk** sleeping behind the tree



The grounding problem

Bruni et al., Distributional Semantics in Technicolor. 2012

- **clover** is blue
- **coffee** is green
- **crows** are white
- **deers** are yellow
- **flour** is black
- **grass** is purple
- the **sky** is green
- **violins** are blue



<http://stockup.net/2014/cute-deer-silhouette.html> | <http://www.wildlife-printables.com/Croco-Silhouettes-Templates.html>

Language grounding

Bruni et al., Multimodal Distributional Semantics. 2014

Language grounding

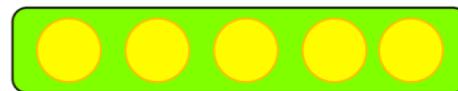
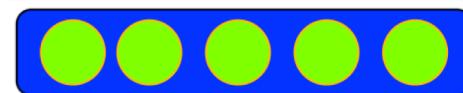
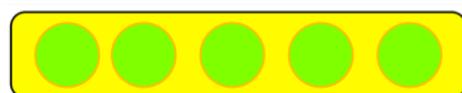
Bruni et al., Multimodal Distributional Semantics. 2014

Enrichment of pure textual vectors with
complementary information coming from
perceptual visual features.

Language grounding

Bruni et al., Multimodal Distributional Semantics. 2014

Enrichment of pure textual vectors with
complementary information coming from
perceptual visual features.



Language grounding

Bruni et al., Multimodal Distributional Semantics. 2014

Enrichment of pure textual vectors with
complementary information coming from
perceptual visual features.

- Word meaning not learned in isolation
but from usage

Language grounding

Bruni et al., Multimodal Distributional Semantics. 2014

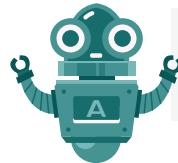
Enrichment of pure textual vectors with
complementary information coming from
perceptual visual features.

- Word meaning not learned in isolation but from usage
- Missing the interactive and dialogical aspects of language

Visual Dialogue

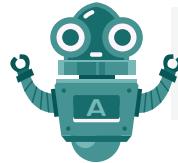


Visual Dialogue



A man and a woman are holding umbrellas

Visual Dialogue

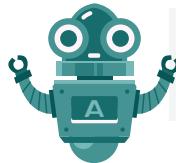


A man and a woman are holding umbrellas

What color is his umbrella?



Visual Dialogue

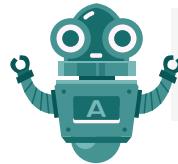


A **man** and a woman are holding umbrellas

What color is **his** umbrella?



Visual Dialogue

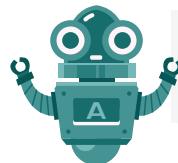


A man and a woman are holding umbrellas

What color is his **umbrella**?



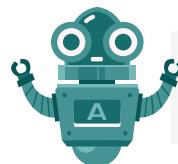
Visual Dialogue



A man and a woman are holding umbrellas

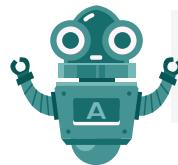


What color is his umbrella?



His umbrella is black

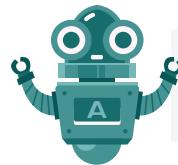
Visual Dialogue



A man and a woman are holding umbrellas



What color is his umbrella?

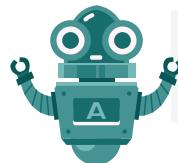
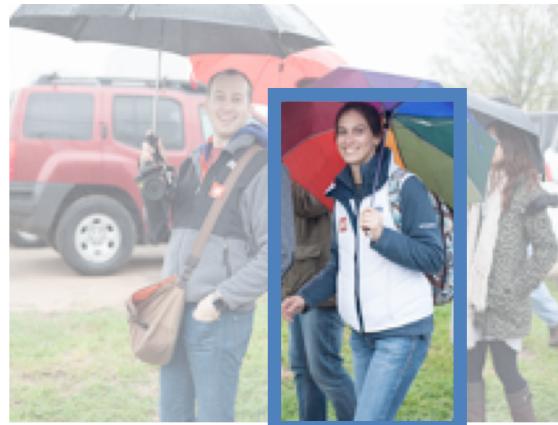


His umbrella is black



What about hers?

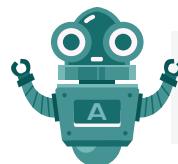
Visual Dialogue



A man and a **woman** are holding umbrellas



What color is his umbrella?

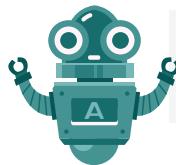


His umbrella is black

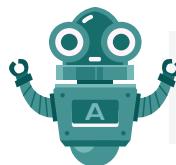


What about **hers**?

Visual Dialogue



A man and a woman are holding umbrellas



His **umbrella** is black

What color is his **umbrella**?



What about **hers**?

