Proceedings Paper:
Alomari, M, Hogg, D and Cohn, A orcid.org/0000-0002-7652-8907 (2018) Natural language learning and grounding for robotic systems. In: Cognitive Processing. 7th International Conference on Spatial Cognition (ICSC 2018), 10-14 Sep 2018, Rome, Italy. Springer Nature , S14-S15.

https://doi.org/10.1007/s10339-018-0884-3

Natural language learning and grounding for robotic Systems

Muhannad Alomari, David Hogg and Anthony Cohn

University of Leeds

In this talk, the problem of bootstrapping knowledge in language and vision for autonomous robots is addressed through novel techniques in grammar induction and word grounding to the perceptual world. The learning is achieved in a cognitively plausible loosely-supervised manner from raw linguistic and visual data and is validated on four datasets.

The learning framework is divided into three processes: (1) perceptual raw data is clustered into a number of Gaussian components to learn the 'visual concepts'; (2) frequent co-occurrence of words and visual concepts are used to learn the language grounding; (3) the learned language grounding and visual concepts are used to induce probabilistic grammar rules to model the language structure. Language grounding is enabled by seeking frequent co-occurrence between words and learned visual concepts using integer programming.

Probabilistic CFG rules are generated to model the language by mapping sentences to learned visual concepts, as opposed to traditional supervised grammar induction techniques where the learning is only made possible by using manually annotated training examples. Although individual aspects of this work have been addressed in previous work, we believe that this is the first time that, in a single system, it has been shown how a string of words with no preassigned syntactic categories, or semantics assigned, these can be grounded to automatically learned perceptual concepts, both static and dynamic, and also to learn a grammar of the language. Moreover this is shown to be possible in an online framework, i.e. incrementally, and without storing all previous examples.