

Grounded language understanding: Varieties of contextual grounding

Christopher Potts

Stanford Linguistics

CS224u: Natural language understanding



Our task: Color description in context

	Context	Utterance
	 	blue
	 	The darker blue one
	 	teal not the two that are more green
	 	dull pink not the super bright one
	 	not any of the regular greens
	 	Purple
	 	blue

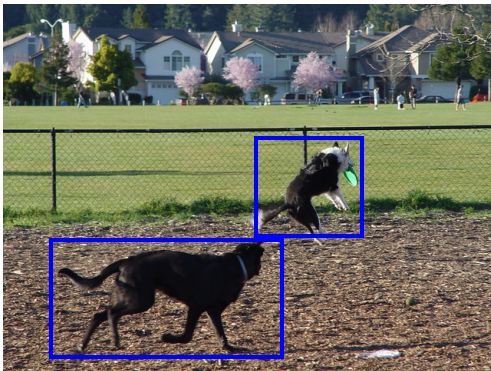
Stanford Colors in Context corpus
(Monroe et al. 2017)

Discriminative image labeling



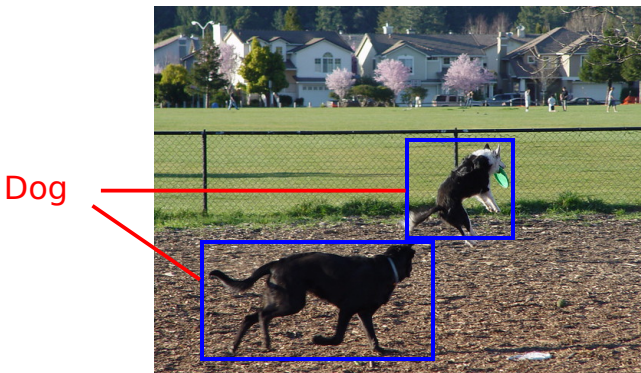
Mao et al. 2016

Discriminative image labeling



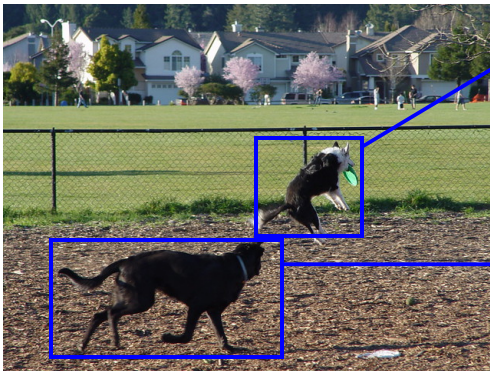
Mao et al. 2016

Discriminative image labeling



Mao et al. 2016

Discriminative image labeling

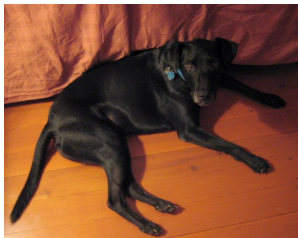


A little dog
jumping and
catching a
frisbee

A big dog
running

Mao et al. 2016

Discriminative image captioning



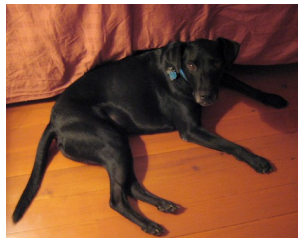
Vedantam et al. 2017; Cohn-Gordon et al. 2018

Discriminative image captioning



Vedantam et al. 2017; Cohn-Gordon et al. 2018

Discriminative image captioning



Vedantam et al. 2017; Cohn-Gordon et al. 2018

Machine translation

She chopped up the tree.



Elle coupa l'arbre.

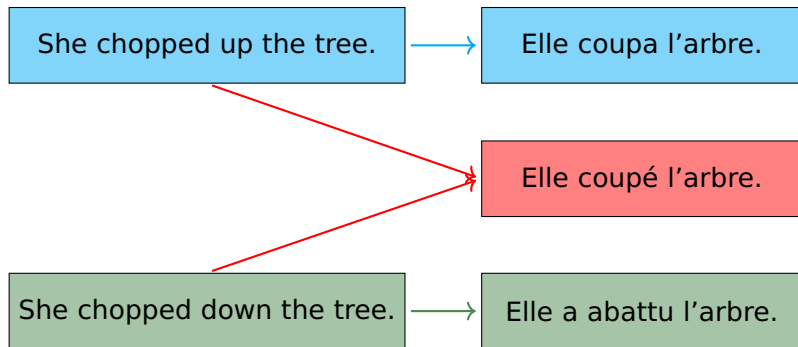
She chopped down the tree.



Elle a abattu l'arbre.

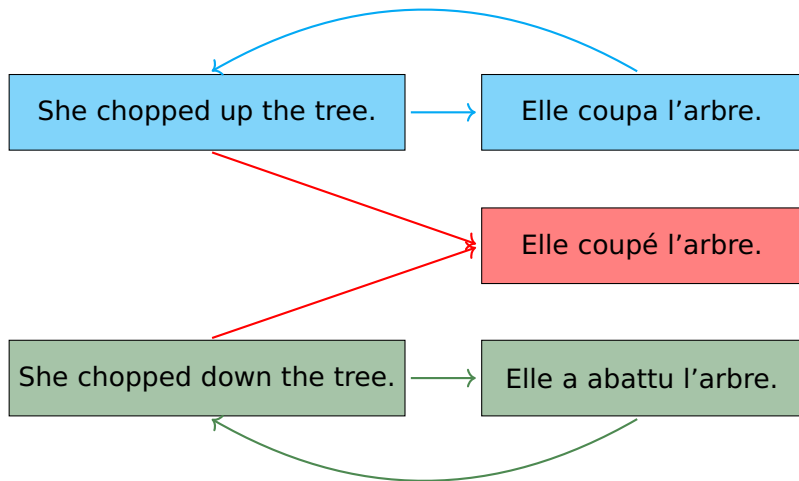
Cohn-Gordon and Goodman 2019

Machine translation



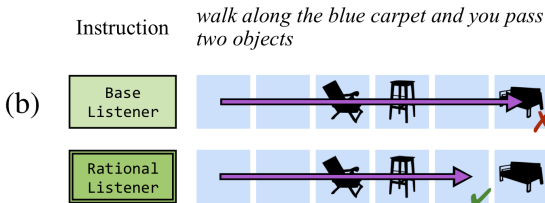
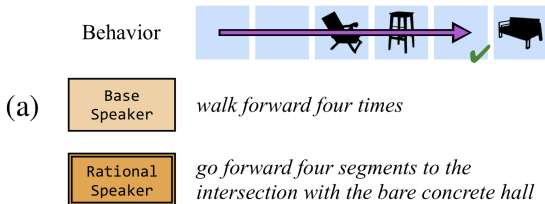
Cohn-Gordon and Goodman 2019

Machine translation

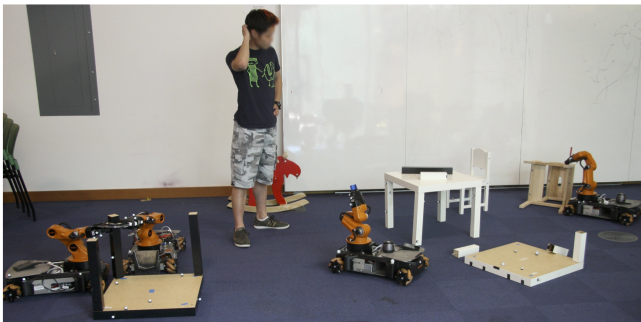


Cohn-Gordon and Goodman 2019

Generating and following instructions

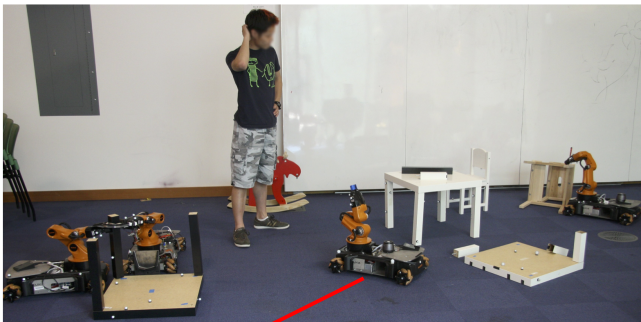


Collaborative problem solving



Tellex et al. 2014

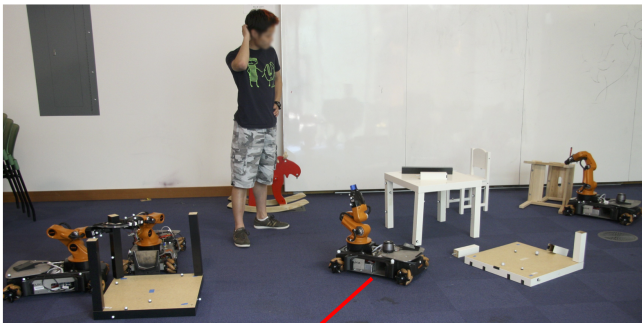
Collaborative problem solving



Help me!

Tellex et al. 2014

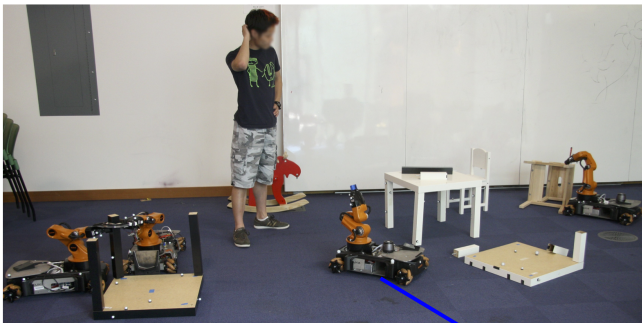
Collaborative problem solving



Hand me
the leg

Tellex et al. 2014

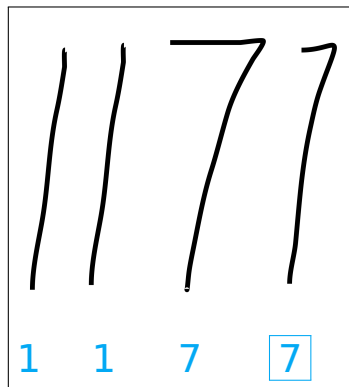
Collaborative problem solving



Hand me the white
leg on the table

Tellex et al. 2014

Optical character recognition



References I

- Reuben Cohn-Gordon and Noah Goodman. 2019. [Lost in machine translation: A method to reduce meaning loss](#). In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 437–441, Minneapolis, Minnesota. Association for Computational Linguistics.
- Reuben Cohn-Gordon, Noah D. Goodman, and Christopher Potts. 2018. Pragmatically informative image captioning with character-level inference. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 439–443, Stroudsburg, PA. Association for Computational Linguistics.
- Daniel Fried, Jacob Andreas, and Dan Klein. 2018. [Unified pragmatic models for generating and following instructions](#). In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 1951–1963, New Orleans, Louisiana. Association for Computational Linguistics.
- Junhua Mao, Jonathan Huang, Alexander Toshev, Oana Camburu, Alan L. Yuille, and Kevin Murphy. 2016. Generation and comprehension of unambiguous object descriptions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 11–20. IEEE.
- Will Monroe, Robert X. D. Hawkins, Noah D. Goodman, and Christopher Potts. 2017. Colors in context: A pragmatic neural model for grounded language understanding. *Transactions of the Association for Computational Linguistics*, 5:325–338.
- Stefanie Tellex, Ross A. Knepper, Adrian Li, Thomas M. Howard, Daniela Rus, and Nicholas Roy. 2014. [Asking for help using inverse semantics](#). In *Proceedings of Robotics: Science and Systems*.
- Ramakrishna Vedantam, Samy Bengio, Kevin Murphy, Devi Parikh, and Gal Chechik. 2017. Context-aware captions from context-agnostic supervision. *arXiv:1701.02870*.