COMP 790.139 (Spring 2017)
Adv. Topics in NLP:
Language Grounding for Robotics

Mohit Bansal

Class Info/Logistics

- COMP 790.139 ‘Language Grounding for Robotics’
- **3 UNITS**
- Instructor: Mohit Bansal (SN258, http://www.cs.unc.edu/~mbansal/)
- Time: Tuesdays 9.30am-12.00pm
- Room: FB008
- Office Hours: Tue 12.00-1.00pm (by appointment), SN258
- Course Webpage: http://www.cs.unc.edu/~mbansal/teaching/robonlp-seminar-spring17.html
- **Course Email**: nlpcomp790unc@gmail.com

About Me

- Asst. Professor, CS, UNC (joined Fall 2016)
- Res. Asst. Professor, TTI-Chicago/UChicago, 2013-2016
- PhD, UC Berkeley, 2008-2013

- Research Interests:
  - Past: Syntactic parsing, coreference resolution, taxonomy induction, world knowledge and commonsense induction
  - Current: Multimodal and embodied semantics (i.e., language with vision and speech, for robotics); human-like language generation and Q&A/dialogue; interpretable and structured deep learning

- Office SN258

- Webpage: http://www.cs.unc.edu/~mbansal/, Email: mbansal@cs.unc.edu

Your Introductions

- Please say your:
  - Name
  - Department/degree/major
  - Research interests (ML/AI/NLP/CV experience? Coding?)
  - Fun fact ☺

(send me a few lines of intro at nlpcomp790unc@gmail.com)

About the Course (and its Goals)

- Research-oriented seminar course! We will read lots of interesting papers, brainstorm, and do fun novel projects!
- Basic NLP and machine learning (some deep learning) experience expected – check fall2016 NLP intro slides!
- We’ll cover several latest research topics in NLP+robotics
- You will read papers in advance and write summaries
- You will present papers in class
- Brainstorm regularly and code + write up fun/novel projects!
- Some lecture(s) on academic/research quality paper writing

Prerequisites

- Graduate research-level class
- Some machine learning and coding experience is definitely expected!
- Moreover, some NLP and RL background is highly recommended
- Projects will require research and coding skills; summaries and paper presentation will require solid ML clarity

Expectations/Grading

- Paper presentation (15%)
- Paper written summaries (25%)
- Class participation, discussion and brainstorming (20%)
- Project reports and presentations (40%)
  (midterm = 15% and final = 25%)

Lateness Policy

- Written summaries are due *before class* by email to nlpcomp790unc@gmail.com
- First paper summary submission will have no late penalties.
- After that, for every week's summary submission, there will be a 25% value reduction per late day.
- Other lateness policies (for projects, etc.) will be sent via email during the semester

Collaboration Policy

- Paper summaries have to be written and submitted individually.
- Projects are encouraged to be done in pairs, with clearly outlined contributions from each team member
- But individual projects are fine too (e.g., if it relates to your current research)

Paper Presentation

- Lead discussion for 1-2 papers on a topic some week (may be done in pairs/groups depending on class size)
- Read related papers and present background to audience
- Present task and model details of given paper
- Present demo’s of related code, etc.
- Ask interesting questions to initiate brainstorming
- Mention some next steps, future work, extension ideas!

Paper Written Summaries

- 0.5-1 page (per paper) write-up for every week’s paper(s)
- Describe the task
- Summarize the method
- Explain the novelty
- Discuss the next steps or potential improvements

Class Participation and Brainstorming

- Audience students expected to take part in lively discussion every paper reading!
- After every topic gets completed (i.e., several papers in 2-3 weeks), we will have a brainstorming and ‘idea-generation’ session!
- Exact details to be announced soon but students expected to submit and discuss novel idea(s) on the whole general topic, e.g., new related task or dataset, new approach to existing task, combinations of tasks/approaches, etc.
- Don’t hesitate to propose fancy ideas ☺, but try to keep them grounded/feasible and think of how to approach them realistically (in terms of datasets, models, speed, memory, etc.)

Project

- This is a ‘Reading, Coding, and Writing’ class!
- Students will pick (early) their favorite topic among latest cutting-edge research topics covered in class
- And will try a novel idea (implementing + extending or original) -- I am happy to discuss details!
- Midterm and final report + presentation (and possibly some updates)
- Might be in pairs/groups depending on final class size
- Use conference style files and aim for conference-quality papers
- Will have some lecture(s) on research-quality paper writing

Reference Books

- **SLP2**: D. Jurafsky & James H. Martin. “Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition”. Prentice Hall, Second Edition, 2009.
- **SLP3**: Some draft chapters of the third edition are available online at https://web.stanford.edu/~jurafsky/slp3/
- **FSNLP**: Chris Manning and Hinrich Schütze, Foundations of Statistical Natural Language Processing, MIT Press. Cambridge, MA: May 1999. http://nlp.stanford.edu/fsnlp/
- **ML Background**: Andrew Ng’s Coursera Machine Learning course https://www.coursera.org/learn/machine-learning
- **Reinforcement Learning**: Sutton & Barto Book: Reinforcement Learning: An Introduction. https://webdocs.cs.ualberta.ca/~sutton/book/the-book.html

Course Syllabus/Topics (tentative)

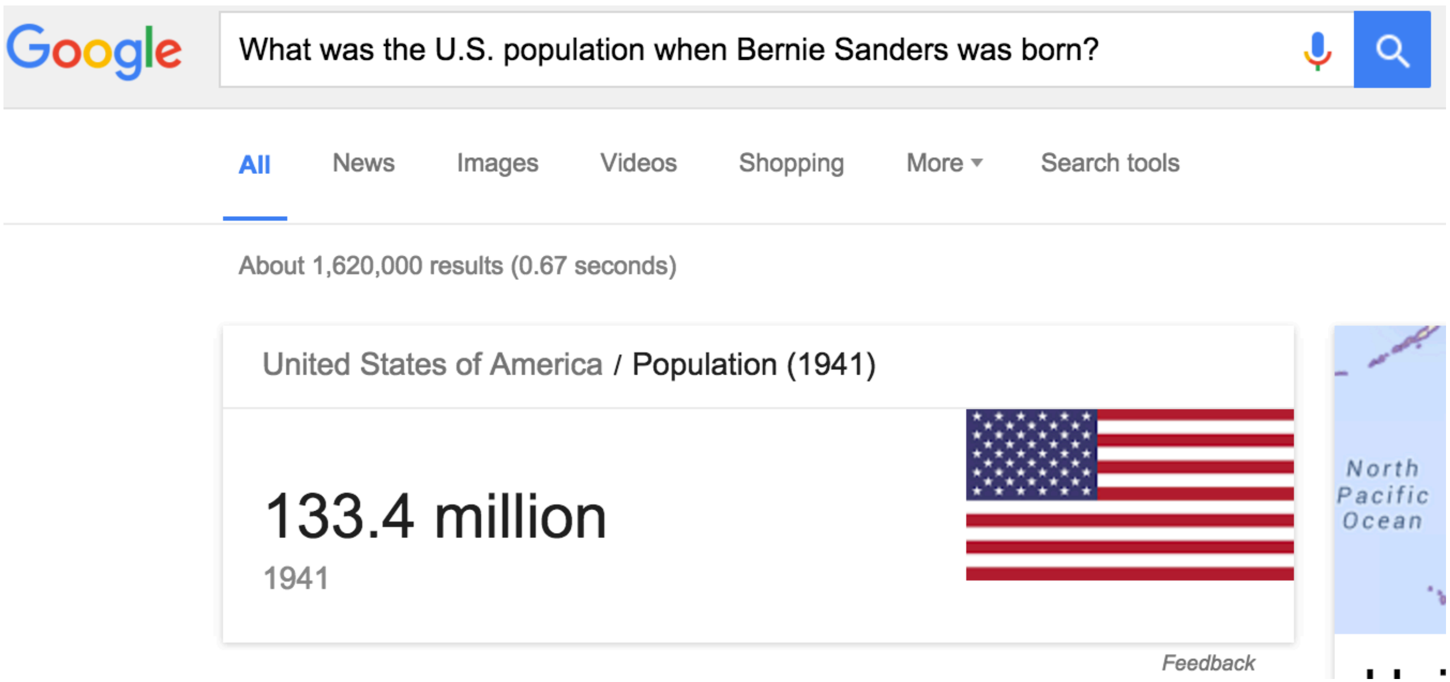
- interpreting and executing verbal instructions for navigation, articulation, manipulation, assembly, skill learning, etc.
- human-robot collaboration and dialogue for learning new subactions, mediating shared perceptual basis, referring expression generation, etc.
- grounding and language learning via dialogue-based and interactive games.
- automatic language generation for embodied tasks.
- grounded reinforcement learning.
- grounded knowledge representations (mapping language to world).
- machine learning models (structured and deep), datasets, and metrics for embodied language.

What is NLP?

- Question answering

What is NLP?

- Question answering



What is NLP?

- Question answering

Google

which countries border the black sea

All Maps Images News Shopping More Search tools

About 2,710,000 results (0.81 seconds)

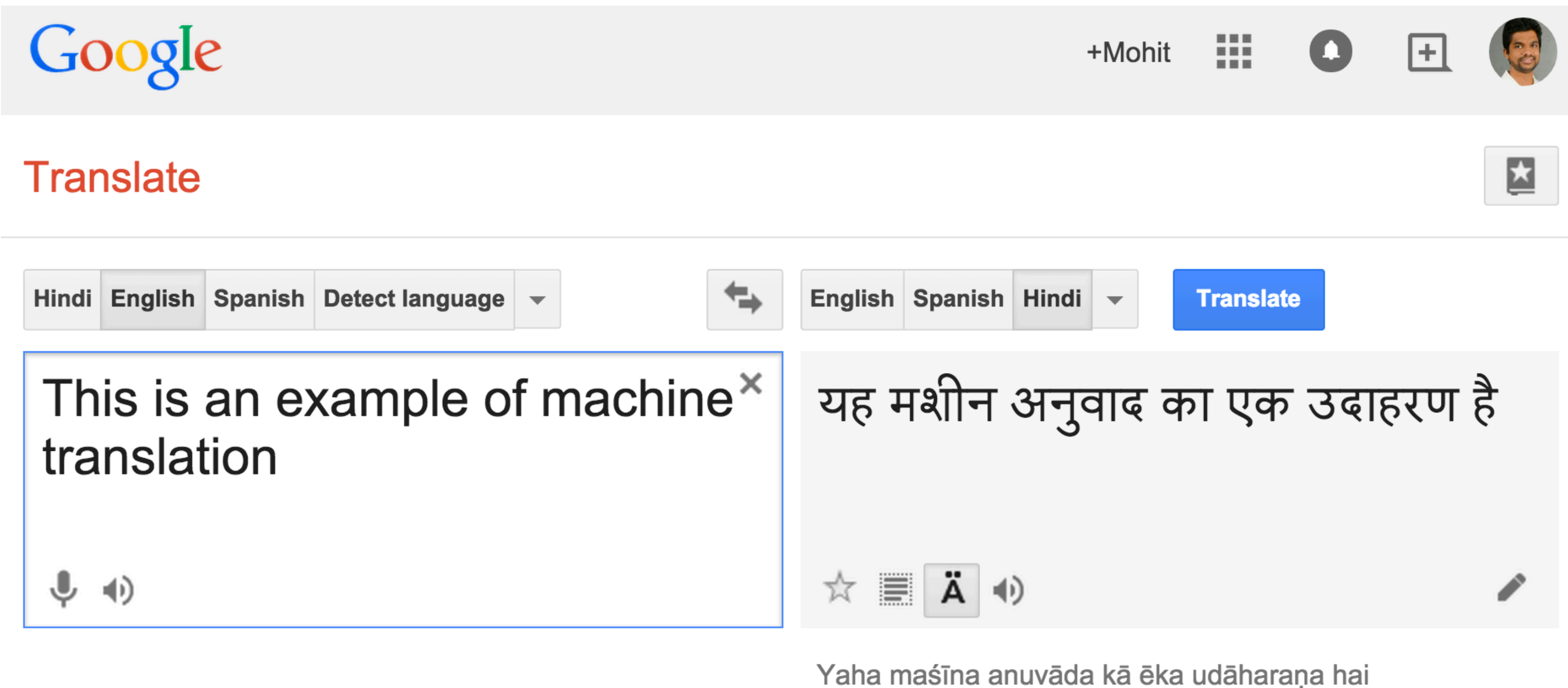
This major inland sea is bordered by six countries — **Romania** and **Bulgaria** to the west; **Ukraine**, **Russia**, and **Georgia** to the north and east; and **Turkey** to the south. Additionally, it is impacted by another 10 nations through the five major rivers that empty into the Black Sea, the largest of which is the Danube River.

Black Sea Geography - College of Earth, Ocean, and Environment
https://www.ceoe.udel.edu/**blacksea**/geography/index.html University of Delaware

About this result • Feedback

What is NLP?

- Machine Translation



What is NLP?

- Sentiment Analysis

Sentiment Analysis with Python NLTK Text Classification

This is a demonstration of **sentiment analysis** using a `NLTK 2.0.4` powered **text classification** process. It can tell you whether it thinks the text you enter below expresses **positive sentiment**, **negative sentiment**, or if it's **neutral**. Using **hierarchical classification**, *neutrality* is determined first, and *sentiment polarity* is determined second, but only if the text is not neutral.

Analyze Sentiment

Language

english

Enter text

It always amazes me how Universal never cares to create anything remotely clever when it comes to their animations, and so once again they come up with a harmless little story that wants to be cute and funny (which it is sometimes) but is only bound to be quickly forgotten.

Enter up to 50000 characters

Analyze

Sentiment Analysis Results

The text is **neg**.

The final sentiment is determined by looking at the classification probabilities below.

Subjectivity

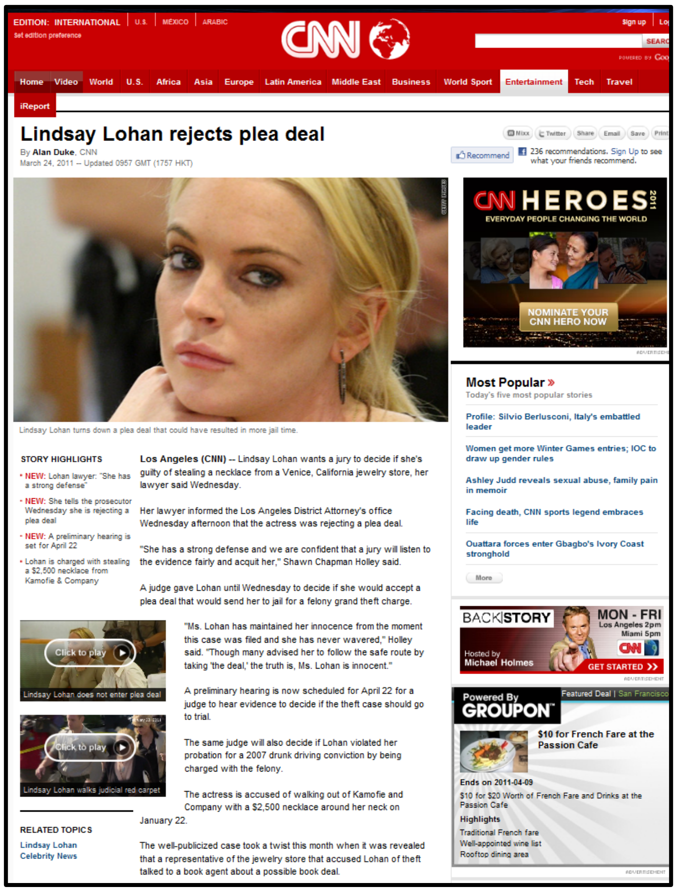
- neutral: 0.3
- **polar: 0.7**

Polarity

- pos: 0.2
- **neg: 0.8**

What is NLP?

- Natural Language Generation: Summarization



EDITION: INTERNATIONAL | U.S. | MÉXICO | ARABIC
Set edition preference

Home | Video | World | U.S. | Africa | Asia | Europe | Latin America | Middle East | Business | World Sport | Entertainment | Tech | Travel | iReport

Lindsay Lohan rejects plea deal

By **Alan Duke**, CNN
March 24, 2011 -- Updated 0957 GMT (1757 HKT)

Lindsay Lohan turns down a plea deal that could have resulted in more jail time.

STORY HIGHLIGHTS

- NEW: Lohan lawyer: "She has a strong defense"
- NEW: She tells the prosecutor Wednesday she is rejecting a plea deal
- NEW: A preliminary hearing is set for April 22
- Lohan is charged with stealing a $2,500 necklace from Kamofie & Company

Los Angeles (CNN) -- Lindsay Lohan wants a jury to decide if she's guilty of stealing a necklace from a Venice, California jewelry store, her lawyer said Wednesday.

Her lawyer informed the Los Angeles District Attorney's office Wednesday afternoon that the actress was rejecting a plea deal.

"She has a strong defense and we are confident that a jury will listen to the evidence fairly and acquit her," Shawn Chapman Holley said.

A judge gave Lohan until Wednesday to decide if she would accept a plea deal that would send her to jail for a felony grand theft charge.

Click to play
Lindsay Lohan does not enter plea deal

Click to play
Lindsay Lohan walks judicial red carpet

"Ms. Lohan has maintained her innocence from the moment this case was filed and she has never wavered," Holley said. "Though many advised her to follow the safe route by taking 'the deal,' the truth is, Ms. Lohan is innocent."

A preliminary hearing is now scheduled for April 22 for a judge to hear evidence to decide if the theft case should go to trial.

The same judge will also decide if Lohan violated her probation for a 2007 drunk driving conviction by being charged with the felony.

The actress is accused of walking out of Kamofie and Company with a $2,500 necklace around her neck on January 22.

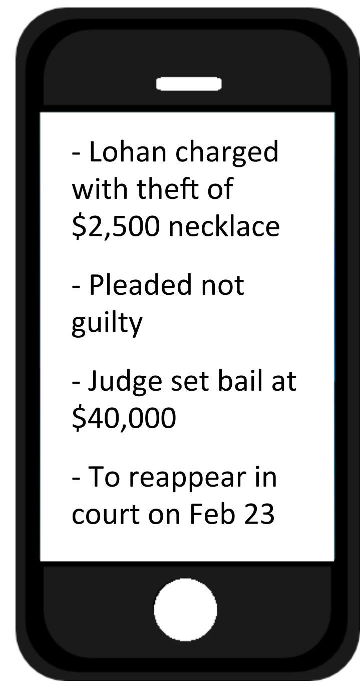
The well-publicized case took a twist this month when it was revealed that a representative of the jewelry store that accused Lohan of theft talked to a book agent about a possible book deal.

RELATED TOPICS
Lindsay Lohan
Celebrity News

Most Popular »
Today's five most popular stories

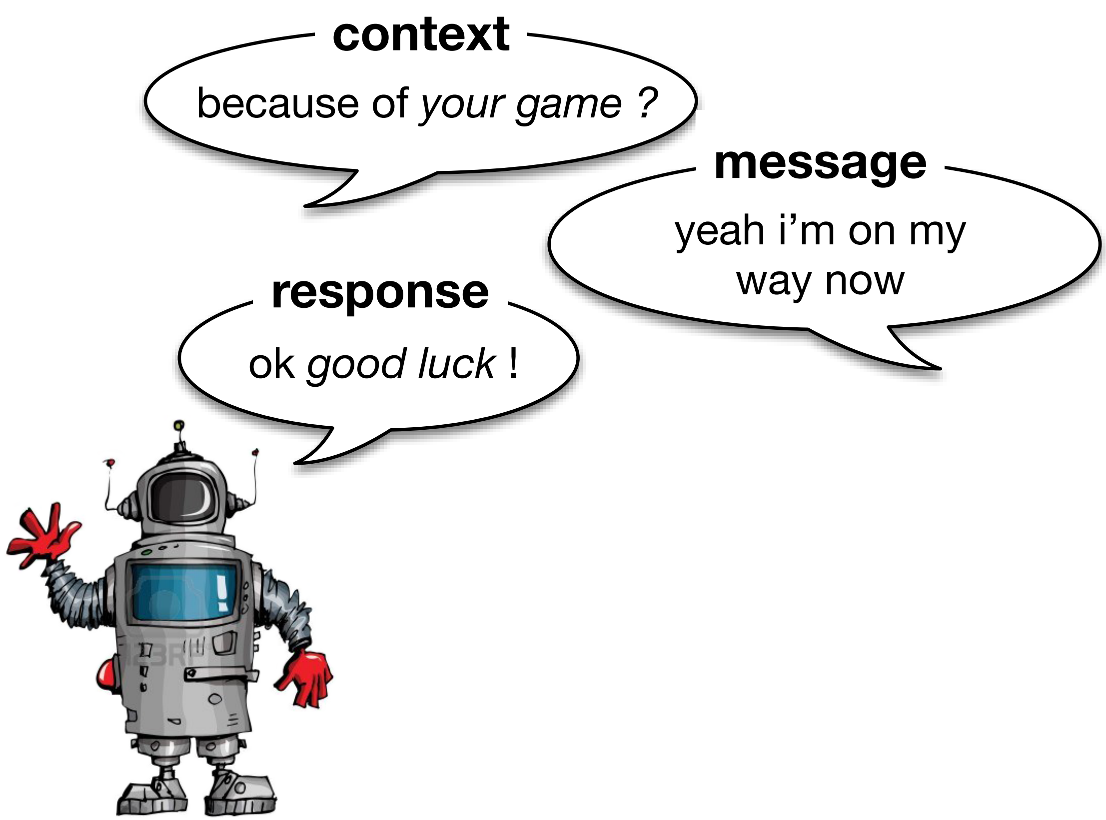
- Profile: Silvio Berlusconi, Italy's embattled leader
- Women get more Winter Games entries; IOC to draw up gender rules
- Ashley Judd reveals sexual abuse, family pain in memoir
- Facing death, CNN sports legend embraces life
- Ouattara forces enter Gbagbo's Ivory Coast stronghold

More



What is NLP?

- Natural Language Generation: Conversation/Dialogue



[Sordoni et al., 2015]

What is NLP?

- Natural Language Generation: Image Captioning

www.cs.toronto.edu/~nitish/nips2014demo/results/8832804.html

Results

Tags

- authors
- scones
- luncheon
- breakfast
- seder

Nearest Caption in the Training Dataset

a man cuts a cake while children sit around at the table , looking on .

Generated Captions

- two people at a table with a cake .
- the two people are having a meal that is in a party .
- a man and two children in a blue table with a cake .
- a man sitting at a table with a bunch of cake on it .
- a man and woman sitting at a table with cake at a party .

back

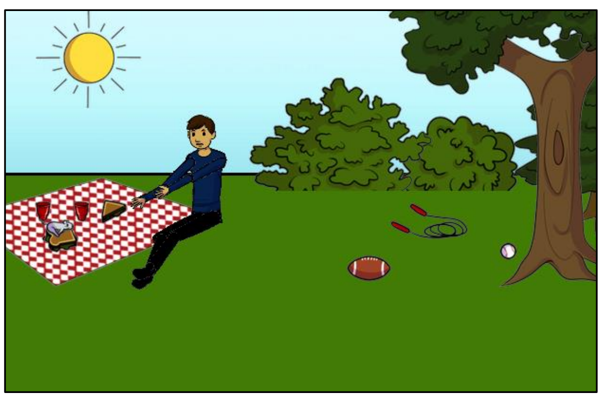
[UToronto]

What is NLP?

- Natural Language Generation: Visual Question Answering

What color are her eyes?
What is the mustache made of?

How many slices of pizza are there?
Is this a vegetarian pizza?



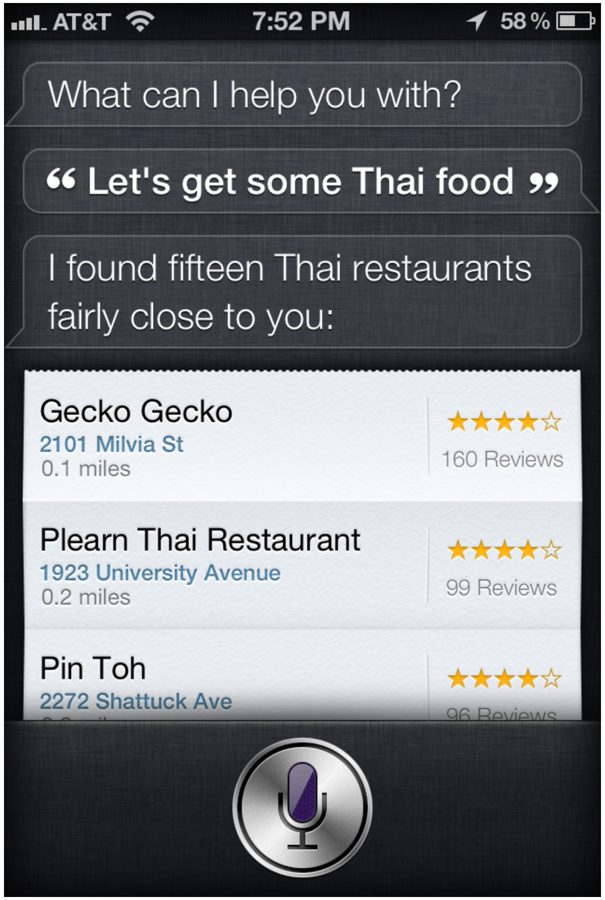
Is this person expecting company?
What is just under the tree?

Does it appear to be rainy?
Does this person have 20/20 vision?

[Antol et al., 2015]

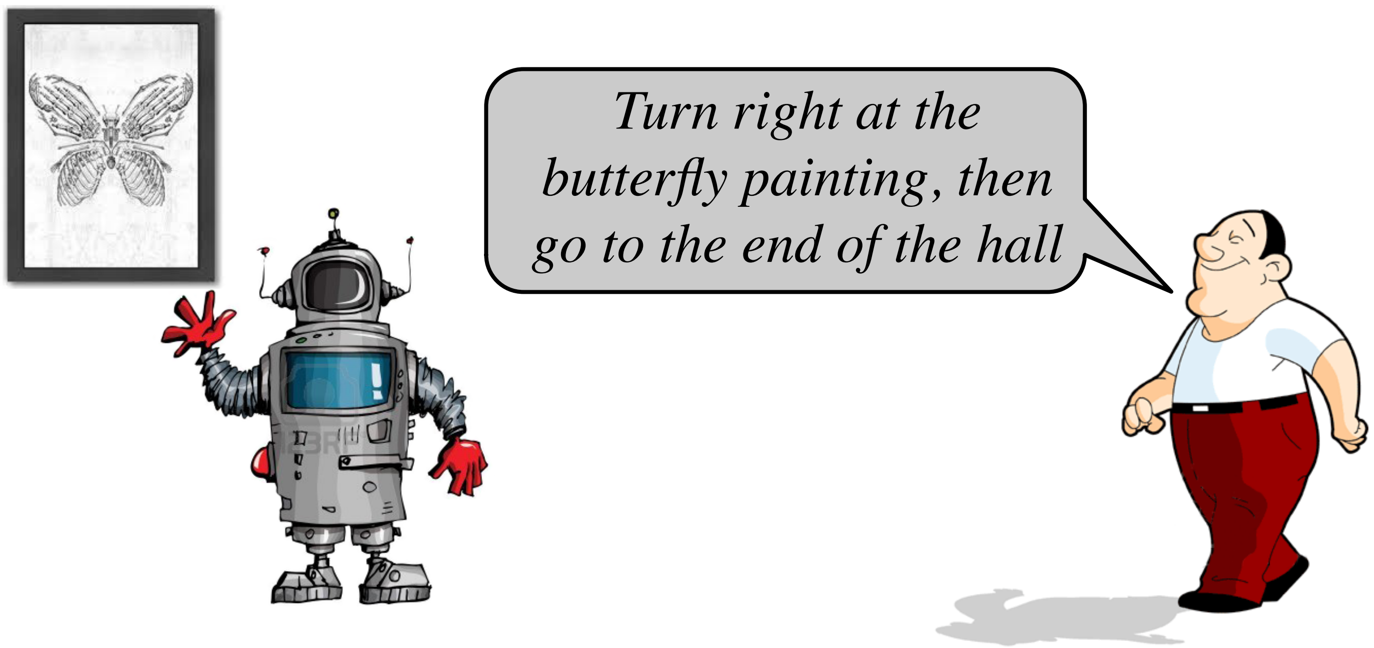
What is NLP?

- Automatic Speech Recognition



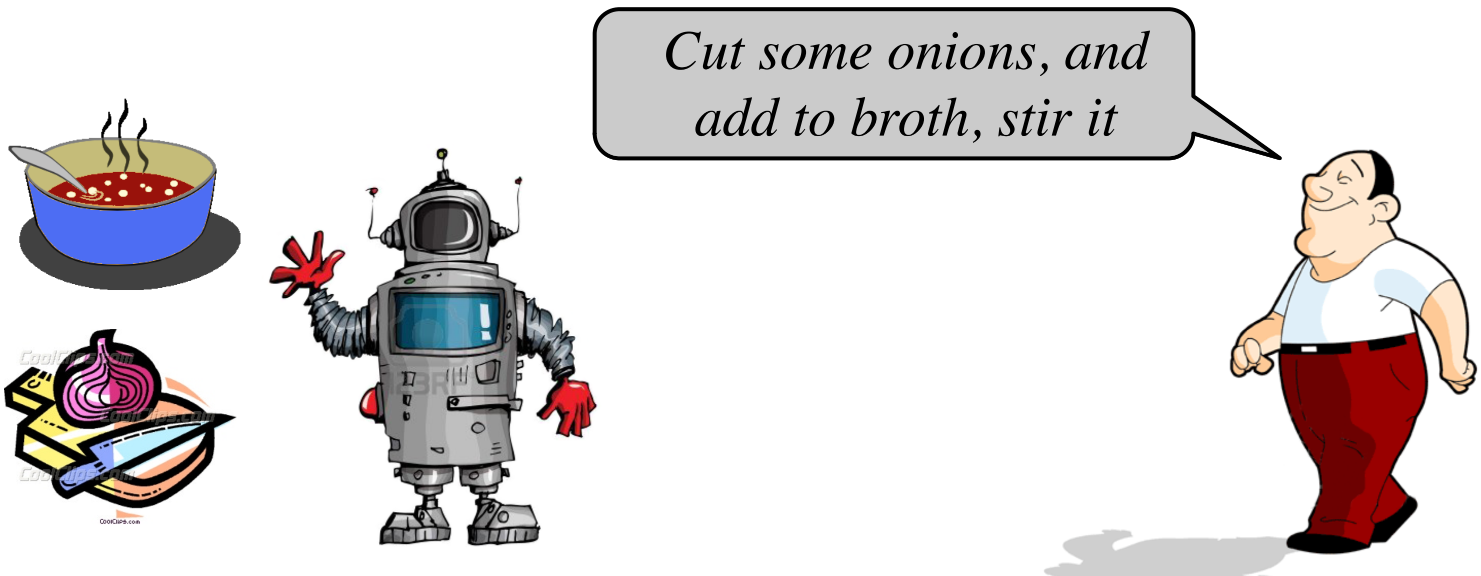
What is NLP?

- Task-based instructions, e.g., navigation, grasping, manipulation, skill learning



What is NLP?

- Task-based instructions, e.g., navigation, grasping, manipulation, skill learning



Some Sample NLP+Robotics Papers

- Navigation Instruction Following

Place your back against the wall of the "T" intersection. Go forward one segment to the intersection with the blue-tiled hall. This interesction [sic] contains a chair. Turn left. Go forward to the end of the hall. Turn left. Go forward one segment to the intersection with the wooden-floored hall. This intersection conatains [sic] an easel. Turn right. Go forward two segments to the end of the hall. Turn left. Go forward one segment to the intersection containing the lamp. Turn right. Go forward one segment to the empty corner.

Figure 1: An example of a route instruction-path pair in one of the virtual worlds from MacMahon, Stankiewicz, and Kuipers (2006) with colors that indicate floor patterns and wall paintings, and letters that indicate different objects. Our method successfully infers the correct path for this instruction.

Learning to interpret natural language navigation instructions from observations. Chen and Mooney. AAAI 2011.
Weakly supervised learning of semantic parsers for mapping instructions to actions. Artzi and Zettlemoyer. TACL 2013.
Listen, Attend, and Walk: Neural Mapping of Navigational Instructions to Action Sequences. Mei, Bansal, and Walter. AAAI 2016

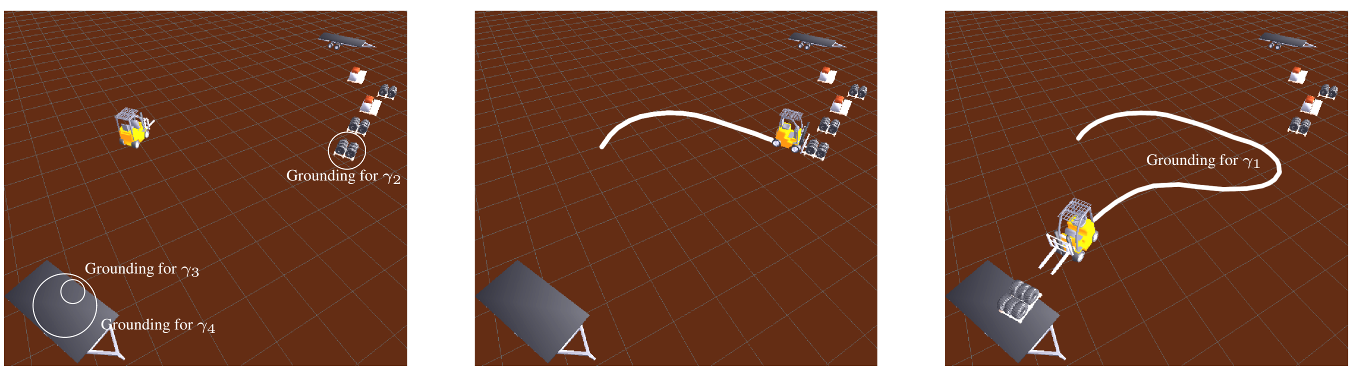
Some Sample NLP+Robotics Papers

- Navigation + Manipulation Instructions

Commands from the corpus
- Go to the first crate on the left and pick it up.
- Pick up the pallet of boxes in the middle and place them on the trailer to the left.
- Go forward and drop the pallets to the right of the first set of tires.
- Pick up the tire pallet off the truck and set it down

(a) Robotic forklift (b) Sample commands

Figure 1: A target robotic platform for mobile manipulation and navigation (Teller et al., 2010), and sample commands from the domain, created by untrained human annotators. Our system can successfully follow these commands.



(a) Object groundings (b) Pick up the pallet (c) Put it on the truck

Figure 4: A sequence of the actions that the forklift takes in response to the command, "Put the tire pallet on the truck." (a) The search grounds objects and places in the world based on their initial positions. (b) The forklift executes the first action, picking up the pallet. (c) The forklift puts the pallet on the trailer.

Understanding Natural Language Commands for Robotic Navigation and Mobile Manipulation. Tellex, Kollar, Dickerson, Walter, Banerjee, Teller, and Roy. AAAI 2011.

Some Sample NLP+Robotics Papers

- Navigation Dialogue

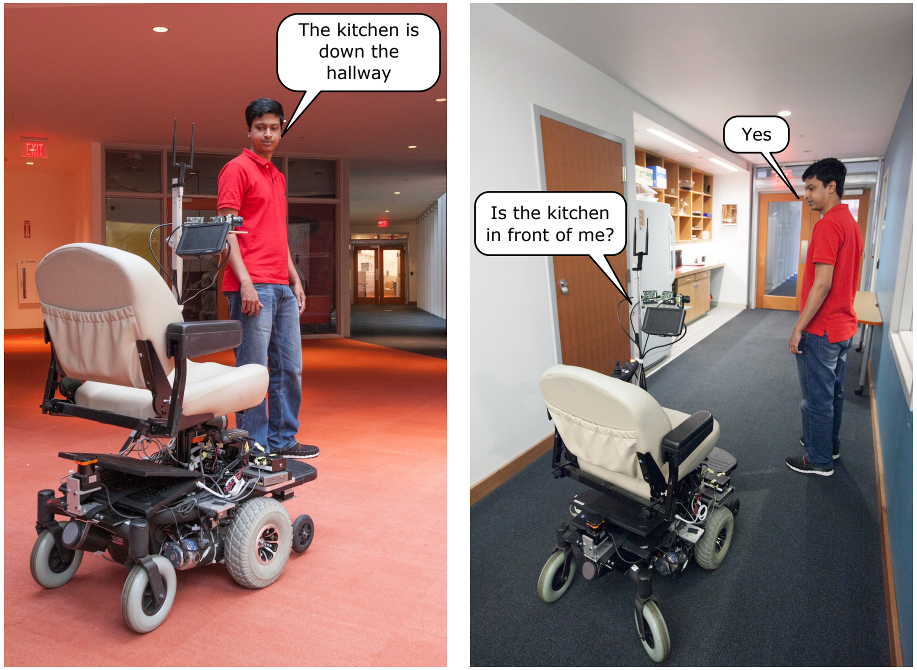


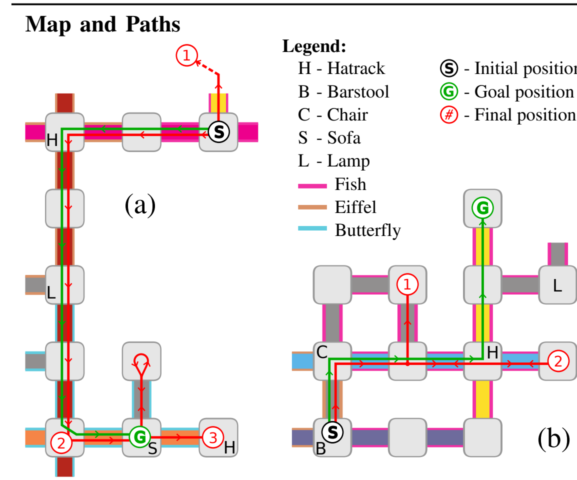
Fig. 1. A user gives a tour to a robotic wheelchair designed to assist residents in a long-term care facility. (Left) The guide provides an ambiguous description of the kitchen's location. (Right) When the robot is near one of the likely locations, it asks the guide a question to resolve the ambiguity.

Information-Theoretic Dialog to Improve Spatial-Semantic Representations. Hemachandra and Walter. IROS 2015.

Some Sample NLP+Robotics Papers

- Navigation Instruction Generation

Fig. 4. Participants' field of view in the virtual world used for the human navigation experiments.



Instructions

(a)	Human	"with your back to the wall turn left. walk along the flowers to the hatrack. turn left. walk along the brick two alleys past the lamp. turn left. move along the wooden floor to the chair. in the next block is a hatrack"
	Ours	"you should have the olive hallway on your right now. walk forward twice. turn left. move until you see wooden floor to your left. face the bench. move to the bench"
(b)	Human	"head toward the blue floored hallway. make a right on it. go down till you see the fish walled areas. make a left in the fish walled hallway and go to the very end"
	Ours	"turn to face the white hallway. walk forward once. turn right. walk forward twice. turn left. move to the wall"

Fig. 8. Examples of paths from the SAIL corpus that ten participants (five for each map) followed according to instructions generated by humans and by our method. Paths in red are those traversed according to human-generated instructions, while paths in green were executed according to our instructions. Circles with an "S" and "G" denote the start and goal locations, respectively.

Navigational Instruction Generation as Inverse Reinforcement Learning with Neural Machine Translation. Daniele, Bansal, and Walter. HRI 2017.

Some Sample NLP+Robotics Papers

- Assembly Instruction Execution

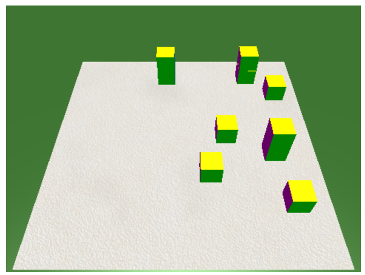
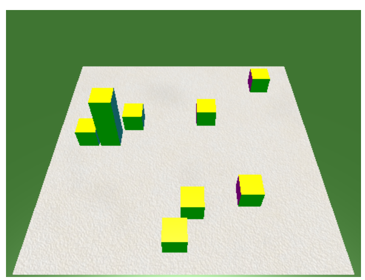
Scene	**Utterance**
	Move the block that is currently located closest to the top left corner to the bottom left of the table, slightly higher than the block in the bottom right corner.
Error:	`7.29 Block lengths`
	Move the block closest to the top left corner so it is above half a block length to the right of the blocks near the lower left corner of the table.
Error:	`0.94 Block lengths`

Table 6: Above are two commands and the worlds they apply to. Below we see the prediction error of our best model.

Natural Language Communication with Robots. Bisk, Yuret, and Marcu. NAACL 2016.

Some Sample NLP+Robotics Papers

- Recipes: Bakebot (http://projects.csail.mit.edu/video/research/robo/bakebot_final.mp4)

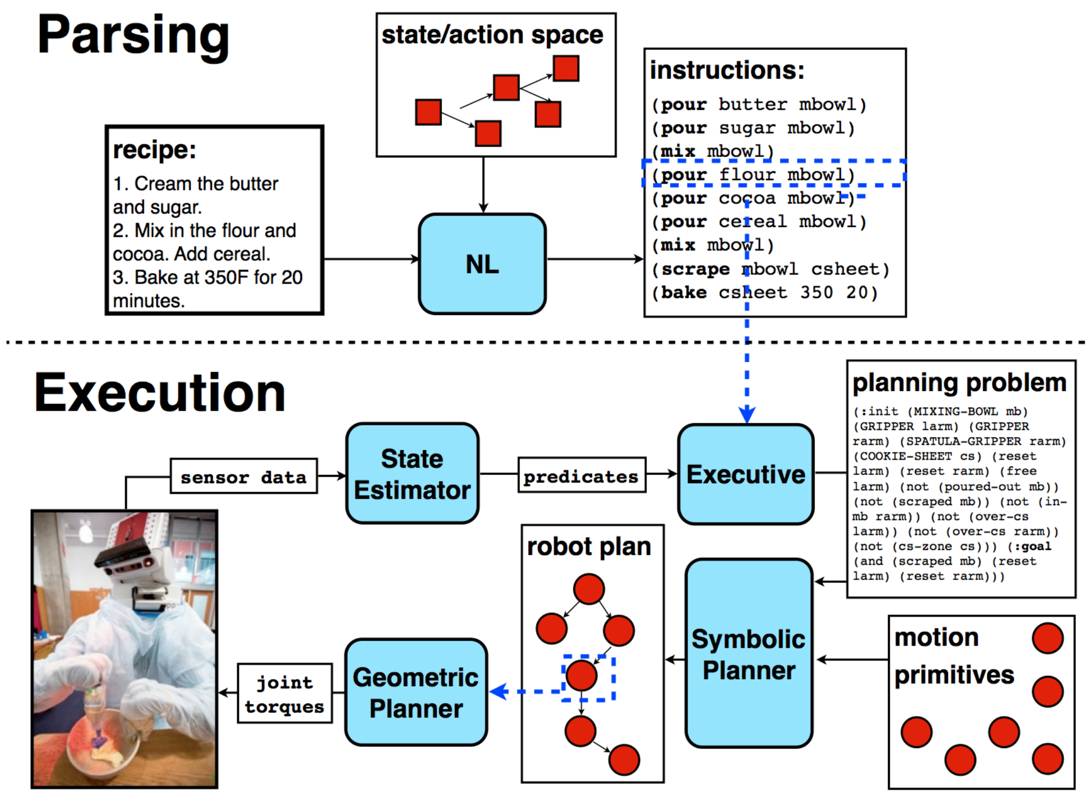


Fig. 3 Architecture of the BakeBot system. The NL system processes the plain text recipe, producing a high-level plan which is sent to the robot. For each instruction in the high-level plan, the motion planner assembles a motion plan and executes it on the PR2 robot.

Interpreting and Executing Recipes with a Cooking Robot. Bollini, Tellex, Thompson, Roy, Rus. ISER 2012.

Some Sample NLP+Robotics Papers

- Recipes: Tell Me Dave (http://tellmedave.cs.cornell.edu/)

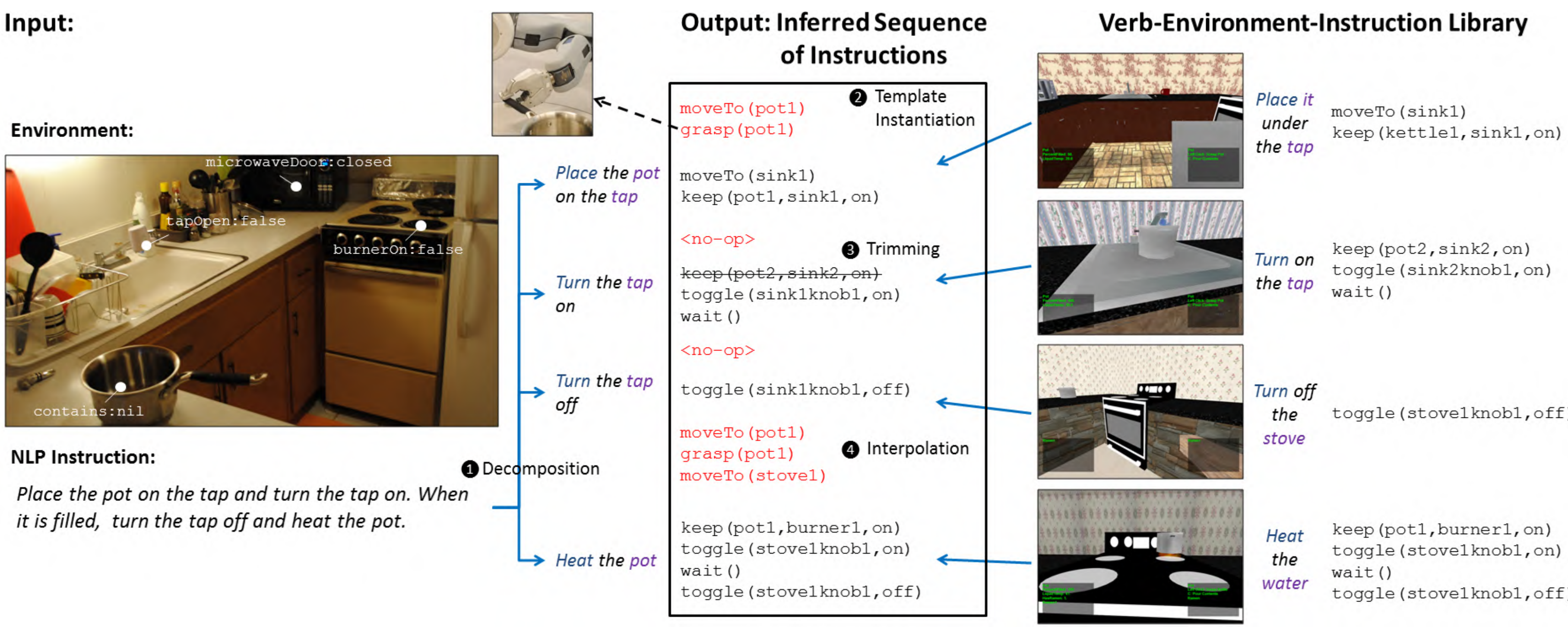


Fig. 1. **Natural Language Instructions to sequence of instructions** for a given new environment. Our approach takes description in natural language and sequences together robotic instructions that are appropriate for a given environment and task. Note that the NL instructions are often ambiguous, and are incomplete, and need to be grounded into the environment.

Tell Me Dave: Context-Sensitive Grounding of Natural Language to Mobile Manipulation Instructions, Misra, Sung, Lee, and Saxena. RSS 2014.

Some Sample NLP+Robotics Papers

- Recipes: Tell Me Dave (http://tellmedave.cs.cornell.edu/)

Fig. 4. **Robot Experiment.** Given the language instruction for making the dessert 'Affogato': *'Take some coffee in a cup. Add icecream of your choice. Finally, add raspberry syrup to the mixture.'*, our algorithm outputs a sequence that the PR2 executes to make the dessert. (Please see the video.)

Tell Me Dave: Context-Sensitive Grounding of Natural Language to Mobile Manipulation Instructions, Misra, Sung, Lee, and Saxena. RSS 2014.

Some Sample NLP+Robotics Papers

- Recipes: RoboBarista (http://robobarista.cs.cornell.edu/)

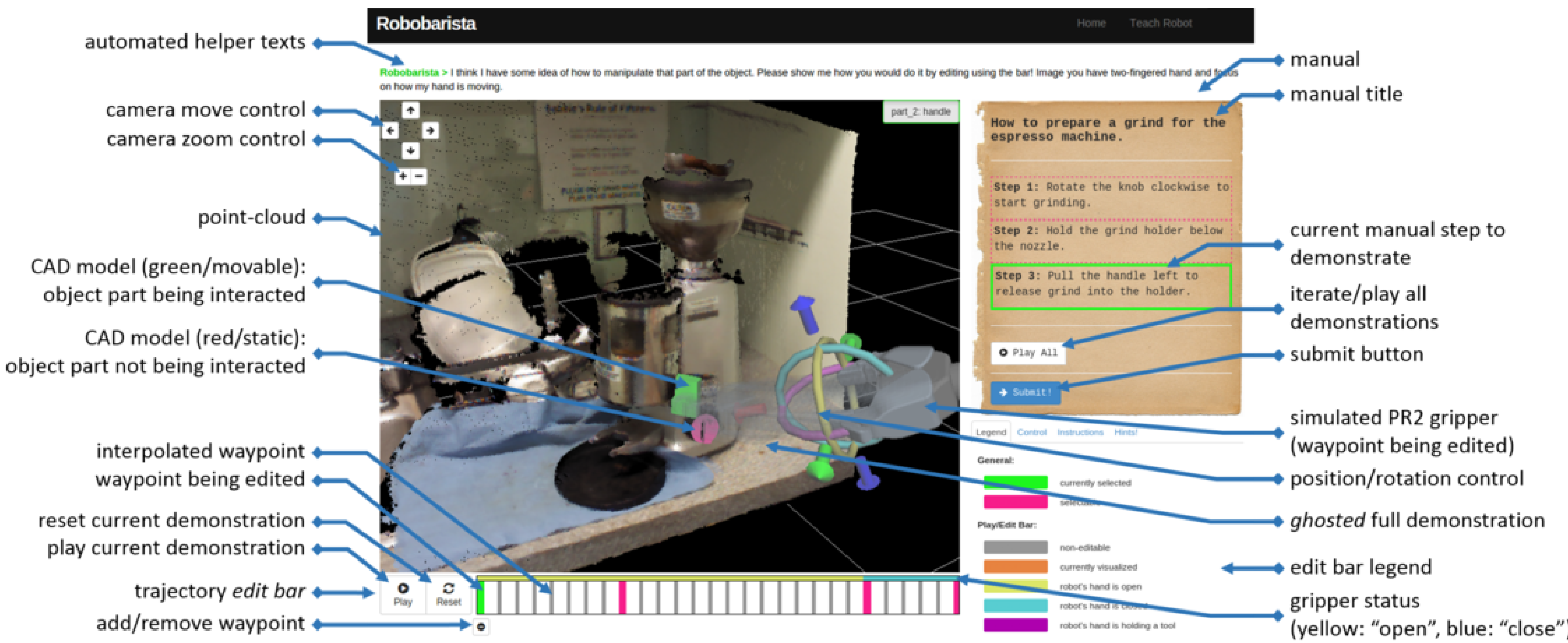


Fig. 5. **Screen-shot of Robobarista,** the crowd-sourcing platform running on Chrome browser. We have built Robobarista platform for collecting a large number of crowd demonstrations for teaching the robot.

Robobarista: Object Part based Transfer of Manipulation Trajectories from Crowd-sourcing in 3D Pointclouds. Sung, Jin, and Saxena. ISRR 2015.

Some Sample NLP+Robotics Papers

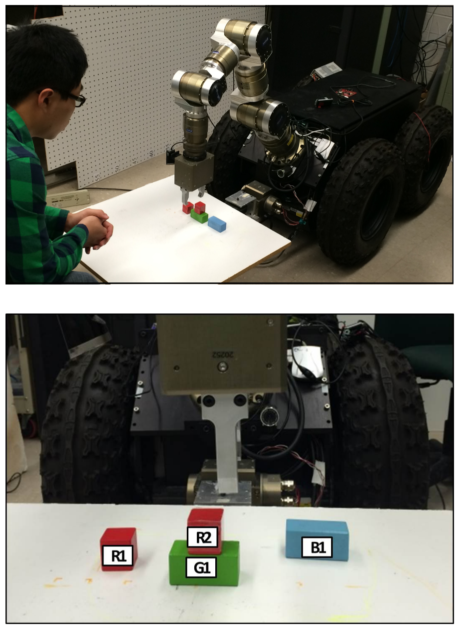
- RoboBarista: http://robobarista.cs.cornell.edu/

Fig. 10. **Examples of transferred trajectories** being executed on PR2. On the left, PR2 is able to rotate the 'knob' to turn the lamp on. In the third snapshot, using two transferred trajectories, PR2 is able to hold the cup below the 'nozzle' and press the 'lever' of 'coffee dispenser'. In the last example, PR2 is frothing milk by pulling down on the lever, and is able to prepare a cup of latte with many transferred trajectories.

Robobarista: Object Part based Transfer of Manipulation Trajectories from Crowd-sourcing in 3D Pointclouds. Sung, Jin, and Saxena. ISRR 2015.

Some Sample NLP+Robotics Papers

- Learning New Actions by Situated Human-Robot Dialogue



H1: *Stack the blue block on the red block on your right.*
R1: *How to stack, could you teach me?*
H2: *Grab the blue block.*
R2: *What do you mean by grab?*
H3: *Open your gripper.*
R3: *Ok.*
H4: *Move to the blue block.*
R4: *Sure.*
H5: *Close gripper.*
R5: *Alright then.*
H6: *Now you achieve the grab action.*
R6: *Ok, got it.*
H7: *Move the blue block to the top of the right red block.*
R7: *Understood.*
H8: *Open gripper.*
R8: *Sure.*
H9: *Ok, now you are done with the stack action.*
R9: *Ok.*

Figure 1: An example setup and dialogue. Objects are marked with labels only for the illustration purpose.

Back to the Blocks World: Learning New Actions through Situated Human-Robot Dialogue. She, Yang, Cheng, Jia, Chai, Xi. SigDial 2014.

Some Sample NLP+Robotics Papers

- Learning To Interpret Language Commands via Dialogue

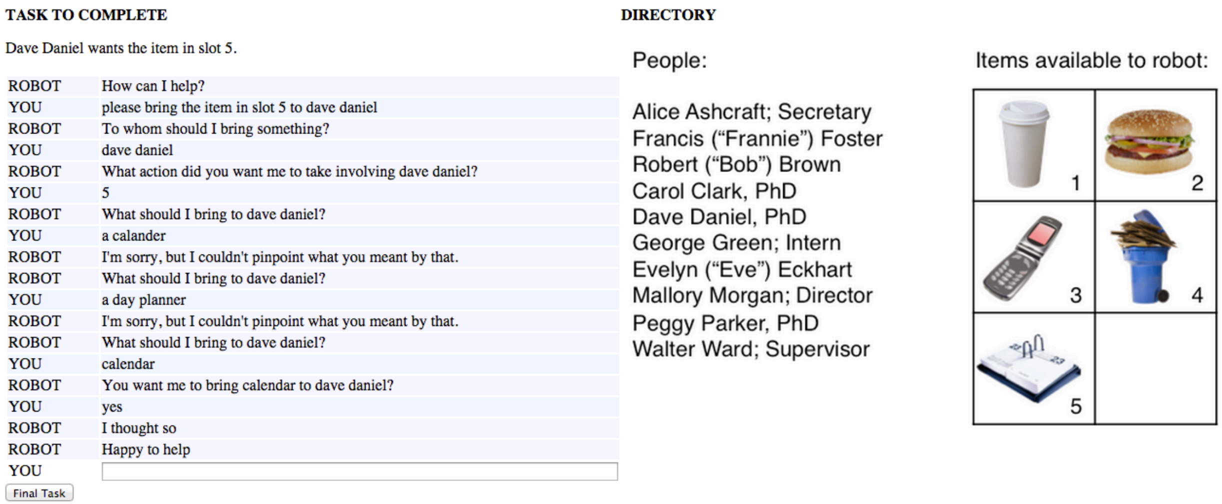


Figure 3: The Mechanical Turk interface for the delivery task. This abridged conversation is from a Turker in training batch 0, when the system had access to only the seed lexicon. Because of this conversation, the agent learned that "calender" and "planner" mean "calendar" during retraining.

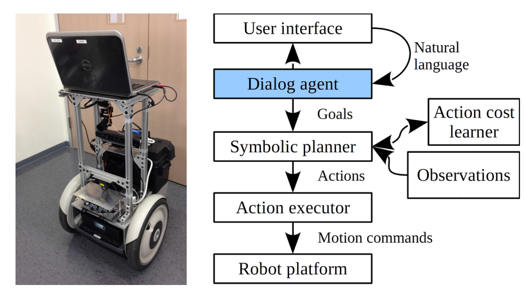


Figure 5: Left: Robot platform (Segbot) used in experiments. **Right:** Segbot architecture, implemented using Robot Operating System (ROS).

Learning to Interpret Natural Language Commands through Human-Robot Dialog. Thomason, Zhang, Mooney and Stone. IJCAI 2015.

Some Sample NLP+Robotics Papers

- Learning Language Games Through Interaction

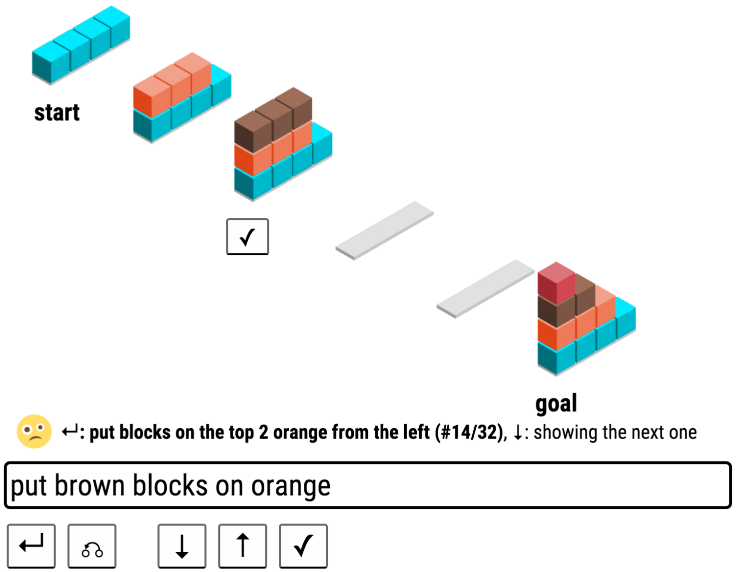
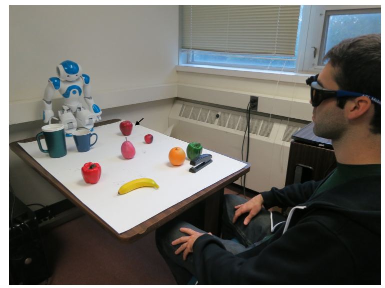


Figure 1: The SHRDLURN game: the objective is to transform the start state into the goal state. The human types in an utterance, and the computer (which does not know the goal state) tries to interpret the utterance and perform the corresponding action. The computer initially knows nothing about the language, but through the human's feedback, learns the human's language while making progress towards the game goal.

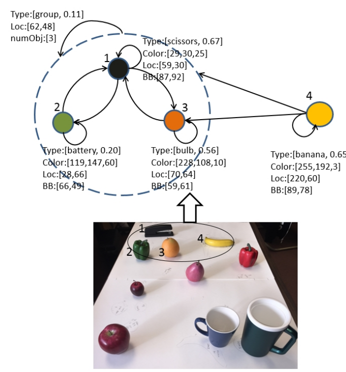
Learning Language Games through Interaction. Wang, Liang, and Manning. ACL 2016.

Some Sample NLP+Robotics Papers

- Collaborative Referring Expression Generation by Dialogue



(a) An example of situated setup.



(b) An example of the robot's internal representation of the shared environment.

R1:	"Do you see a group of two on the left?" *(Robot points to the group)*
H1:	"Yes" *(The group becomes the landmark)*
R2:	"That group I was just talking about, do you see an object in the back that is on the right within that group?" *(Describe the intended object in relation to the landmark)*
H2:	"No" *(Human rejects the description)*
R3:	"That group I was just talking about, do you see an object on the right that is to the right of that group?" *(Robot chooses another description strategy)*
H3:	"Yes"

(c) An example of collaborative referential process.

Collaborative Models for Referring Expression Generation in Situated Dialogue. Fang, Doering, and Chai. AAAI 2014.

Next Week

- Website will be updated with initial paper list very soon
- Once paper list is up, start volunteering for paper presentation!
- We might do pairs/groups of students depending on class size
- Project details and dates will be announced soon