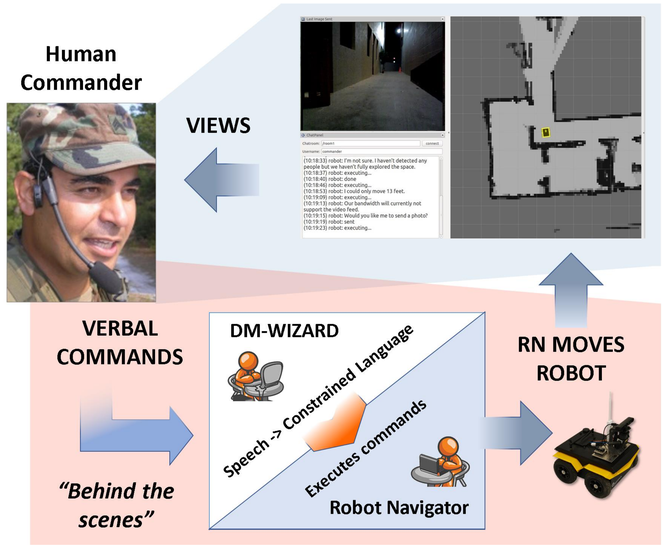
Towards Efficient Human-Robot Dialogue Collection: Moving Fido into the Virtual World

Cassidy Henry[1], Pooja Moolchandani[1], Kimberly A. Pollard[1], Claire Bonial[1], Ashley Foots[1], Cory Hayes[1], Ron Artstein[2], Clare Voss[1], David Traum[2] and Matthew Marge[1]

1 U.S. Army Research Laboratory, Adelphi, MD 20783 USA 2 USC Institute for Creative Technologies, Playa Vista, CA 90094 USA email: cassidy.r.henry.ctr@mail.mil

Introduction

- **Human language** is a natural means of communication when interacting with robots, and allows for **effective teaming**
  - allows bi-directional information exchange, with the benefit of **familiarity and flexibility** for humans
- *Focal task:* Collaborative search-and-navigation with remote human teammate and on-site robot, "Fido"
- **Multi-phase, iterative** human-robot situated dialogue collection effort, the "Bot Language" Project



- Wizard-of-Oz set-up with **two human wizards** standing in as robot AI to collect data for training an initial automated system.
  - **Dialogue Manager:** intermediary, sends typed communication
  - **Robot Navigator:** moves robot based on DM instruction
- Alternate between **physical and simulated robot**, because experiments with physical robots are much more labor intensive.

Phase	DM	RN	Enviro	Gaze/ MultiSense	# Participants
Expt 1	WoZ, Typed	WoZ, Joystick	Real	No	10
Expt 2	WoZ, GUI	WoZ, Joystick	Real	No	10
Expt 3	WoZ, GUI	WoZ, Joystick	Simulated	Yes	30+
Expt 4	Automated	Automated	Simulated	Yes	30+
Expt 5	Automated	Automated	Real	Yes	10+

- Two focused efforts to address challenges of data collection: **corpus creation** of interactions and **building simulated** experimental setup

APPROVED FOR PUBLIC RELEASE. UNCLASSIFIED

Challenges in Human-Robot Data Collection

- **Large amounts of data** are needed to capture natural variation in language as well as sufficient training data for automation
- Requires **substantial resources** of human labor, environment, time, and network/operation for **just one participant** at a time

Annotating Human-Robot Dialogue Data

- Data from Experiment 1: **basis for corpus**
  - 10 participants, ~10.5 hours of audio, and 1668 commands
- **Four message streams** *(two audio speech streams, two typed streams)* from **three speaker roles**:
  *Commander (**CMD**), Robot Navigator (**RN**), and Dialogue Manager (**DM**)*

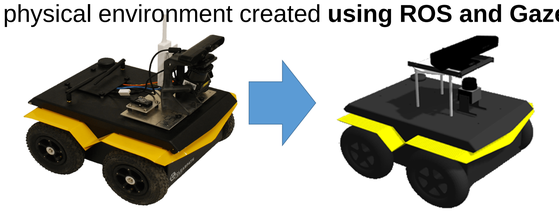
Commander	DM->CMD	DM->RN	Robot Navigator
	Which way should I turn?		
turn right forty five degrees			
		turn right 45, image	
	executing...		
			done
	sent		
go to the first orange cone on your right and send me a photo			
		move to the first orange cone on right, image	
	executing...		
			done
	sent		
are you able to move that orange cone in front of you			
	No, I'm not able to manipulate objects.		

- Initial corpus processing – **transcribing** CMD and RN speech and **alignment** of the four streams to enable analysis
- Annotations to mark **dialogue moves, structure, and relations**
- DM's typed responses from Experiment 1 **clustered and mapped** to GUI buttons for use by DM in Experiment 2:

- Utterance data currently being annotated and analyzed to **automate** dialogue management and language understanding processing modules

Bringing Fido into the Virtual World

- Simulation **reduces resource requirements** and facilitates greater data collection
- **Developed on same OS** as robot to facilitate comparisons and data transition between physical and virtual robots
  - physical environment created **using ROS and Gazebo**

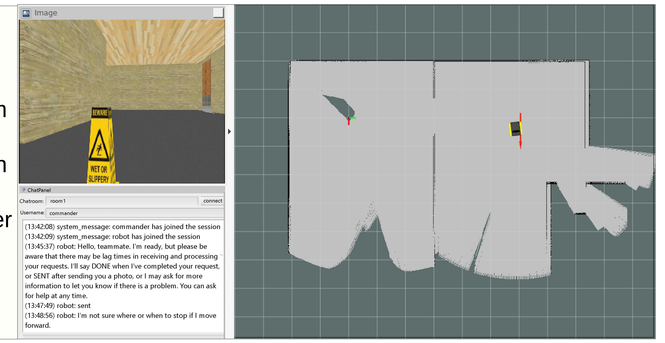


High-Fidelity Virtual *Clearpath Robotics Jackal* equipped with same sensors as physical platform

Map Building: SLAM | RGB Virtual Camera
Map Population: LIDAR | provides Robot POV

- **Point and click navigation** was included alongside virtual robot. Along with the GUI, this reduces wizard labor to one person

In simulation, CMD will perform the same tasks and see a screen using the same layout as in earlier experiments (*right*)



Advantages of Simulation

- **Freedom from resource limitations** including reduction of human labor by simplifying controls and setup
- Rapid, massive **parallel data collection** in multiple operational environments
- **Eliminates risk** of physical damage to robot or environment

Conclusion

- **Bot Language** aims to provide more natural ways for humans to interact and communicate with robots
- To facilitate this goal through our multi-phase experimental process, we seek to **iteratively move towards automation** of tasks involved in data collection and physical experiment runs
- Because our simulated robot runs on the same platform as the physical robot, we are able to **validate** simulated results in a physical environment after completed data collection