

NASSLLI @ USC 2022

Multiparty and Multi-floor dialogue structure

Lecture 4.1: Multifloor Dialogue Structure

David Traum

Institute for Creative Technologies

University of Southern California

traum@ict.usc.edu

<https://people.ict.usc.edu/~traum/>

Outline

- Annotation Review
- Multifloor Dialogue Structure Annotation Scheme
- Botlanguage Project
- Botlanguage Annotations
- Context
- “Stop!”
- Conclusions

Dialogue Structure Annotation for Multi-Floor Interaction



*David Traum, Cassidy Henry, Stephanie Lukin, Ron Artstein,
Felix Gervitz, Kimberly A. Pollard, Claire Bonial, Su Lei,
Clare R. Voss, Matthew Marge, Cory J. Hayes, Susan G. Hill*

Types of Dialogue Structure (Traum & Nakatani 1999)

Structure Content

- **Intentional**
- Linguistic
- **Relational/Rhetorical**
- Attentional State
- Turn-taking/floor management
- Grounding
- Participant structure

Structure Granularity

- Micro – within a single turn
- **Meso – short subdialogue**
- Macro – full conversation

Multi-floor Botlanguages Annotations: Meso-level Dialogue Structure

Structure Types

- Intentional:
Transaction Units – smallest unit of specified and performed action, including all dialogue needed to accomplish this
- Relational/Rhetorical :
Relations between utterances within a transaction

Annotations

- TUs: cluster of utterances
 - Not necessarily sequential
- Relations: Label 2nd part utterance with
 - Antecedent
 - Relation type

Example:

- **Customer:** I'd like a cheeseburger
- **Waiter:** one cheeseburger.
- **Waiter:** (placing burger in bag) here you go.
- **Customer:** thanks!
- **Waiter:** would you like fries with that?
- **Customer:** Sure, a large one please!
- **Waiter:** (placing fries box in bag): one large fries.



1. multi-floor dialogue: 2018 annotation schema

Transaction Unit (TU):
a group containing the
initiation and (potential)
fulfillment
of an intent.

- TU 1
1. **Customer:** I'd like a cheeseburger
 2. **Waiter:** one cheeseburger.
 3. **Waiter:** (placing burger in bag) here you go.
 4. **Customer:** thanks!
- TU 2
5. **Waiter:** would you like fries with that?
 6. **Customer:** Sure, a large one please!
 7. **Waiter:** (placing fries box in bag): one large fries.

Traum et al. 2018, LREC



1. multi-floor dialogue: 2018 annotation schema

Relations: describe the structure between pairs of utterances within a TU

- **Antecedent:** the utterance that a subsequent utterance is addressing (e.g., 2 addressing 1)
- **Relation-Type:** relationship between utterance and antecedent (e.g., Acknowledgment)

1. **Customer:** I'd like a cheeseburger
2. **Waiter:** one cheeseburger. Acknowledgement
3. **Waiter:** (placing burger in bag) here you go. Acknowledgement
4. **Customer:** thanks! 3rd turn feedback
5. **Waiter:** would you like fries with that? Answer
6. **Customer:** Sure, large please!
7. **Waiter:** (placing fries in bag): one large fries. Acknowledgement

Traum et al. 2018, LREC

Floor and Participant Structure

Participants and Floors

- Single floor Dyadic (A,B)
- Single floor Multiparty: (A,B,C,...)
- Multiple floors (with different sets of participants): {(A B), (C D E)}

Interactions between Floors

- Same purpose, distinct participants
- Co-located, observable
 - Participants play different roles for different floors (e.g. active participant vs overhearer)
- Some Shared participant(s)
 - multi-communicating (Rentch et al)
- **Multi-floor dialogue:**
 - Same purpose
 - Some Multi-communicating participant(s)
 - Content flows across floors



1. Multi-floor dialogue: introduction

Conversational floor: shares common set of speakers and observers

Multi-floor Dialogue: high-level dialogue purposes are the same, and some content is shared, but other aspects (participant structure, turn-taking expectations) are distinct



Time
↓

I D	Conversational Floor 1		Conversational Floor 2		
	Woman	Server	Server	Cook	Drink Server
1	I'll have a cheeseburger and a small coke				
2		Ah no ah, no coke, pepsi			
3	pepsi				
4			One cheeseburger one pepsi		
5				cheeseburger	
6					pepsi

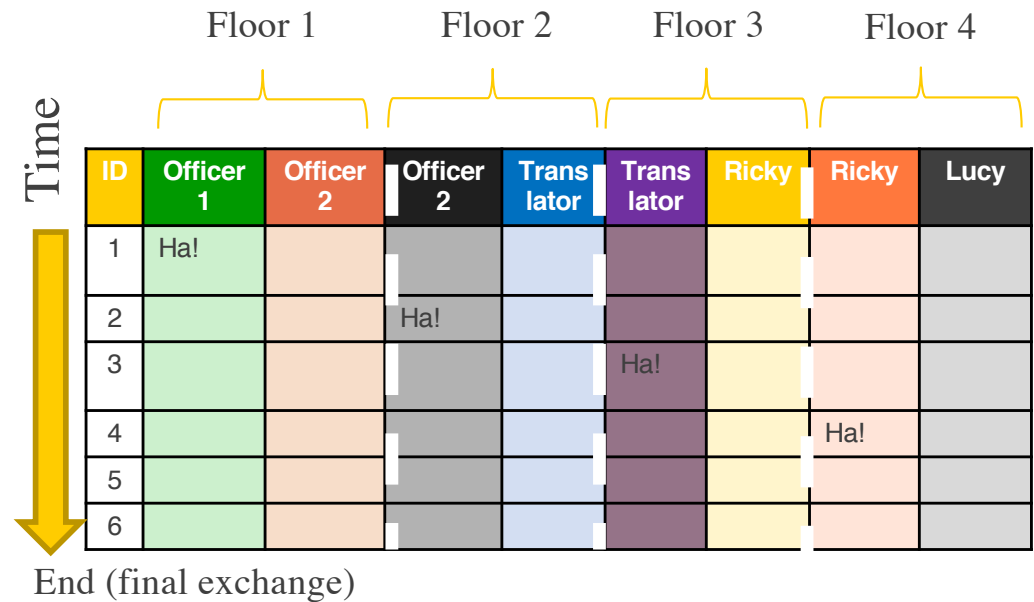
Traum et al. 2018, LREC



1. Multi-floor dialogue: introduction

Conversational floor: shares common set of speakers and observers

Multi-floor Dialogue: high-level dialogue purposes are the same, and some content is shared, but other aspects (participant structure, turn-taking expectations) are distinct



Traum et al. 2018, LREC



1. multi-floor dialogue: 2018 annotation schema

Relation Super-Types

Expansions - relate utterances produced by the *same* participant within the *same* floor (4 Subtypes)

Responses - relate utterances by *different* participants in the *same* floor (24 Subtypes)

Translations - relate utterances in *different* floors (4 Subtypes)

Example Subtypes:

1. Customer→Waiter: I'll have a cheeseburger
2. Customer→ Waiter: and a small coke *Continue*

1. Customer→ Waiter: a small coke, please
2. Waiter→Customer: no coke, pepsi *Ack cant*

1. Customer→ Waiter: I'll have a cheeseburger
2. Waiter→Cook: Cheeseburger!! *Translation-right*

Traum et al. 2018, LREC

Relations by type (1)

Expansions

- a) Continue
- b) (self-) Correction
- c) Link-next
- d) Summarization

Translation

- a) Translation<from,to>
- b) Partial
- c) Quotation
- d) Comment

Relations by type (2) Responses

- a. **Processing:** positive feedback at perception level
- b. **acknowledgement:** positive feedback of understanding
- c. **clarification:** negative feedback of understanding
- d. **question-response**
- e. **reciprocal response:** e.g. "hello" -> "hello"
- f. **3rd turn feedback:** response to feedback
- g. **other**

Response sub-relations

acknowledgment

- ack-done
- ack-doing
- ack-wilco
- ack-understand
- ack-try
- ack-unsure
- ack-cant

clarification

- req-clar
- clar-repair
- missing info
- nack
- req-repeat
- clar-repeat

question-response

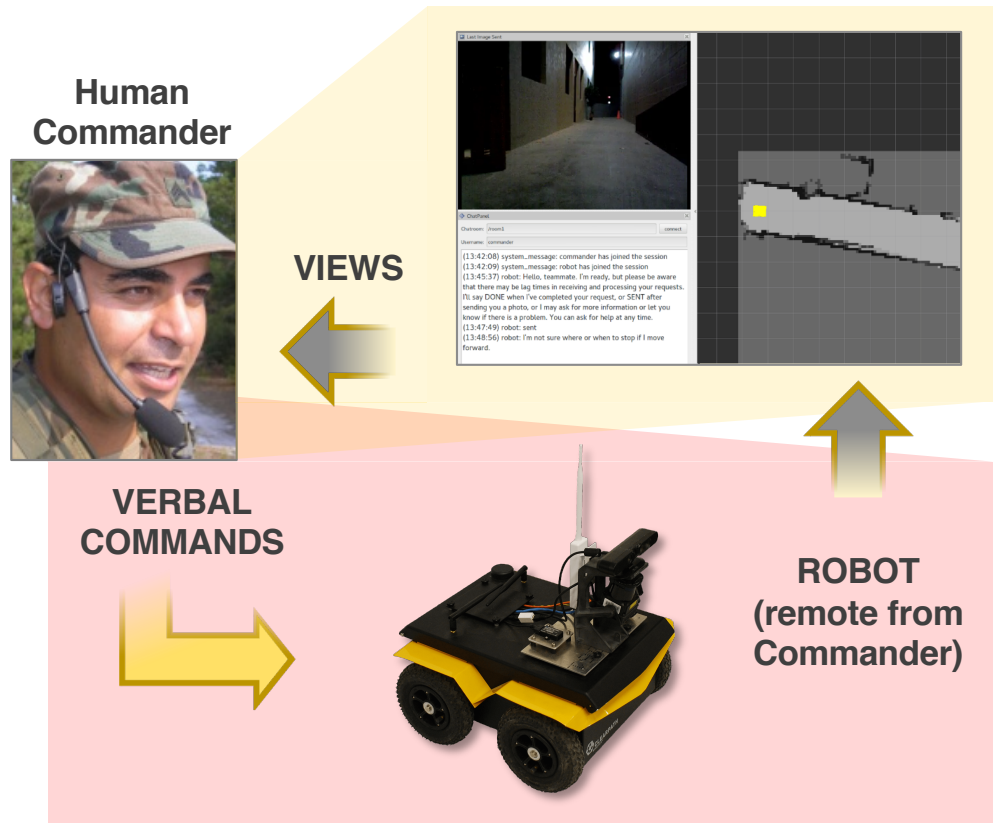
- answer
- Non-Answer-Response (NAR)

Domain: Human-Robot Collaboration



Remote reconnaissance task

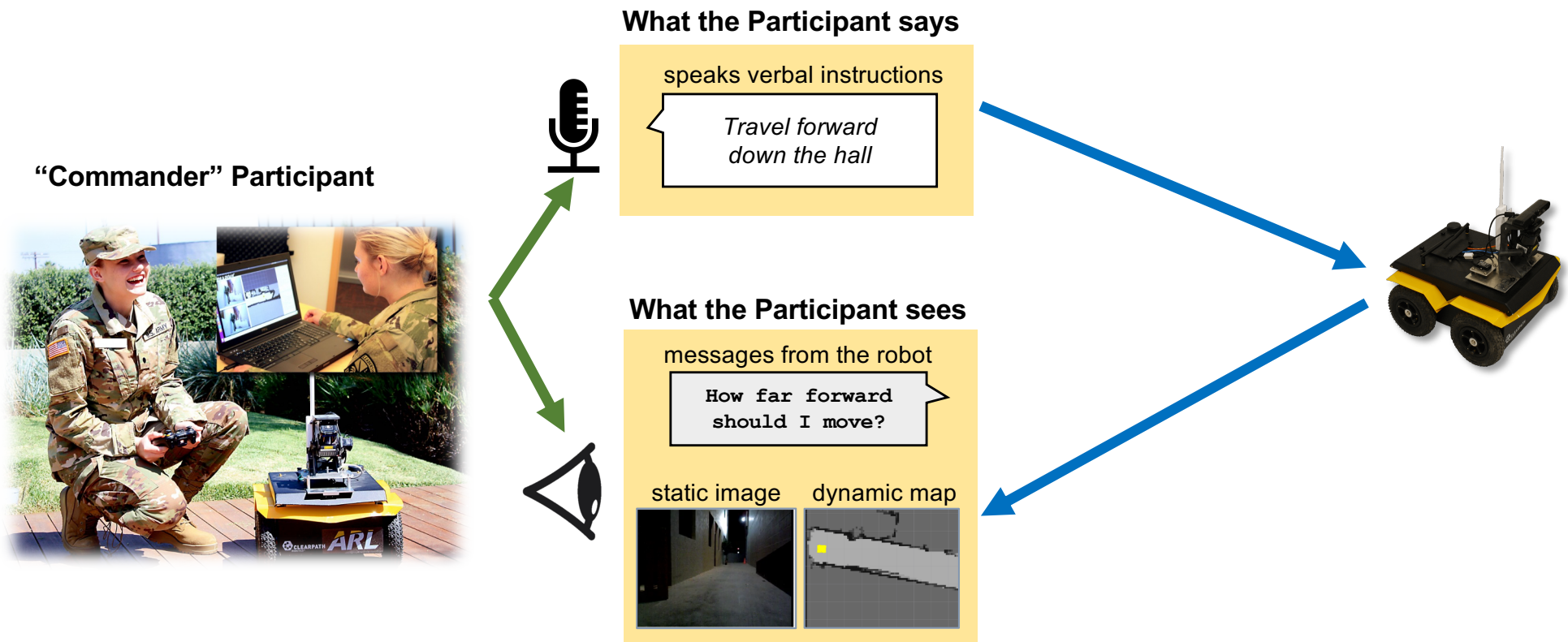
- Unfamiliar environment
- Bandwidth limitations
- User and robot not co-present
- What would the human users want to say?
 - Need to collect a corpus in order to train and evaluate the system.
- How would users naturally collaborate with this robot teammate?



(Marge et al., 2016, IEEE RO-MAN)

16

2. situated dialogue: human-robot dialogue corpus



*Situated Corpus of Understanding Transactions:
Marge et al., RO-MAN 2016*



2. situated dialogue: human-robot dialogue corpus

last photo sent by Robot

The screenshot shows a chat window with a photo of a robot and a LIDAR map. The photo shows a white robot with a yellow sensor on top. The LIDAR map shows a grid with a yellow square indicating the robot's position. The chat window shows a list of messages:

```
Chatroom: /room1 connect
Username: commander
(11:40:38) robot: executing...
(11:40:45) robot: done
(11:41:28) robot: executing...
(11:41:40) robot: done
(11:43:00) robot: executing...
(11:43:13) robot: done
(11:43:20) robot: sent
(11:43:34) robot: executing...
(11:43:42) robot: sent
(11:43:54) robot: executing...
(11:44:13) robot: sent
(11:44:47) robot: executing...
(11:44:49) robot: sent
```

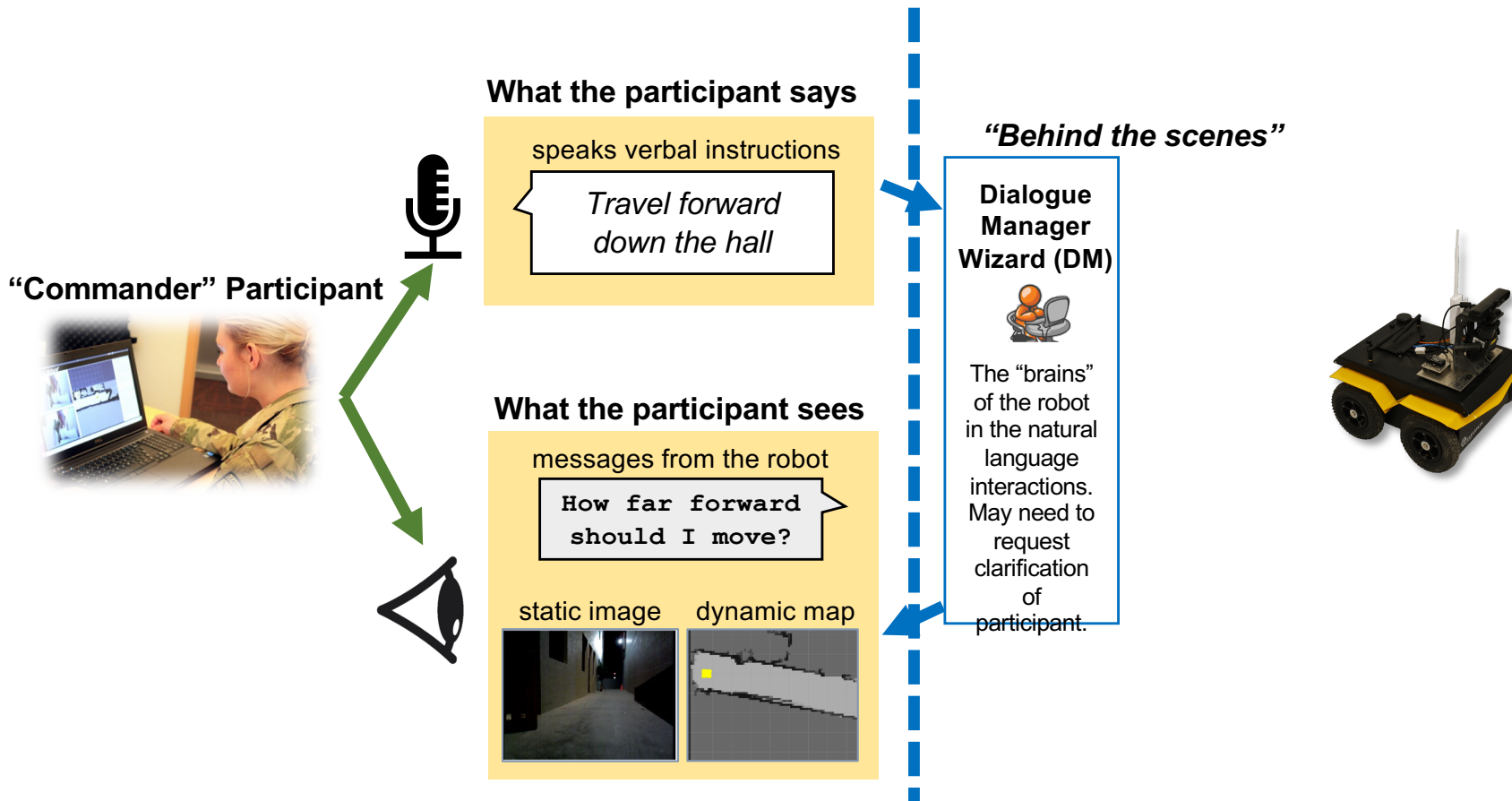
Robot's replies

Robot's LIDAR map of searched area
(LIDAR: Light Detection And Ranging)



WHAT THE PARTICIPANT SEES

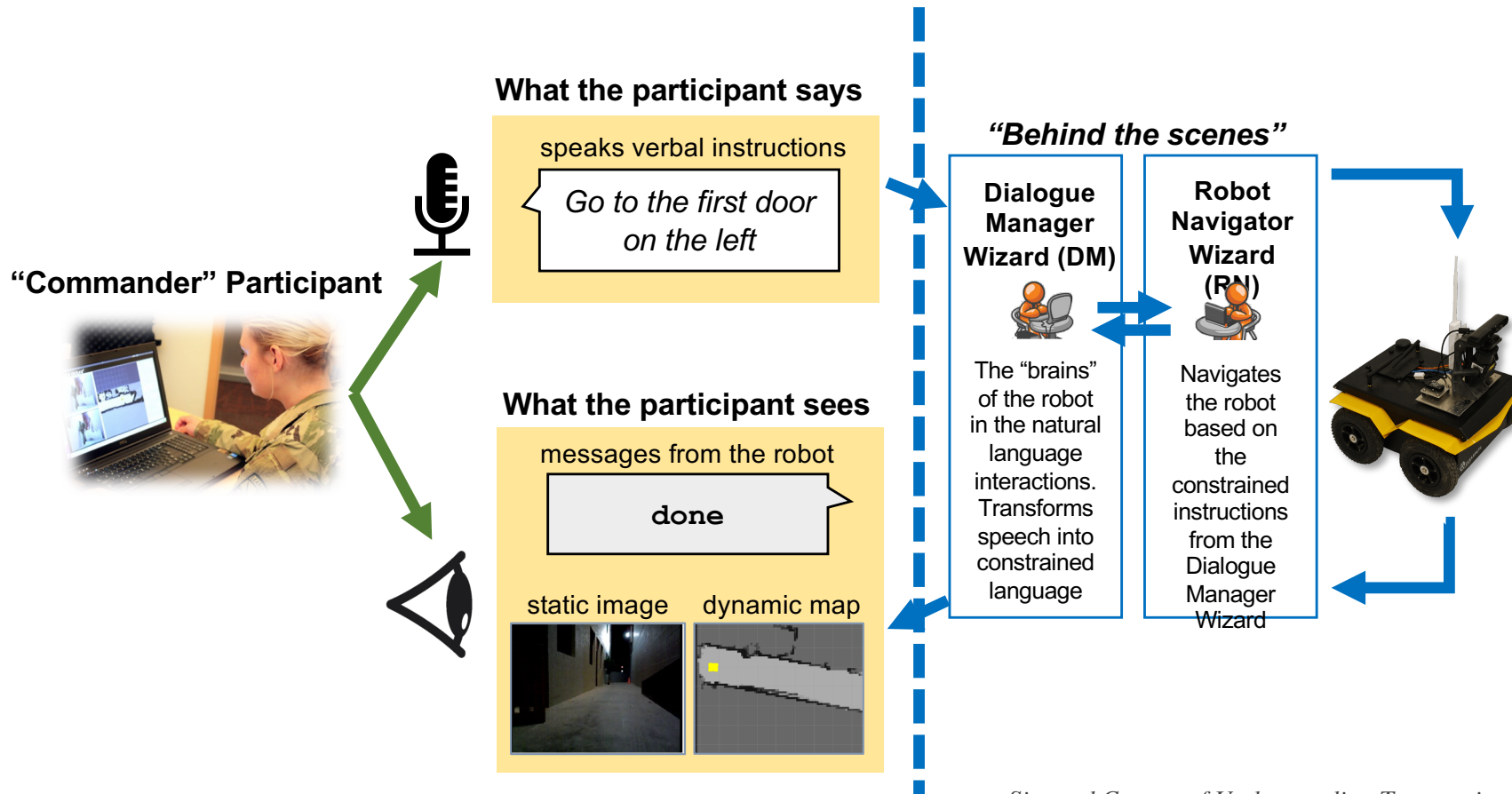
2. situated dialogue: human-robot dialogue corpus



*Situated Corpus of Understanding Transactions:
Marge et al., RO-MAN 2016*



2. situated dialogue: human-robot dialogue corpus



*Situated Corpus of Understanding Transactions:
Marge et al., RO-MAN 2016*



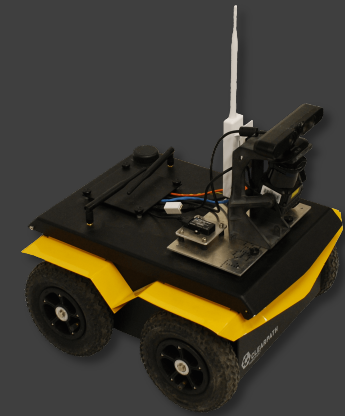
Example Interaction



DM



RN

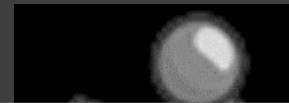




Commander



CMD

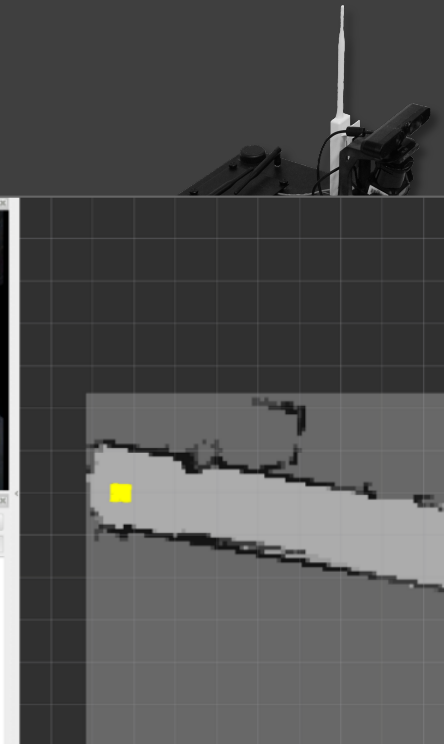


ChatPanel

Chatroom: /room1

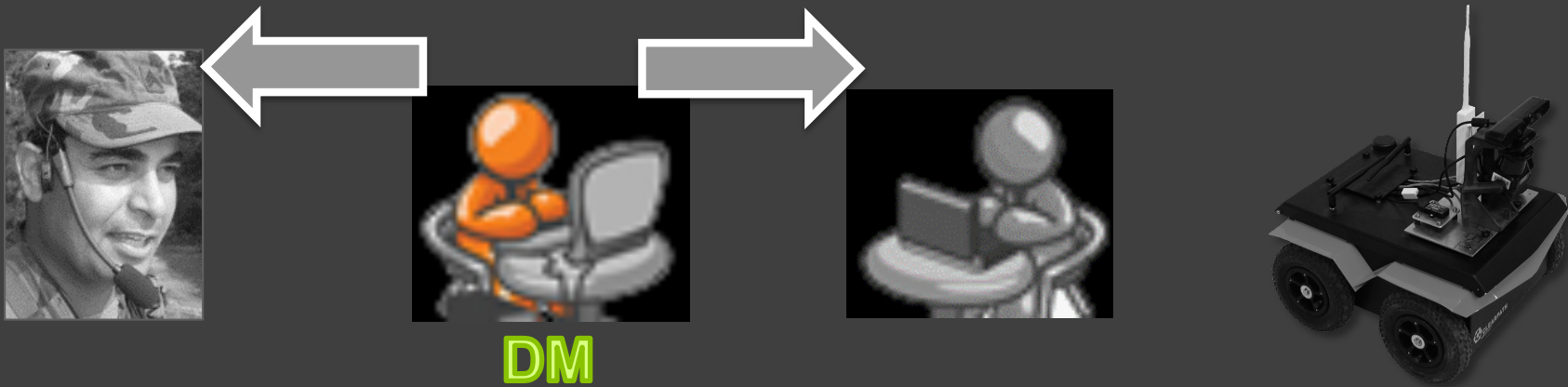
Username: commander

(13:42:08) system_message: commander has joined the session
(13:42:09) system_message: robot has joined the session
(13:45:37) robot: Hello, teammate. I'm ready, but please be aware that there may be lag times in receiving and processing your requests. I'll say DONE when I've completed your request, or SENT after sending you a photo, or I may ask for more information or let you know if there is a problem. You can ask for help at any time.
(13:47:49) robot: sent
(13:48:56) robot: I'm not sure where or when to stop if I move forward.





Wizard #1 – Dialogue Manager

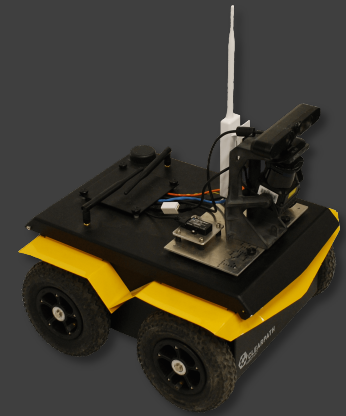
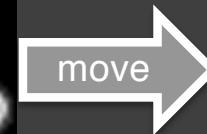




Wizard #2 – Robot Navigator

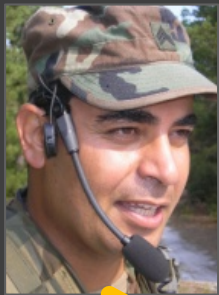


RN





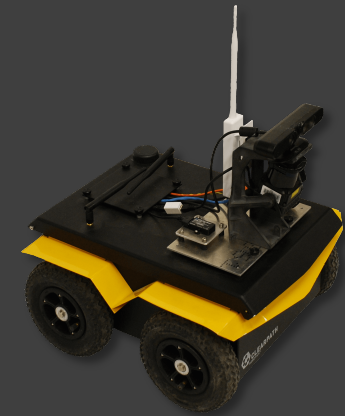
Example Interaction



DM



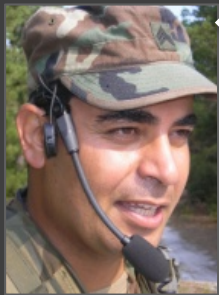
RN



Proceed forward



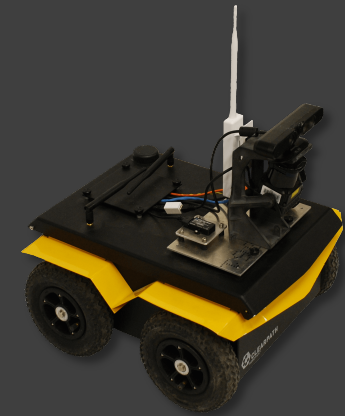
Example Interaction



DM



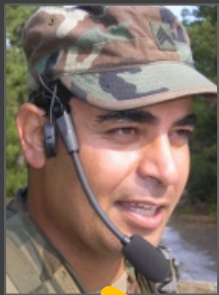
RN



How far? You can tell me to move to an object that you see, or a distance



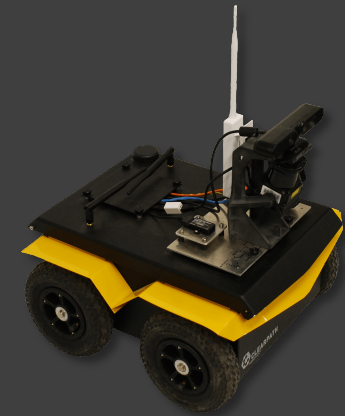
Example Interaction



DM



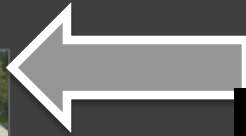
RN



Proceed forward
three feet



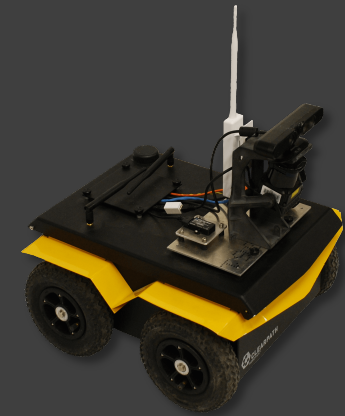
Example Interaction



DM



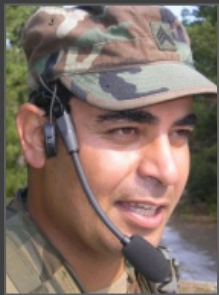
RN



Executing...



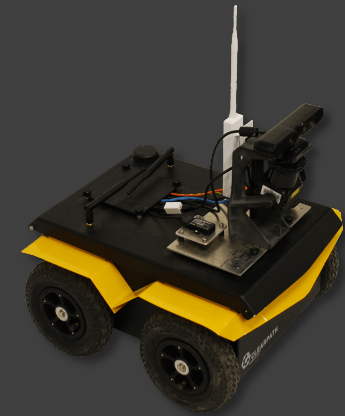
Example Interaction



DM



RN



move forward
three feet



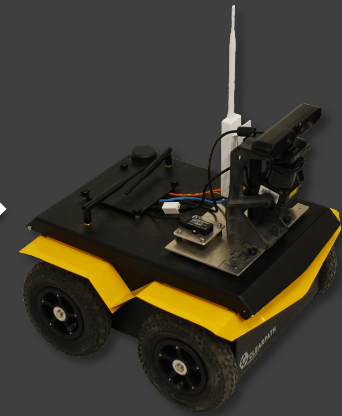
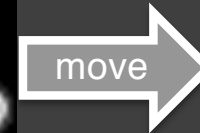
Example Interaction



DM



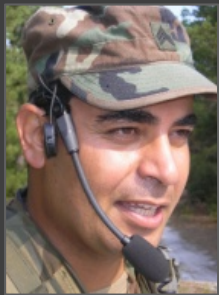
RN



moves robot forward 3 feet



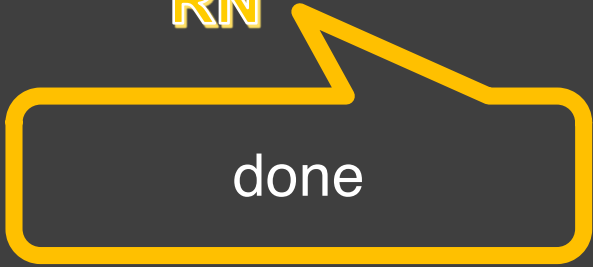
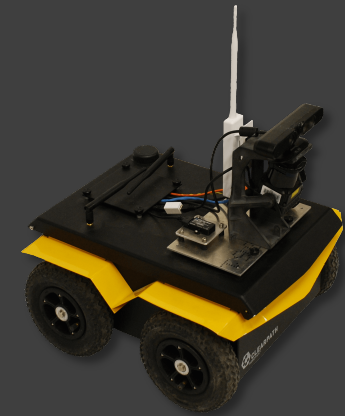
Example Interaction



DM

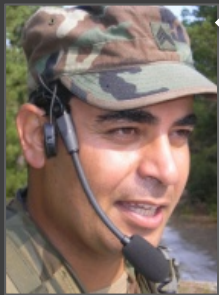


RN





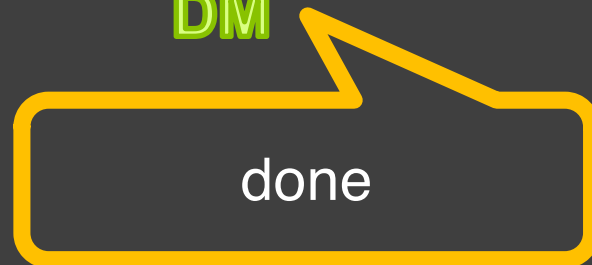
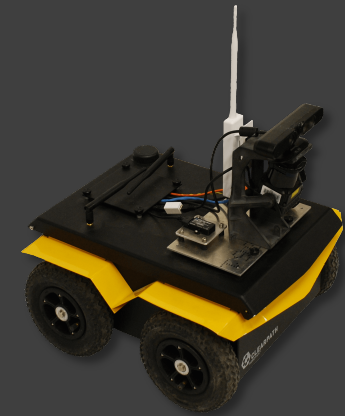
Example Interaction



DM



RN



done



Data - Transcripts

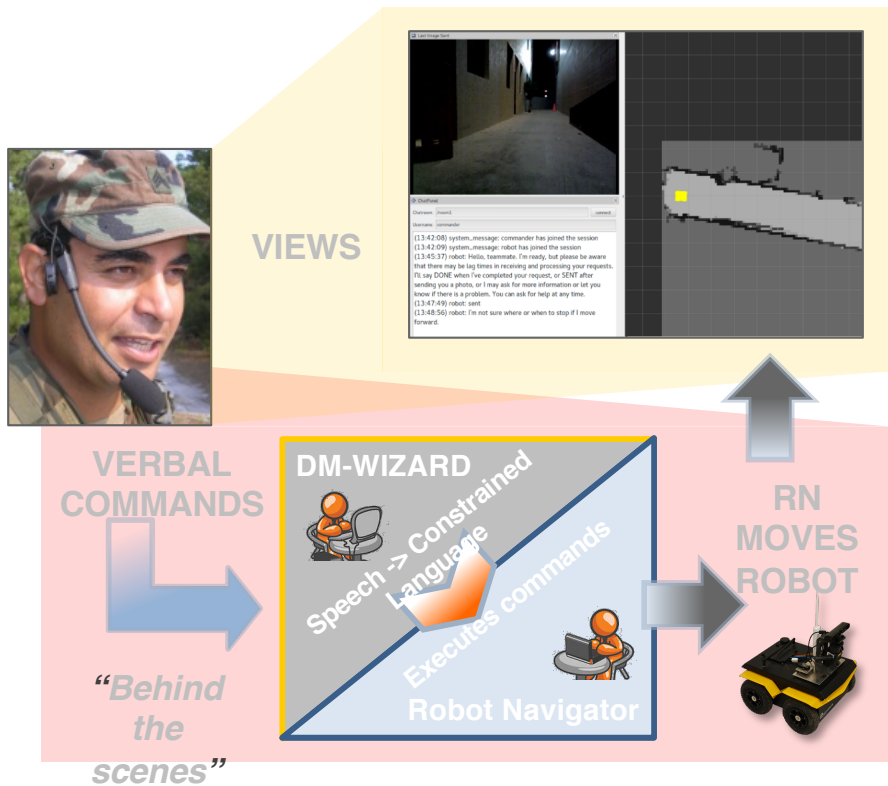
- Time aligned transcripts of 4 data streams
 - 2 audio streams
 - CMD and RN
 - 2 text streams
 - DM->CMD, DM->RN
- Two conversational floors present

Commander (Audio Stream 1)	DM->Commander (Chat Room 1)	DM->RN (Chat Room 2)	RN (Audio Stream 2)
face the <u>doorway</u> on your right			
and take a picture			
	there's a door ahead of me on the right and one just behind me on the right. which would you like me to face?		
the door ahead of you on the right			
		move to face the door ahead of you on the right, image	
	executing...		
			image sent
	sent		





Multifloor Setup



Commander (Audio Stream 1)	DM->Commander (Chat Room 1)	DM->RN (Chat Room 2)	RN (Audio Stream 2)
face the doorway on your right			
and take a picture			
		there's a door ahead of me on the right and one just behind me on the right. which would you like me to face?	
the door ahead of you on the right			
		move to face the door ahead of you on the right, image	
	executing...		
	sent		image sent



Left floor: CMD, DM



Commander (Audio Stream 1)	DM->Commander (Chat Room 1)	DM->RN (Chat Room 2)	RN (Audio Stream 2)
face the doorway on your right			
and take a picture			
there's a door ahead of me on the right and one just behind me on the right. which would you like me to face?			
the door ahead of you on the right			
		move to face the door ahead of you on the right, image	
	executing...		
	sent		image sent

Right Floor: DM, RN



Commander (Audio Stream 1)	DM->Commander (Chat Room 1)	DM->RN (Chat Room 2)	RN (Audio Stream 2)
face the doorway on your right			
and take a picture			
there's a door ahead of me on the right and one just behind me on the right. which would you like me to face?			
the door ahead of you on the right			
		move to face the door ahead of you on the right, image	
	executing...		
	sent		image sent



DM translates (to) left and right



Commander (Audio Stream 1)	DM->Commander (Chat Room 1)	DM->RN (Chat Room 2)	RN (Audio Stream 2)
face the doorway on your right			
and take a picture			
	there's a door ahead of me on the right and one just behind me on the right. which would you like me to face?		
the door ahead of you on the right			
		move to face the door ahead of you on the right, image	
	executing...		
	sent		image sent



2. situated dialogue: human-robot dialogue corpus



move forward

behind the scenes...

go forward 3 feet

processing. . .

You can tell me to move a certain distance or to move to an object

processing. . .

move forward 3 feet

moving. . .

done

done



2. situated dialogue: human-robot dialogue corpus



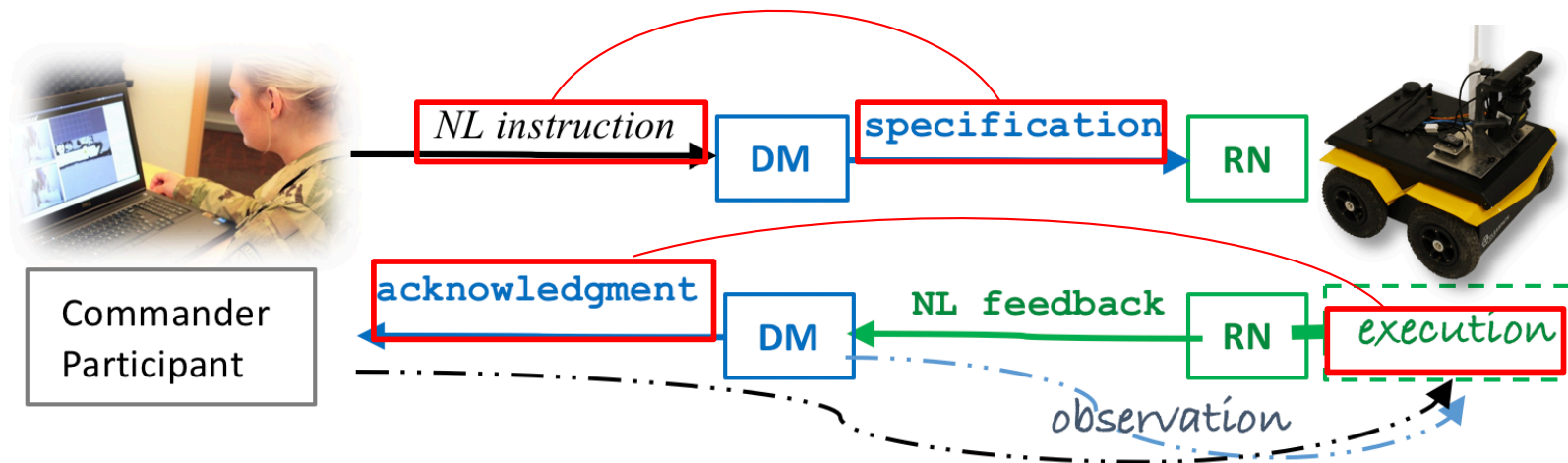
left floor right floor

ID	Participant (Audio Stream 1)	DM -> Participant (Chat Room 1)	DM-> RN (Chat Room 2)	RN (Audio Stream 2)
1	move forward			
2		processing...		
3		You can tell me to move a certain distance or to move to an object		
4	go forward 3 feet			
5		processing...		
6			move forward 3 feet	
7		moving...		
8				done
9		done		
10	what do you see			
11			send image	
12				sent
13		sent		

Time



2. situated dialogue: training data strengths and weaknesses



- Training data establishes bi-directional associations between NL, execution behavior



2. situated dialogue: training data strengths and weaknesses

The screenshot displays the ScoutBot demo interface, divided into three main sections:

- Commander ASR Transcript:** Shows a conversation log with a question: "Anybody asks Commander".
- Dialogue Manager:** Displays a list of responses matching the answer words, including actions like "send image", "turn to face", "move forward", and "turn right/left".
- Dialogue Management Utterances to Commander:** A large empty box for the Commander's output.
- Dialogue Management Utterances to Navigation:** A large empty box for the Navigation system's output.

Labels on the right side of the interface indicate the floor and system components:

- Left Floor: Commander - DM
- Right Floor: DM - RN

At the bottom of the interface, a text box reads: "Dialogue Management Classifier output to Commander and Navigation".

ScoutBot demo: Lukin et al, ACL 2018

Corpus Statistics



Basics

- **60** dialogues
 - **20** participants
 - **3** dialogues each
 - **~20** hours
- **11454 Total Utterances**
 - **3,573** from commanders
 - **5,154** from DM
 - **2,727** from RN

Dialogue Structure Annotations

- **2,230** Transaction Units
- **11,058** Relations
- **644** Unique TU Tree structures
 - Classified into **5** types

Frequent Relations



Type	Subtype	#	%
Translation		4282	39
	Translate-r	2355	21
	Translate-l	1911	17
	comment	21	<1
Expansion		1583	14
	Continue	1175	11
	Link-next	337	3
	correction	50	<1
	summarize	20	<1

Type	Subtype	#	%
Response		5193	47
	acknowledge	3998	36
	clarification	569	5
	processing	315	3
	Question-response	212	2
	other	48	<1
	3 rd turn feedback	37	<1
	reciprocal	14	<1

Structural Types of Transaction Units (TUs)



- **Minimal TU:** single instruction, acks, no repair
- **Extended-Link TU:** multiple instructions, with expansions
- **Repair TU:** contains at least one repair
 - successfully resolved or
 - abandoned
- **QA TU:** starts with question & response rather than instruction
 - simple question,
 - later instruction
- **Other TU:** none of the above (e.g. no response or translation)

Example minimal TU



Utt #	Left Floor		Right Floor		Annotations		
	Commander	DM→CMD	DM→RN	RN	TU #	Antecedent	Relation
1	move forward three feet				1		
2		ok			1	1	ack-wilco
3			move forward 3 feet		1	1	translation-r
4				done	1	3	ack-done
5		I moved forward 3 feet			1	4	translation-l

Example Extended-Link TU



Utt #	Left Floor		Right Floor		Annotations		
	Commander	DM→CMD	DM→RN	RN	TU	Ant	Rel
1	face west				1		
2	and take a photo				1	1	continue
3			face west, photo		1	2*	translation-r
4		executing...			1	2*	ack-doing
5				image sent	1	3	ack-done
6		sent			1	5	translation-l

Example Q&A TUS



Utt #	Left Floor		Right Floor		Annotations		
	Commander	DM→Commander	DM→RN	RN	TU	Ant	Rel
1	how many window openings do you see in front of you				1		
2		three			1	1	answer
3	do you see a yellow flashlight				2		
4		processing...			2	3	processing
5		I'm not sure			2	3	answer
6		If you describe an object, you can help me to learn what it is.			2	3	non-answer response

Example Other TUS



Utt #	Left Floor		Right Floor		Annotations		
	Commander	DM→Commander	DM→RN	RN	TU	Ant	Rel
1	i'm ready				1		
2		I'm also ready			1	1	Reciprocal-response
3		Would you like me to send a picture so you can see the room?"			2		



Frequency of TU Structures (% of corpus)

- Minimal TU (48%)
- Extended-Link TU (26%)
- Repair TU (11%)
 - 9% successfully resolved
 - 2% abandoned
- QA TU (~5%)
 - 4% simple question
 - 1% lead to instruction
- Other TU (11%)

Applications of Annotated Data



- Examination of Dialogue Structure Overlap (Henry et al WiNLP 2018)
- Stylistic differences across individuals and conditions (Lukin et al Sigdial 2018)
- Automating NLU and dialogue management (Gervits et al ACL 2018 Demo)



USC Institute for
Creative Technologies



U.S. ARMY COMBAT CAPABILITIES DEVELOPMENT COMMAND –
ARMY RESEARCH LABORATORY

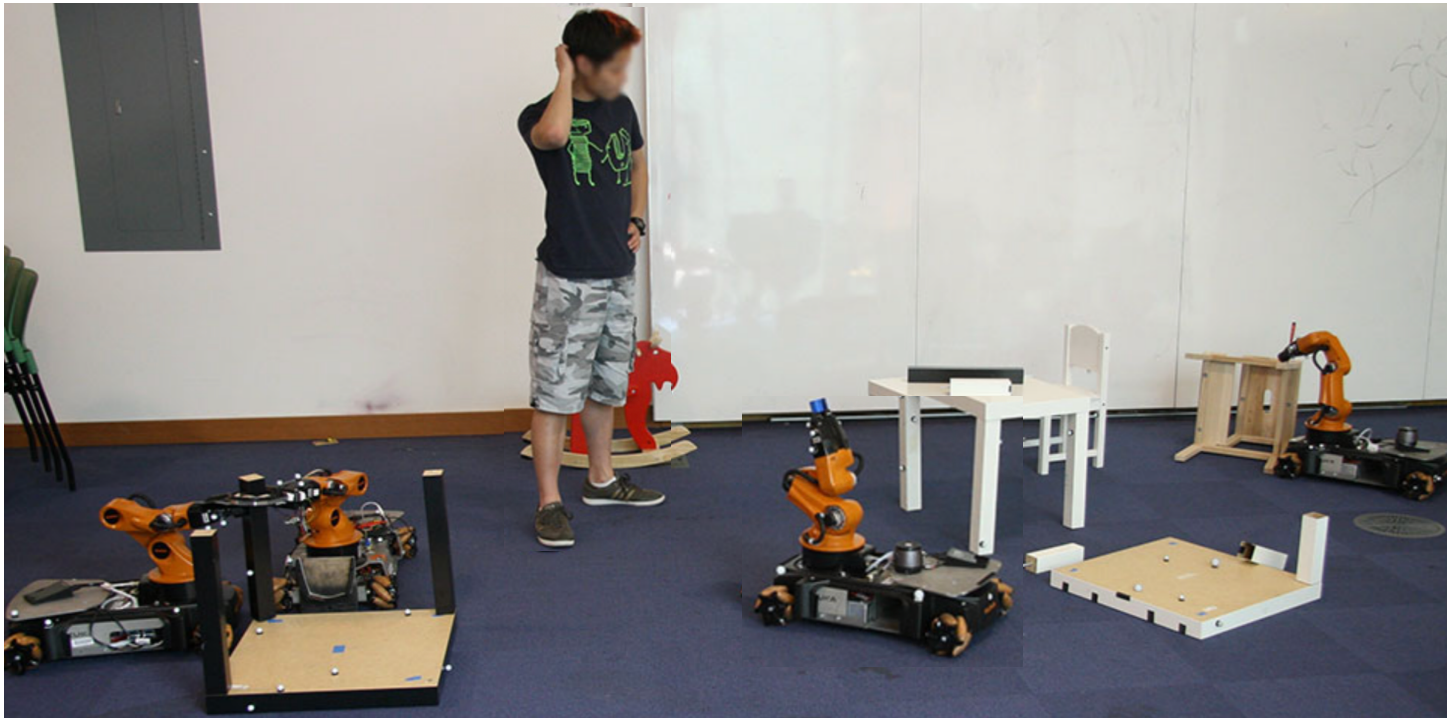
Context Is Key: Annotating Situated Dialogue Relations in Multi-floor Dialogue

- Claire Bonial¹,
- Mitchell Abrams², Anthony L. Baker³, Taylor Hudson³, Stephanie M. Lukin¹,
- David Traum⁴ & Clare R. Voss¹
- ¹ U.S. Army DEVCOM ARL
- ² Institute for Human and Machine Cognition (IHMC)
- ³ Oak Ridge Associated Universities (ORAU)
- ⁴ USC Institute for Creative Technologies



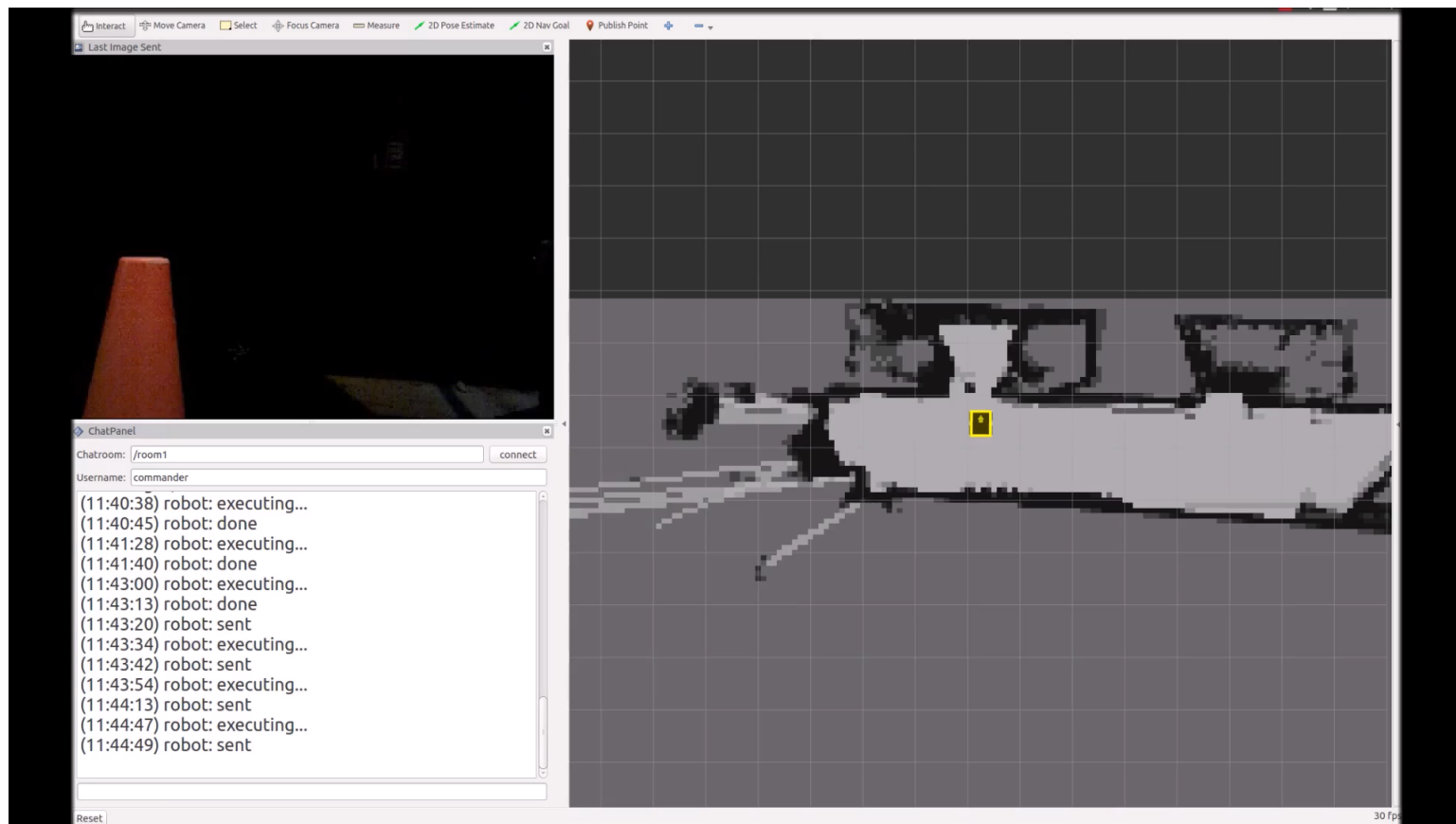
2. situated dialogue: human-robot dialogue corpus

Goal: Extend our *multi-floor* dialogue annotation schema to account for features of *situated* dialogue—interpretation draws upon info from physical environment, conversational history, robot’s physical form, etc.



Picture Credit: Knepper et al. 2015

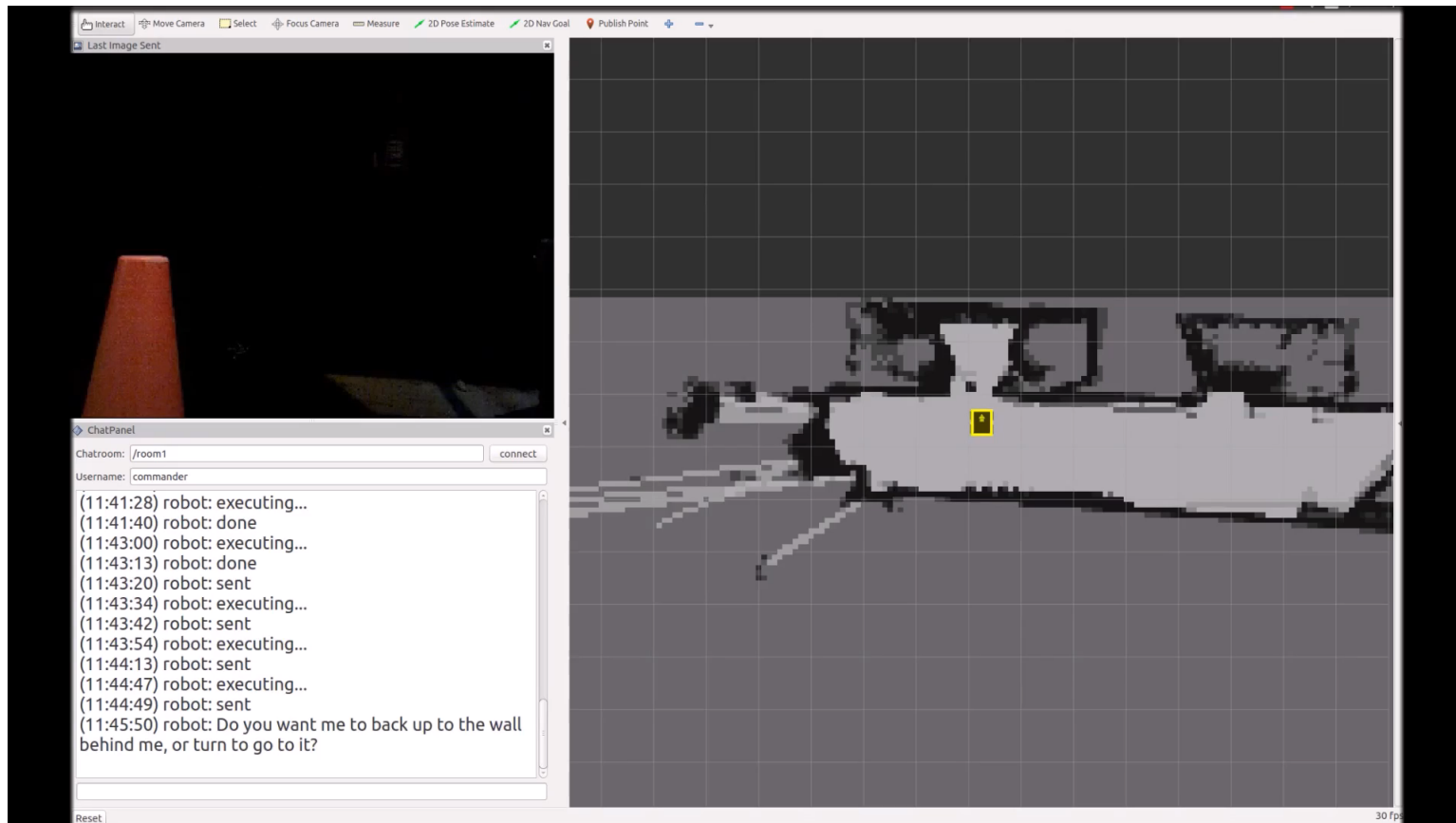
2. situated dialogue: human-robot dialogue corpus



“go to the wall behind you, face north and then take a picture”



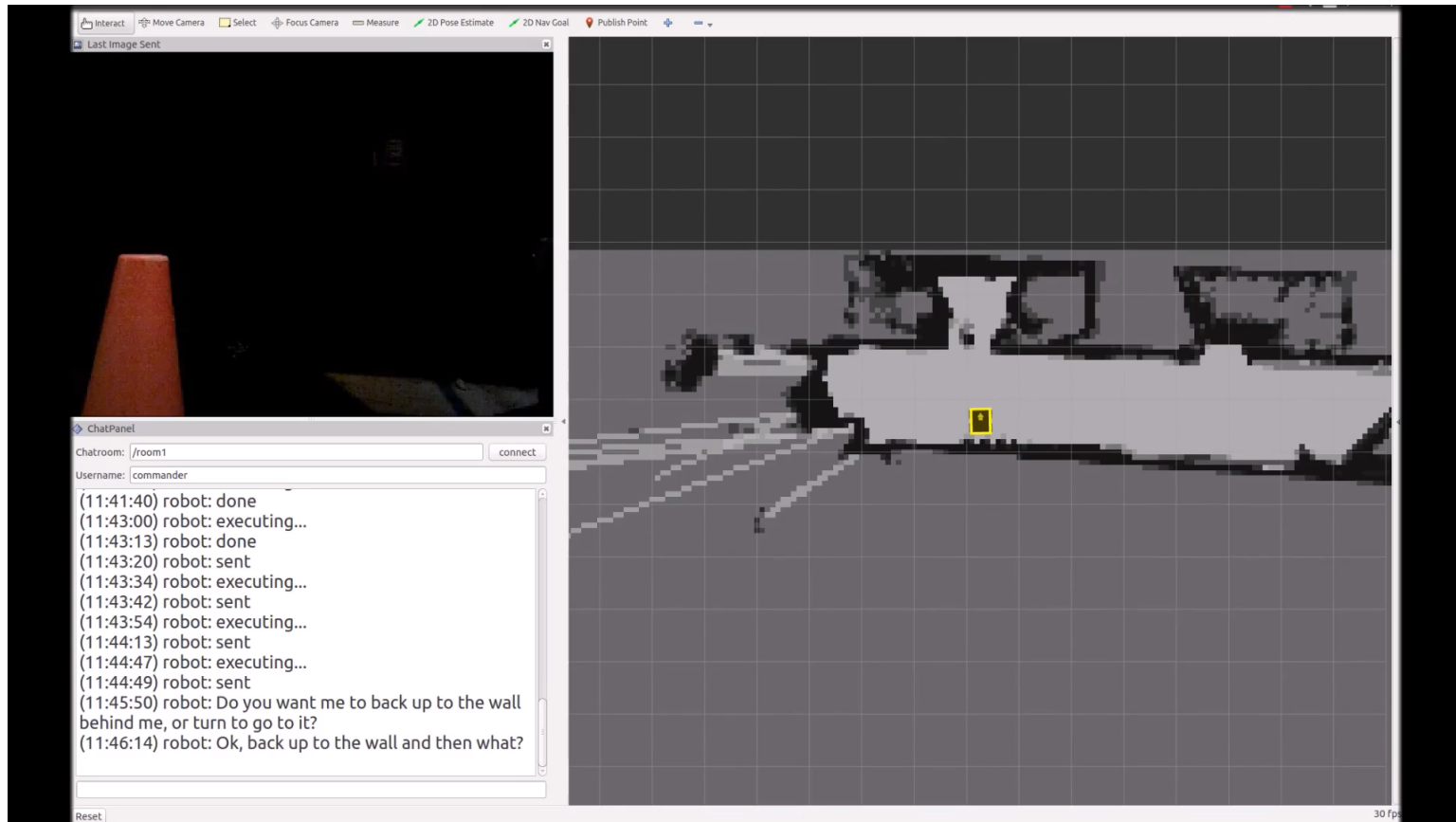
2. situated dialogue: human-robot dialogue corpus



“back up”



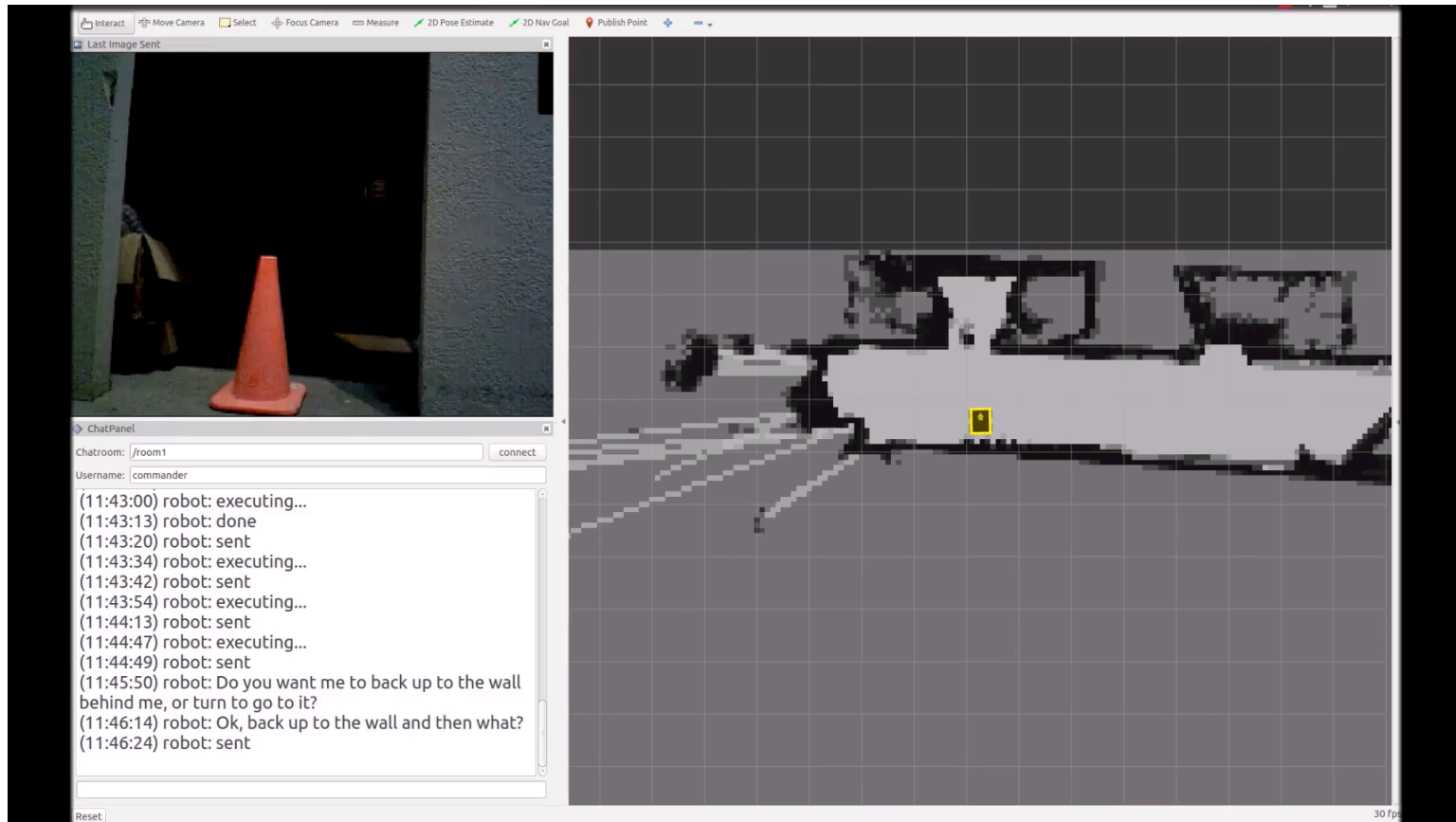
2. situated dialogue: human-robot dialogue corpus



“take a picture”



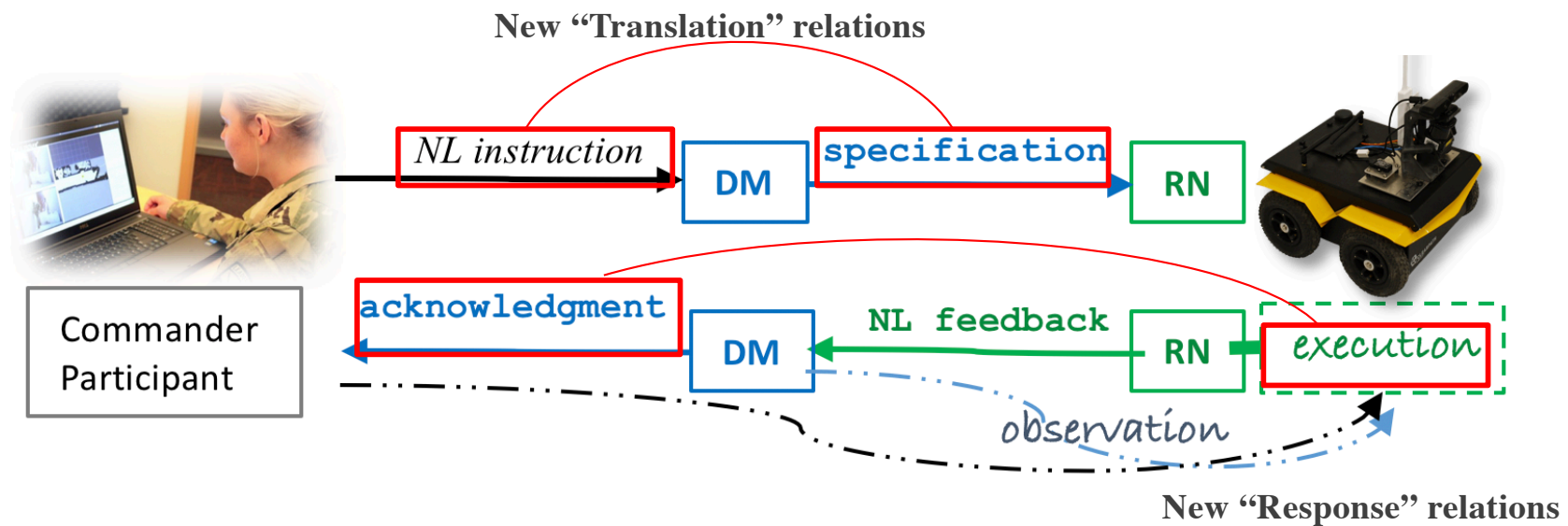
2. situated dialogue: human-robot dialogue corpus



“go to the other door”



2. situated dialogue: training data strengths and weaknesses



- Some **associations** between NL, execution behavior are only valid in particular certain **situated** contexts

Modifications to annotation schema



Relation Super-Types

Expansions - relate utterances produced by the *same* participant within the *same* floor (4 Subtypes)

Responses - relate utterances by *different* participants in the *same* floor (24 Subtypes)

(26 Subtypes) 1. Ack-doing-prep
2. Ack-wilco-prep

Translations - relate utterances in *different* floors (4 Subtypes)

(10 Subtypes) 1. Translation-r-direct
2. Translation-r-landmark
3. Translation-r-situated
4. Translation-r-default
5. Translation-r-history
6. Translation-r-contextual

Example Subtypes:

1. Customer→Waiter: I'll have a cheeseburger
2. Customer→ Waiter: and a small coke *Continue*

1. Customer→ Waiter: a small coke, please
2. Waiter→Customer: here you go *Acknowledgement-done*

1. Customer→ Waiter: I'll have a cheeseburger
2. Waiter→Cook: Cheeseburger!! *Translation-right*

Traum et al. 2018, LREC



3. Schema extensions: landmark and direct translation extensions

TU	ID	Participant (Audio Stream 1)	DM -> Participant (Chat Room 1)	DM-> RN (Chat Room 2)	RN (Audio Stream 2)	Ante- cedent	Relation- Type
1	1	go through the doorway directly in front of you					
1	2	and take a photo				1	continue
1	3		processing. . .			2*	processing
1	4			move into Conf Room		1	translation-r- landmark
1	5			then		4	link-next
1	6			send image		2	translation-r- direct
1	7		moving. . .			1	ack-doing
1	8				uh done and sent	6*	ack-done
1	9		done, sent			8	translation-l

Direct Translations: Uses the same or synonymous words, where the translation is applicable in any physical or conversational context.

Landmark Translations: Refers to a unique landmark name known only to members of the right floor.

3. Schema extensions : situated and default translation extensions



TU	ID	Participant (Audio Stream 1)	DM -> Participant (Chat Room 1)	DM-> RN (Chat Room 2)	RN (Audio Stream 2)	Ante- cedent	Relation- Type
1	1	turn east ninety degrees					
1	2	and travel three feet				1	continue
1	3		processing. . .			2*	processing
1	4			turn left 90 degrees		1	translation-r-situated
1	5			then...		4	link-next
1	6			move forward 3 feet		2	translation-r-default
1	7		turning...			1	ack-doing
1	8		moving...			2	
1	8				done	6*	ack-done
1	9		done			9	translation-l

Situated Translations: Synonymous with original instruction only in the current physical context, but does not specify a unique landmark.

Default Translations: Supplements information by relying on some default assumption related to a robot behavior or capability



3. Schema extensions : History translation extension

TU	ID	Participant (Audio Stream 1)	DM -> Participant (Chat Room 1)	DM-> RN (Chat Room 2)	RN (Audio Stream 2)	Ante- cedent	Relation- Type
1	1		You often ask for images at the end of movement instructions. Should I send one each time?				
1	2	yes				1	offer-accept
2	3	back up five feet					
2	4			back up 5 feet		3	translation-r-direct
2	5			send image		3	translation-r-history
2	7		executing...			3	ack-doing

History Translations: All or part of the translation is only relevant given the dialogue history, in which it was established that a certain instruction should be interpreted in a particular way.



3. Schema extensions : contextual translation, preparatory acknowledgments

TU	ID	Participant (Audio Stream 1)	DM -> Participant (Chat Room 1)	DM-> RN (Chat Room 2)	RN (Audio Stream 2)	Antecedent	Relation-Type
1	1	take a picture of the wall on your left					
1	3		processing. . .			1	processing
1	4			move to left wall		1	translation-r-contextual-partial
1	6			send image		4	continue
1	7		moving. . .			1	ack-doing-prep
1	8				done and sent	6*	ack-done
1	9		done, sent			8	translation-l

Underspecified, Contextual Translations: Draws upon situational or conversational context, but precisely what contextual information is being used is unclear, underspecified, or there are two or more factors.

Acknowledgment – Doing/Will-Comply Preparation: Speaker understands the command and a preparation step required for compliance with the command is underway (doing) or will be done (will-comply).

4. Annotated corpus: Analysis of type frequencies



Corpus Overview:

- 168 human-robot dialogues annotated, validated
- Total of 40,873 relations

Super-Type Relations Frequencies:

- 36.4% Acknowledgment super-type
- 36.5% Translation super-type

New Relation Types Frequencies:

- 70% Direct Translations (no situated language)
- 30% have situated language
- New preparatory acknowledgments have small, but critical impact

Relation	
	Translation-r
	Direct
	Direct-partial
<i>situated</i>	Contextual
	Contextual-partial
	Landmark
	Landmark-partial
	Situated
	Situated-partial
	History
	History-partial
	Default
	Default-partial
	Updated Ack Types
	Will-comply
	Doing
<i>situated</i>	Will-comply-prep
	Doing-prep

Table 5: Frequencies and % of updated relations.

4. Annotated corpus: Inter-annotator agreement



Reliability measured through
Inter-Annotator Agreement (IAA)

- Comparable or higher IAA than original, unmodified schema
- New annotation categories are clearly identifiable

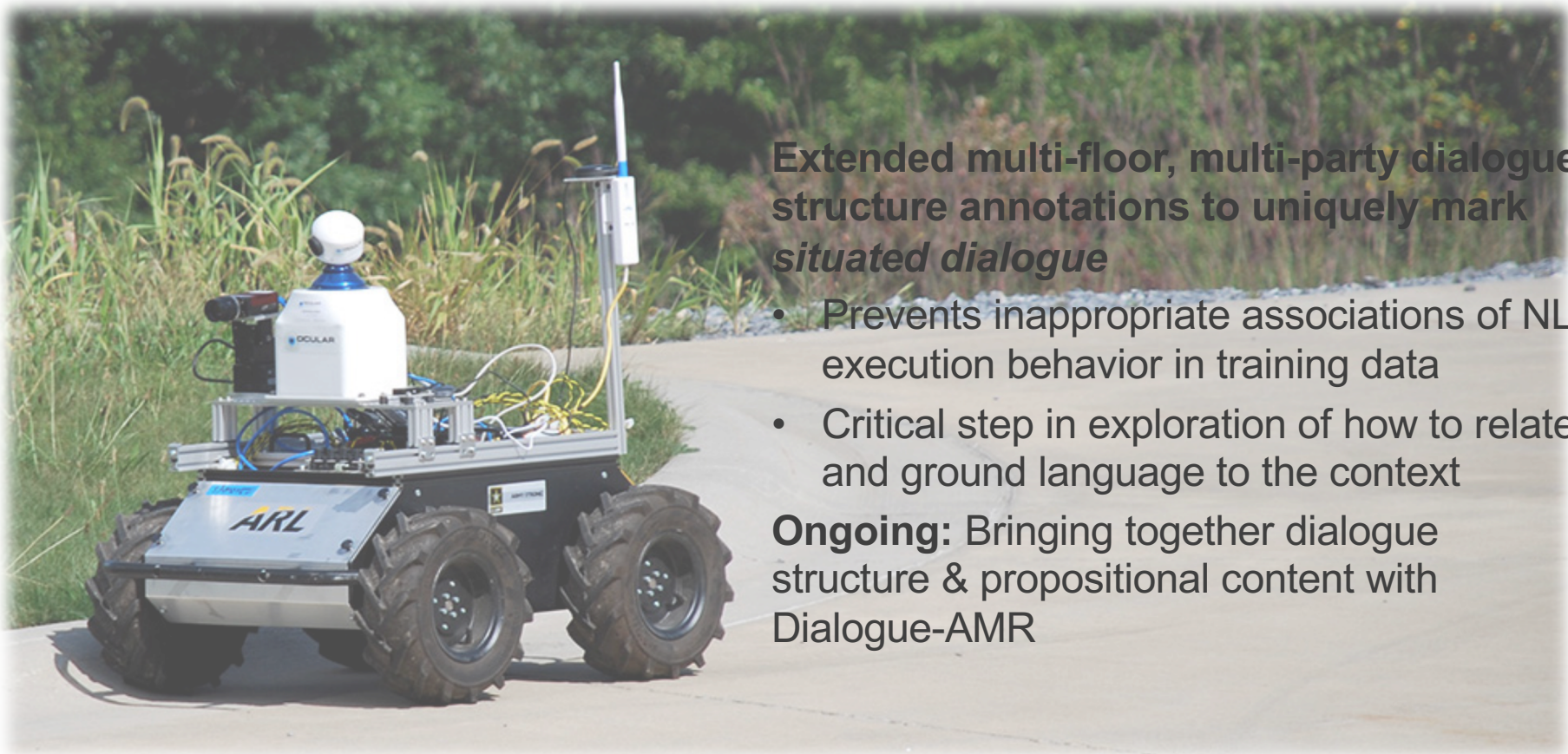
Markable Type	Agreement		Distance Metric
	Unmodified Schema	Modified Schema	
Antecedents	0.72–0.82	0.79- 0.94	Nominal ^a
Relation Types	0.77–0.89	0.83- 0.93	Nominal ^a
Transaction Units	0.48– 0.93	0.65-0.85	MASI ^b

^aKrippendorff (1980) ^bPassonneau (2006)

Table 6: IAA of the original, unmodified schema of Traum et al. (2018) and our modified schema.



5. Conclusions & future work



Extended multi-floor, multi-party dialogue structure annotations to uniquely mark *situated dialogue*

- Prevents inappropriate associations of NL - execution behavior in training data
- Critical step in exploration of how to relate and ground language to the context

Ongoing: Bringing together dialogue structure & propositional content with Dialogue-AMR

Dialogue-AMR: Bonial et al, LREC-2020

Making Sense of “STOP”

- Some meanings of saying “STOP” to another
 1. Action – slow down and halt motion
 2. Pause - Cease execution of current action (temporarily)
 3. Cancel current action
 4. Cancel future action

Possible Responses to “stop”

- A. perform a “stopping” action, to terminate current velocity
- B. halt current execution of an action (and later do something unrelated)
- C. pause current execution of an action (and resume the action later)
- D. pause execution and resume a slightly altered action after a correction has been specified
- E. ignore the command as redundant with what has already been done (or already planned to do)
- F. explain or request clarification when the command seems inappropriate or unclear
- G. refrain from repeating a previous or current action (that might or might not currently be planned to do again)

Temporal Positioning of “Stop” in instruction-execution-grounding sequence

1. As part of the initial instruction, prior to beginning execution
2. During the grounding of the instruction; for example, when the operator has specified part but not all of the instruction sequence, or if the addressee is clarifying, questioning, or negotiating some aspects, such as a termination point or manner
3. After the instruction has been given and accepted, but before execution has begun
4. During execution, when part has been performed and part remains unperformed
5. After execution, but before the conversational partners ground the fact that execution has (successfully or unsuccessfully) terminated
6. After it has been agreed that the action has terminated (seems unrelated to the instruction, perhaps like the first case, relating to a new instruction)

Temporal Positioning of “Stop” in In Botlanguage Multi-floor dialogue

1. As part of the initial instruction, prior to beginning execution
2. During the grounding of the instruction;
3. After the instruction has been given, but before execution has begun
 - 3.1 After the instruction has been given by commander but before it has been translated by DM to the RN
 - 3.2 After the instruction has been translated by DM, but before it has been carried out by RN
4. During execution, when part has been performed and part remains unperformed
5. After execution, but before the conversational partners ground the fact that execution has (successfully or unsuccessfully) terminated
 - 5.1 After the RN finished executing, but before the RN has acknowledged completion
 - 5.2 After the RN has acknowledged completion but before the DM has translated the completion.
6. After it has been agreed that the action has terminated (seems unrelated to the instruction, perhaps like the first case, relating to a new instruction)

Example: unsuccessful execution

<i>Floor 1</i>			<i>Floor 2</i>	
#	User	DM > User	DM >RN	RN
82	um go <pause. 33>go straight			
83		How far should I move forward?		
84	five feet			
85		ok		
86			move forward 5 feet	
87	okay stop			
88			stop	
89				done, I could move about 4
90		done		
91	go east, go east five feet			

Table 3: Evidence from unsuccessful execution of the original instruction (line 89) and re-use of structure of the original instruction indicates “stop” precedes a *change* repair strategy.

25x	279.71	robot turn forty five degrees right				8		
26	286.29	and continue to second doorway				8	25	continue
27A	298.25			turn 45 right,		8	25	translatio n-r-direct
27B				move forward to second doorway		8	26	translatio n-r- default
28	304.15		executing...			8	26*	ack-doing
29	314.21	robot stop				8	26*	correctio n
30	322.35			stop		8	29	translatio n-r-direct
31	333.6	robot take a photo				9		

<i>Floor 1</i>			<i>Floor 2</i>	
#	User	DM > User	DM >RN	RN
68	keep moving until you see your next shovel			
69		I think you are more familiar with shovels than I am.		
70	move forward <pause .41> ten feet			
71		processing...		
72			move forward 10 feet	
73		I will move forward as far as I can, ok?		
74	stop			
75				done, it was nine
76		done		
77	do you see any shovels near you?			

Table 4: The failed execution of the command underway (line #76) when stop is issued, as well as the speaker's multiple attempts at producing an instruction that reflects their intent, provide evidence that "stop" (line #74) is an edit marker, providing the opportunity for a *fresh start* repair in line 77.

<i>Floor 1</i>			<i>Floor 2</i>	
#	User	DM > User	DM > RN	RN
69	move forward ten feet			
70			move forward 10 feet	
71		executing...		
72				done
73	stop			
74		done		
75	take a picture			
76			.	

		Repair Status		
		No evidence of repair - halt motion	Evidence of repair - change strategy	Evidence of repair - fresh start strategy
Timing	Original instruction	38	0	0
	During grounding	2	7	0
	After grounding, before execution	0	0	0
	During execution	59	7	5
	After execution, before grounding termination	2	0	0
	After grounding termination	12	0	0

Table 2: Corpus counts of “stop”, as characterized along the dimensions of the timing of issuance and the status as a potential edit marker, signaling repair to come, either change or fresh start repair strategies.

Conclusions:

- Many aspects to discourse/dialogue structure
 - Types of phenomena
 - Types of structural relationships
- Scaling of Interactional Richness:
 - Discourse -> Dialogue -> Multiparty Dialogue
 - > Multifloor Dialogue
- Some aspects of structure only revealed in richer situations
- Most phenomena apply in richer situations, but some generalizations may not (e.g. global intentional structure)
- Still more work to do in all areas, but especially multi-party/multi-floor