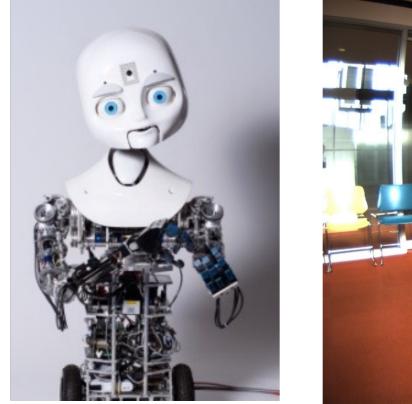
#### Natural Language and Spatal Reasoning

**Stefanie Tellex** 

MIT Media Lab Ph.D. Thesis Defense

# Where should the robot go?

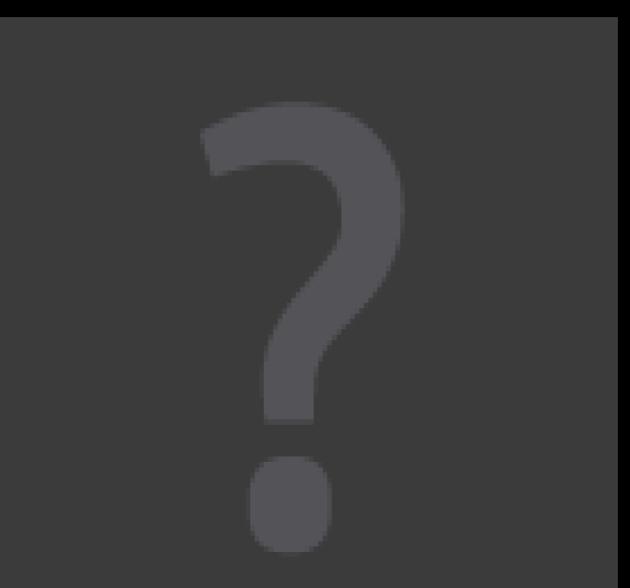
With your back to the windows, walk straight through the door near the elevators. Continue to walk straight, going through one door until you come to an intersection just past a whiteboard. Turn left, turn right, and enter the second door on your right (sign says "Administrative Assistant").





# Show me people walking across the kitchen.

# Show me people walking across the kitchen.



# Motivation

- What semantic structures can enable a system to understand and use spatial language in realistic situations?
- Applications
  - Information surveillance and reconnaissance.
  - Dialog interface to navigation devices.
  - Supervisory control of robots.

# Outline

- Spatial Prepositions
- Spatial Description Clauses
- Direction Understanding
- Spatial Language Video Retrieval
- Conclusion

# **Related Work**

- Cognitive Semantics
  - Talmy (2005), Landau and Jackendoff (1993), Regier and Carlson (2001), Siskind (2001)
- Language and Robots
  - MacMahon et. al (2006), Hsiao et. al (2008), Skubic et. al (2004)
- Video Retrieval
  - Katz et. al (2004), Ivanov and Wren (2005), Fleischman et. al (2006), Naphade et. al (2006)

#### **Spatial Prepositions**

### **Spatial Prepositions in English**

about, above, across, after, against, along, alongside, amid(st), among(st), around, at, atop, behind, below, beneath, beside, between, betwixt, beyond, by, down, from, in, inside, into, near, nearby, off, on, onto, opposite, out, outside, over, past, through, throughout, to, toward, under, underneath, up, upon, via, with, within, without, far from, in back of, in between, in front of, in line with, on top of, to the left of, to the right of, to the side of

(from Landau and Jackendoff, 1993)

the, of, and, to, a, in, that, is, was, he, for, it, with, as, his, on, be, at, by, i, this, had, not, are, but, from, or, have, an, they, which, one, were, you, all, her, she, there, would, their, we, him, been, has, when, who, will, no, more, if, out, its, so, up, said, what, about, than, into, them, can, only, other, time, new, some, could, these, two, may, first, then, do, any, like, my, now, over, such, our, man, me, even, most, made, after, also, well, did, many, before, must, years, back, through, much, where, your, way, down, should, because, long, each, just, state, people, those, too, how, Mr., little, good, world, make, very, year, still, see, own, work, men, day, get, here, old, between, both, life, being, under, three, never, know, same, last, another, while, us, off, might, great, states, go, come, since, against, right, came, take, used, himself, few, house, American, use, place, during, high, without, again, home, around, small, however, found, mrs, part, school, thought, went, say, general, once, upon, every, left, war, don't, does, got, united, number, hand, course, water, until, always, away, public, something, fact, less, through, far, put, head, think, called, set, almost, enough, end, took, government, night, yet, system, better, four, nothing, told, eyes, city, going, president, why, days, present, point, didn't, look, find, asked, second, group, later, next, room, social, business, knew, program, give, half, side, face, toward, white, five, let, young, form, given, per, order, large, several, national, important, possible, rather, big, among, case, often, early, john, things, looked, ever, become, best, need, within, felt, along, children, saw, church, light, power, least, family, development, interest, others, open, thing, seemed, want, area, god, members, mind, help, country, service, turned, door, done, law, although, whole, line, problem, sense, certain, different, kind, began, thus, means, matter, perhaps, name, times, york, itself, action, human, above, week, company, free, example, hands, local, show, history, whether, act, either, gave, death, feet, today, across, body, past, guite, taken, anything, field, having, seen, word, car, experience, I'm, money, really, class, words, already, college, information, tell, making, sure, themselves, together, full, air, shall, held, known, period, keep, political, real, miss, probably, century, question, seems, behind, cannot, major, office, brought, special, whose, boy, cost, federal, economic, self, south, problems, heard, six, study, ago, became, moment, run, available, job, street, result, short, west, age, change, position, board, individual, reason, 10

the, of, and, to, a, in, that, is, was, he, for, it, with, as, his, on, be, at, by, i, this, had, not, are, but, from, or, have, an, they, which, one, were, you, all, her, she, there, would, their, we, him, been, has, when, who, will, no, more, if, out, its, so, up, said, what, about, than, into, them, can, only, other, time, new, some, could, these, two, may, first, then, do, any, like, my, now, over, such, our, man, me, even, most, made, after, also, well, did, many, before, must, years, back, through, much, where, your, way, down, should, because, long, each, just, state, people, those, too, how, Mr., little, good, world, make, very, year, still, see, own, work, men, day, get, here, old, between, both, life, being, under, three, never, know, same, last, another, while, us, off, might, great, states, go, come, since, against, right, came, take, used, himself, few, house, American, use, place, during, high, without, again, home, around, small, however, found, mrs, part, school, thought, went, say, general, once, upon, every, left, war, don't, does, got, united, number, hand, course, water, until, always, away, public, something, fact, less, through, far, put, head, think, called, set, almost, enough, end, took, government, night, yet, system, better, four, nothing, told, eyes, city, going, president, why, days, present, point, didn't, look, find, asked, second, group, later, next, room, social, business, knew, program, give, half, side, face, toward, white, five, let, young, form, given, per, order, large, several, national, important, possible, rather, big, among, case, often, early, john, things, looked, ever, become, best, need, within, felt, along, children, saw, church, light, power, least, family, development, interest, others, open, thing, seemed, want, area, god, members, mind, help, country, service, turned, door, done, law, although, whole, line, problem, sense, certain, different, kind, began, thus, means, matter, perhaps, name, times, york, itself, action, human, above, week, company, free, example, hands, local, show, history, whether, act, either, gave, death, feet, today, across, body, past, guite, taken, anything, field, having, seen, word, car, experience, I'm, money, really, class, words, already, college, information, tell, making, sure, themselves, together, full, air, shall, held, known, period, keep, political, real, miss, probably, century, question, seems, behind, cannot, major, office, brought, special, whose, boy, cost, federal, economic, self, south, problems, heard, six, study, ago, became, moment, run, available, job, street, result, short, west, age, change, position, board, individual, reason, 11

# Modeling Spatial Prepositions

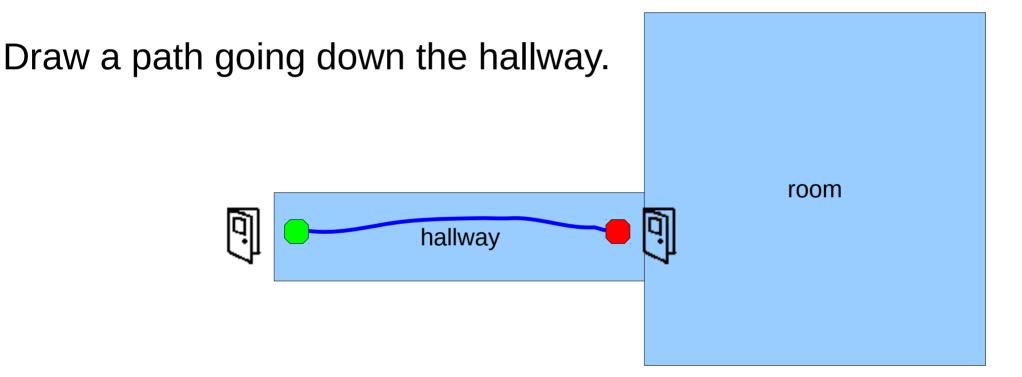
- Functions that take geometric arguments.
- Classifiers for "to," "across," "towards," "through," "around," etc.
- Library of features.
- Train and test on labeled examples.

# Learning Spatial Prepositions

Probability of *past* Examples of *past* **Features** f1 0.8 . . . fn f1 0.95 . . . fn f1 0.5 . . . fn f1 0.67 e.g., the . . . e.g., naïve Bayes, fn minimum decision tree, SVM distance between f1 the robot and the 0.1 . . . landmark fn

#### Across Video

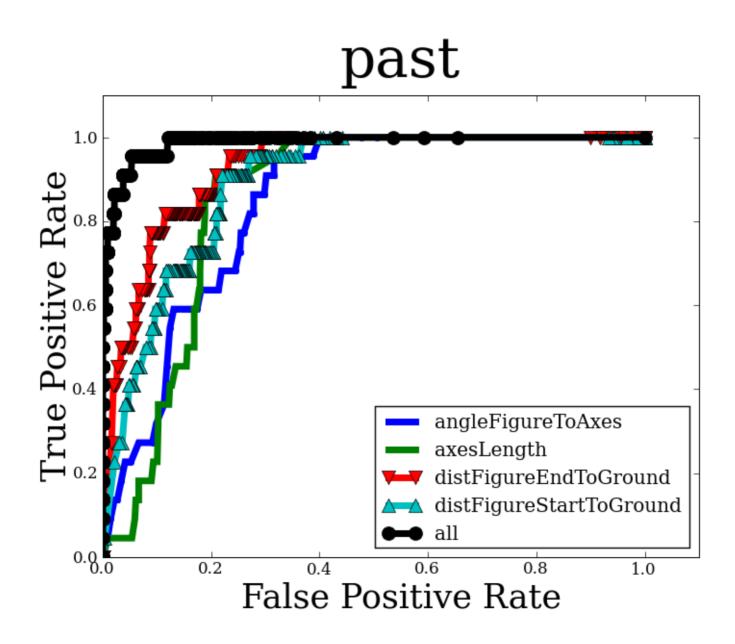
## Corpus of Examples of Spatial Relations

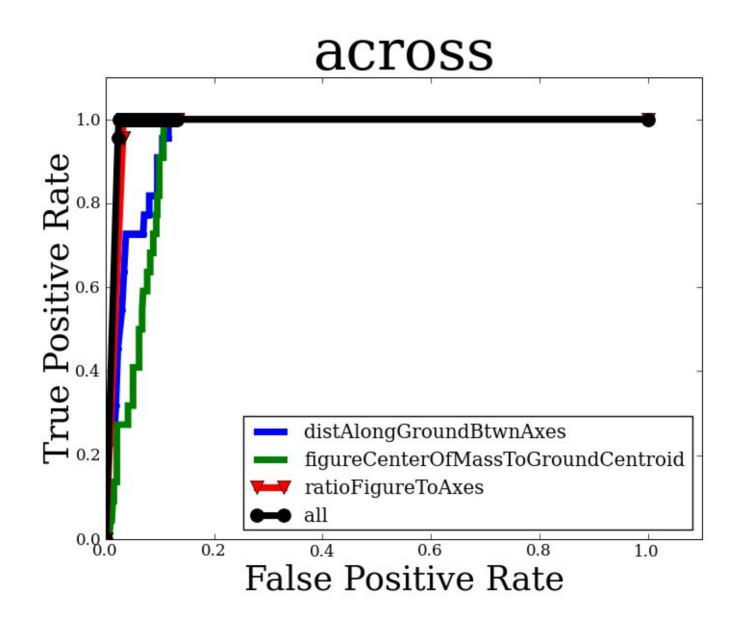


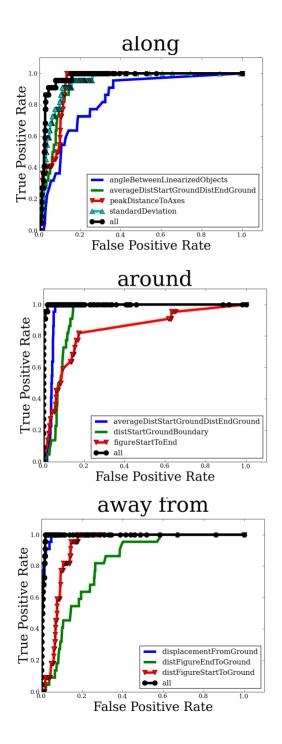
# Training Classifiers for Spatial Prepositions

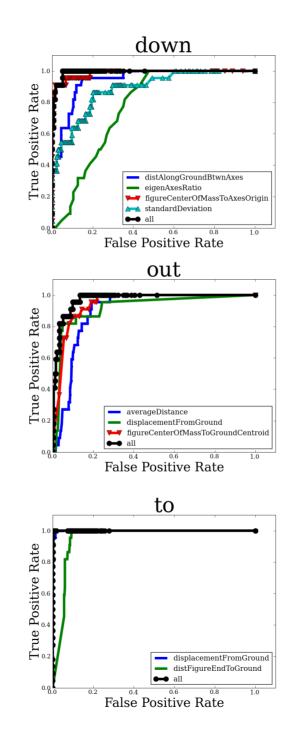
- Binary classifiers.
- Positive examples of "across" are negative examples of "to."

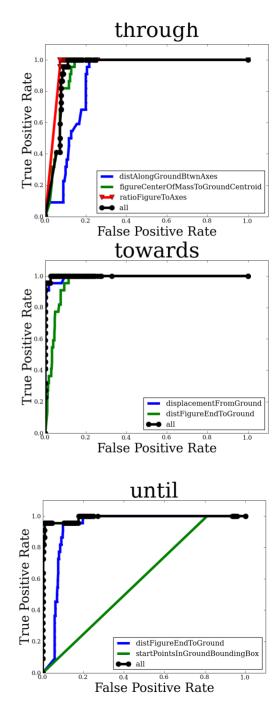
- Some exceptions (e.g., "to" and "towards.")











#### Phrase-based Video Retrieval



#### Annotation Task

Complete the sentence: "The person is going..."

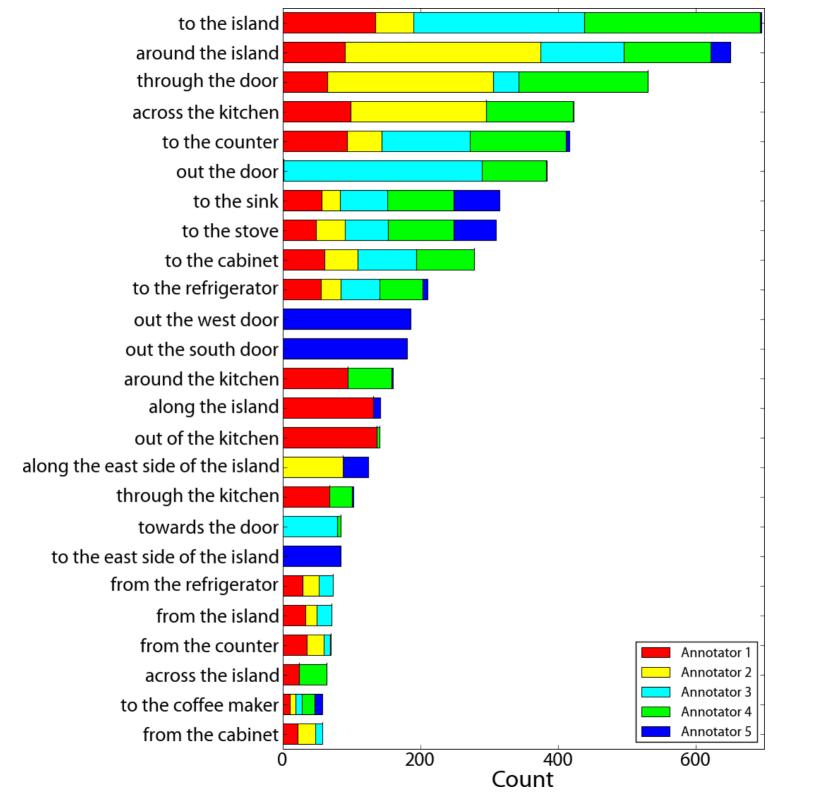


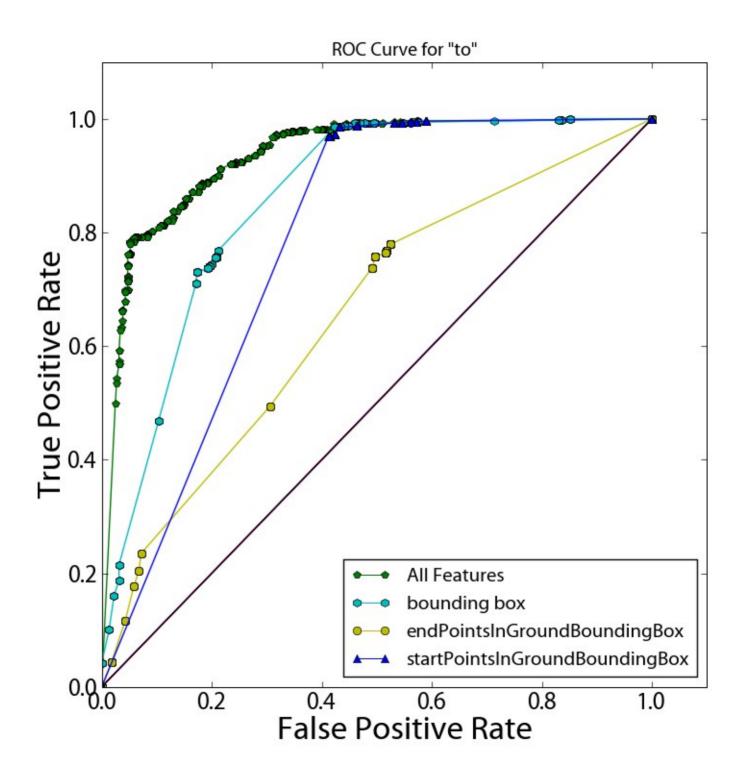


### Corpus

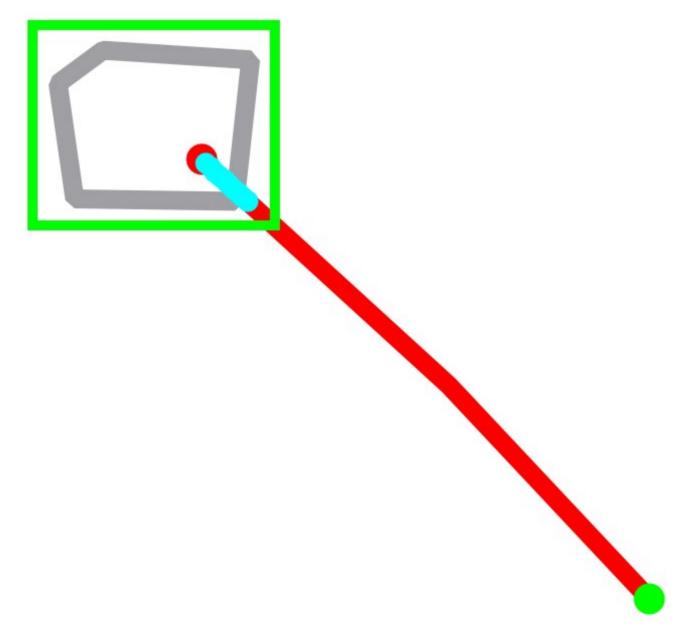


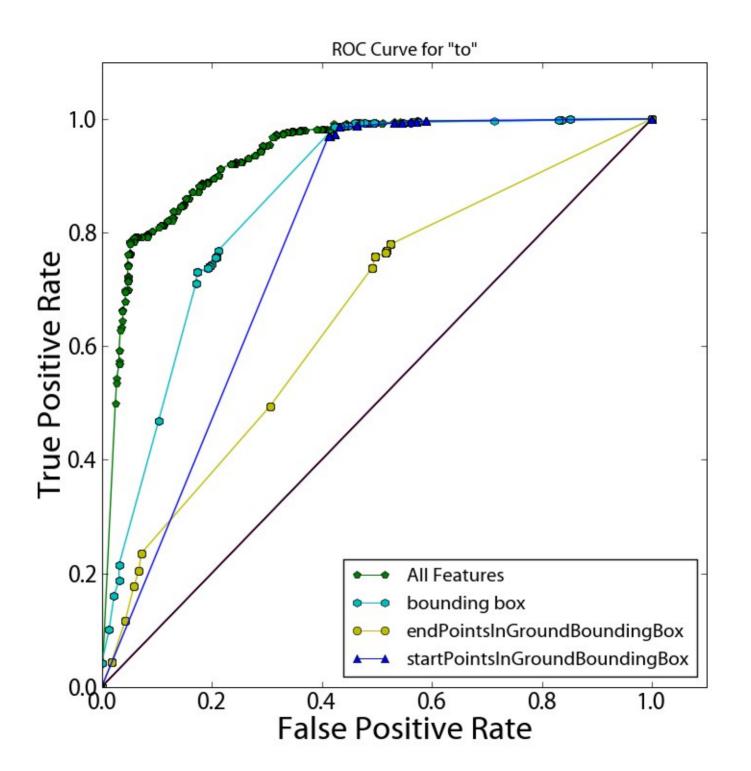
to the counter. along the east side of the island. from the refrigerator. to the cabinet. across the kitchen.





#### endPointsInGroundBoundingBox









#### Full model: 0.99 Bounding boxes: 0.76





#### Full model: 0.10 Bounding boxes: 0.63

## Results

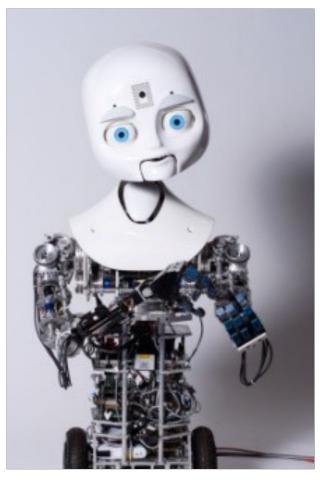
- Library of features for spatial prepositions.
- Models for eleven prepositions.
- Phrase-based video retrieval system.

# Outline

- Spatial Prepositions
- Spatial Description Clauses
- Direction Understanding
- Spatial Language Video Retrieval
- Conclusion

## **Two Different Domains**

Direction
 Understanding

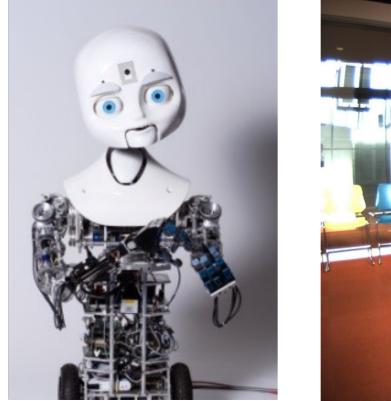


Video Retrieval



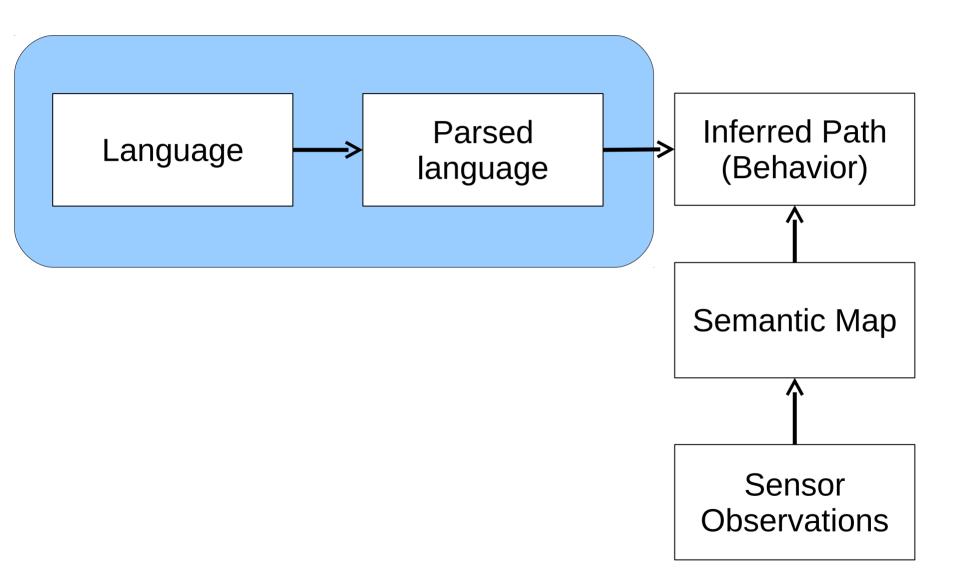
# Where should the robot go?

With your back to the windows, walk straight through the door near the elevators. Continue to walk straight, going through one door until you come to an intersection just past a whiteboard. Turn left, turn right, and enter the second door on your right (sign says "Administrative Assistant").



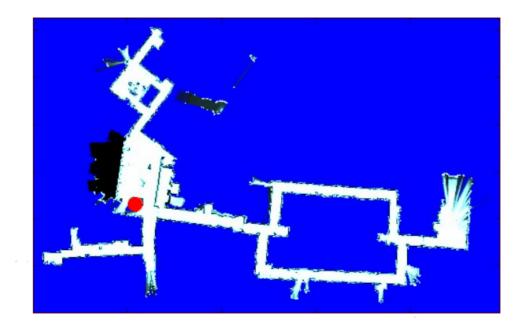


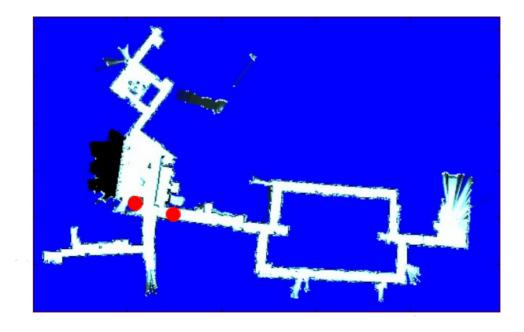
### System

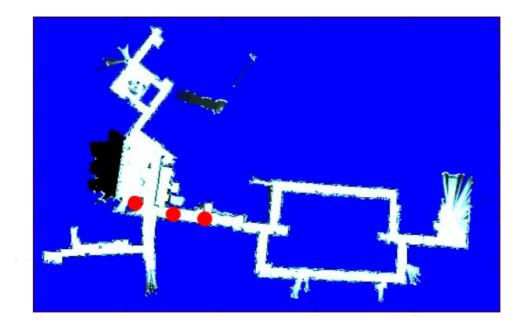


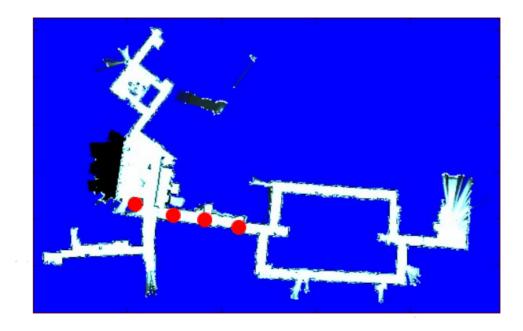
# Structure of Natural Language Directions

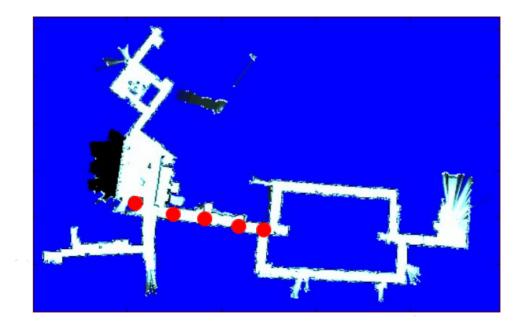
Sequential

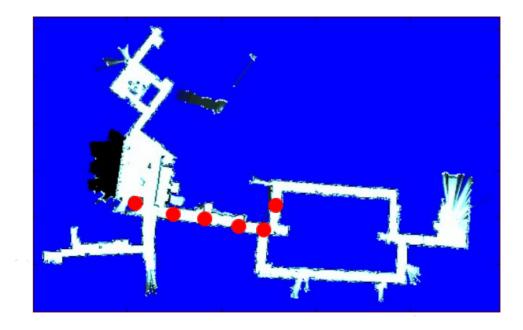


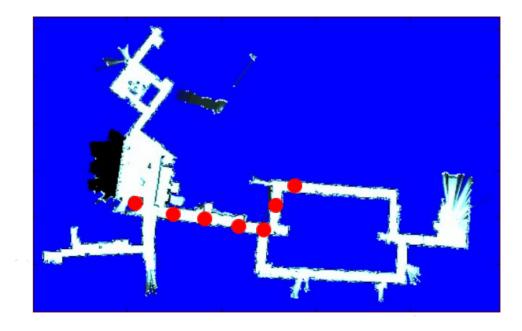


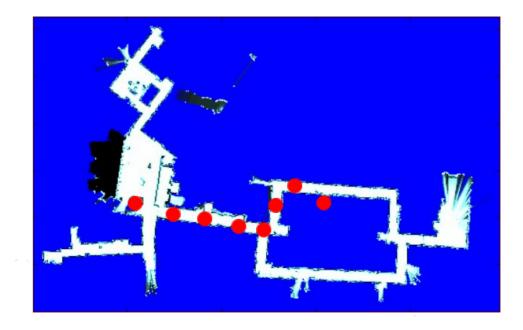












# Structure of Natural Language Directions

- Sequential.
- Directive verbs.

### **Directive Verbs**

# Structure of Natural Language Directions

- Sequential.
- Directive verbs.
- Landmarks.

#### Landmarks

# Structure of Natural Language Directions

- Sequential.
- Directive verbs.
- Landmarks.
- Spatial relations.

# **Spatial Relations**

# Structure of Natural Language Directions

- Sequential.
- Directive verbs.
- Landmarks.
- Spatial relations.

Go through the door.

• Figure: <you>

- Figure: <you>
- Verb: Go

- Figure: <you>
- Verb: Go
- Spatial relation: through

- Figure: <you>
- Verb: Go
- Spatial relation: through
- Landmark: the door.

Continue to walk straight, going through one door until you come to an intersection just past a white board.

Verb: Continue to walk Spatial Relation: straight

Continue to walk straight, going through one door until you come to an intersection just past a white board.

Verb: going Spatial Relation: through Landmark: one door

Continue to walk straight, going through one door until you come to an intersection just past a white board.

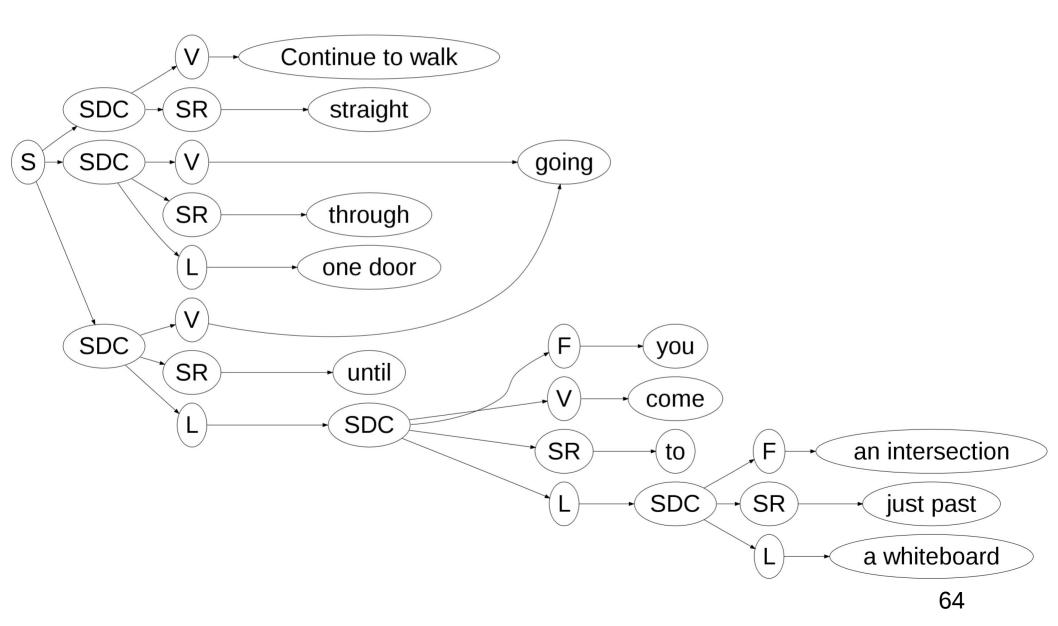
Verb: going Spatial Relation: until Landmark: you come to an intersection just past a white board

Continue to walk straight, going through one door until you come to an intersection just past a white board.

Figure: you Verb: come Spatial Relation: to Landmark: an intersection just past a white board

Continue to walk straight, going through one door until you come to an intersection just past a white board.

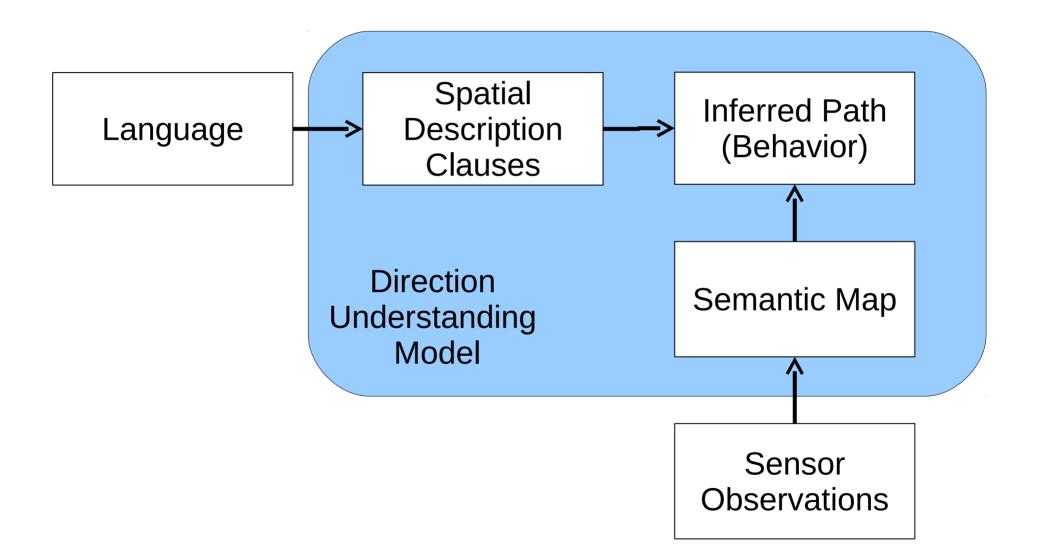
Figure: an intersection Spatial Relation: just past Landmark: a white board



# Outline

- Spatial Prepositions
- Spatial Description Clauses
- Direction Understanding
- Spatial Language Video Retrieval
- Conclusion

# System



## Model

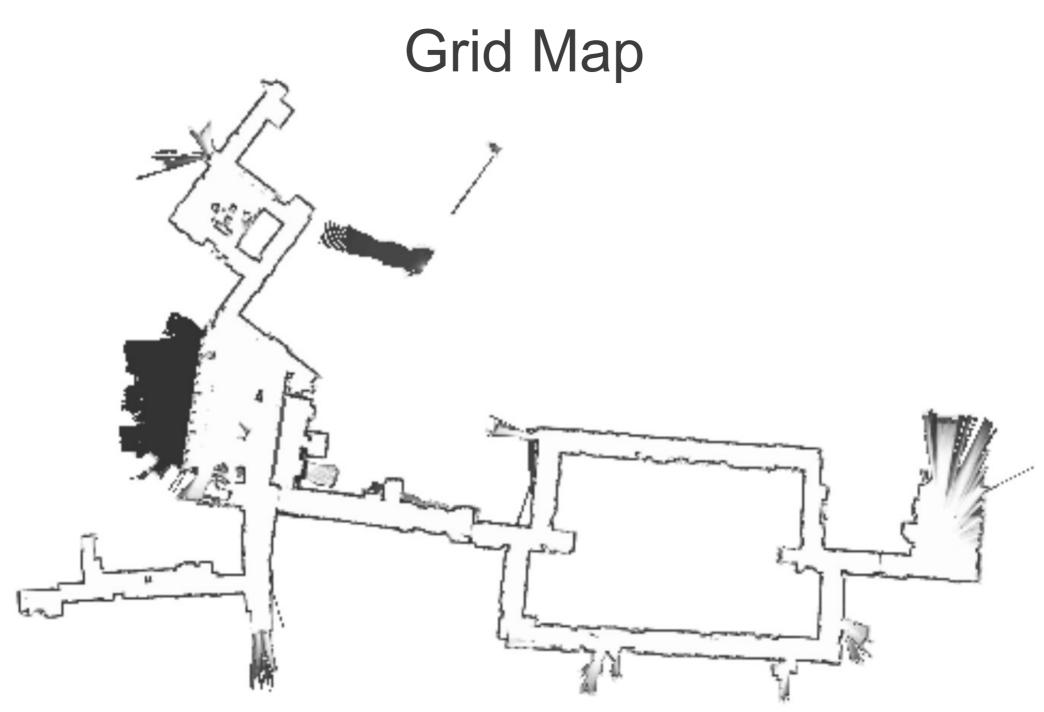
*argmax*<sub>path</sub> p(path, directions, observations)

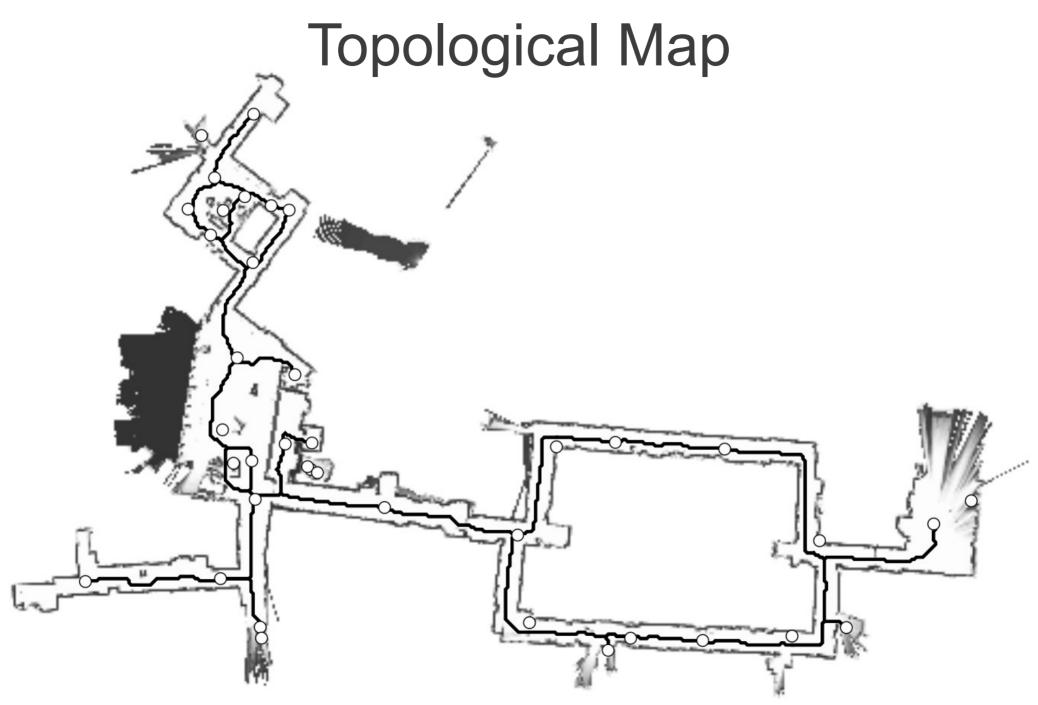
# Model

p(path, directions, observations) =

p(directions | path, observations) p(path) p(observations)(independence assumptions)

 $p(sdc_1...sdc_M|vp_1...vp_{M+1}, o_1...o_K) p(vp_1...vp_{M+1}) p(o_1...o_K)$ (topological map of viewpoints, and a sequence of SDCs)





# Model

p(path, directions, observations) =

p(directions | path, observations) p(path) p(observations)(path and observations are independent)  $p(sdc_1...sdc_M | vp_1...vp_{M+1}, o_1...o_K) p(vp_1...vp_{M+1}) p(o_1...o_K)$ (topological map of viewpoints, and a sequence of SDCs)  $\prod_{i=1}^{M} p(sdc_i | vp_i, vp_{i+1}, o_1...o_K) \prod_{i=1}^{M} p(vp_{i+1} | vp_1...vp_i) p(o_1...o_K)$ (Paths are Markov)

# Probability of Directions Given a Possible Movement

 $p(sdc_i|vp_i,vp_{i+1},o_1...o_K)$ 

 $= p(figure_i, verb_i, spatial\_relation_i, landmark_i | vp_i, vp_{i+1}, o_1 ... o_K)$ (The structure of a spatial description clause.)

# **Spatial Description Clause**

Go through the door.

- Figure: <you>
- Verb: Go
- Spatial relation: through
- Landmark: the door.

# Probability of Directions Given a Possible Movement

 $p(sdc_i|vp_i,vp_{i+1},o_1...o_K)$ 

 $= p(figure_i, verb_i, spatial\_relation_i, landmark_i | vp_i, vp_{i+1}, o_1 ... o_K)$ (The structure of a spatial description clause.)

 $= p(figure_i | vp_i, vp_{i+1}, o_1 \dots o_K) p(verb_i | vp_i, vp_{i+1}) \times p(spatial\_relation_i | landmark_i, v_i, v_{i+1}, o_1 \dots o_K) \times p(landmark_i | v_i, v_{i+1}, o_1 \dots o_K)$ 

(Factor according to certain independence assumptions.)

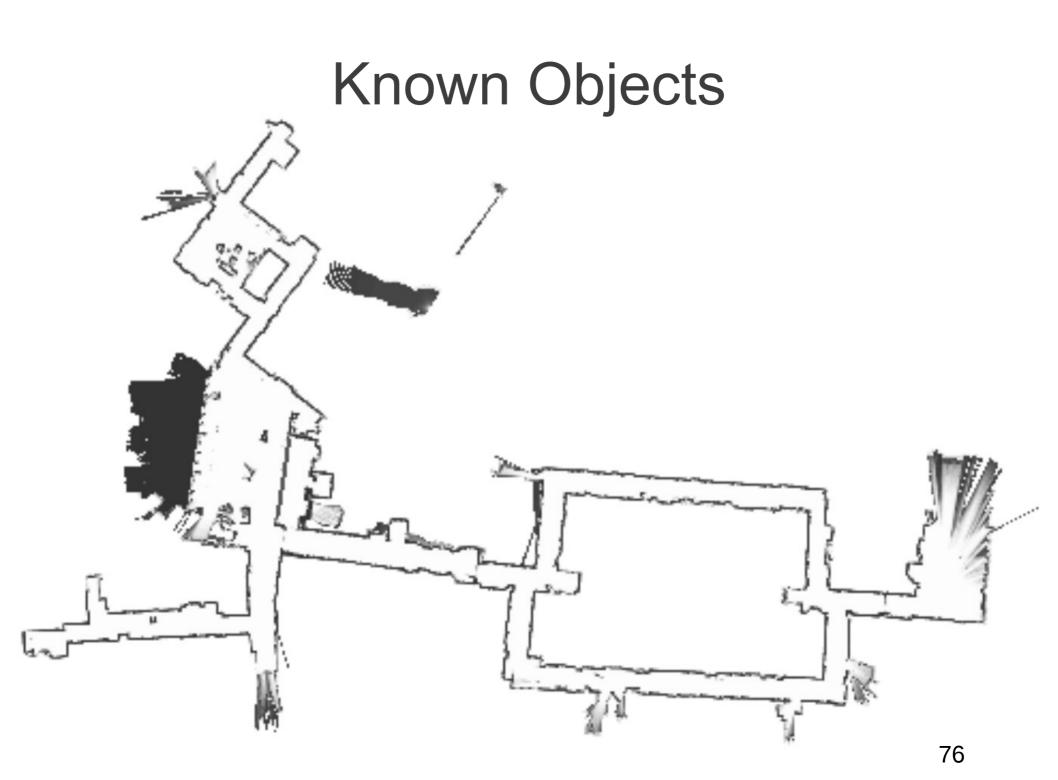
# Probability of Directions Given a Possible Movement

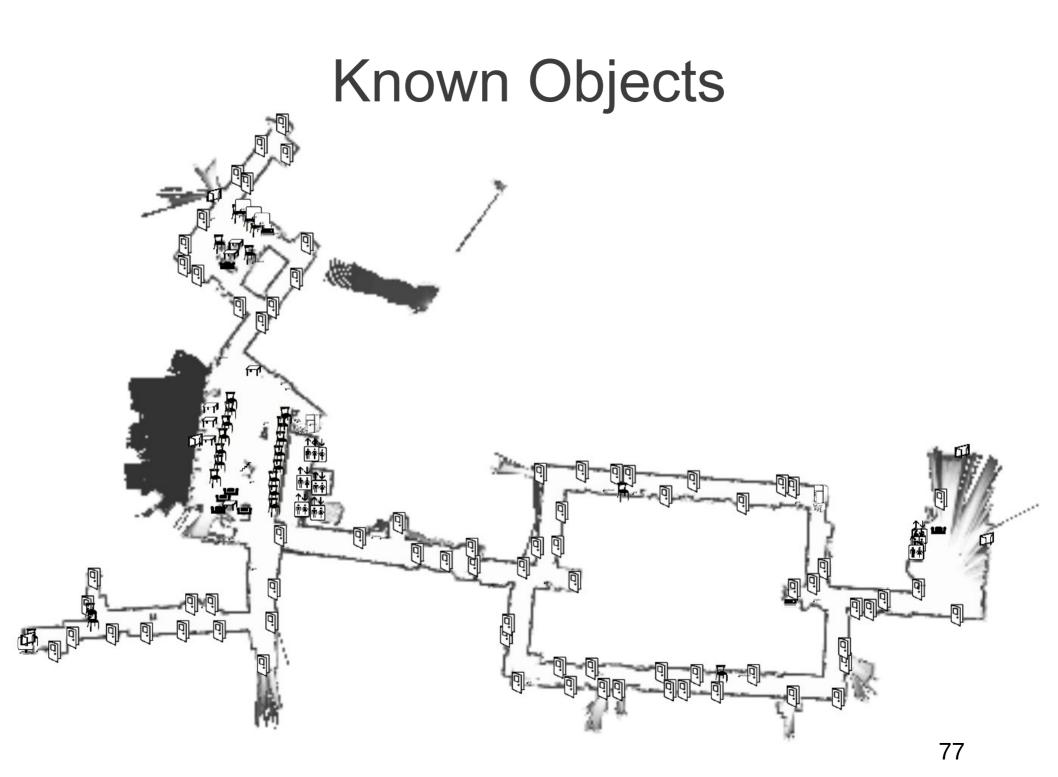
 $p(sdc_i|vp_i,vp_{i+1},o_1...o_K)$ 

 $= p(figure_i, verb_i, spatial\_relation_i, landmark_i | vp_i, vp_{i+1}, o_1 ... o_K)$ (The structure of a spatial description clause.)

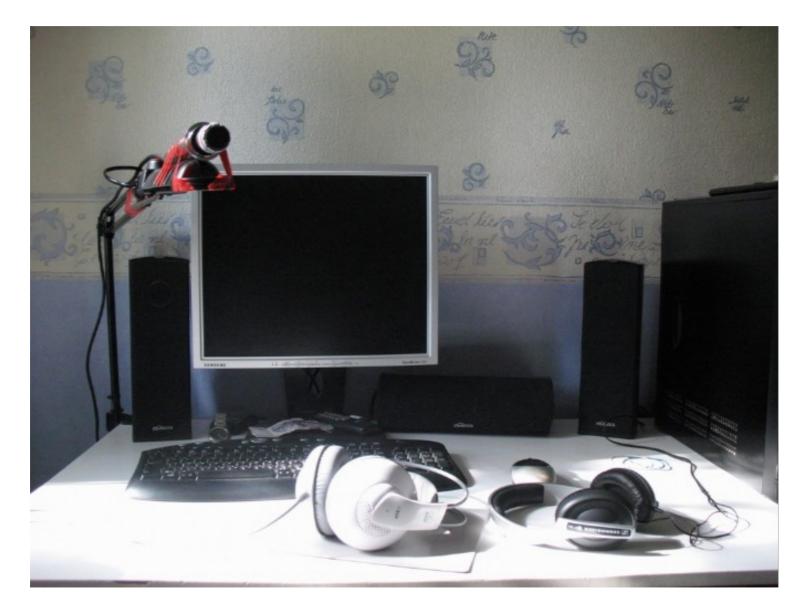
 $= p(figure_i | vp_i, vp_{i+1}, o_1 \dots o_K) p(verb_i | vp_i, vp_{i+1}) \times p(spatial\_relation_i | landmark_i, v_i, v_{i+1}, o_1 \dots o_K) \times p(landmark_i | v_i, v_{i+1}, o_1 \dots o_K)$ 

(Factor according to certain independence assumptions.)





#### **Contextual relationships**



Tags include: desktop, monitor, computer, keyboard, and mouse (Kollar and Roy, 2009)

# Probability of Directions Given a Possible Movement

 $p(sdc_i|vp_i,vp_{i+1},o_1...o_K)$ 

 $= p(figure_i, verb_i, spatial\_relation_i, landmark_i | vp_i, vp_{i+1}, o_1 ... o_K)$ (The structure of a spatial description clause.)

 $= p(figure_i | vp_i, vp_{i+1}, o_1 \dots o_K) p(verb_i | vp_i, vp_{i+1}) \times p(spatial\_relation_i | landmark_i, v_i, v_{i+1}, o_1 \dots o_K) \times p(landmark_i | v_i, v_{i+1}, o_1 \dots o_K)$ 

(Factor according to certain independence assumptions.)

# Probability of Directions Given a Possible Movement

 $p(sdc_i|vp_i,vp_{i+1},o_1...o_K)$ 

 $= p(figure_i, verb_i, spatial\_relation_i, landmark_i | vp_i, vp_{i+1}, o_1 ... o_K)$ (The structure of a spatial description clause.)

 $= p(figure_{i} | vp_{i}, vp_{i+1}, o_{1}...o_{K}) p(verb_{i} | vp_{i}, vp_{i+1}) \times \frac{p(spatial\_relation_{i} | landmark_{i}, v_{i}, v_{i+1}, o_{1}...o_{K})}{p(landmark_{i} | v_{i}, v_{i+1}, o_{1}...o_{K})}$ 

(Factor according to certain independence assumptions.)

# Learning Spatial Relations

Probability of *past* Examples of *past* Features f1 0.8 . . . fn f1 0.95 . . . fn f1 0.5 . . . fn f1 0.67 e.g., the . . . e.g., naïve Bayes, fn minimum decision tree, SVM distance between f1 the robot and the 0.1 . . . landmark fn

# Probability of Directions Given a Possible Movement

 $p(sdc_i|vp_i,vp_{i+1},o_1...o_K)$ 

 $= p(figure_i, verb_i, spatial\_relation_i, landmark_i | vp_i, vp_{i+1}, o_1 ... o_K)$ (The structure of a spatial description clause.)

 $= p(figure_i | vp_i, vp_{i+1}, o_1 \dots o_K) p(verb_i | vp_i, vp_{i+1}) \times p(spatial\_relation_i | landmark_i, v_i, v_{i+1}, o_1 \dots o_K) \times p(landmark_i | v_i, v_{i+1}, o_1 \dots o_K)$ 

(Factor according to certain independence assumptions.)

# Verbs in Route Directions

- Almost always "left," "right," or "straight."
  - "Continue to walk straight."
  - "Take a right at the photocopier."
- Exceptions
  - "Follow the wall into the small kitchen."
  - "Orient yourself with your back to the computers."

# Probability of Directions Given a Possible Movement

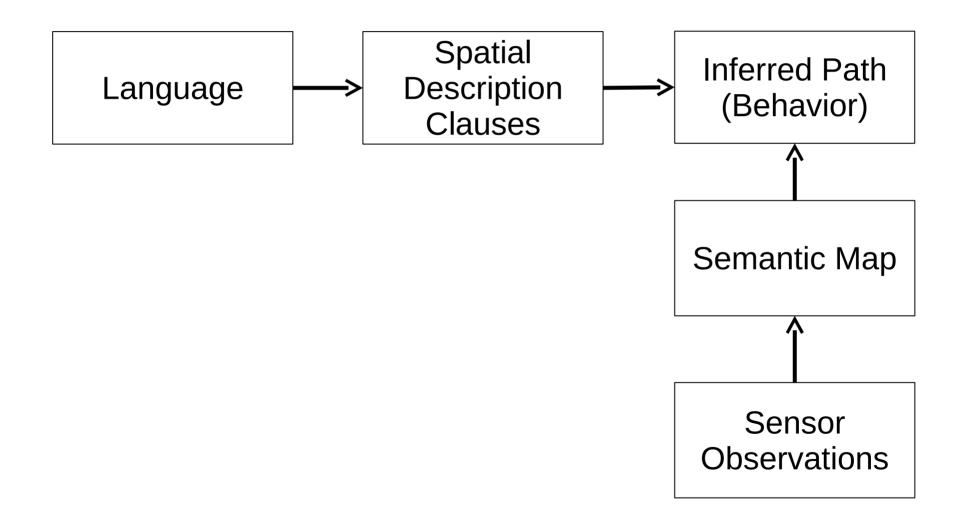
 $p(sdc_i|vp_i,vp_{i+1},o_1...o_K)$ 

 $= p(figure_i, verb_i, spatial\_relation_i, landmark_i | vp_i, vp_{i+1}, o_1 ... o_K)$ (The structure of a spatial description clause.)

 $= p(figure_i | vp_i, vp_{i+1}, o_1 \dots o_K) p(verb_i | vp_i, vp_{i+1}) \times p(spatial\_relation_i | landmark_i, v_i, v_{i+1}, o_1 \dots o_K) \times p(landmark_i | v_i, v_{i+1}, o_1 \dots o_K)$ 

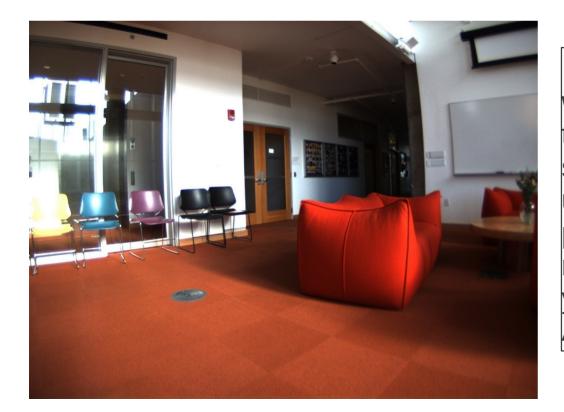
(Factor according to certain independence assumptions.)

## System



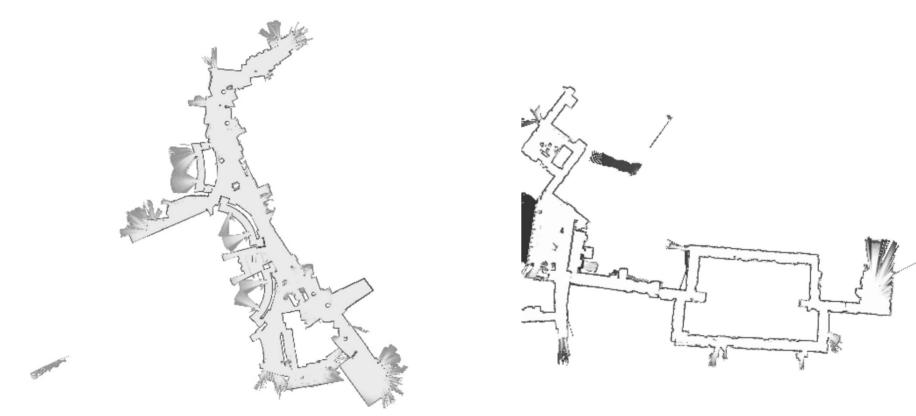
### Corpus of Natural Language Directions

- Subjects asked to write directions as if they were giving them to a person.
- 30 subjects, 300 directions, two indoor environments.



With your back to the windows, walk straight through the door near the elevators. Continue to walk straight, going through one door until you come to an intersection just past a whiteboard. Turn left, turn right, and enter the second door on your right (sign says "Administrative Assistant").

### Corpus of Natural Language Directions



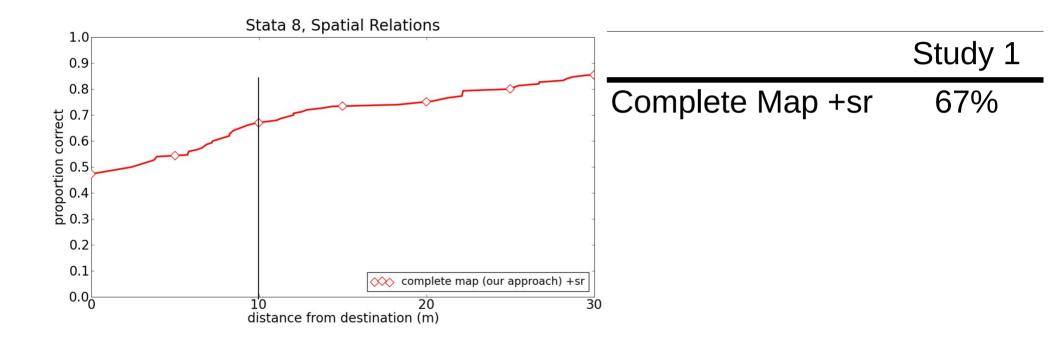
## Task

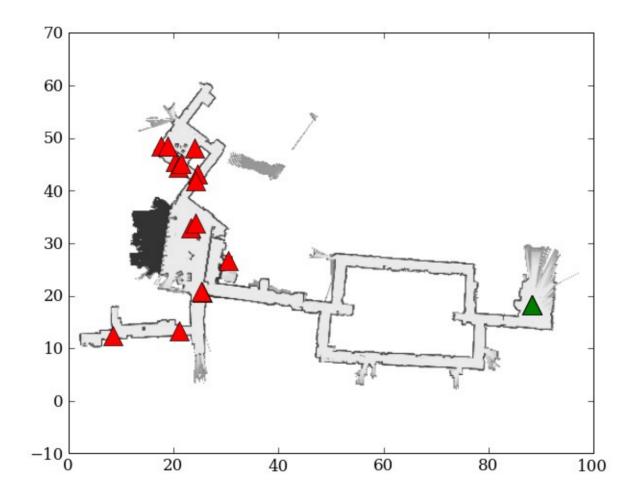
Given the text of the directions, a starting location, and a semantic map of the environment, follow the directions to the correct destination.

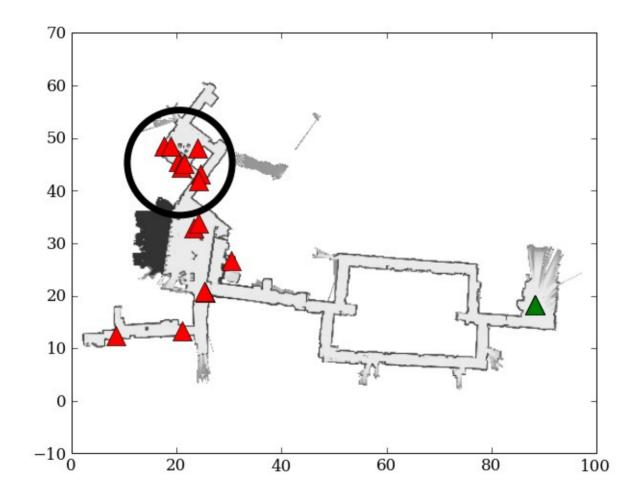
	Study 1
Humans (All directions)	85%

	Study 1
Humans (All directions)	85%
Our approach	67%

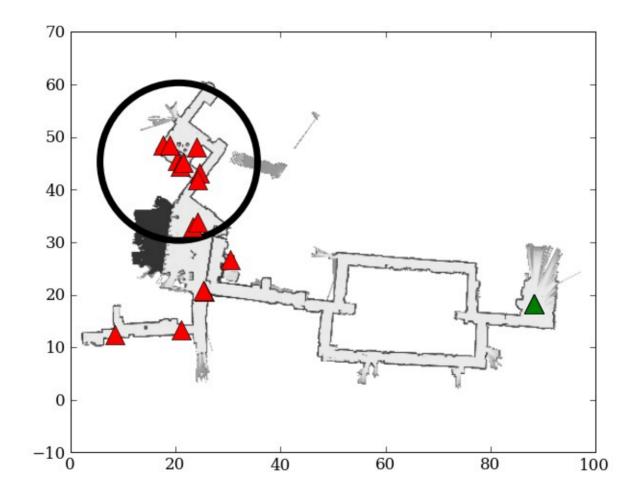
## **Spatial Relations**





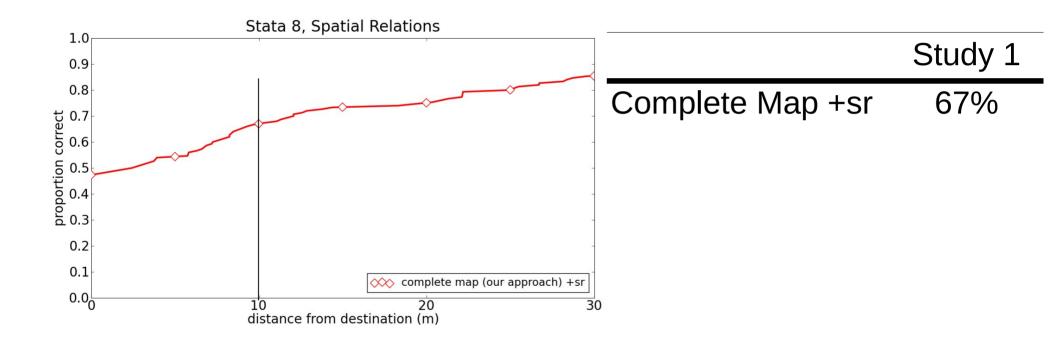


Proportion correct at 10 meters: 0.53

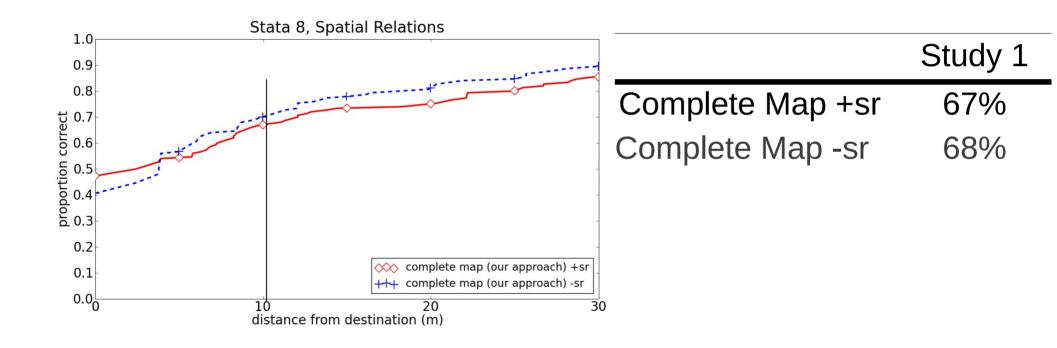


Proportion correct at 15 meters: 0.67

## **Spatial Relations**



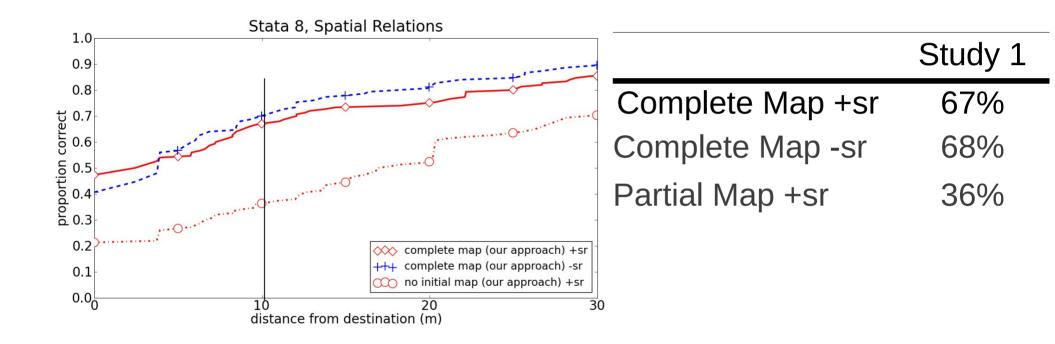
## **Spatial Relations**



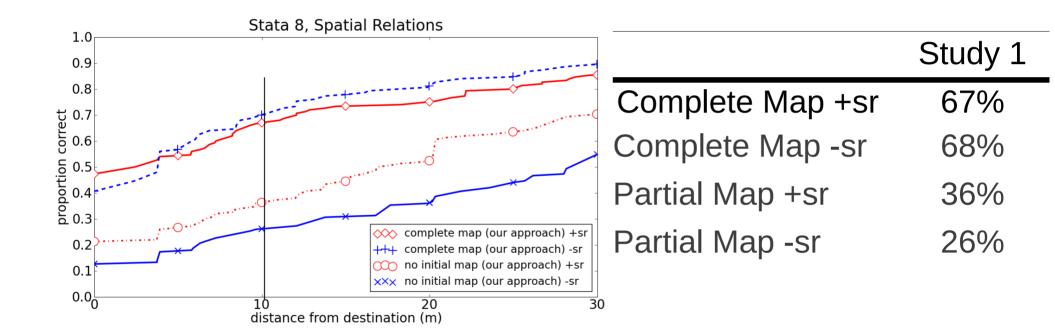
# Inference Methods

- Complete map.
  - Find the globally best path using a complete map of the environment.
- Partial map.
  - Greedy algorithm that uses local information from the environment.

## **Spatial Relations**

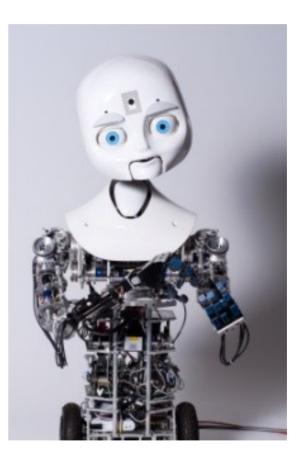


## **Spatial Relations**

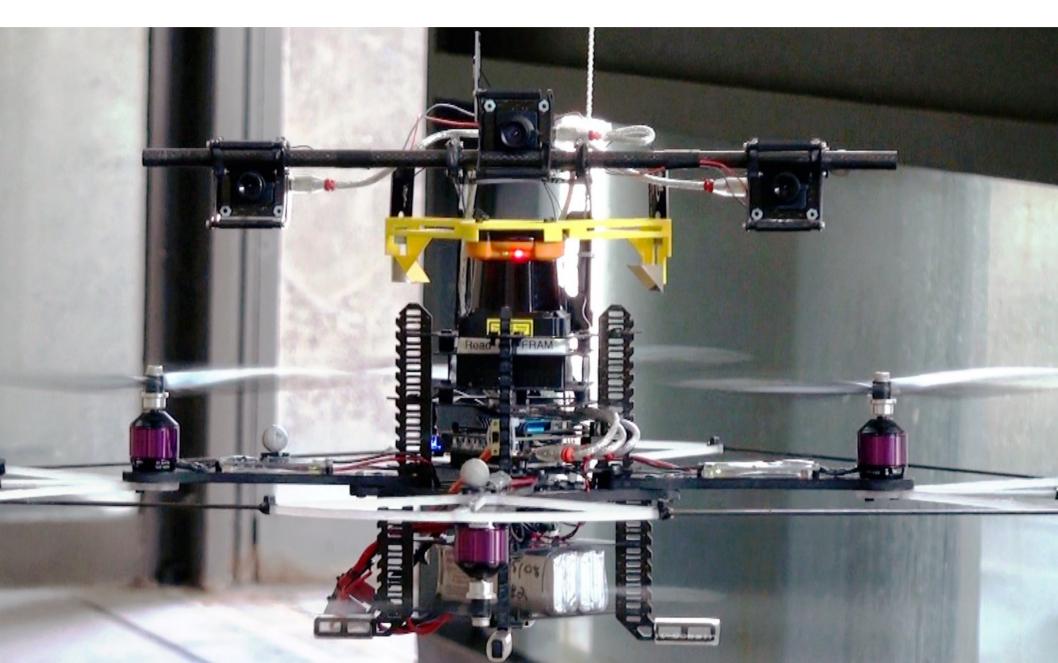


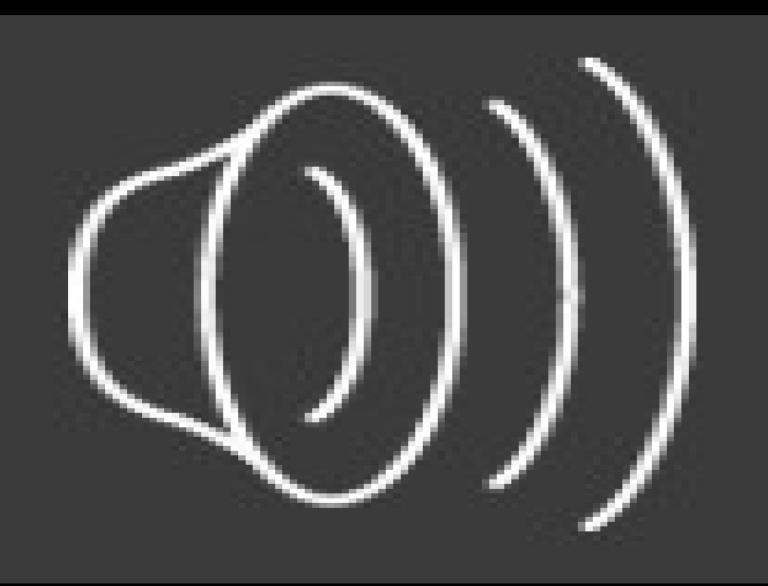
### Platforms





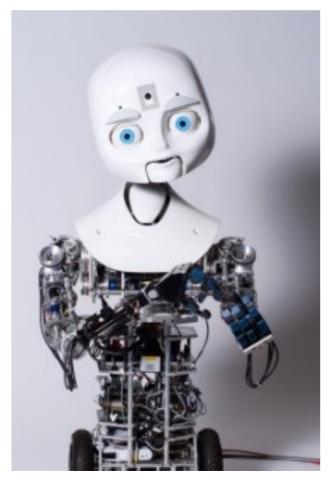
### **Robotic Helicopter**



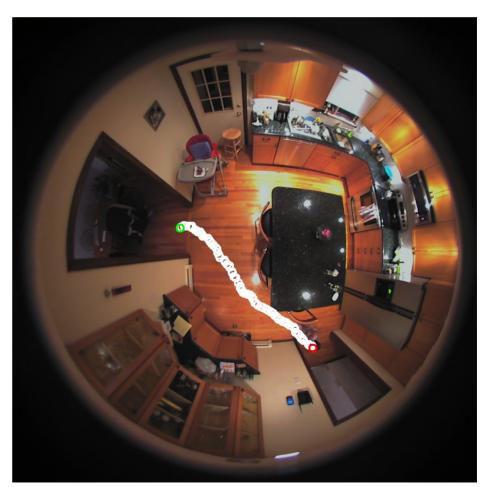


# **Two Different Domains**

Direction
 Understanding



Video Retrieval



# Outline

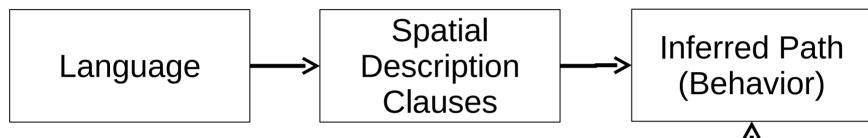
- Spatial Prepositions
- Spatial Description Clauses
- Direction Understanding
- Spatial Language Video Retrieval
- Conclusion

#### Show me people going across the kitchen.

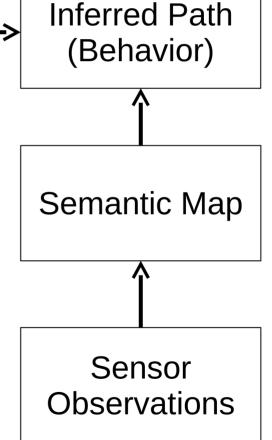


People walking from the hallway door, around the island, to the kitchen sink

# System



- Directive understanding.
- Video retrieval/event recognition.



# Corpus

- Watch a video clip.
- Write a natural language description of the movement of a person in the clip, so that another person could draw the trajectory.
- Not asked to simplify language in any way, or use a particular vocabulary.
- 696 descriptions from fourteen annotators.



"The person walked from the couches in the living room to the dining room table."

"The woman entered the dining room from the living room."

"She walks from the hallway into the dining room and stands by the side of the dining room table that is nearest to the kitchen."

"The person walked from the couch in the living to the dining table in the dining room."

"The person enters the dining room from the living room and goes to the table near the entrance to the kitchen."

"She starts in the living room and walks to in front of the desk."

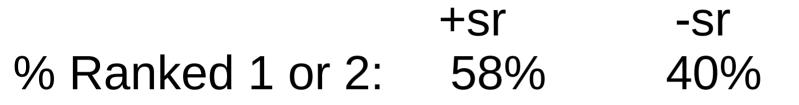
"The person enters the dining room from the stairway or living room area. She goes to the long side of the table nearest to the kitchen doorway."

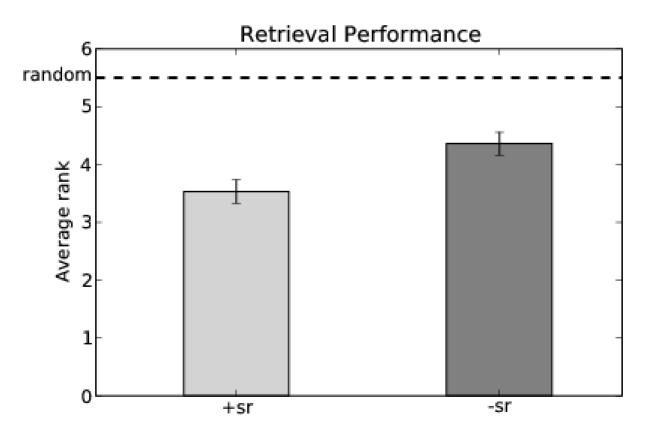
"The person walks from the left-bottom side of the dining room table over tot he (*sic*) shelves."

#### **Evaluation Metric**

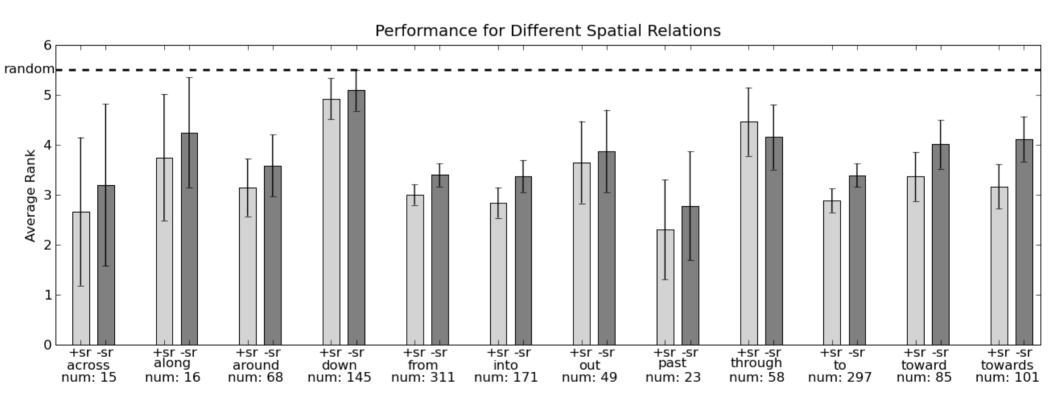
- Take a trajectory.
- Add in 9 other random clips.
- Rank all ten clips according to the original description.
- Report average rank of the original clip.

#### Performance With and Without Spatial Relations

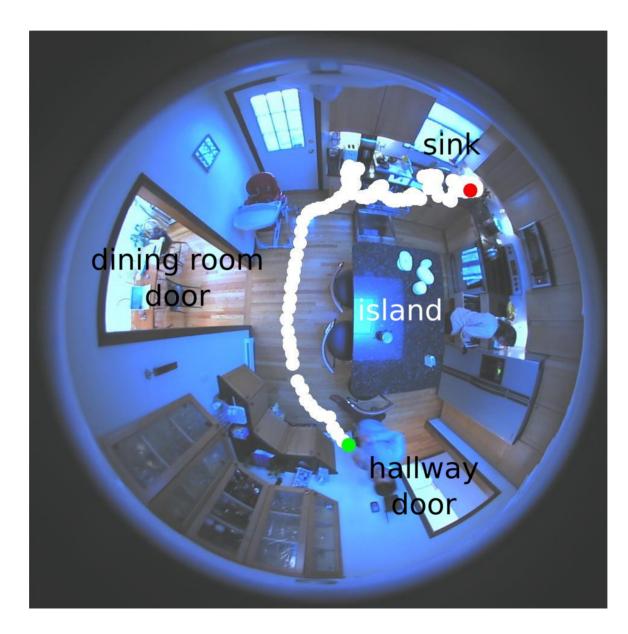




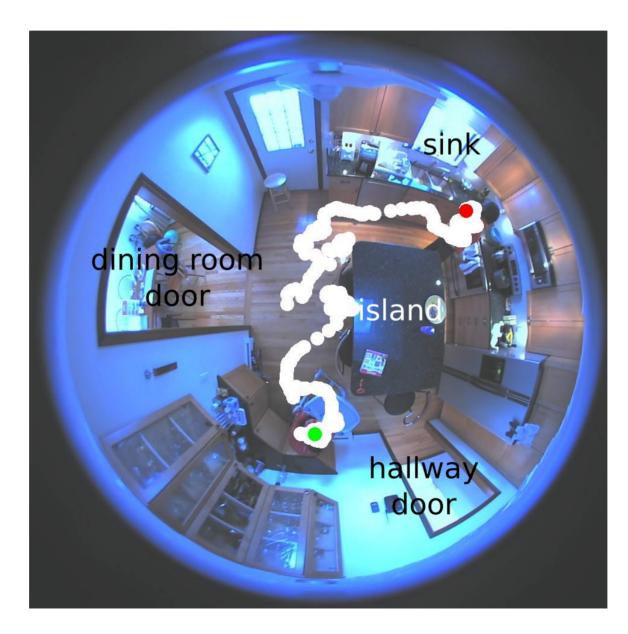
#### Performance With and Without Particular Spatial Relations



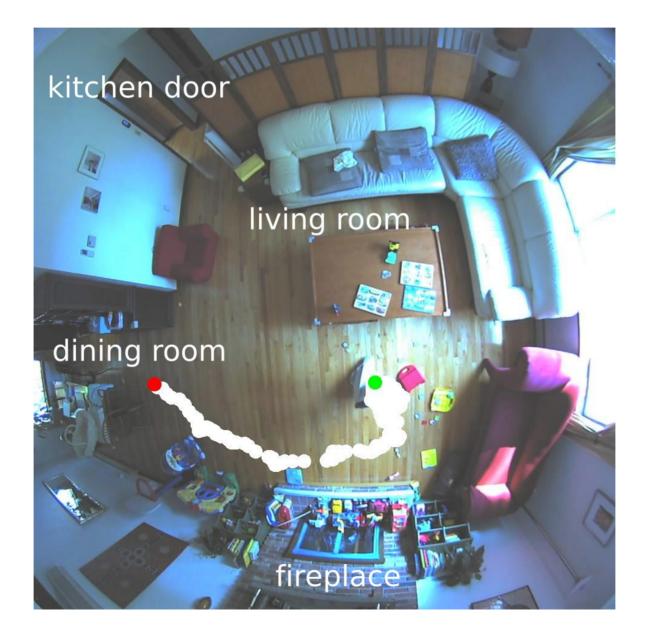
# People walking from the hallway door, around the island, to the kitchen sink



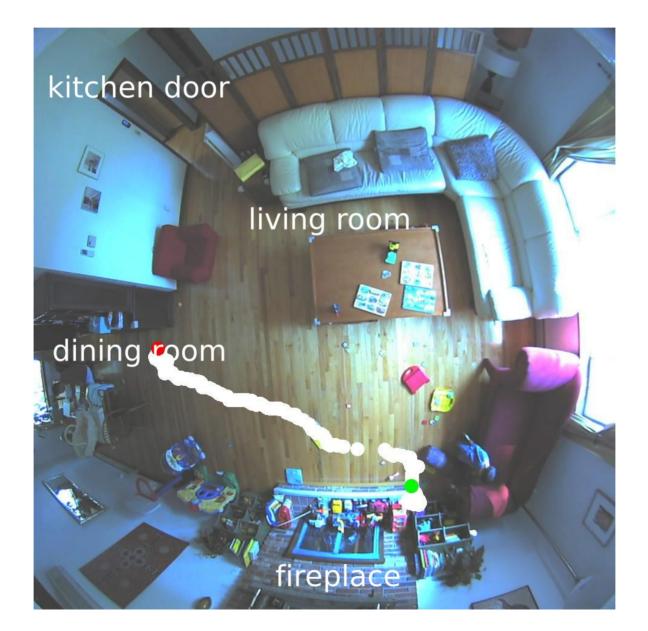
# People walking from the hallway door, around the island, to the kitchen sink



# She walks past the fireplace and stands by the bookshelf.



# She walks past the fireplace and stands by the bookshelf.



#### Results

- Effective video retrieval system that takes arbitrary spatial language.
- Corpus-based evaluation shows robustness.

# Outline

- Spatial Prepositions
- Spatial Description Clauses
- Direction Understanding
- Spatial Language Video Retrieval
- Conclusion

## Future Work

- From paths to plans.
  - "Empty the dishwasher."
  - "Unload the truck."
- Dialogue.
  - "Drive to the door near the elevators."
  - "Where are the elevators?"
- Less supervision.

## Contributions

- Developed a library of features that capture the semantics of spatial prepositions.
- Analyzed which features work best for each preposition.
- Spatial Description Clauses to decompose spatial language discourse.
- Systems for understanding natural language directions and spatial language video retrieval.
- Corpus-based evaluations show robustness.

#### Acknowledgements Piotr Mitros

Tom Kollar Deb, Yuri, Boris, Cynthia, Gerry Sussman, Nick Roy Cogmac Linda Peterson Karina, Mutsumi, Aaron, Necsys Gremio, Dave, Carie, Kai-yuh, Lin Ania & Seth & Moby Staci & Seth, Shannon & Jon, Scott & Onna Grandma Patches and Grandma Nicholas Mom and Dad

# What's New Since Proposal

- Analysis of what features for spatial prepositions.
- Models for time-based verbs like "meet" and "follow."
- Extending it to 3d. (Helicopter video!)
- Spatial language video retrieval.

#### Video Retrieval Metrics

Query	Avg. Precision
People coming out of the bathroom.	0.833
People walking into the baby's bedroom.	0.917
People walking down the hall.	0.967
People walking around the table. in the living room	1.000
People walking into the kitchen.	1.000
People walking out of the kitchen.	0.704
People walking from the hallway door, around the island, to the kitchen sink.	0.583
Mean Average Precision	0.858

# Related Work – Cognitive Semantics

- Talmy. (2005) The Fundamental System of Spatial Schemas in Language.
- Landau and Jackendoff. (1993) "What" and "Where" in Spatial Language and Spatial Cognition.
- Regier. (1992) The Acquisition of Lexical Semantics for Spatial Terms: A Connectionist Model of Perceptual Categorization.
- Regier and Carlson. (1991) Grounding Spatial Language in Perception: An Empirical and Computational Investigation.
- Siskind. (1993) Grounding the Lexical Semantics of Verbs in Visual Perception Using Force Dynamics and Event Logic.

# Related Work – Direction Understanding

- Hsiao et. al (2008) Conversational table-top robot.
- Skubic et. al (2002) Natural language commands for robots.
- MacMahon et. al (2005) Following natural language directions.
- Matuszek et. al (2010) Learning to follow natural language directions.
- Levit and Roy (2007) Spatial language dialog understanding.

### Related Work – Video Retrieval

- Katz et al. (2004) "Show me all the cars leaving the garage."
- Ivanov and Wren (2006) Query by example trajectory.
- Fleischman et al. (2006) "Show me people making coffee."
- Naphade et al. (2006) Large-Scale Concept
  Ontology for Multimedia

- CRF chunker tags each word.
- Trained from annotated data.
- Features (tri-grams)
  - $\circ$  Words
  - Part-of-speech tags
  - Label sequence
- Greedy algorithm groups tagged words together.

Continue to walk straight, going through one door until you come to an intersection just past a white board.

Continue to walk straight, going through one door until you come to an intersection just past a white board.

V: Continue to walk straight

Continue to walk straight, going through one door until you come to an intersection just past a white board.

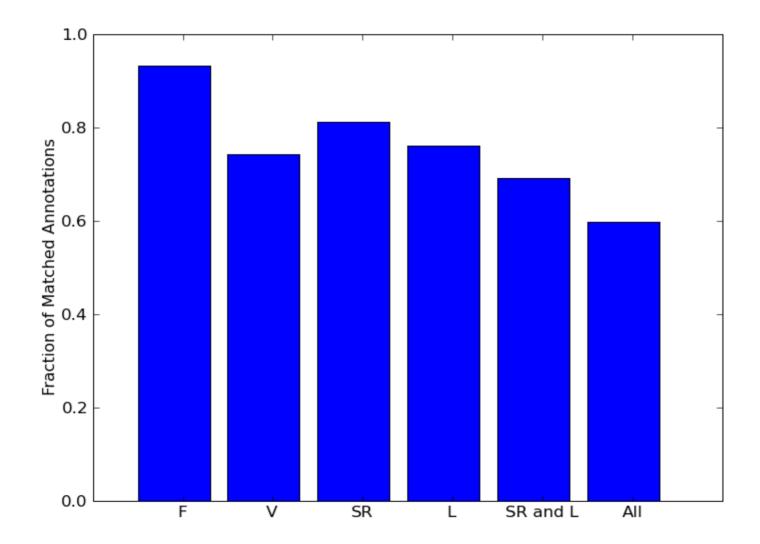
V: going, SR: through, L: one door

Continue to walk straight, going through one door until you come to an intersection just past a white board.

SR: until, L: you come to an intersection

Continue to walk straight, going through one door until you come to an intersection just past a white board.

SR: just past, L: a white board



## Contributions

- Create a library of features for grounding spatial prepositions.
- Analyze which features perform best for specific prepositions.
- Compare semantics of spatial prepositions in two different domains.
- Model for understanding spatial language discourse.

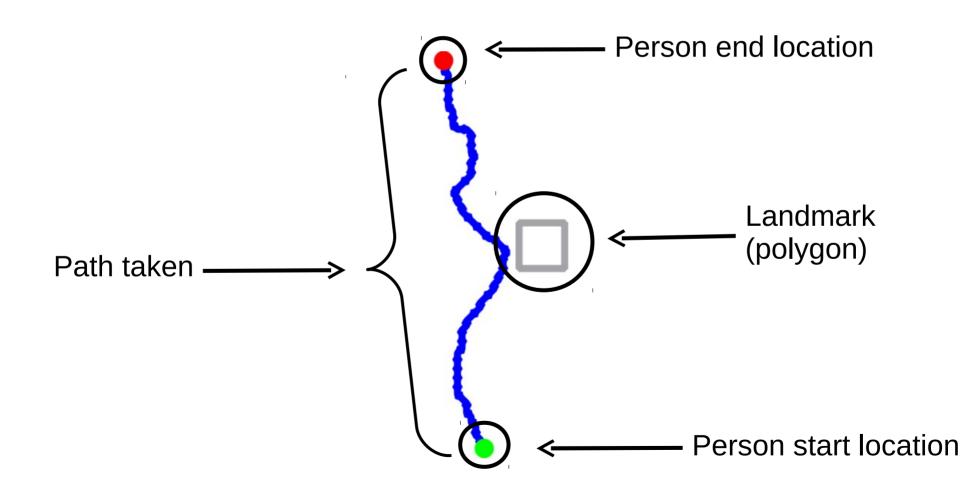
#### **Spatial Prepositions in English**

about, above, across, after, against, along, alongside, amid(st), among(st), around, at, atop, behind, below, beneath, beside, between, betwixt, beyond, by, down, from, in, inside, into, near, nearby, off, on, onto, opposite, out, outside, over, past, through, throughout, to, toward, under, underneath, up, upon, via, with, within, without, far from, in back of, in between, in front of, in line with, on top of, to the left of, to the right of, to the side of

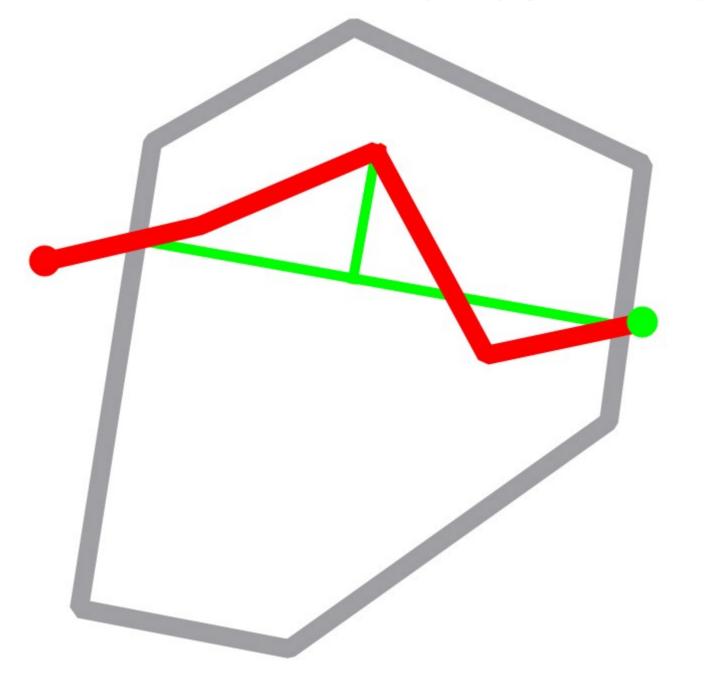
(from Landau and Jackendoff, 1993)

With your back to the windows, walk straight through the door near the elevators. Continue to walk straight, going through one door until you come to an intersection just past a whiteboard. Turn left, turn right, and enter the second door on your right (sign says "Administrative Assistant"). the windows, walk the elevators. Continue to walk you come an intersection turn right, and enter the second door "Administrative Assistant"). the door , going one door a whiteboard. Turn left, (sign says

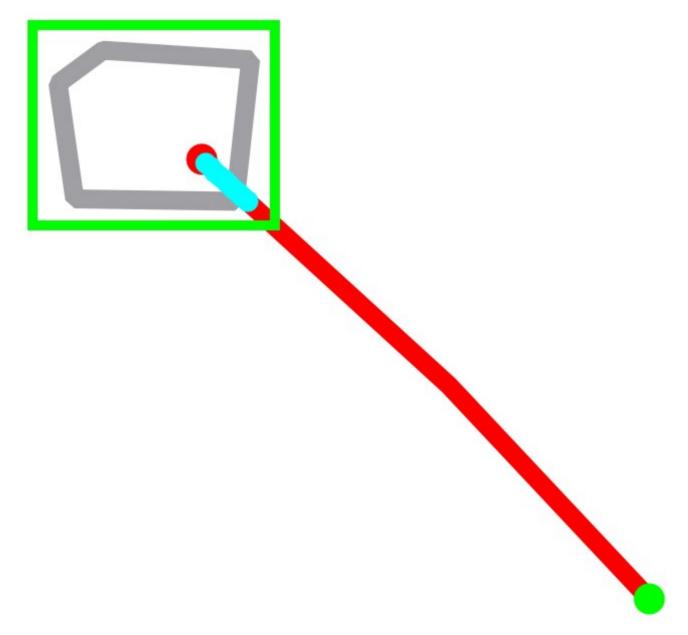
#### "Go past the chair."



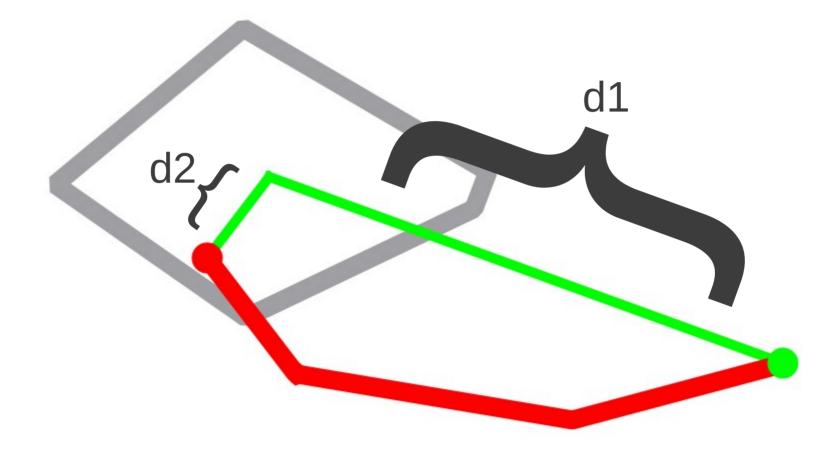
#### peakDistanceToAxes = max(dist(figure, axes))

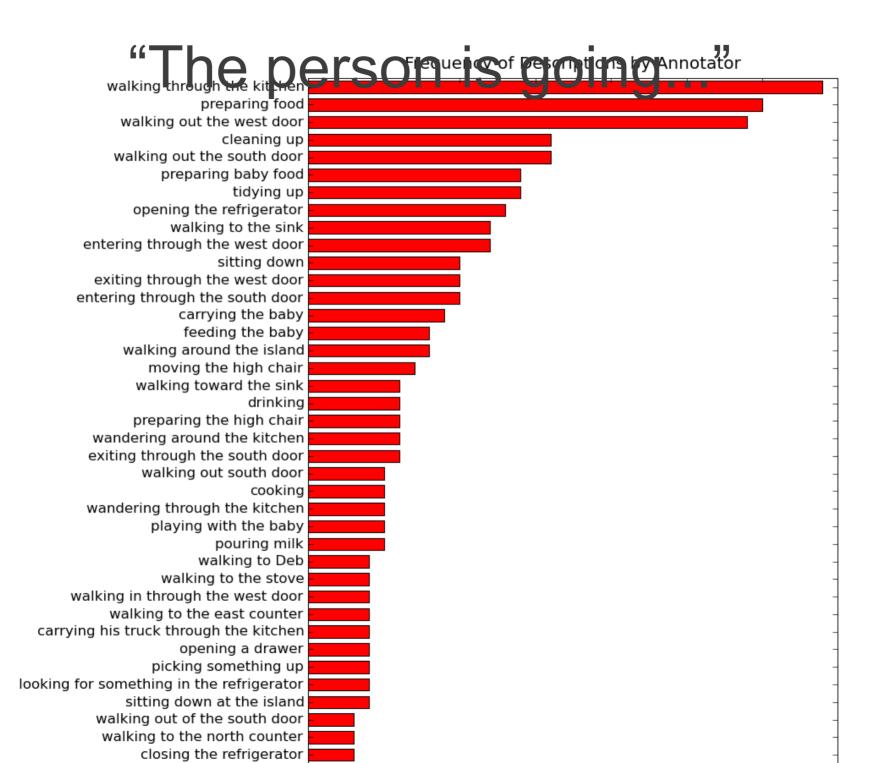


#### endPointsInGroundBoundingBox

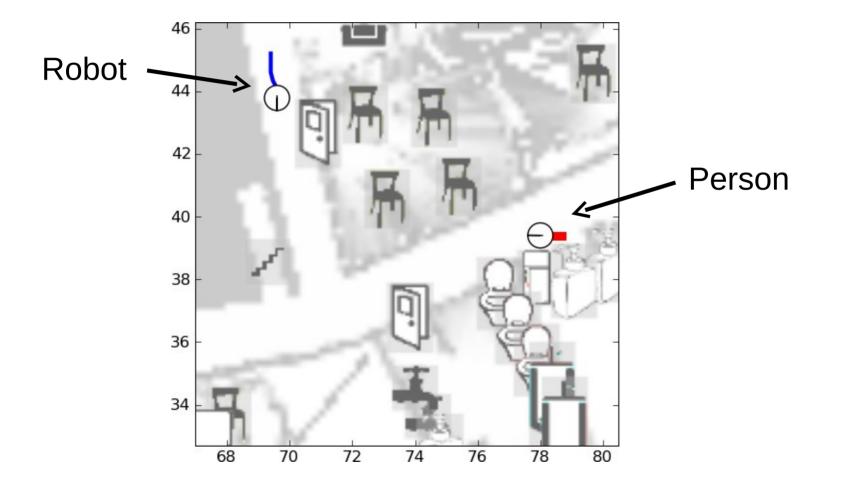


#### displacementFromGround=d2 - d1

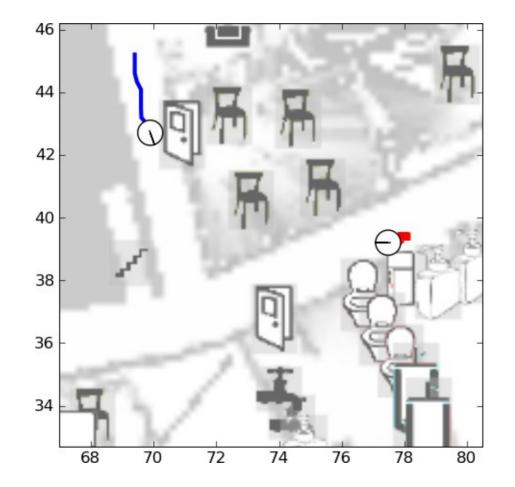




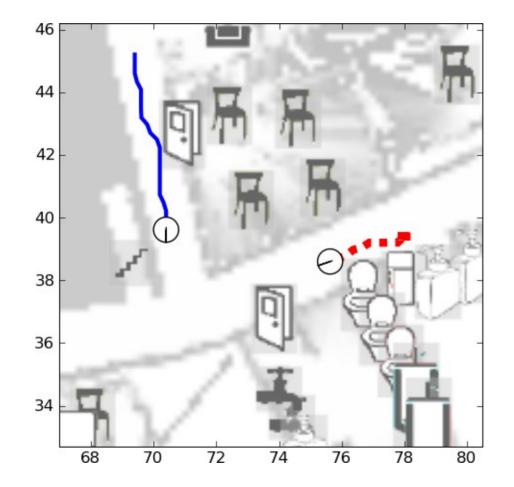
#### Meet the person.



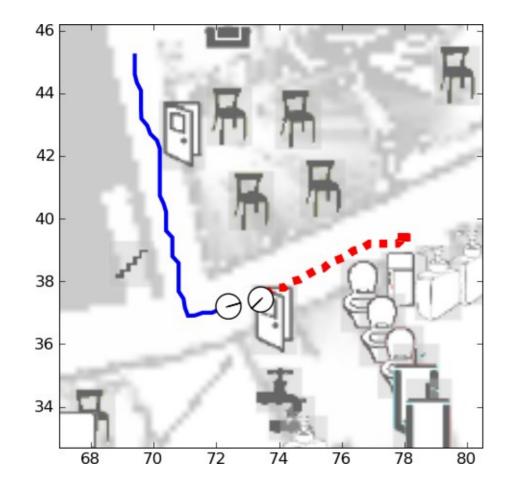
#### Meet the person.

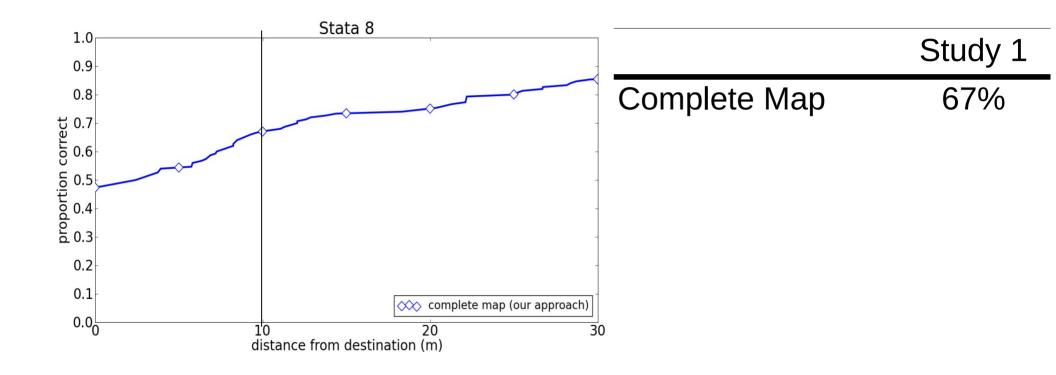


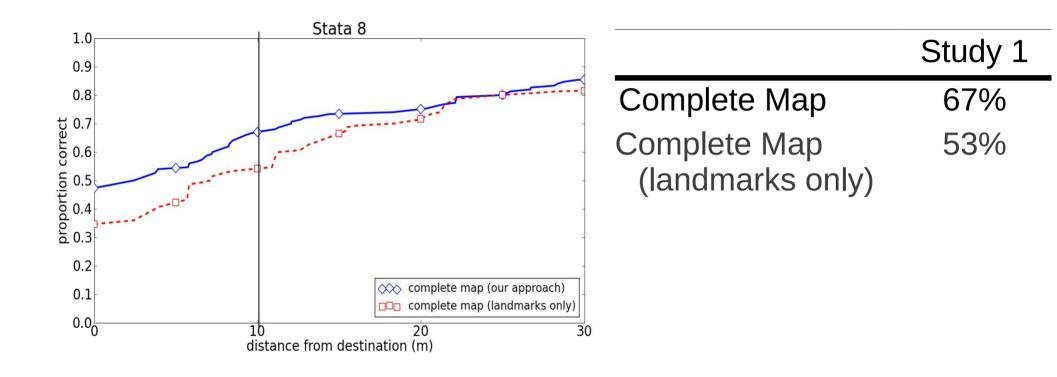
#### Meet the person.

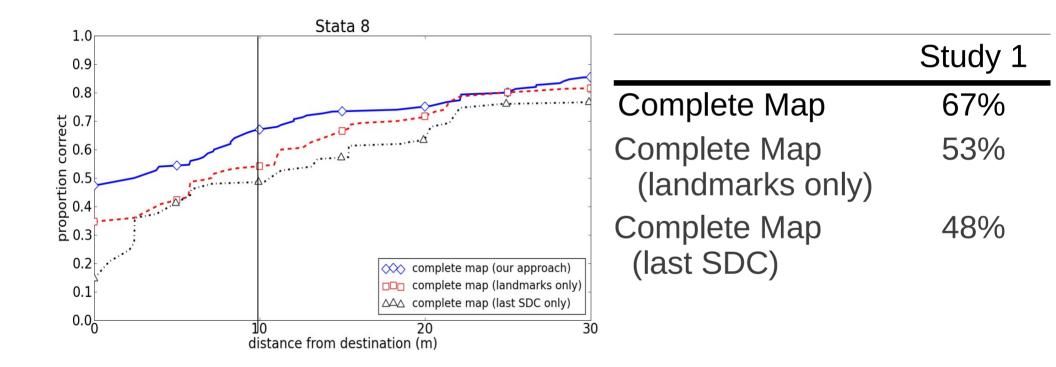


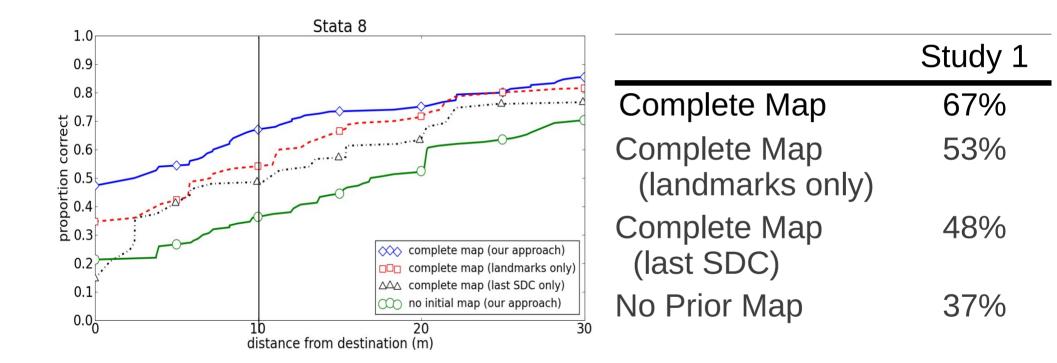
#### Meet the person.



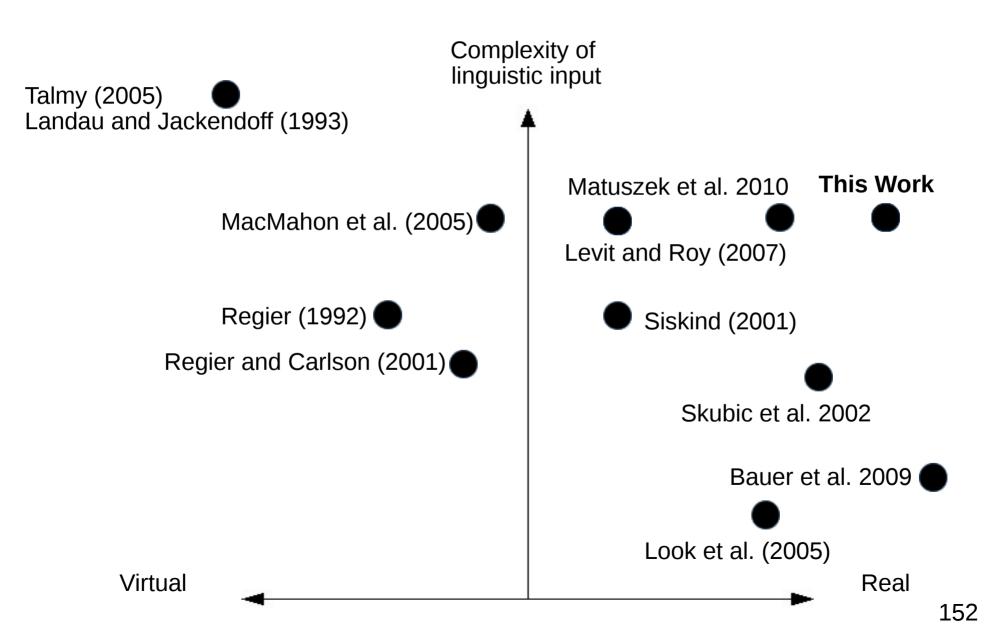






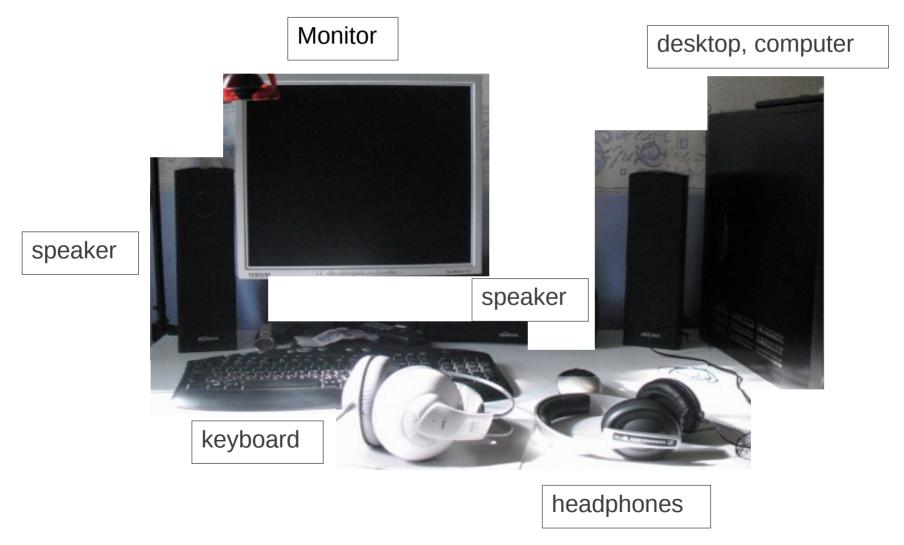


### **Related Work**



Show me people going across the kitchen.

#### Observations



fawn megatheriid shoebill bellmagpie birchrod nummery sunbonnet promptbox cobia horsecar aegis hanging downywoodpecker landingplace bald-facedhornet westerntoad boskopman leadpencil lory telephonebook seaboat cygnusolor salamandrasalamandra uropsilussoricipes morningecat drosky chimneycorner golfequipment deathma grk sanddollar landingstrip addressingmachine ramdisk hussamonkey swivel phoneline akingsnake arctocephalusphilippi maskingpiece storagemedium steeringsystem broiles astohel ladyfish pilling erisecanal circuitry cassok morobird tankship personnelipouto seabat mountainbike wishingcap okapiajohnstoni bannister second puddingwife jamjar tether whitepelican decapterusmacarellus jinrikisha tarsiussyrichta badmintonracquet basketweave europeansole collegiatedictionary heteroscelusineanus drinkingglass longtailweasel briefoaseomputer goshewk readhouse sheetande dourse basteture dourse sheetande collegiatedictionary heteroscelusineanus drinkingglass longtailweasel briefoaseomputer goshewk readhouse sheetande dourse basteture dourse sheetande collegiatedictionary heteroscelusineanus drinkingglass longtailweasel briefoaseomputer goshewk readhouse sheetande dourse sheetande dourse sheetande collegiatedictionary heteroscelusineanus drinkingglass longtailweasel briefoaseomputer goshewk readhouse sheetande dourse there withing for some source and the sheetande dourse sheetande dourse there withing for some source and the sheetande collegiatedictionary heteroscelusineanus drinkingglass longtailweasel briefoaseomputer goshewk readhouse sheetande dour parts for some source and the source and

Em spacecapsule chlorophonëusnigrifrons rivulus Sanchil frontprojector holidayresort tiptruck gr fretwork summons wage-lit remnant pound linot fraquail equuscaballus lygomorph pen-talledtresh d blade wineskim wage unskellunge tibetanteri rayon pinto dyke airthermometer stationsofthec sursopeannightjar mouflon diagram possum quipu w sunssubligarius wren romger backdrop wine power

#### initagram for tiptrook gr point remnant pound link mossellunge tibetatiering mermender stationsofthe med gram possum quipu w mer backdrop wine power 1- tailedkangaroo oldgiory bellpush hatchet ridgepole holocentrusasčensionis lifeline backpack neandertalhan hoodedseal chamatorquata ground-shaker polyzoan chickenhawk artificialhorizon fin chute-the owner bloodclam jurymast najanaja workhorse gellinula se citellurichardsoni hula-hoog phonebooth photocell medline conomysymmison veloco telephonesystem ravehook pseudaletiaunipuncta silver haddbaard fluteglass silievalve mastotemesslectrodominious razorfish chaetdoon salvelinusalpinus blueher alleuvrend

Schuler yoshawk rokamuse shekamun paspo oječina commonlouse sportinghouse drawkniše k rumner projection safehouse plowhorse teacloth g intake excost nailfile electricary beckethem unionjack cryostat borkodak loopknot cowshark fashionglate hydroid coffeeshop hairgel winch conicprojection truesparrow handhelddrill ser nch newelpost plowshare yellowfin playhouse as

commeshoe swine macroelemustemmineki navingmaterial oceanitesoceanicus laburinthofminos telescone swathe stabiliser garment blinder spider taphouse rubberstamp boxendwrench spottedweakfish babirusa sciaenid triptych bearclaw semiconductor dsl impulseturbine kingoftheherring fishingboat ter sträwman knit cornerstone gasturbine wadingpool americaneagle pitumuspinetorum birdie springgun retrenchment kakatoegalerita jumbojet pneumoenceph planetarugear roll-onroll-off alligatorsnappingturtle wasp uplandsandpiper guahaug wash aiguilette embouchure chevron carolinawren pronghorn stripedsguirrel pocketgopher silkstöcking cetorhinusmaximus resuscitator switchblade hardboard critter fix-itshop cou otable potholder rubicolarupicola removabledisk ässagai iceax riverorawn cremationchamber cruiseliner vibratingreed skimmer corbiegable birddog tarponatlanticus fluinggecko hibachi sheetrock monotreme shower eg nixindhowl lander dozer snakenit runner dwarfn gasolinebomb shootingbox waterbottle zalophuslobatus deathadder storeriagccipitamaculata ukilotherm nacifichalibut displaucase culdelamne drainn nbrasebook warragal scunbozoan telemeter greekfret crematorium greenturtle navement stens meringue rranhicmaterial bathrobe textbook restorative aint eschrichtiusrobustus denartmentstore locket templeofartemis elopssaurus blacktopping oilpainting titmouse falloutshelter whirlaway banister electric-arcfurnace ailuropodamelanoleuca americanm thundersnake motorcoach burncenter orkquilt dressing shade boulevard woodtick bubble-jetprinter) minterface evanorometer siding future stammel tensimeter darko guadalupefurseal letteropener greenwing foul-weathergear twin powerpack bicuclerack boar mousetrap boat haemulonparra entrant dini: knobble rotarmactuator tocsin bandedgecko daro. suchiatrichospital indiancobra weatherglass masonjar marmeflute canillarutuhin portraitlens hirudomedicinalis italiangreuhound summerredbird cuclops ropetow torsk jackhammer wateroistol pulr conventicle rhodesianridgeback engine cashdispenser falsie crevallejack ai wire nekinese cream collie commonnewt ual nuclearweapon elasticband corker brandysnifter microchip enchiridion mitrebox operatinginstructions saree calciumlight biped salvut leastbittern landir oocketcomb mightumouse hamster bradawl cameralens action icenick watercart airhole oilmaint handset odometer mexican tone swan'sdown adobehouse scan bogy gristmill bidet safetyrail scad longpillow whitetigshark chancellery broadbill sika squeezebox gearlever miner'scat craniate bronze dress pneumaticityre tapestry axis beehouse officer'smess compoundient seasing redbat strawworm speedboat sk w duplicidentata hockeupuck medicineball

#### Corpus

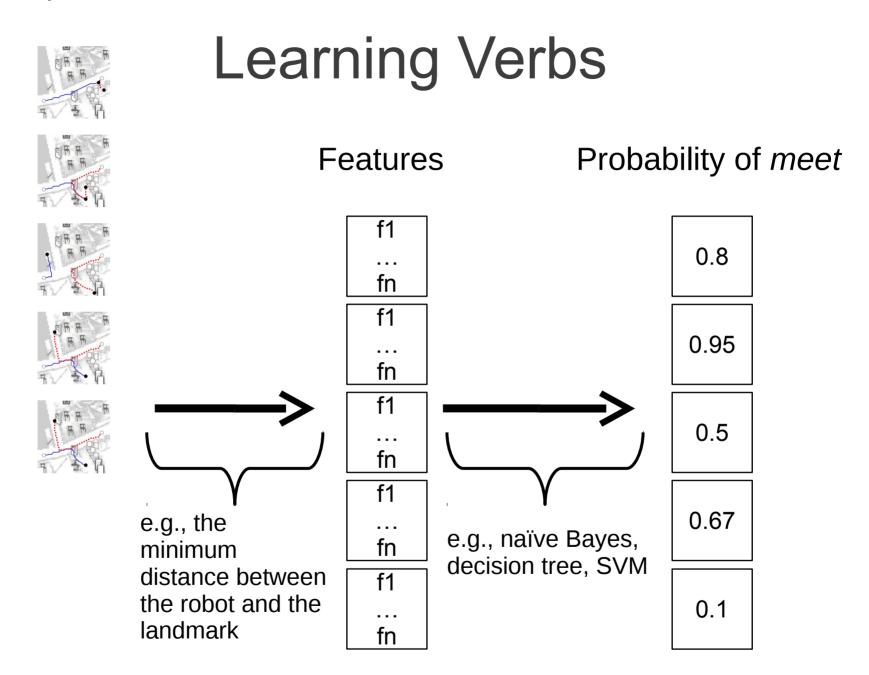


Positive example of "to the counter" Negative example of "across the counter"

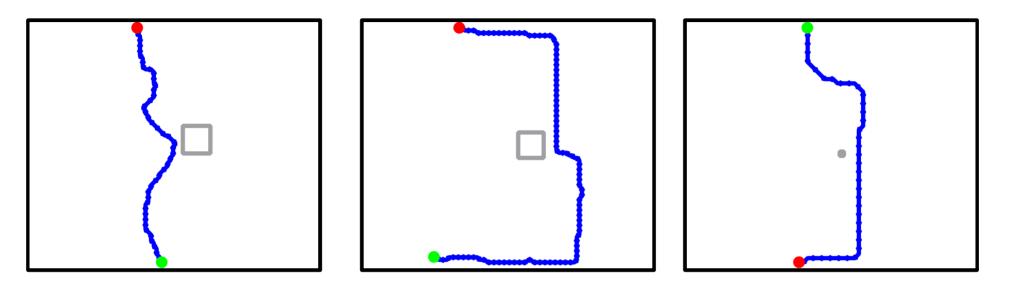
# Verbs in More General Commands

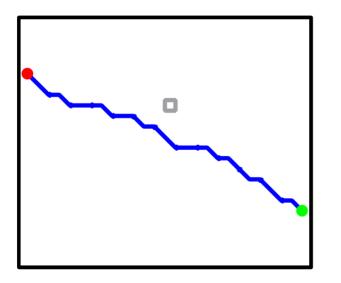
- "Bring the CHW folks from the windy stairs to 395."
- "Meet Javier at the Dreyfus entrence and bring him to the 33x neighborhood."
- "If visitors try to go to the wrong room, chase them to the right room."

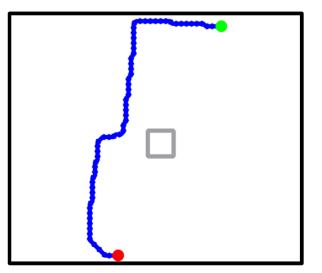
#### Examples of *meet*



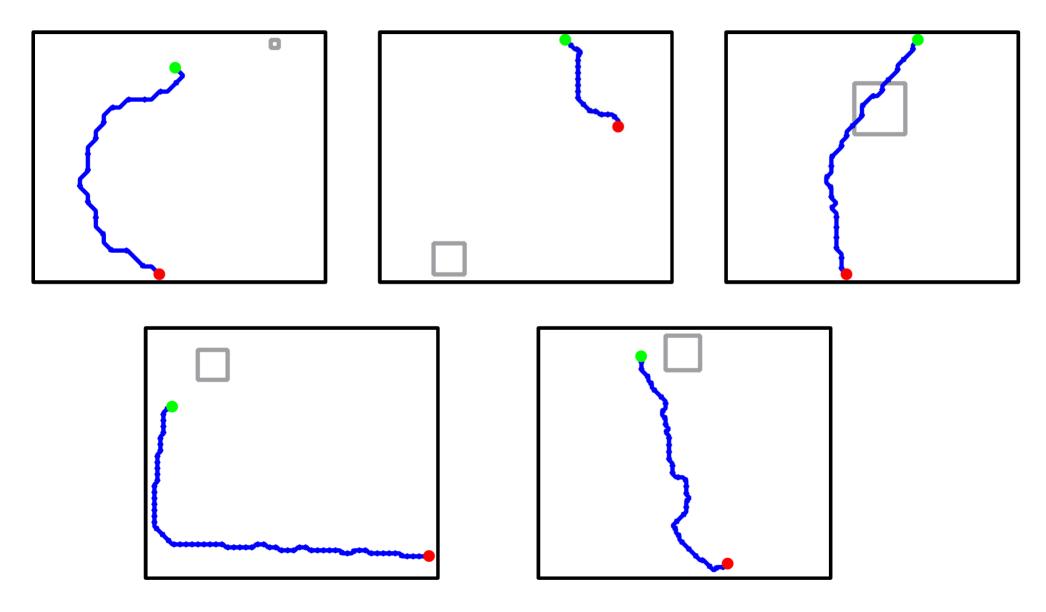
# High Scoring Examples of "past"

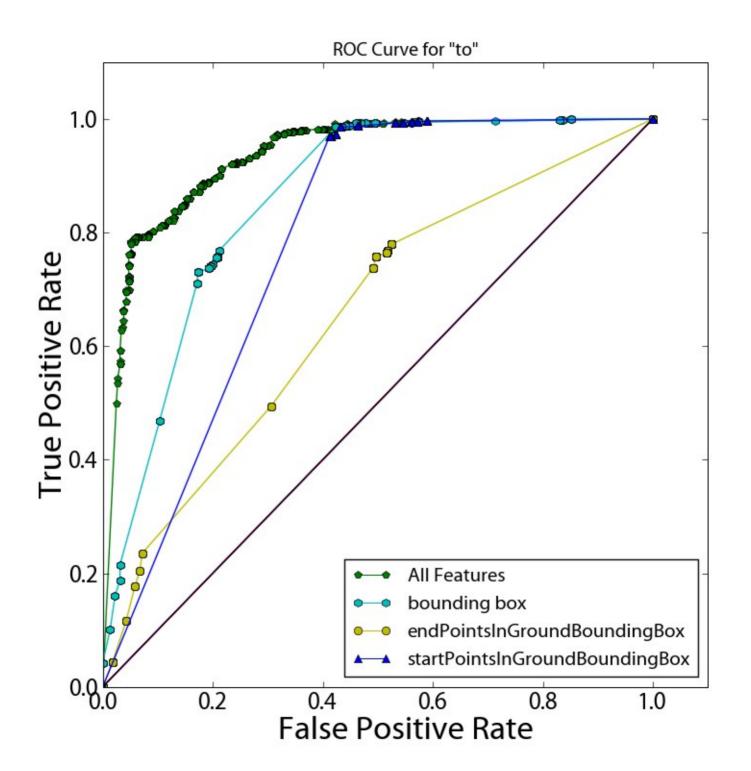




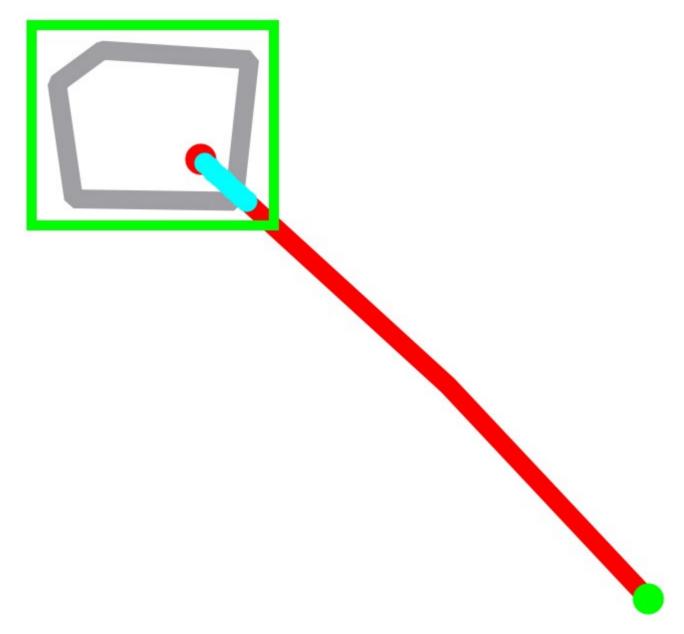


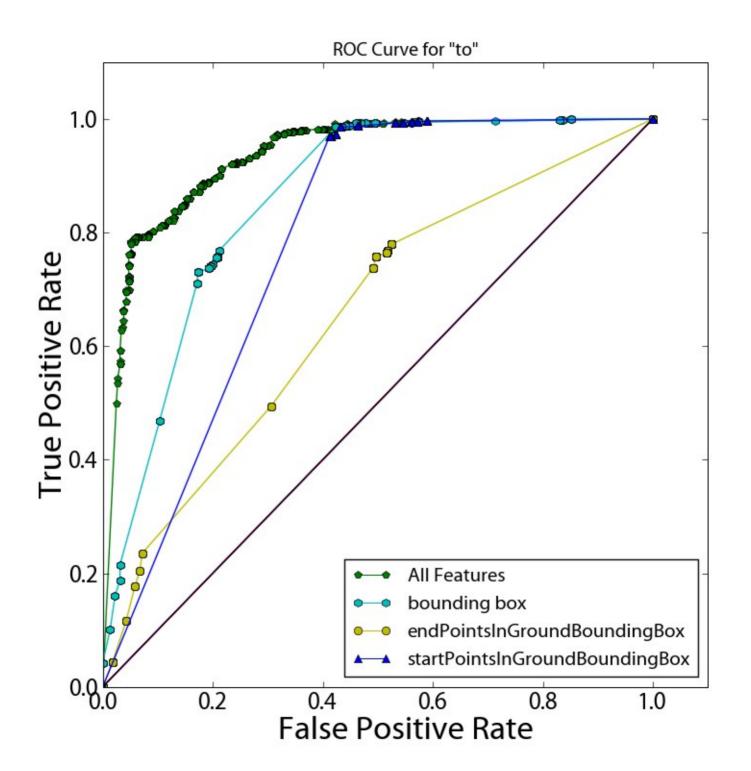
# Low Scoring Examples of "past"





#### endPointsInGroundBoundingBox









#### Full model: 0.99 Bounding boxes: 0.76





Full model: 0.10 Bounding boxes: 0.63



### Grounding

#### He's going across the kitchen.

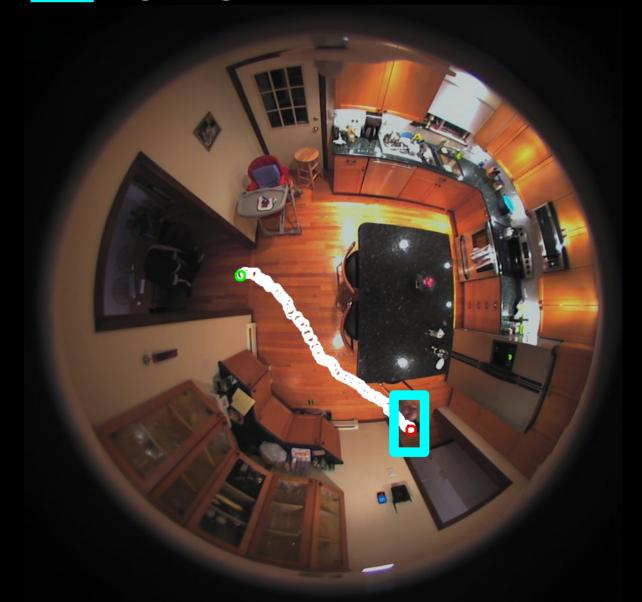


### Grounding

#### He's going across the kitchen.



#### Grounding He's going across the kitchen.



### Grounding He's going across the kitchen.



# Grounding He's going across the kitchen.

