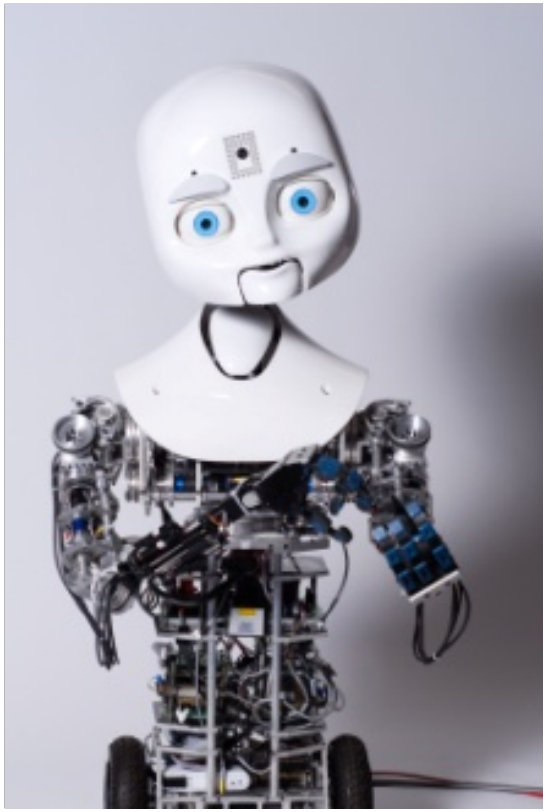
Natural Language and Spatial Reasoning

Stefanie Tellex

What is he doing?

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").



What semantic structures can enable a system to understand and use spatial language in realistic situations?

Methodology

- Pick a subset of language.
- Collect a corpus focusing on that subset.
- Make a program that understands language in the corpus.
- See how well the program works.
- Analyze why the program works.

Methodology

- Pick a subset of language.
- Collect a corpus focusing on that subset.
- Make a program that understands language in the corpus.
- See how well the program works.
- Analyze why the program works.
- Do this in more than one domain.

Outline

- Problem Statement
- Spatial Prepositions
- Spatial Language Video Retrieval
- Direction Understanding
- Next Steps
- Contributions

Spatial Language is Pervasive

“Where is the bus stop?”

“Take a right at the elevators.”

“Through the years you've never let me down.”

“This book jumps across many different topics.”

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.

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, quite, 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,

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, quite, 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,

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").

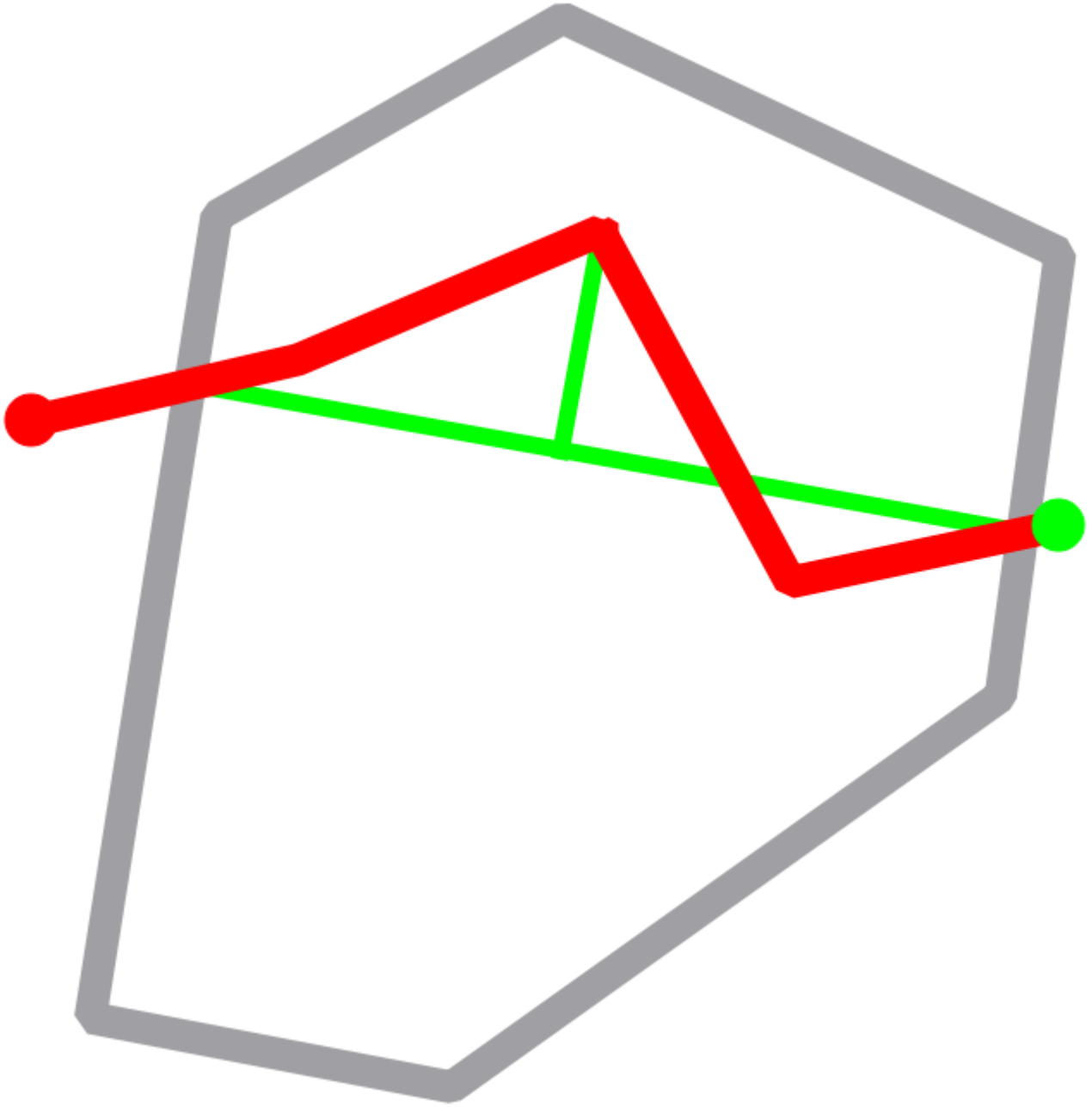
the windows, walk the door the elevators. Continue to walk , going one door you come an intersection a whiteboard. Turn left, turn right, and enter the second door (sign says "Administrative Assistant").

Modeling Spatial Prepositions

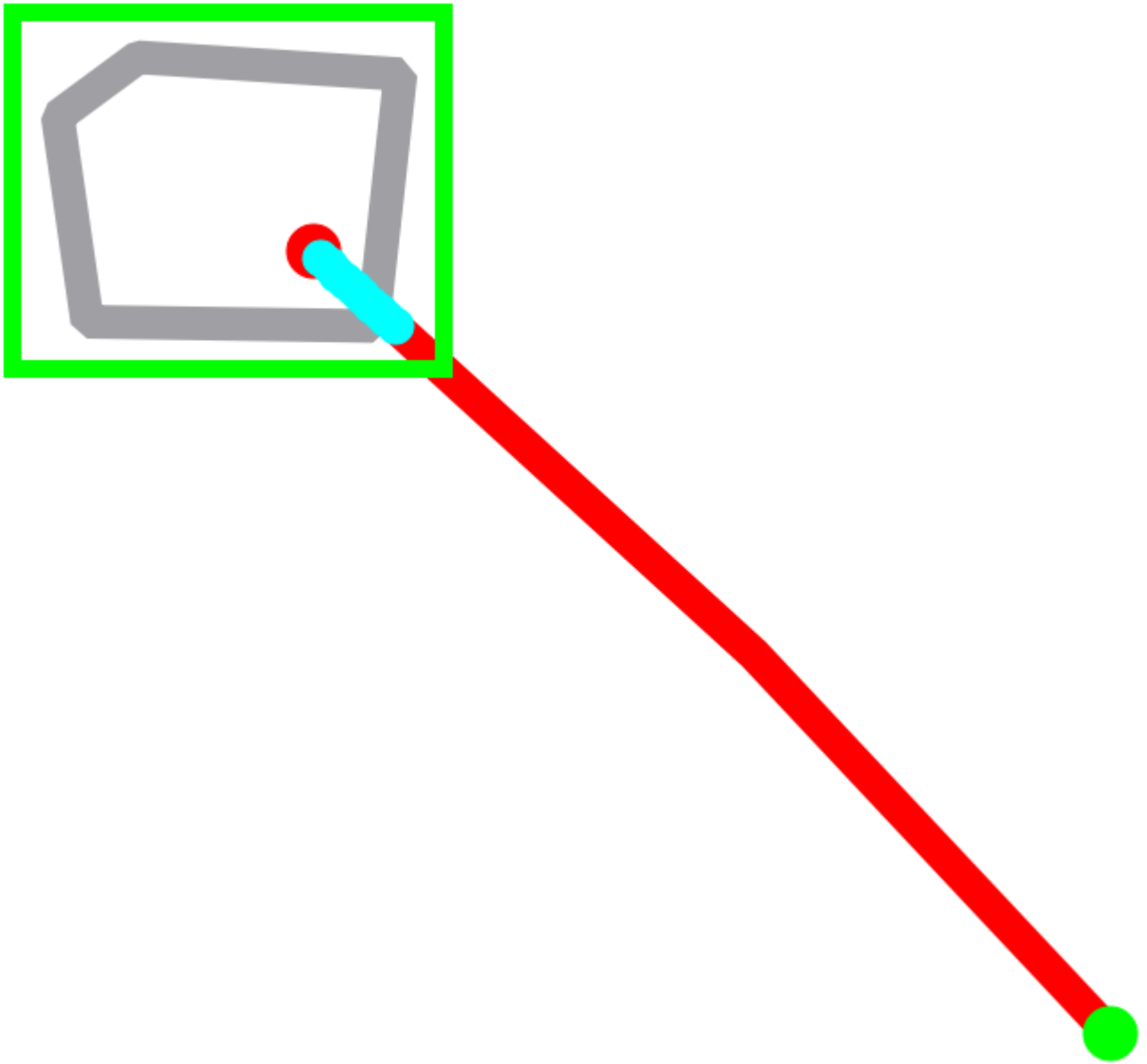
- Functions that take geometric arguments.
- Classifiers for “to,” “across,” “towards,” “through,” “around,” etc.
- Library of features.
- Train and test on labeled examples.

Across Video

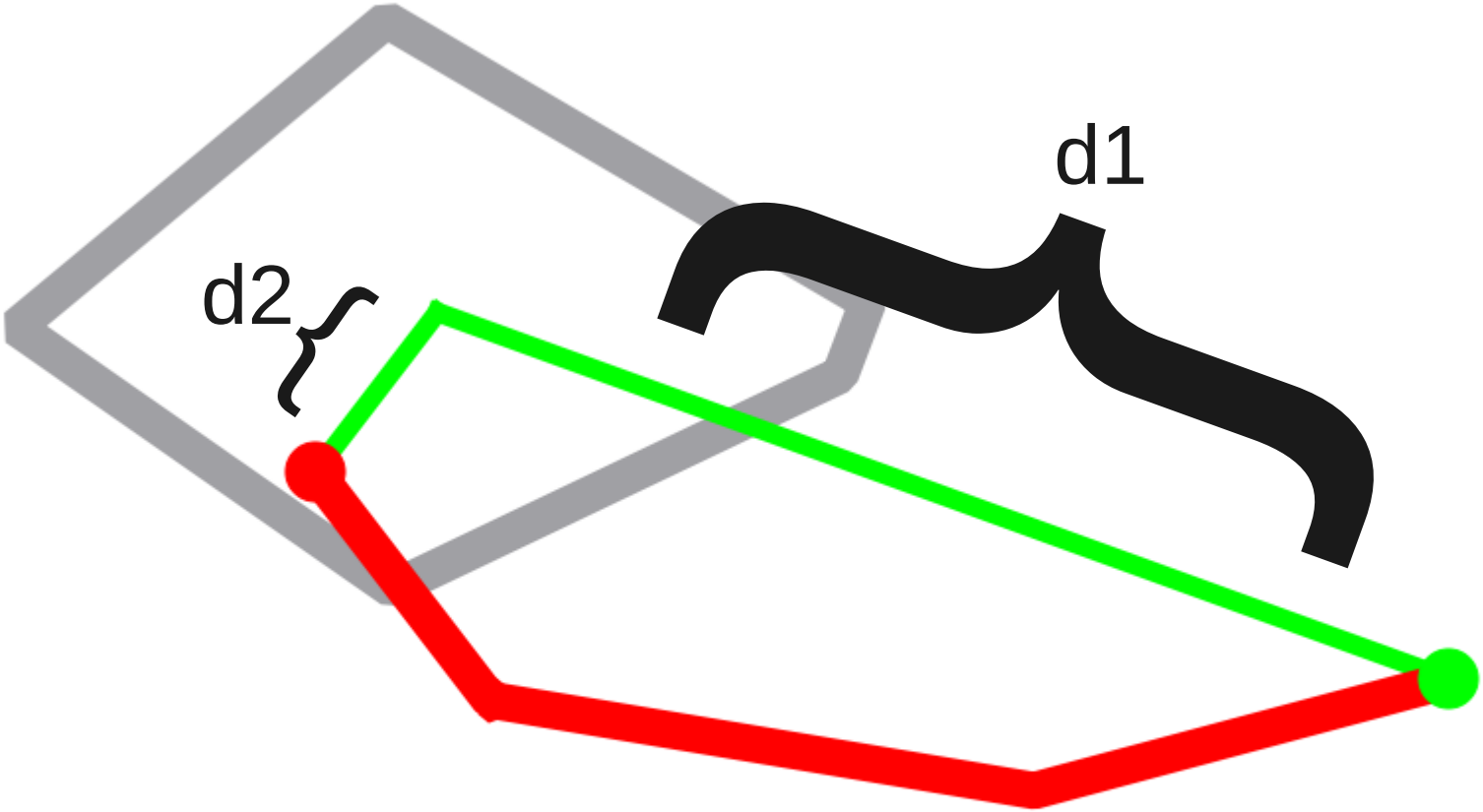
peakDistanceToAxes = max(dist(figure, axes))



endPointsInGroundBoundingBox



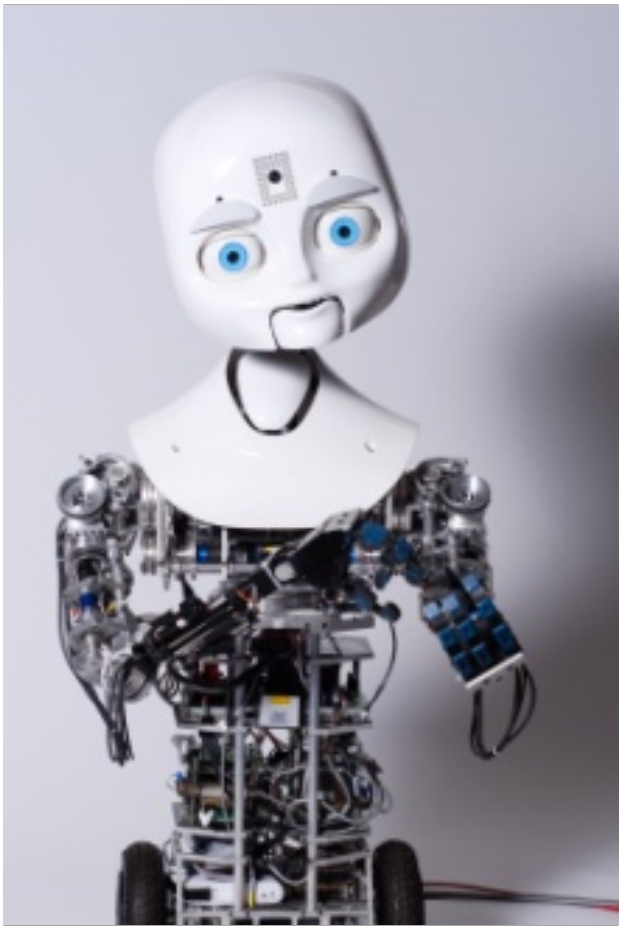
displacementFromGround=d2 - d1



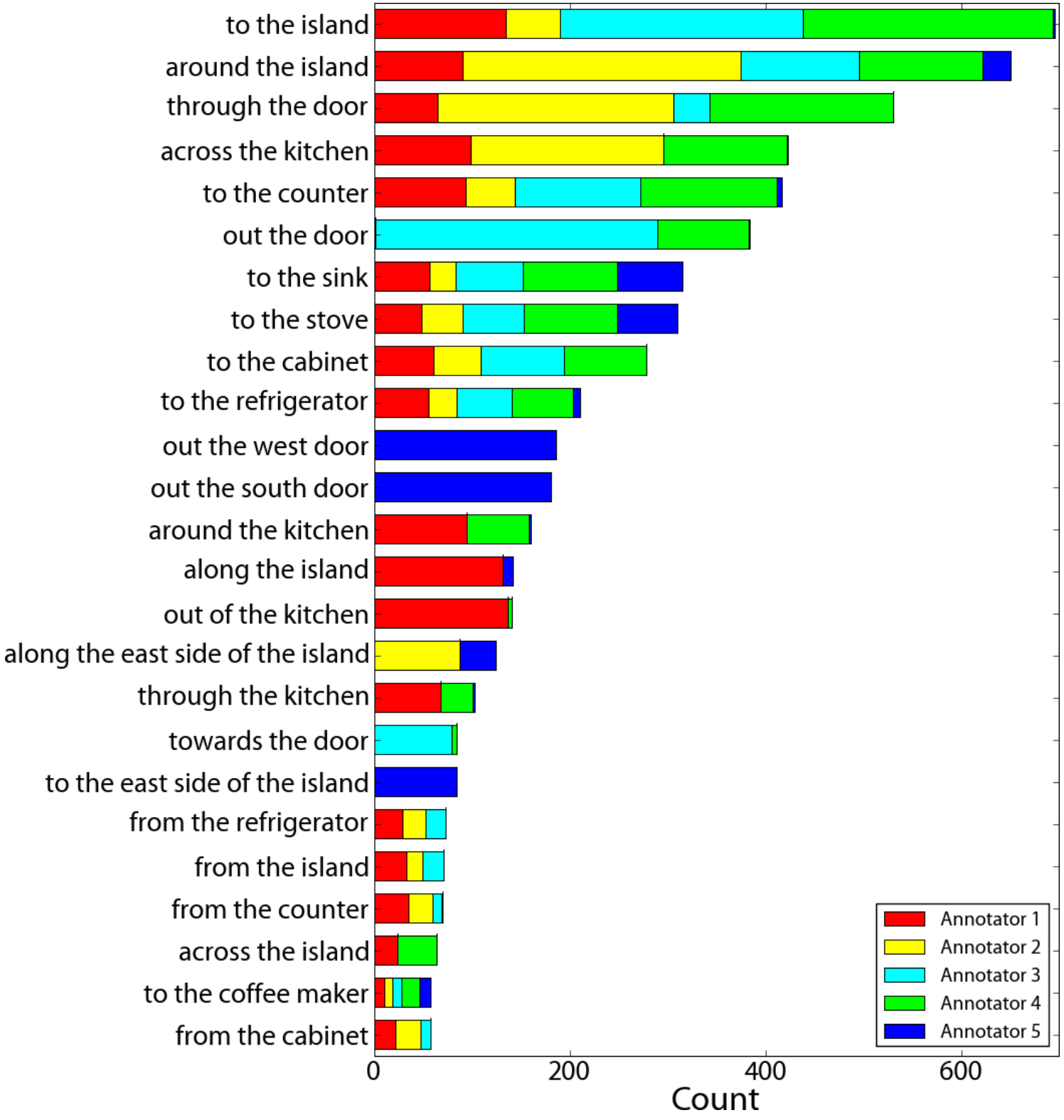
Two Different Domains

- Video Retrieval

- Direction Understanding



The person is going...


to the island
around the island
through the door
across the kitchen
to the counter
out the door
to the sink
to the stove
to the cabinet
to the refrigerator
out the west door
out the south door
around the kitchen
along the island
out of the kitchen
along the east side of the island
through the kitchen
towards the door
to the east side of the island
from the refrigerator
from the island
from the counter
across the island
to the coffee maker
from the cabinet
Annotator 1
Annotator 2
Annotator 3
Annotator 4
Annotator 5
0
200
400
600
Count

Corpus

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

Training Classifiers for Spatial Prepositions

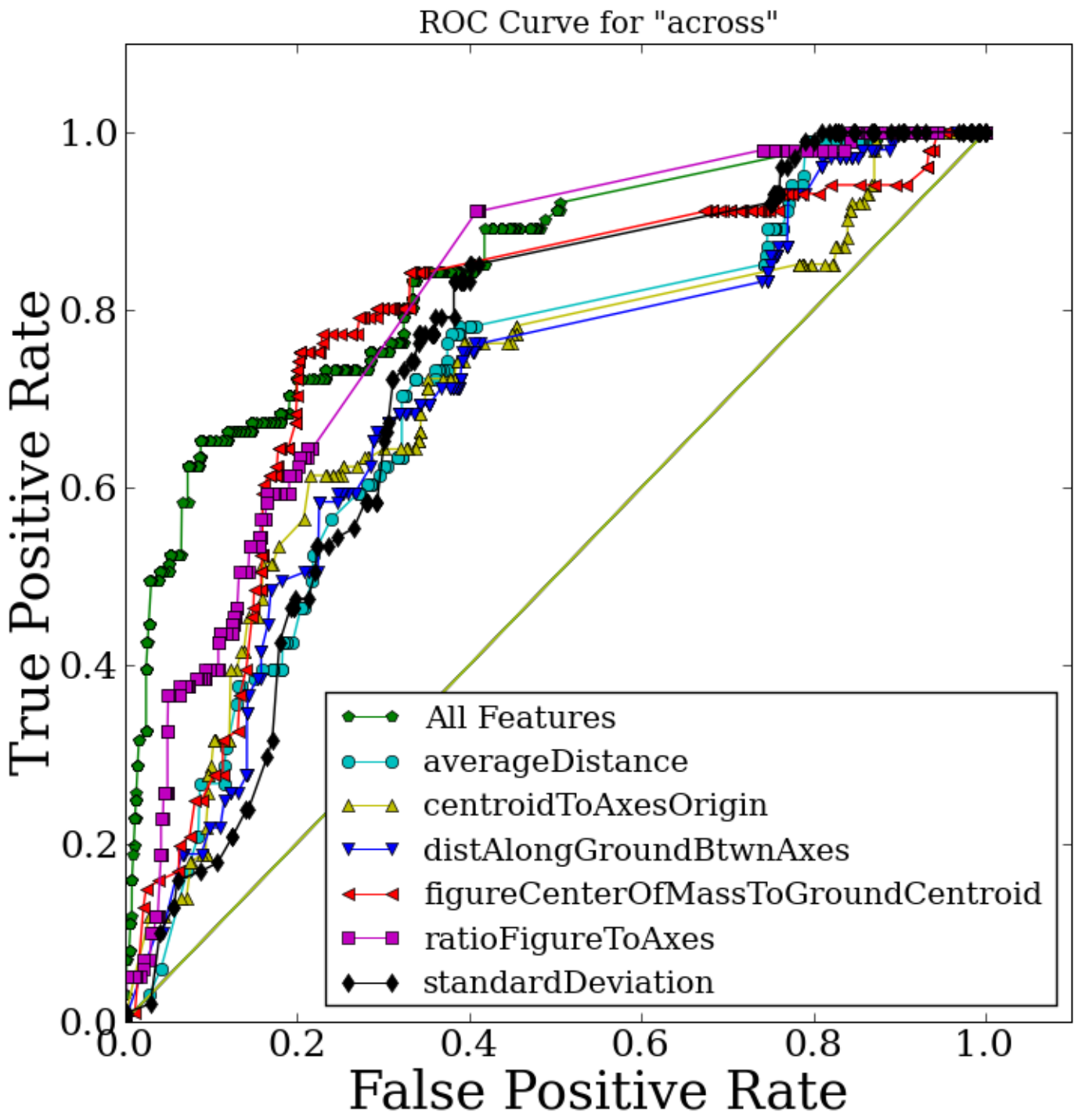
Binary classifiers.

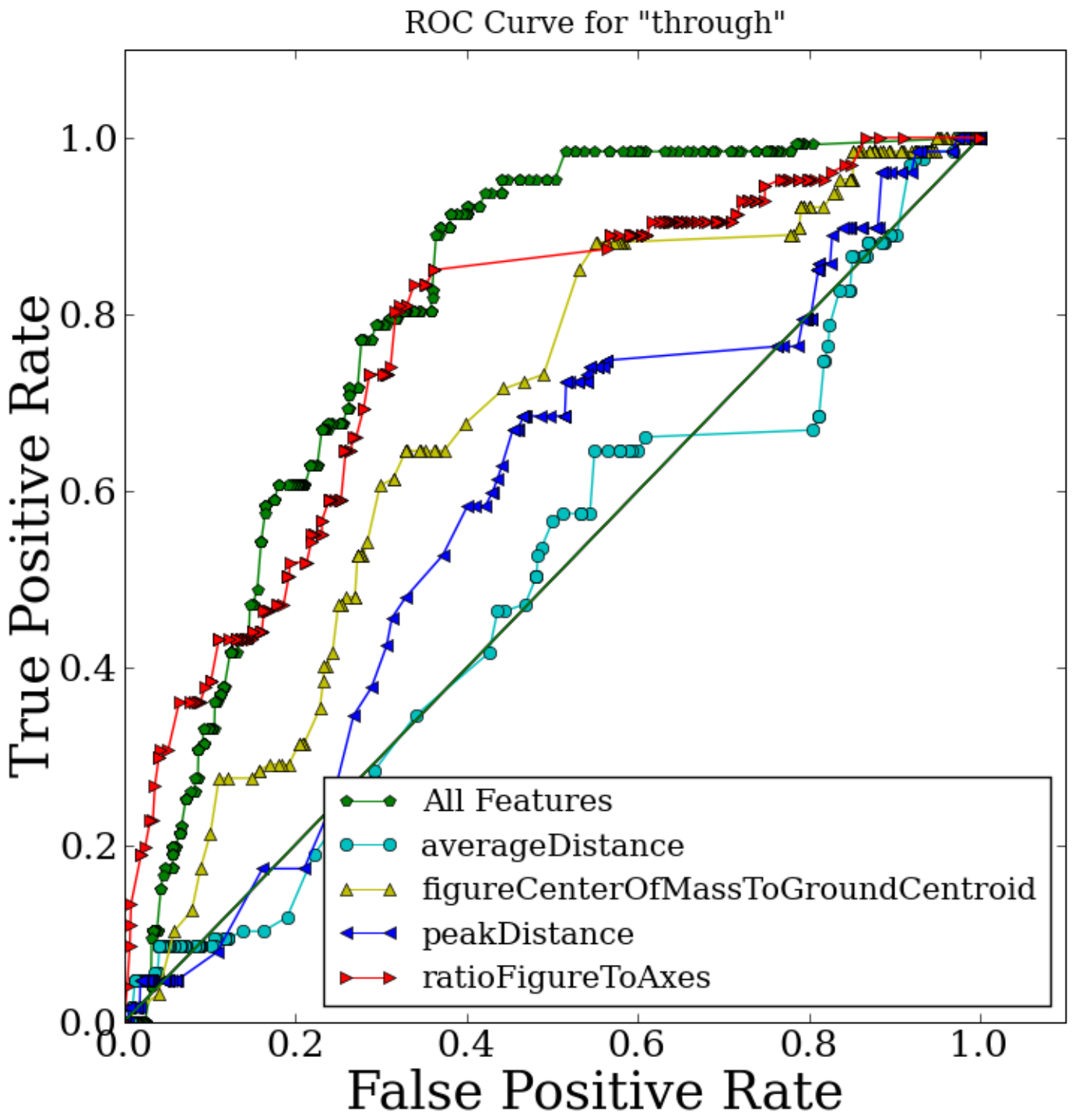
Positive examples of “across” are negative examples of “to.”

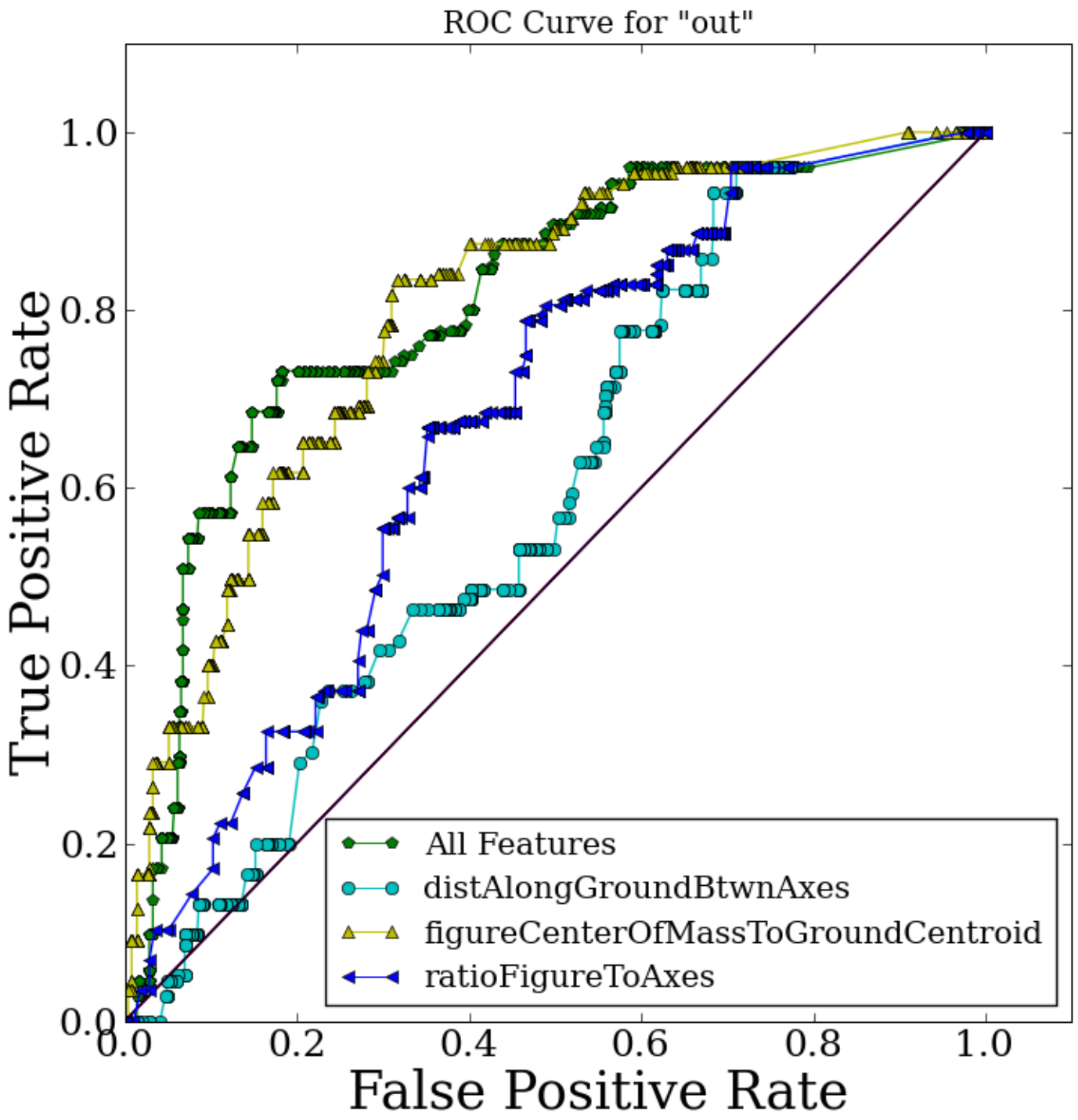
- Some exceptions (e.g., “to” and “towards.”)

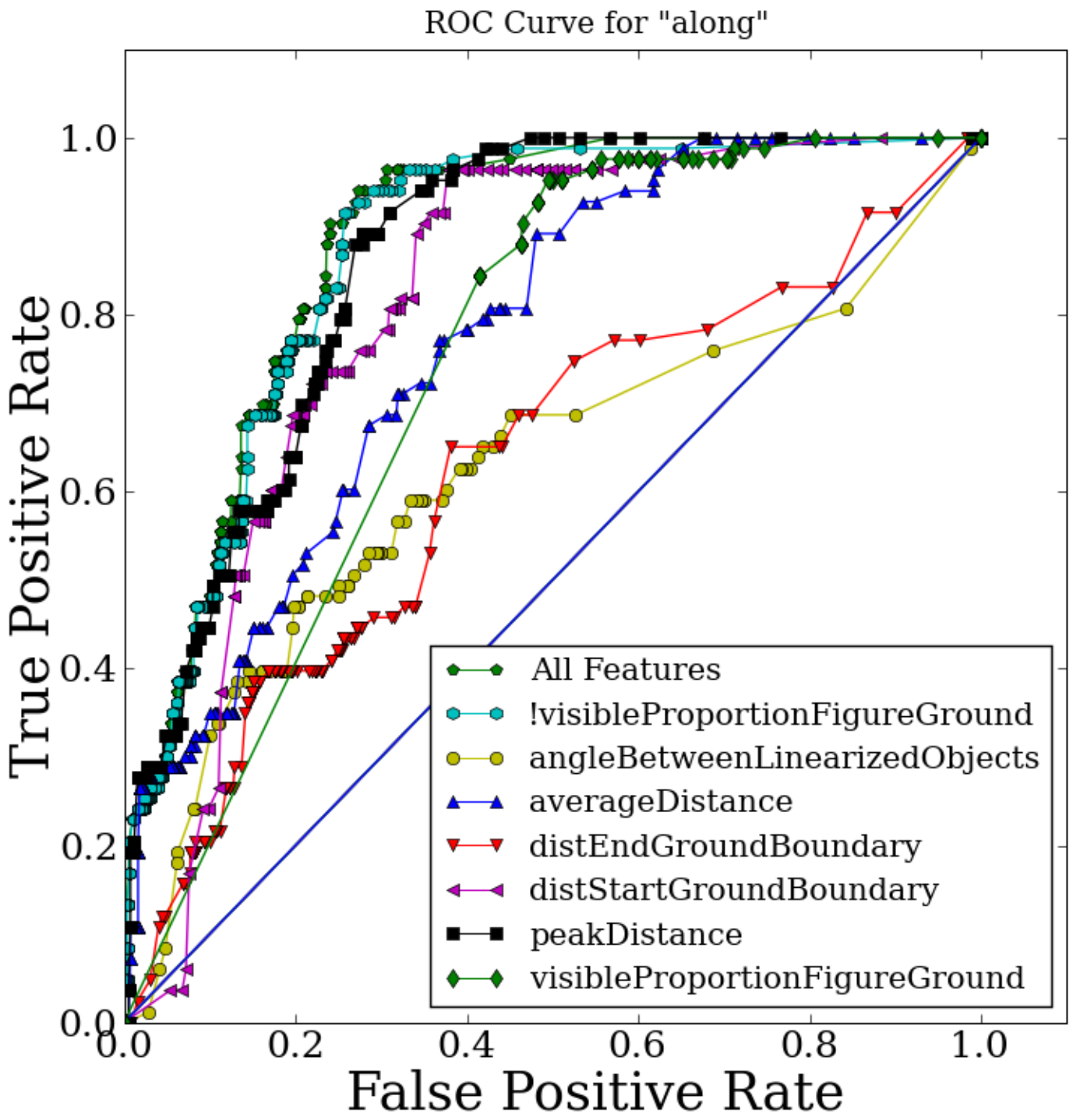
Corpus

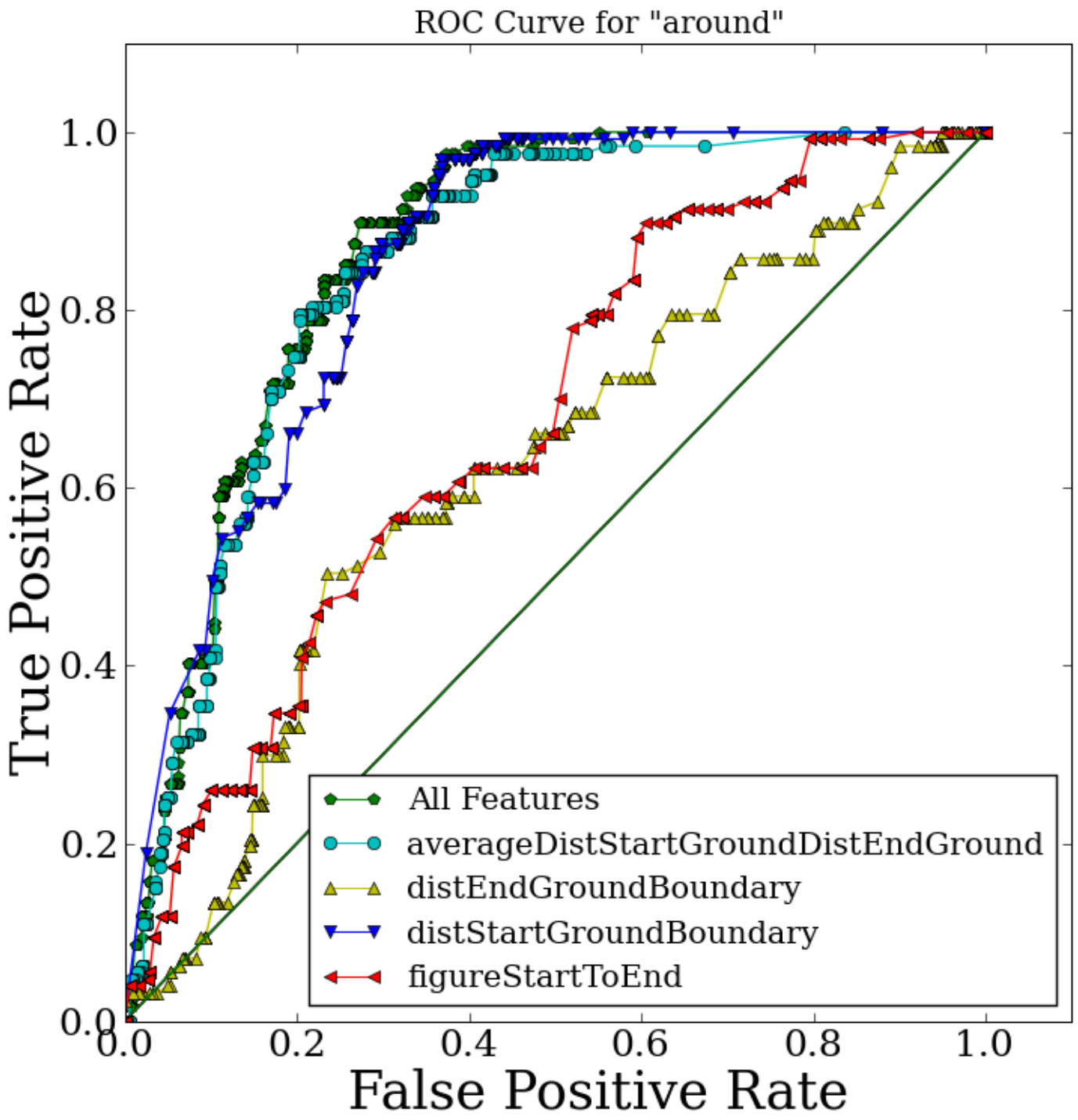
Positive example of “to the counter”
Negative example of “across the counter”

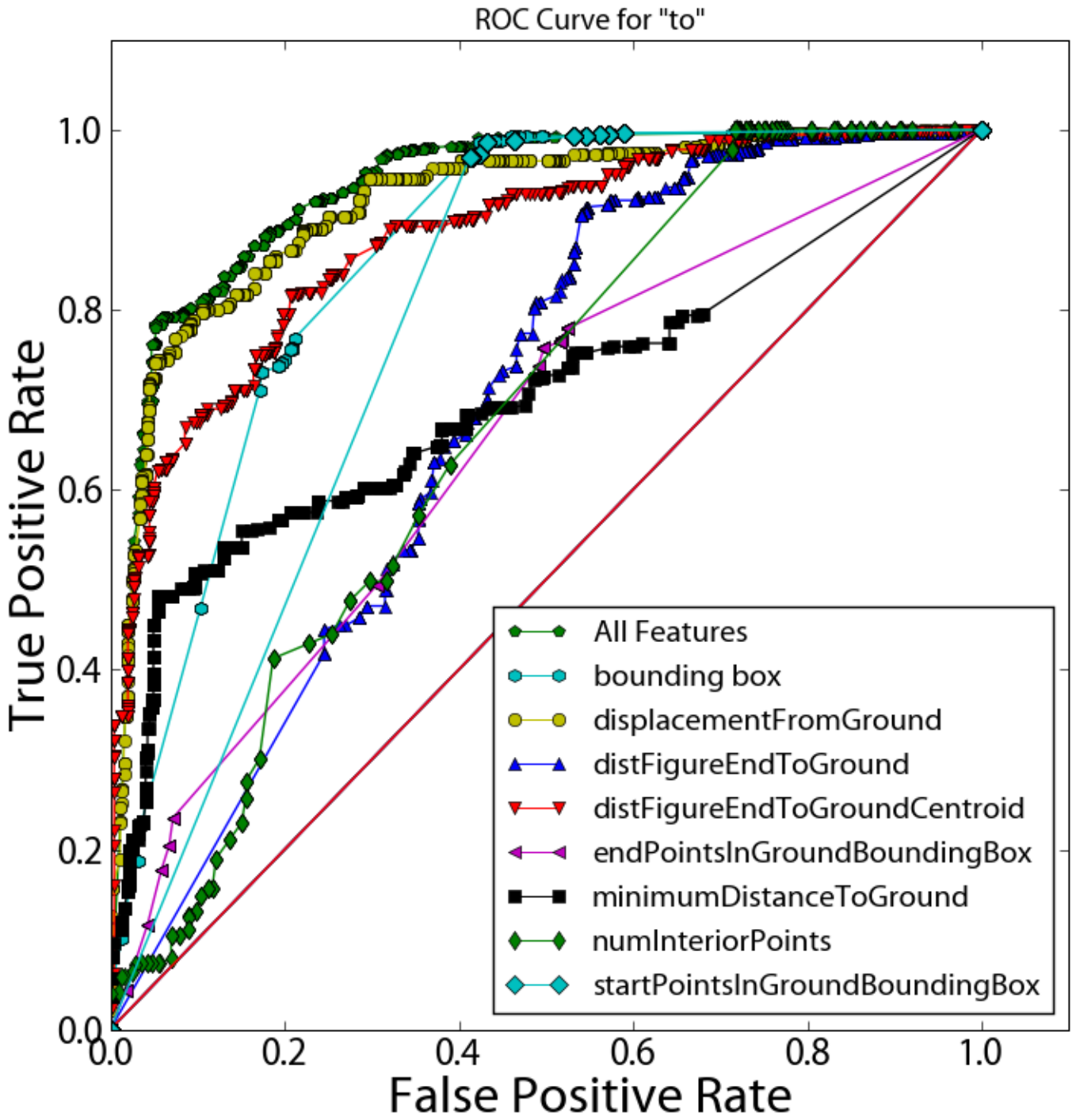

ROC Curve for "across"
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
All Features
averageDistance
centroidToAxesOrigin
distAlongGroundBtwnAxes
figureCenterOfMassToGroundCentroid
ratioFigureToAxes
standardDeviation

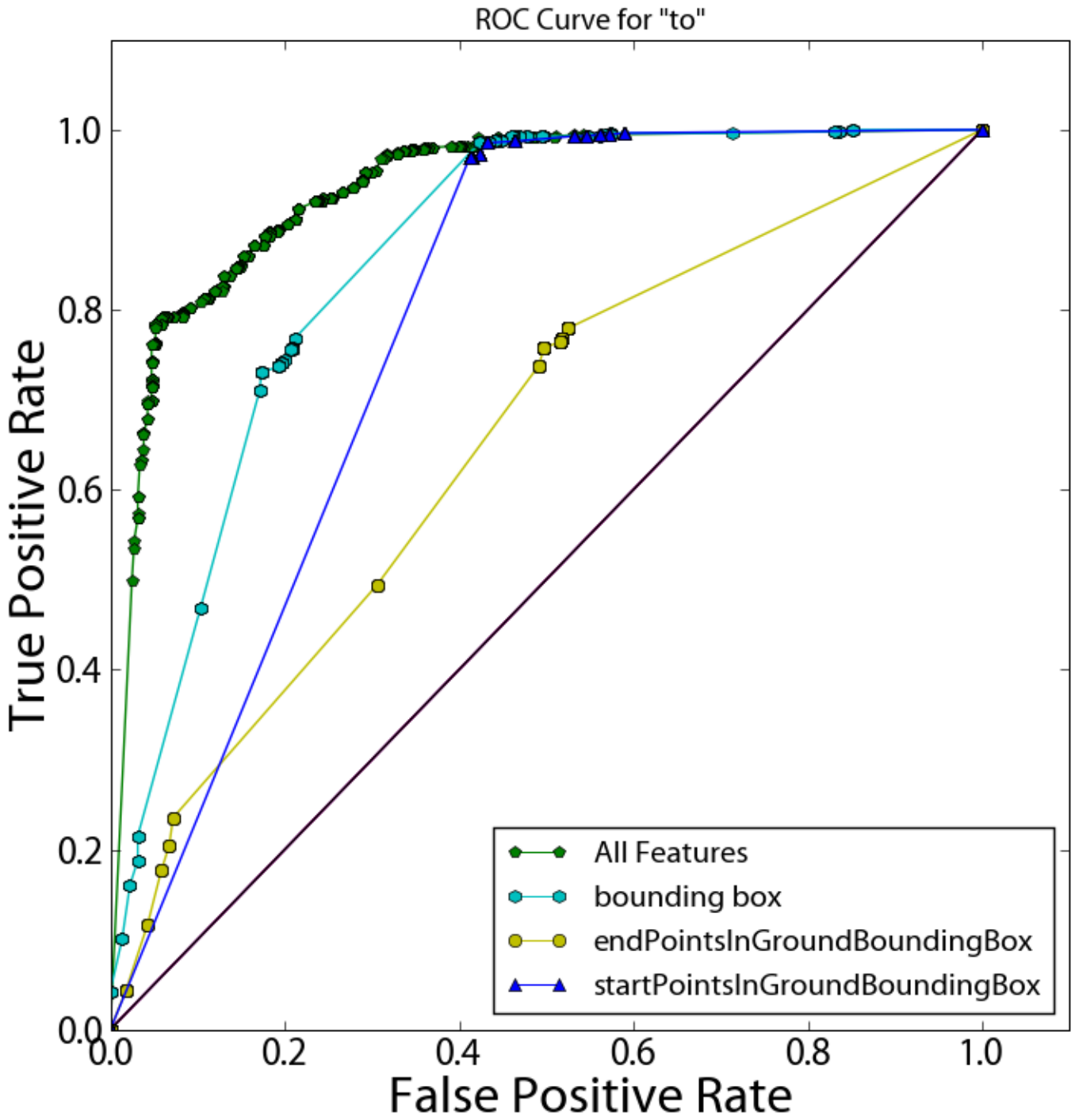

ROC Curve for "through"
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
All Features
averageDistance
figureCenterOfMassToGroundCentroid
peakDistance
ratioFigureToAxes


ROC Curve for "out"
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
All Features
distAlongGroundBtwnAxes
figureCenterOfMassToGroundCentroid
ratioFigureToAxes

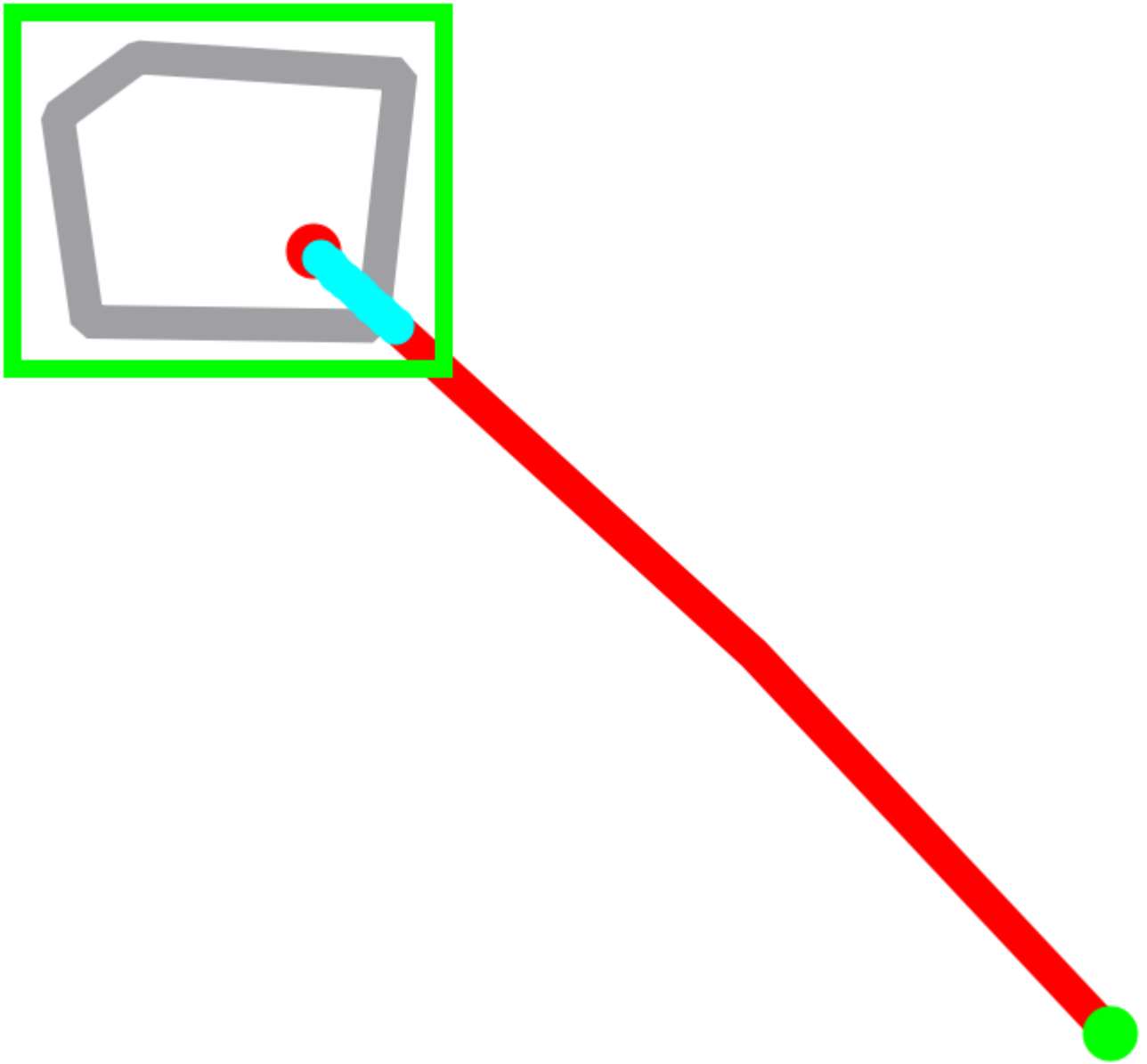

ROC Curve for "along"
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
All Features
!visibleProportionFigureGround
angleBetweenLinearizedObjects
averageDistance
distEndGroundBoundary
distStartGroundBoundary
peakDistance
visibleProportionFigureGround


ROC Curve for "around"
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
All Features
averageDistStartGroundDistEndGround
distEndGroundBoundary
distStartGroundBoundary
figureStartToEnd


ROC Curve for "to"
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
All Features
bounding box
displacementFromGround
distFigureEndToGround
distFigureEndToGroundCentroid
endPointsInGroundBoundingBox
minimumDistanceToGround
numInteriorPoints
startPointsInGroundBoundingBox


ROC Curve for "to"
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
All Features
bounding box
endPointsInGroundBoundingBox
startPointsInGroundBoundingBox

endPointsInGroundBoundingBox



“to the sink”

Full model: 0.99
Bounding boxes: 0.76

“to the sink”

Full model: 0.10
Bounding boxes: 0.63

Results

- Modeled six spatial prepositions.
- Evaluated on a corpus of natural language descriptions of video clips.
- Analyzed which features work.

Limitations

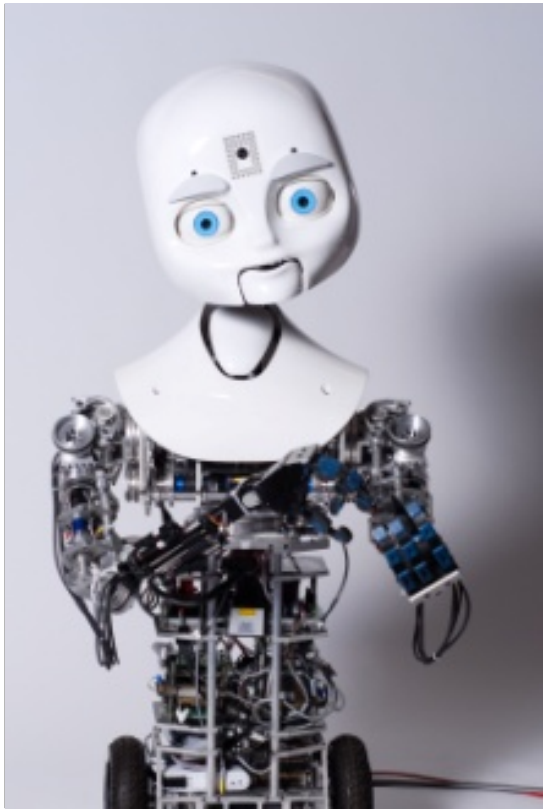
- Phrases, not sentences or paragraphs.
- All in one room.
- Only six spatial prepositions.

Direction Understanding

Joint work with Thomas Kollar

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").

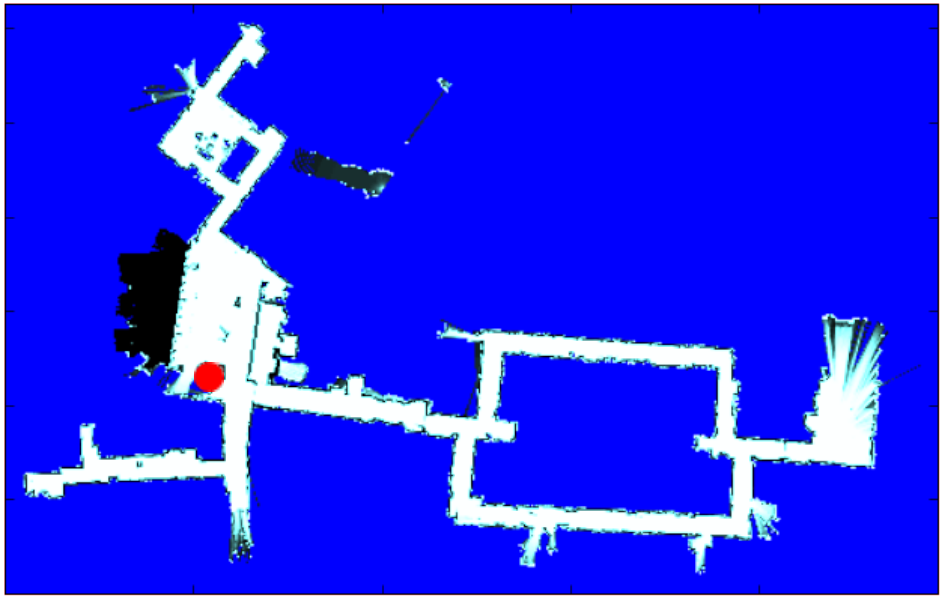


Structure of Natural Language Directions

- Sequential

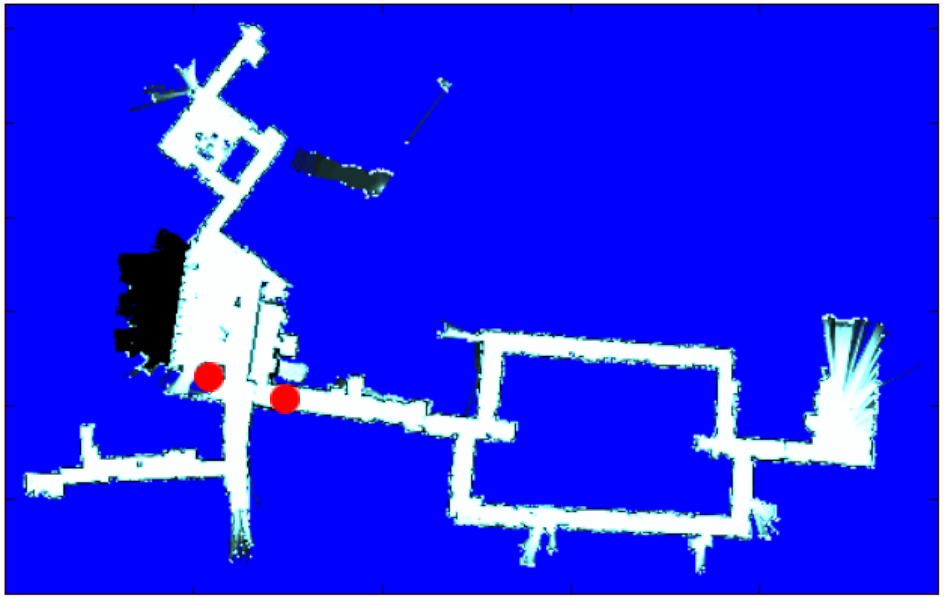
Sequential

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 white board. Turn left, turn right, and enter the second door on your right (sign says "Administrative Assistant").



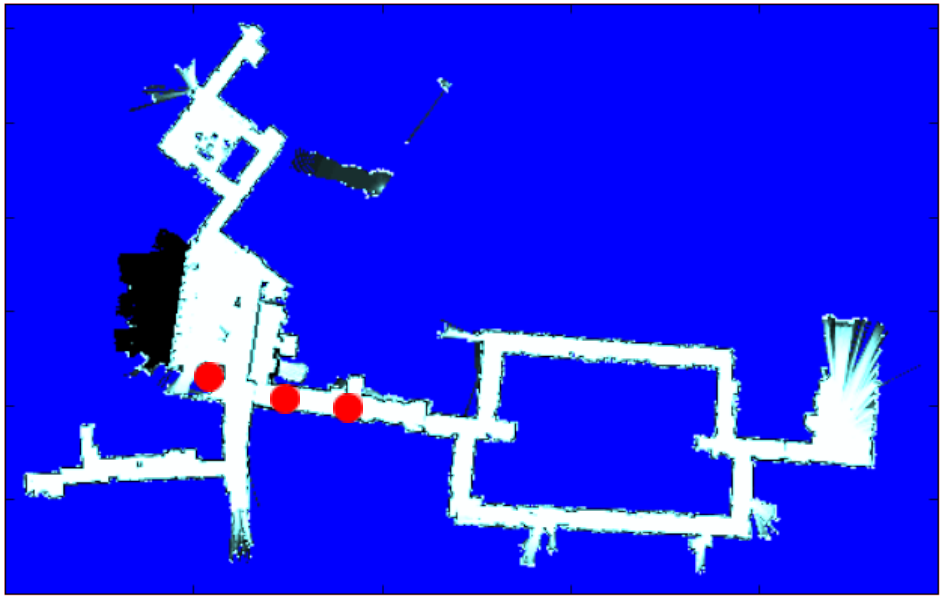
Sequential

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 white board. Turn left, turn right, and enter the second door on your right (sign says "Administrative Assistant").



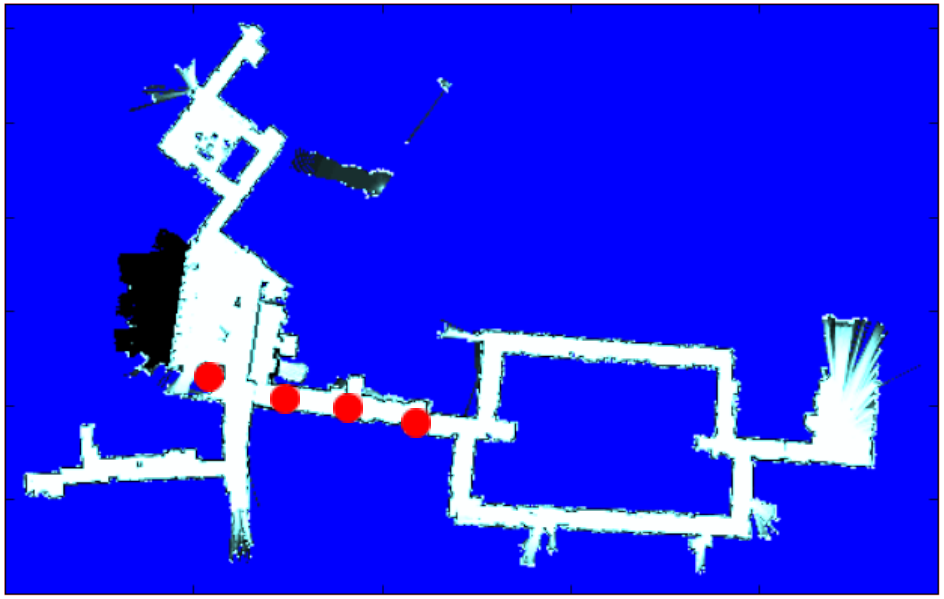
Sequential

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 white board. Turn left, turn right, and enter the second door on your right (sign says "Administrative Assistant").



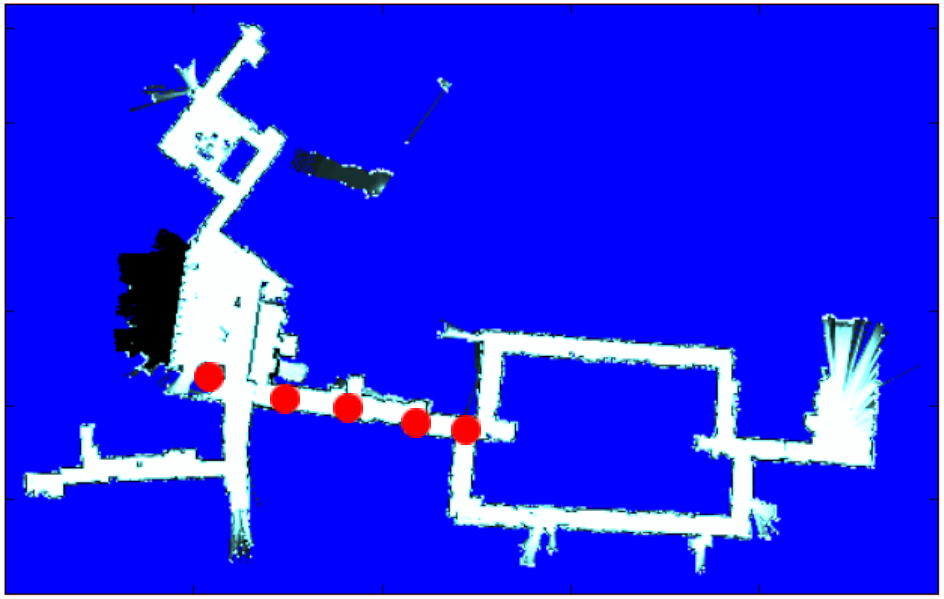
Sequential

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 white board. Turn left, turn right, and enter the second door on your right (sign says "Administrative Assistant").



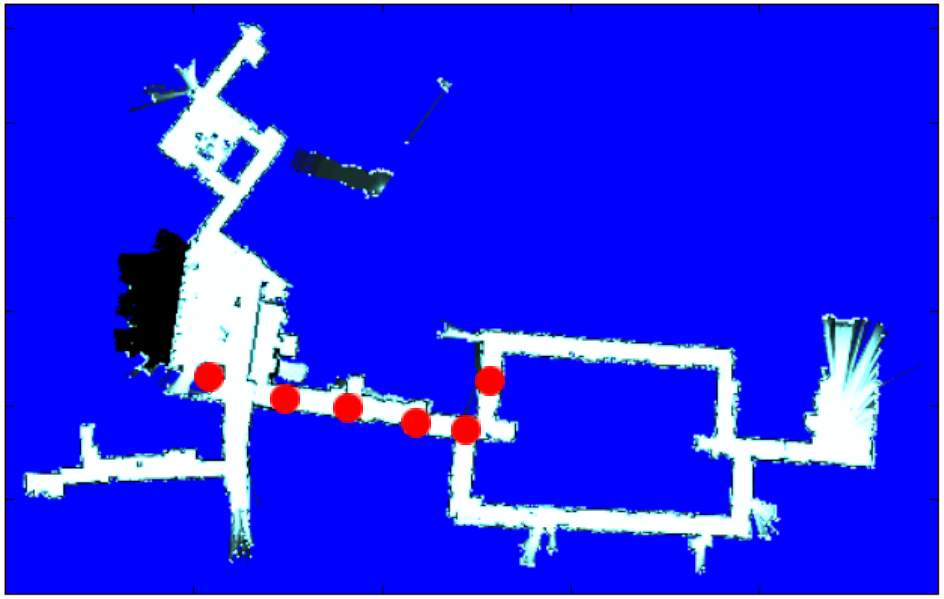
Sequential

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 white board. Turn left, turn right, and enter the second door on your right (sign says "Administrative Assistant").



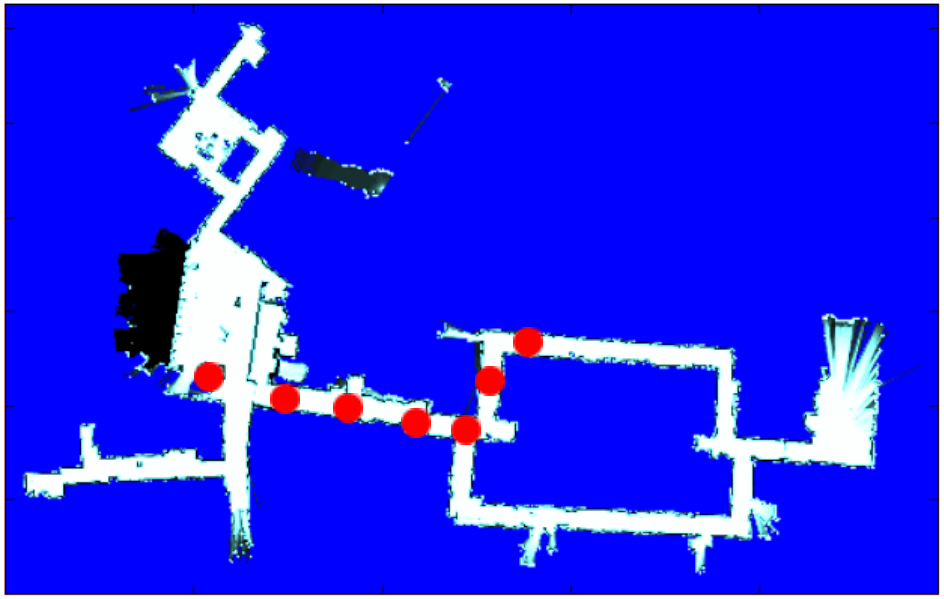
Sequential

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 white board. Turn left, turn right, and enter the second door on your right (sign says "Administrative Assistant").



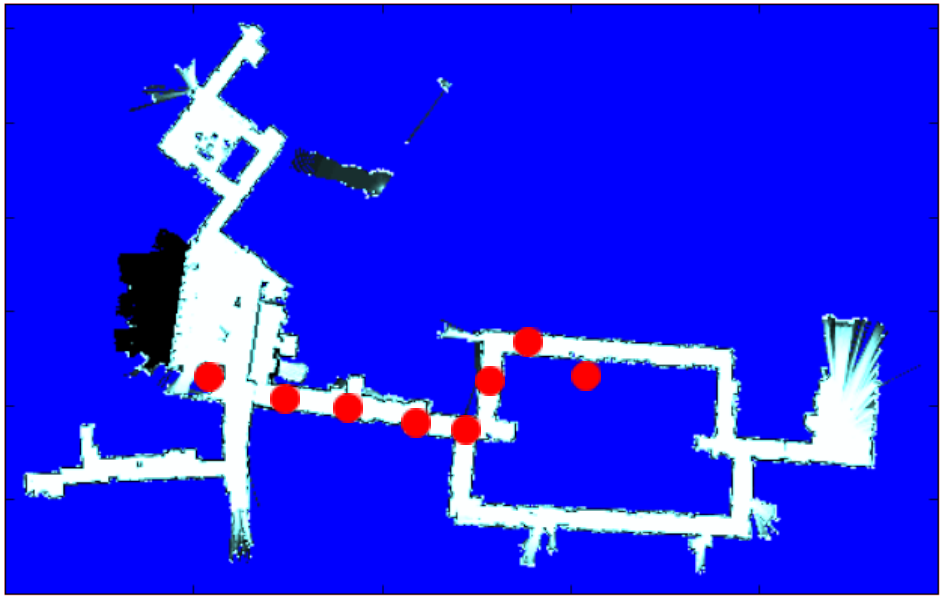
Sequential

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 white board. Turn left, turn right, and enter the second door on your right (sign says "Administrative Assistant").



Sequential

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 white board. Turn left, turn right, and enter the second door on your right (sign says "Administrative Assistant").



Structure of Natural Language Directions

- Sequential.
- Directive verbs.

Directive Verbs

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 white board. Turn left, turn right, and enter the second door on your right (sign says "Administrative Assistant").

Structure of Natural Language Directions

- Sequential.
- Directive verbs.
- Landmarks.

Landmarks

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 white board. Turn left, turn right, and enter the second door on your right (sign says "Administrative Assistant").

Structure of Natural Language Directions

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

Spatial Relations

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 white board. Turn left, turn right, and enter the second door on your right (sign says "Administrative Assistant").

Structure of Natural Language Directions

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

Spatial Description Clause

Go through the door.

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

Spatial Description Clause

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

V: Continue to walk, SR: straight

Spatial Description Clause

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

Spatial Description Clause

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

V: going, SR: until, L: you come to an intersection just past a white board

Spatial Description Clause

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

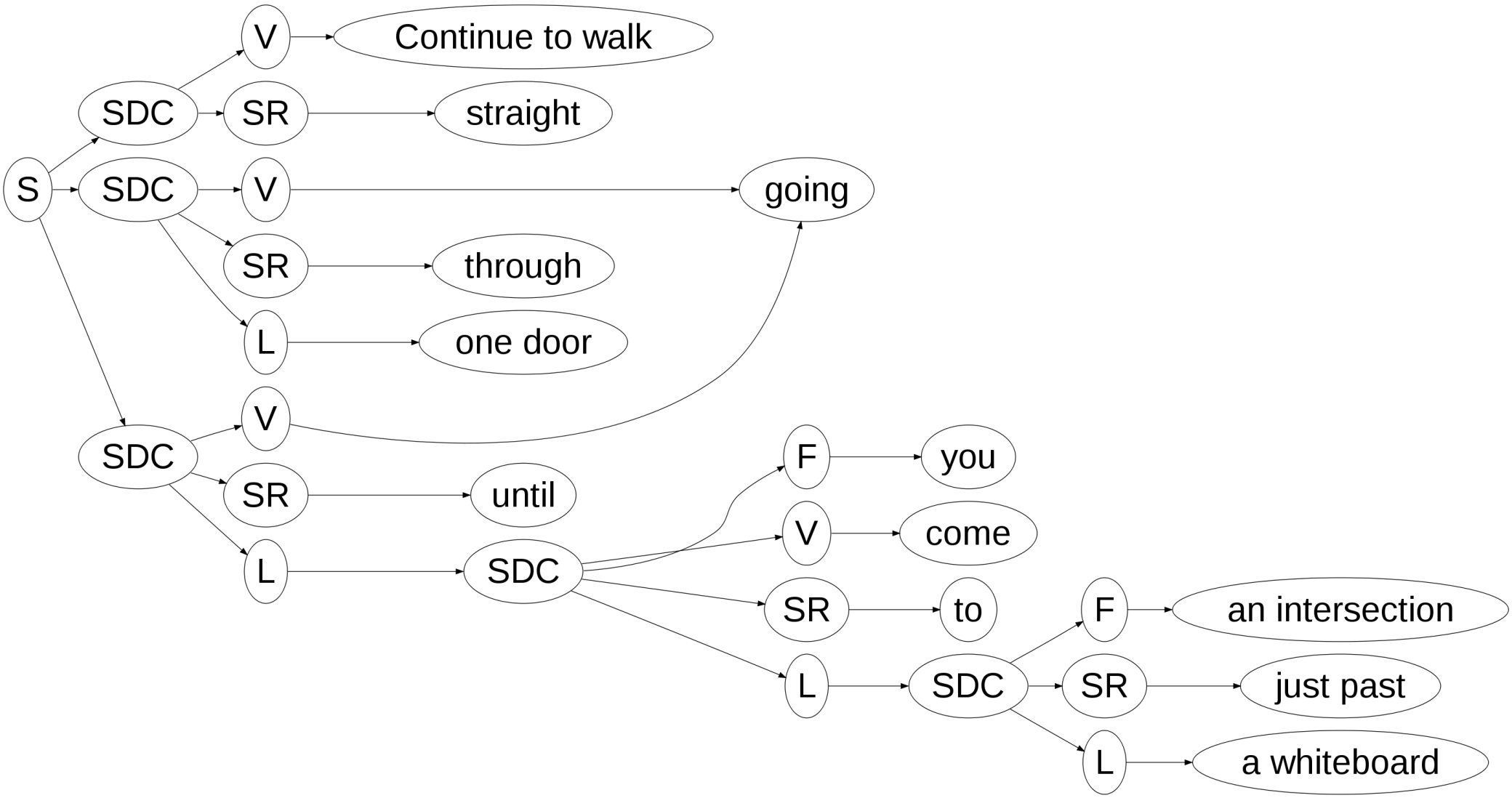
F: you, V: come, SR: to, L: an intersection just past a white board

Spatial Description Clause

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

F: an intersection, SR: just past, L: a white board

Spatial Description Clause



Model

$$argmax_{path}\, p(path, directions, observations)$$

Model

$$p(path, directions, observations) =$$

$$p(directions | path, observations)\, p(path, observations)$$

(definition of conditional probability)

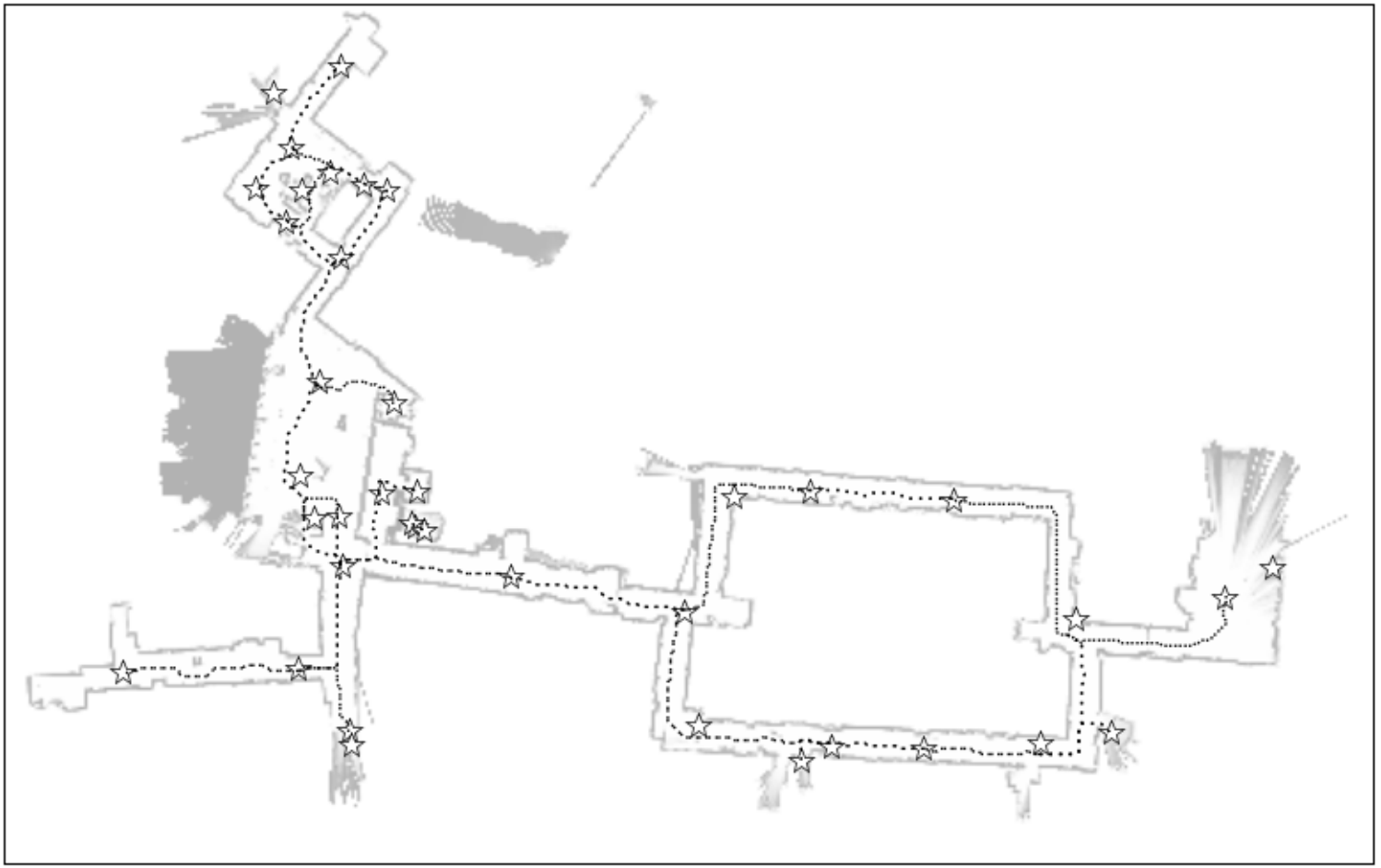
$$p(directions | path, observations)\, p(path)\, p(observations)$$

(path and observations are independent)

$$p(sdc_1 \ldots sdc_M | vp_1 \ldots vp_{M+1}, o_1 \ldots o_K)\, p(vp_1 \ldots vp_{M+1})\, p(o_1 \ldots o_K)$$

(topological map of viewpoints, and a sequence of SDCs)

Topological Map



Model

$$p(path, directions, observations) =$$

$$p(directions | path, observations)\, p(path, observations)$$

(definition of conditional probability)

$$p(directions | path, observations)\, p(path)\, p(observations)$$

(path and observations are independent)

$$p(sdc_1 \ldots sdc_M | vp_1 \ldots vp_{M+1}, o_1 \ldots o_K)\, p(vp_1 \ldots vp_{M+1})\, p(o_1 \ldots o_K)$$

(topological map of viewpoints, and a sequence of SDCs)

$$\prod_{i=1}^{M} p(sdc_i | vp_i, vp_{i+1}, o_1 \ldots o_K) \prod_{i=1}^{M} p(vp_{i+1} | vp_1 \ldots vp_i)\, p(o_1 \ldots o_K)$$

(Paths are Markov)

Model

$$p(path, directions, observations) =$$

$$p(directions | path, observations)\, p(path, observations)$$

(definition of conditional probability)

$$p(directions | path, observations)\, p(path)\, p(observations)$$

(path and observations are independent)

$$p(sdc_1 \ldots sdc_M | vp_1 \ldots vp_{M+1}, o_1 \ldots o_K)\, p(vp_1 \ldots vp_{M+1})\, p(o_1 \ldots o_K)$$

(topological map of viewpoints, and a sequence of SDCs)

$$\prod_{i=1}^{M} p(sdc_i | vp_i, vp_{i+1}, o_1 \ldots o_K) \prod_{i=1}^{M} p(vp_{i+1} | vp_1 \ldots vp_i)\, p(o_1 \ldots o_K)$$

(Paths are Markov)

Probability of Directions Given a Possible Movement

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

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

(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

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

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

(The structure of a spatial description clause.)

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

(Factor according to certain independence assumptions.)

Probability of Directions Given a Possible Movement

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

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

(The structure of a spatial description clause.)

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

(Factor according to certain independence assumptions.)

Observations

Contextual relationships

Tags include: desktop, monitor, computer, keyboard, and mouse

Probability of Directions Given a Possible Movement

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

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

(The structure of a spatial description clause.)

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

(Factor according to certain independence assumptions.)

Probability of Directions Given a Possible Movement

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

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

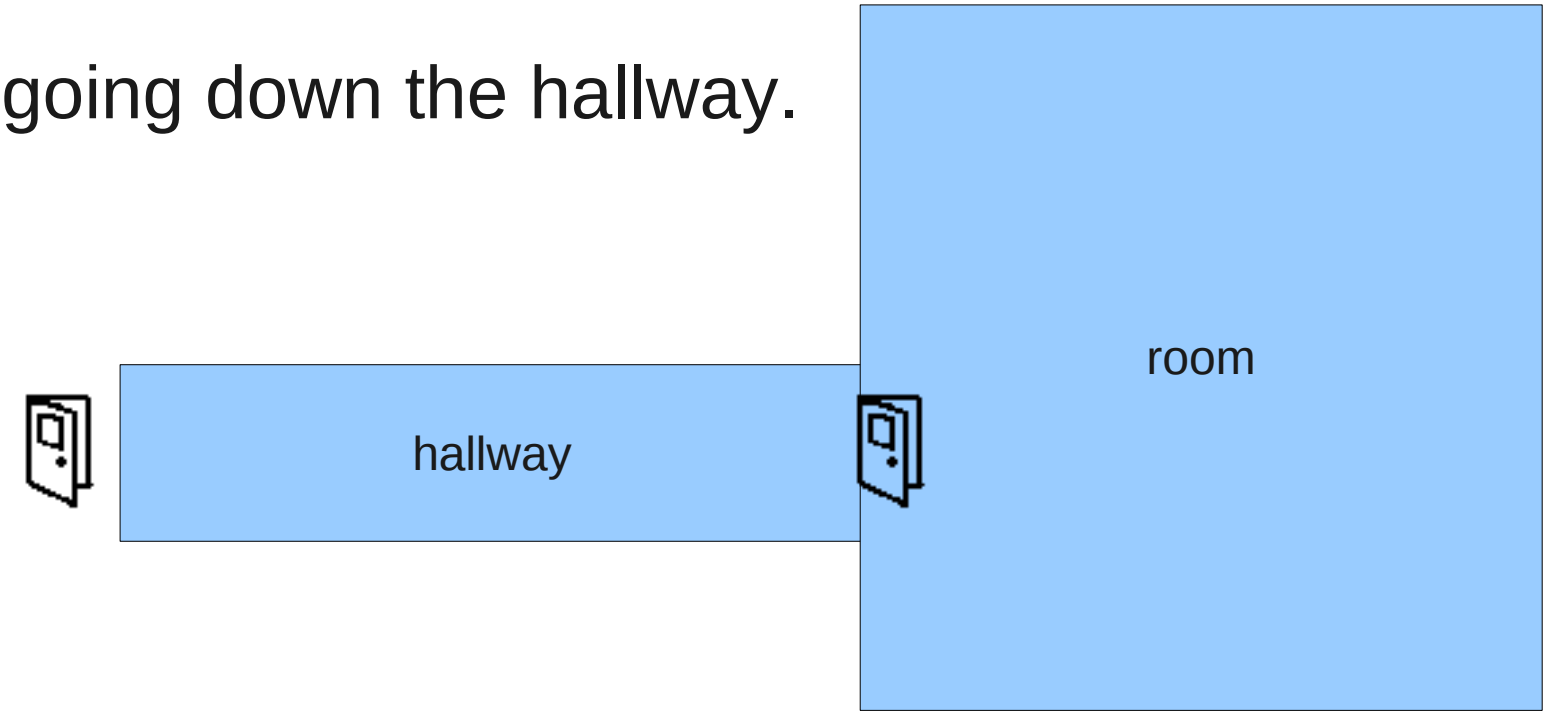
(The structure of a spatial description clause.)

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

(Factor according to certain independence assumptions.)

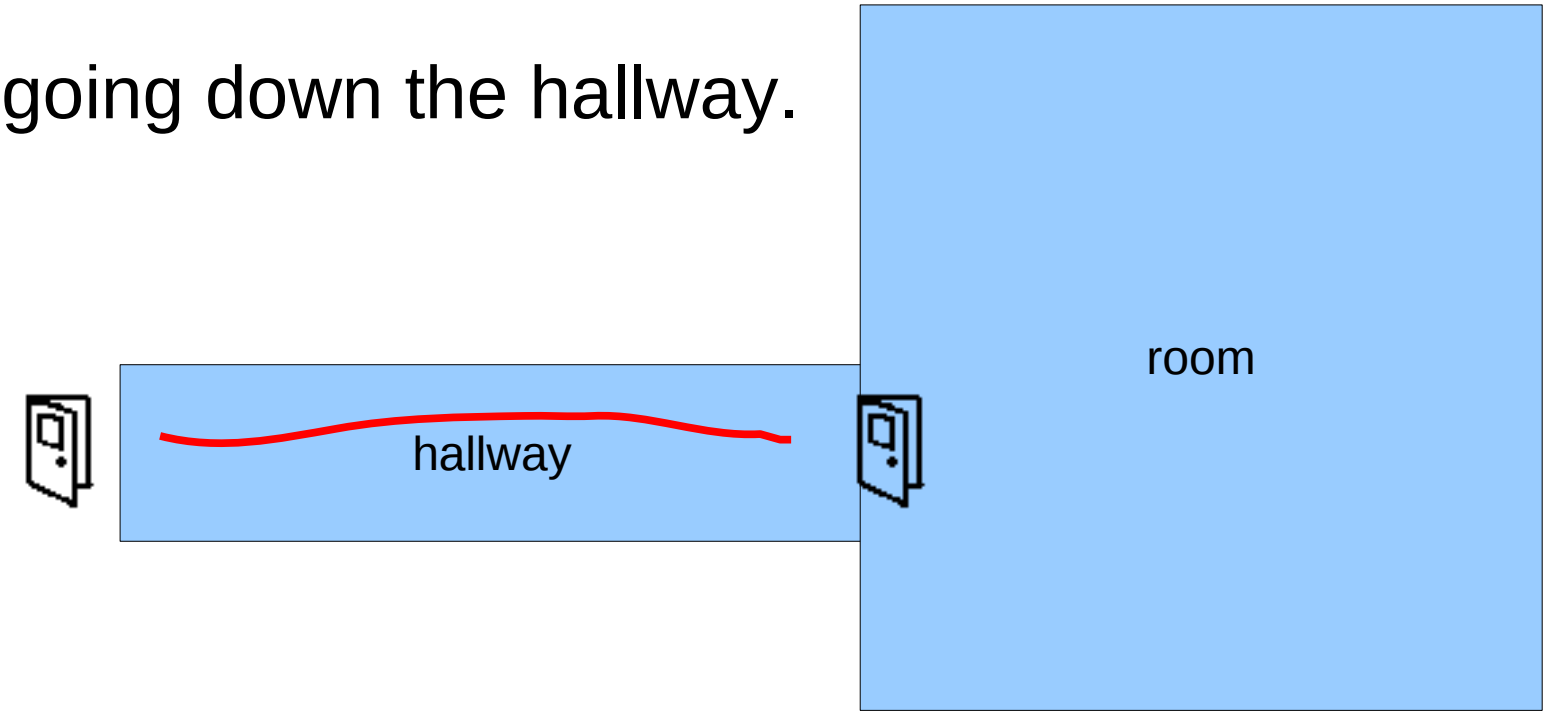
Learning Distributions for Spatial Relations

Draw a path going down the hallway.

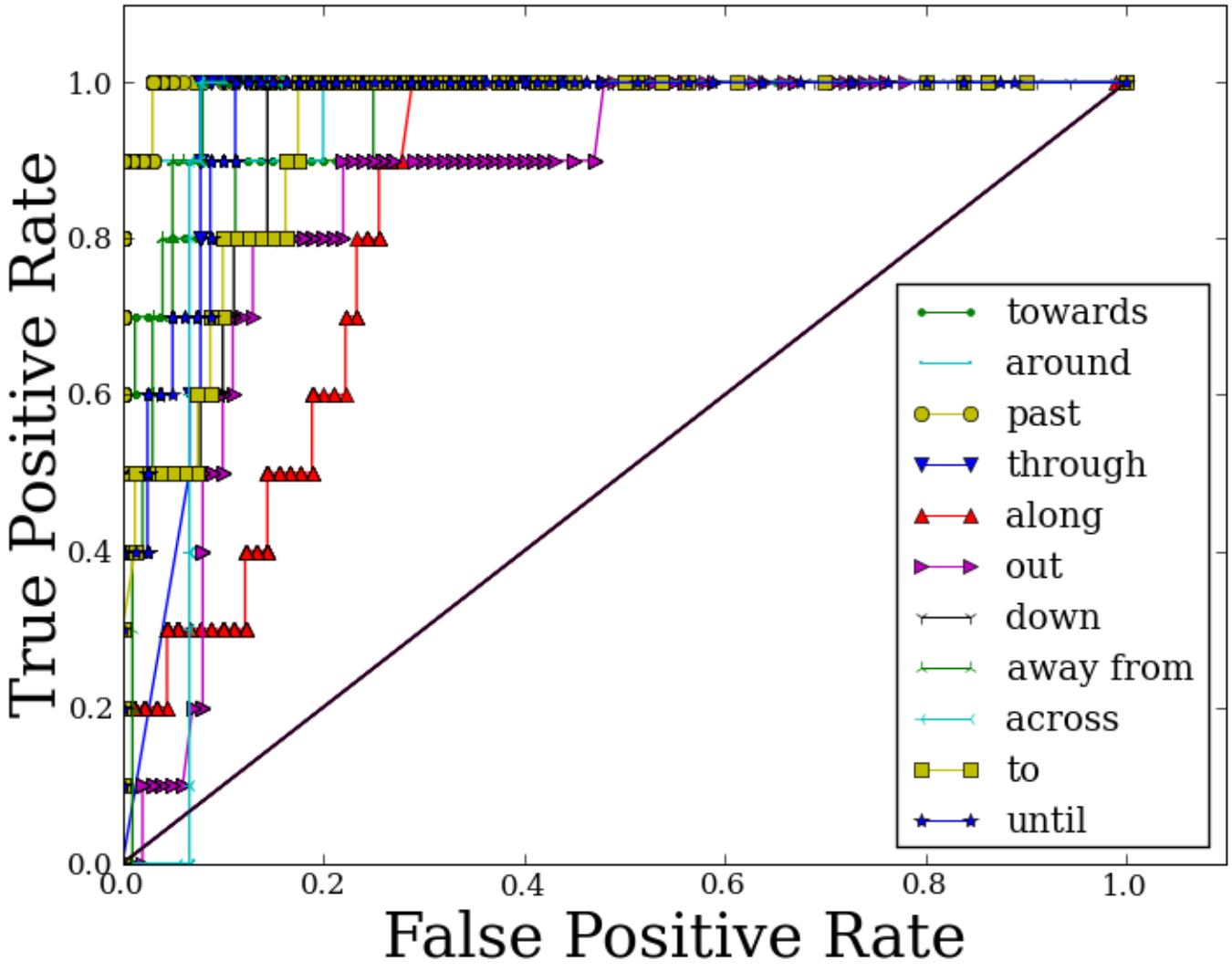


Learning Distributions for Spatial Relations

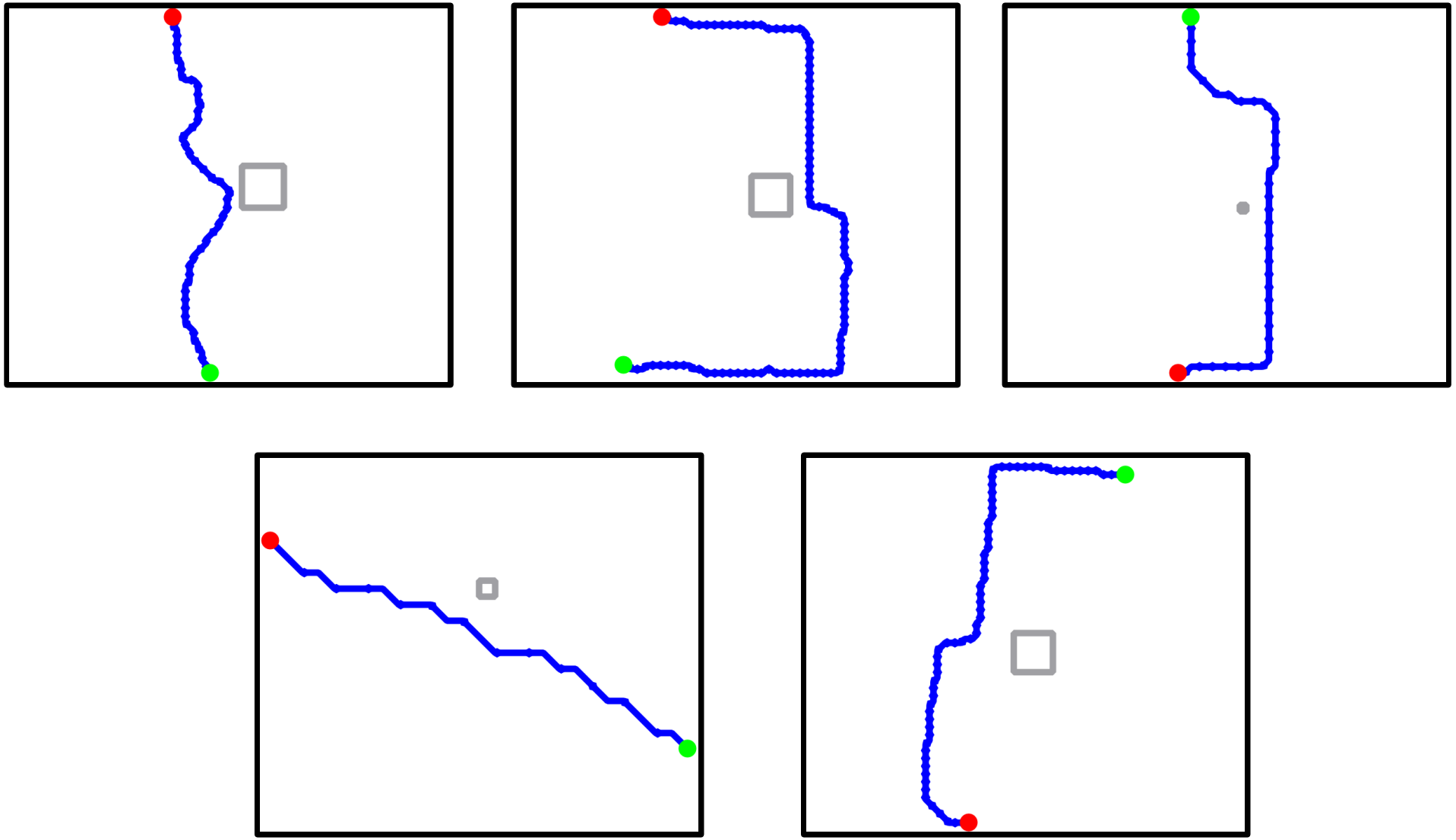
Draw a path going down the hallway.



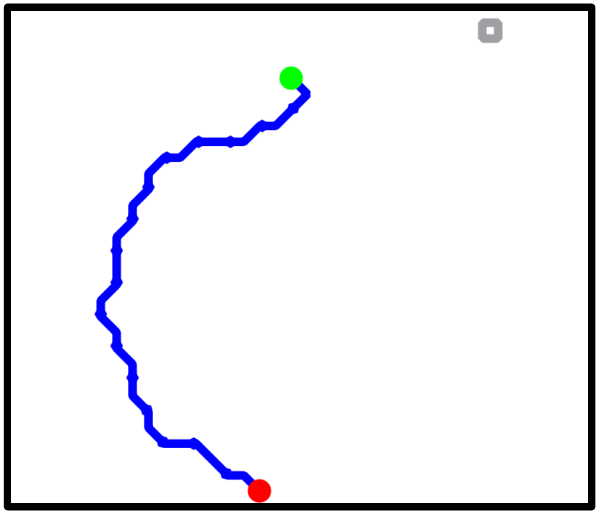
Performance of Classifiers

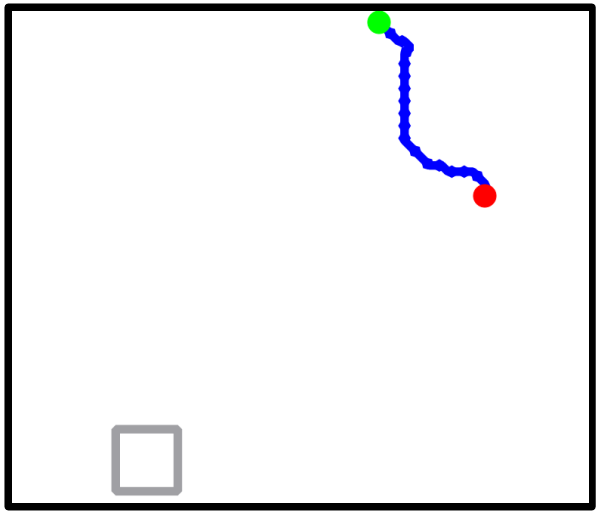


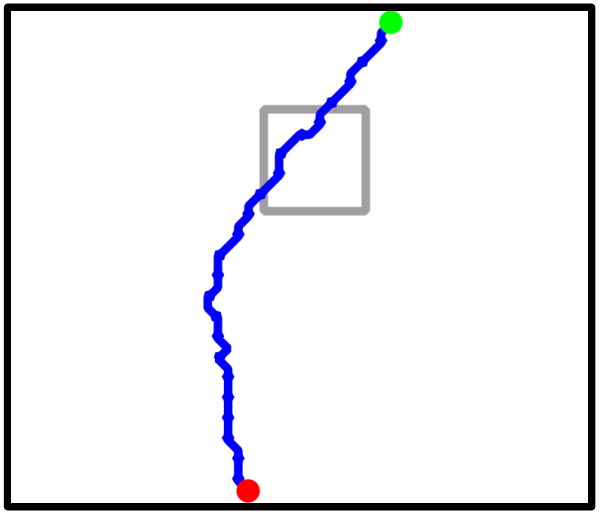
High Scoring Examples of “past”

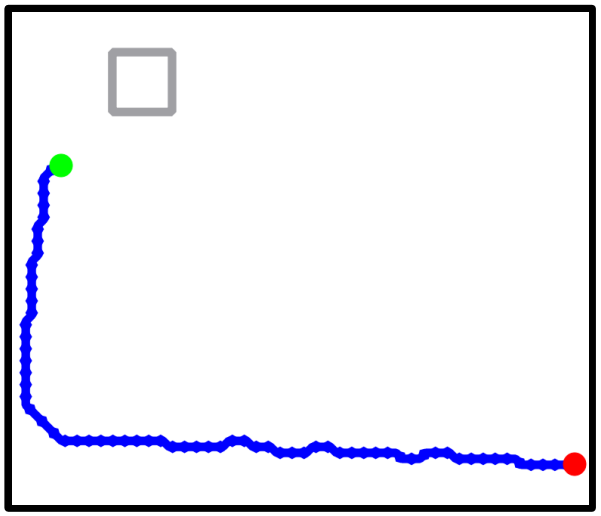


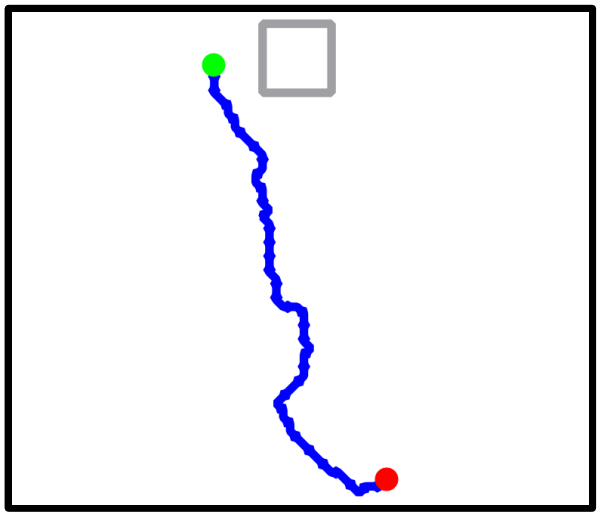
Low Scoring Examples of “past”





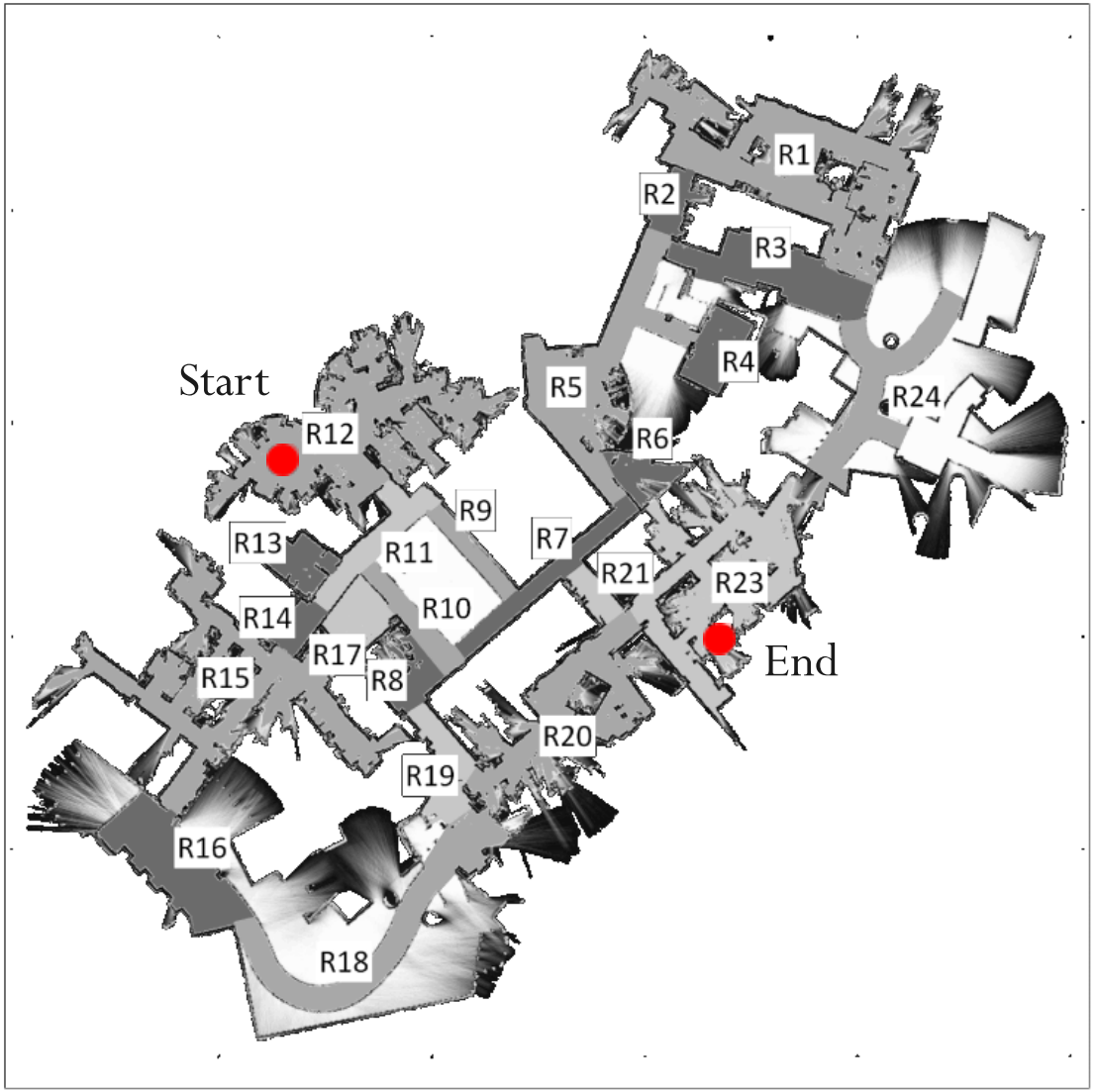






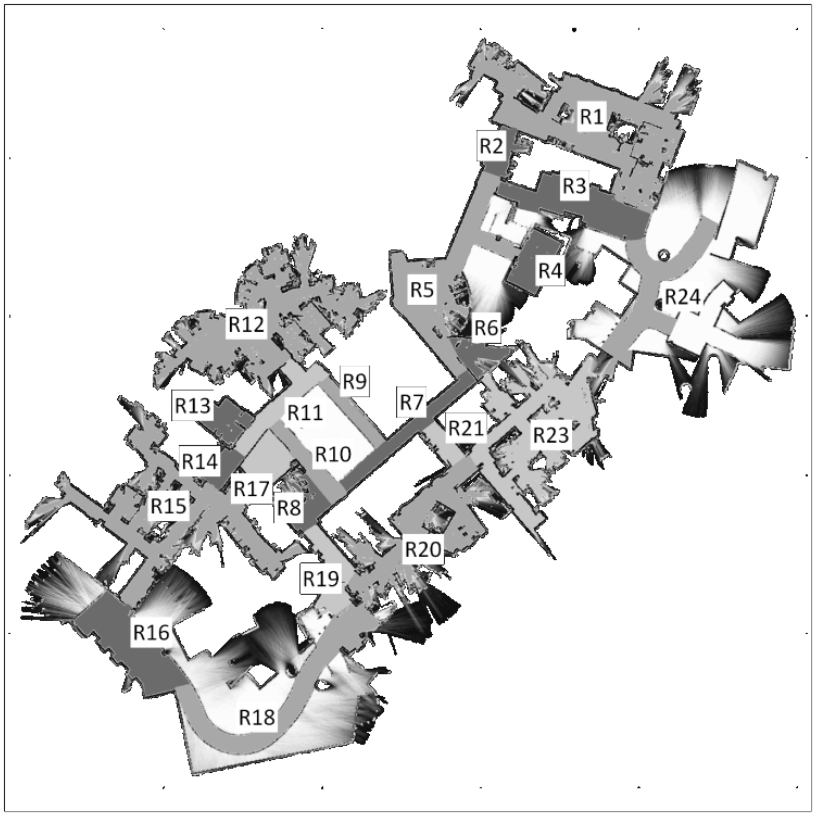
Corpus of Natural Language Directions

Instructions:
Write, in whatever way is natural to you, directions from the start label to the end label, with the goal of communicating this route to another person who does not have a map.



Corpus of Natural Language Directions

- Study 1 (Stata 3)
  - 10 people, 16 routes
- Study 2 (Stata 8)
  - 15 people, 10 routes
- Study 3 (Stata 1)
  - 15 people, 10 routes
- Study 4 (outdoors)
  - ongoing



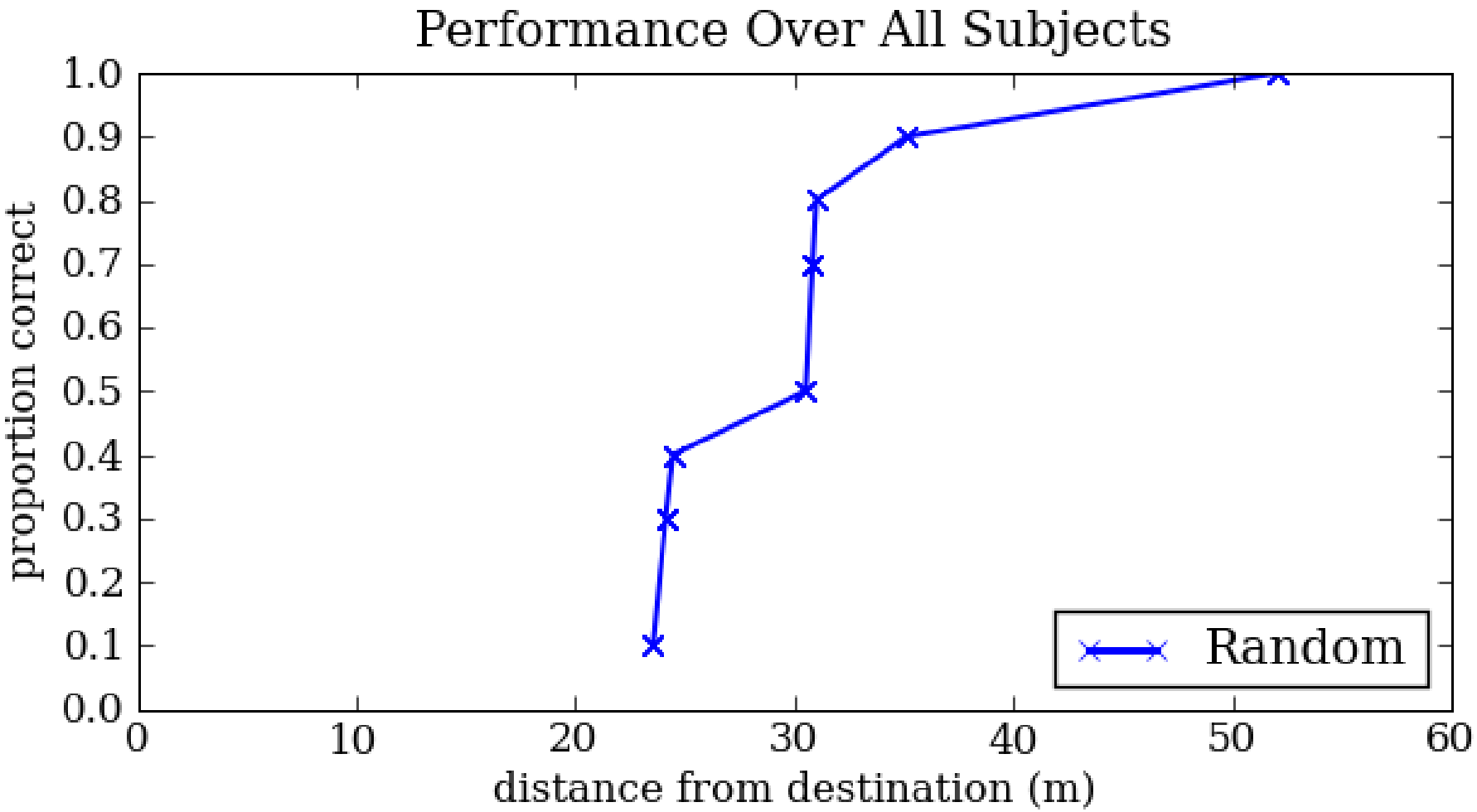
Human Performance

	Stata 3	Stata 8	Stata 1
All directions	63%	85%	86%
Best Direction Giver	87%	100%	100%
Worst Direction Giver	43%	30%	20%

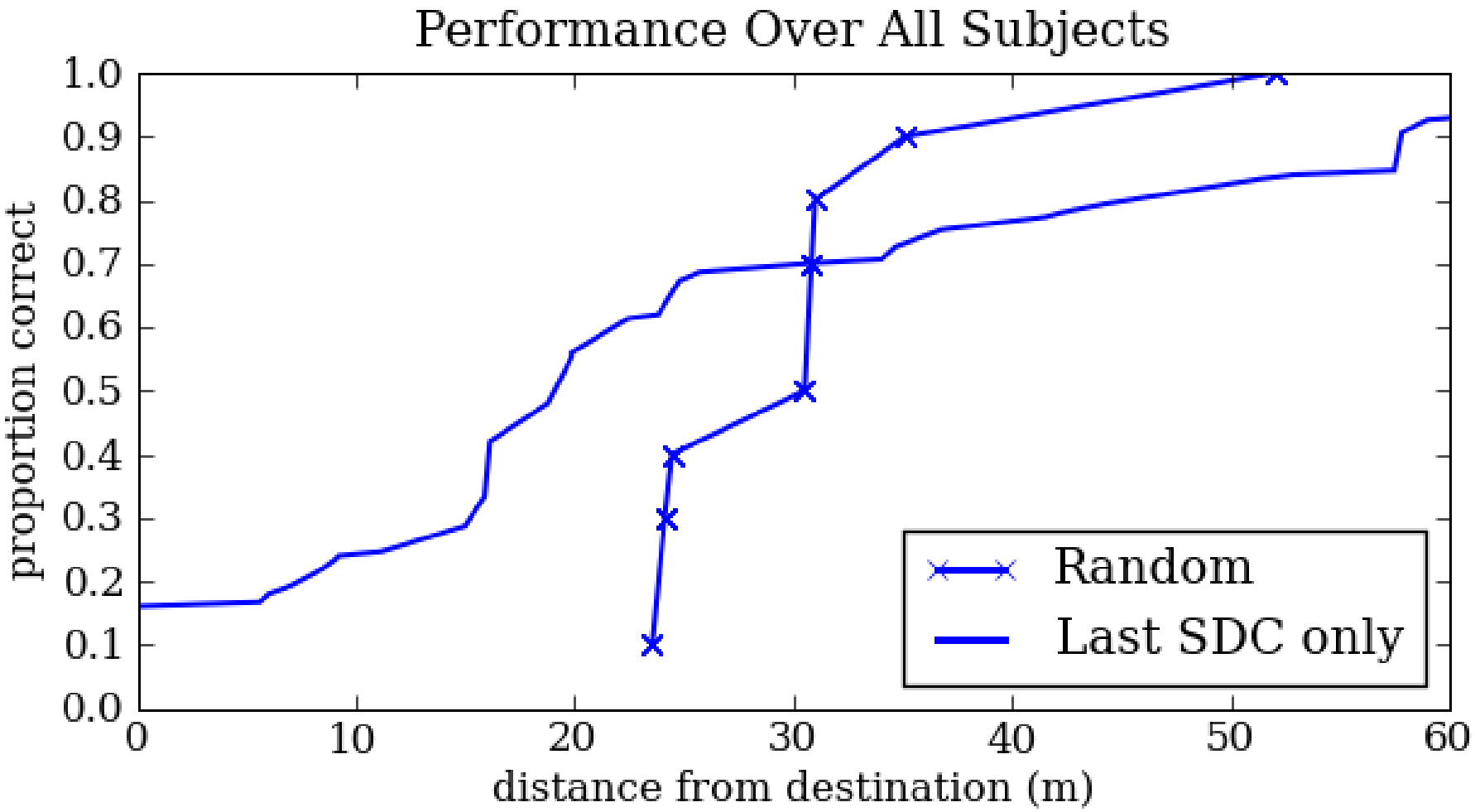
Inference

- Global.
  - Viterbi-style algorithm.
  - Last-SDC only.
  - Landmarks only.
- Local.
- Random.

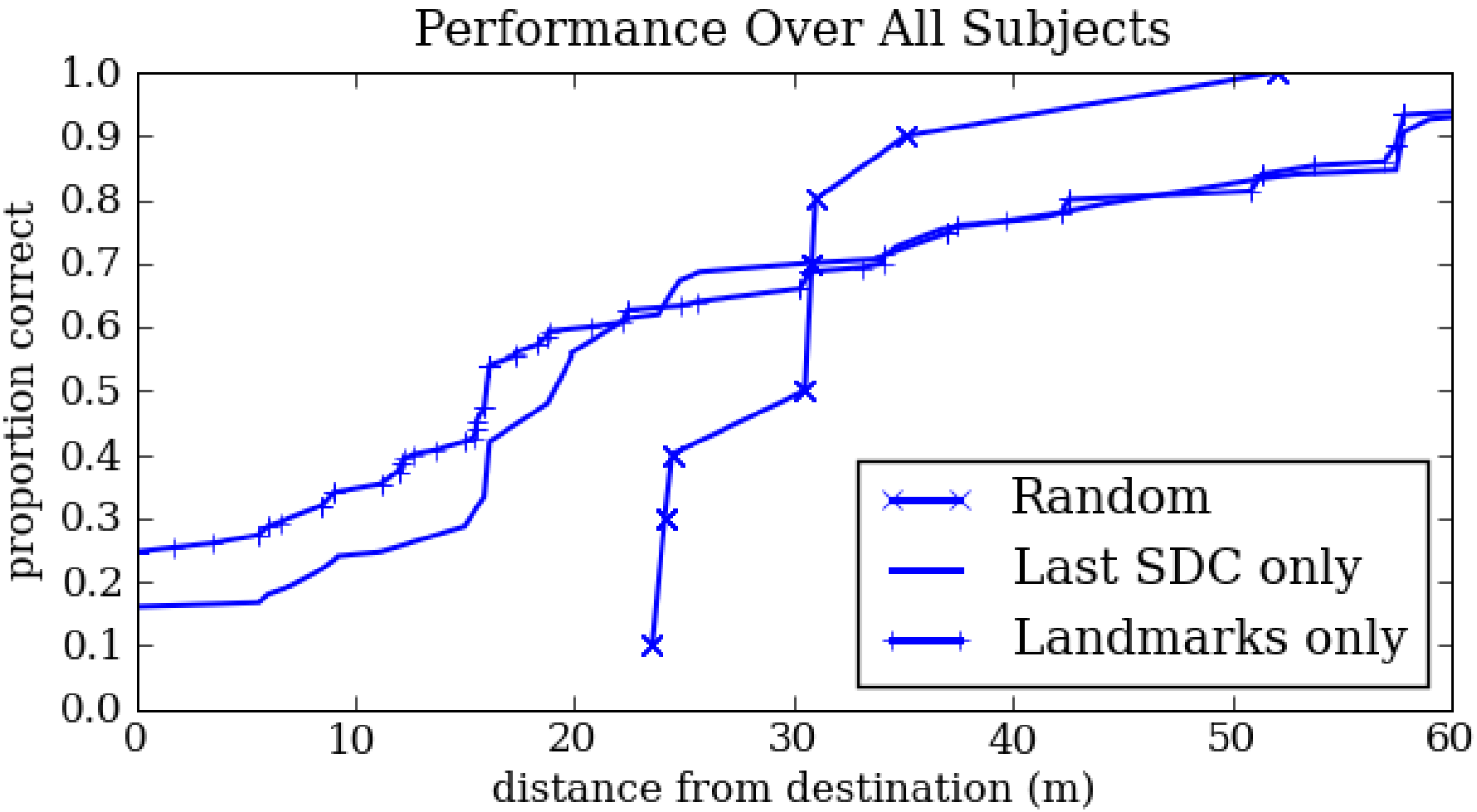
Comparison to Baselines



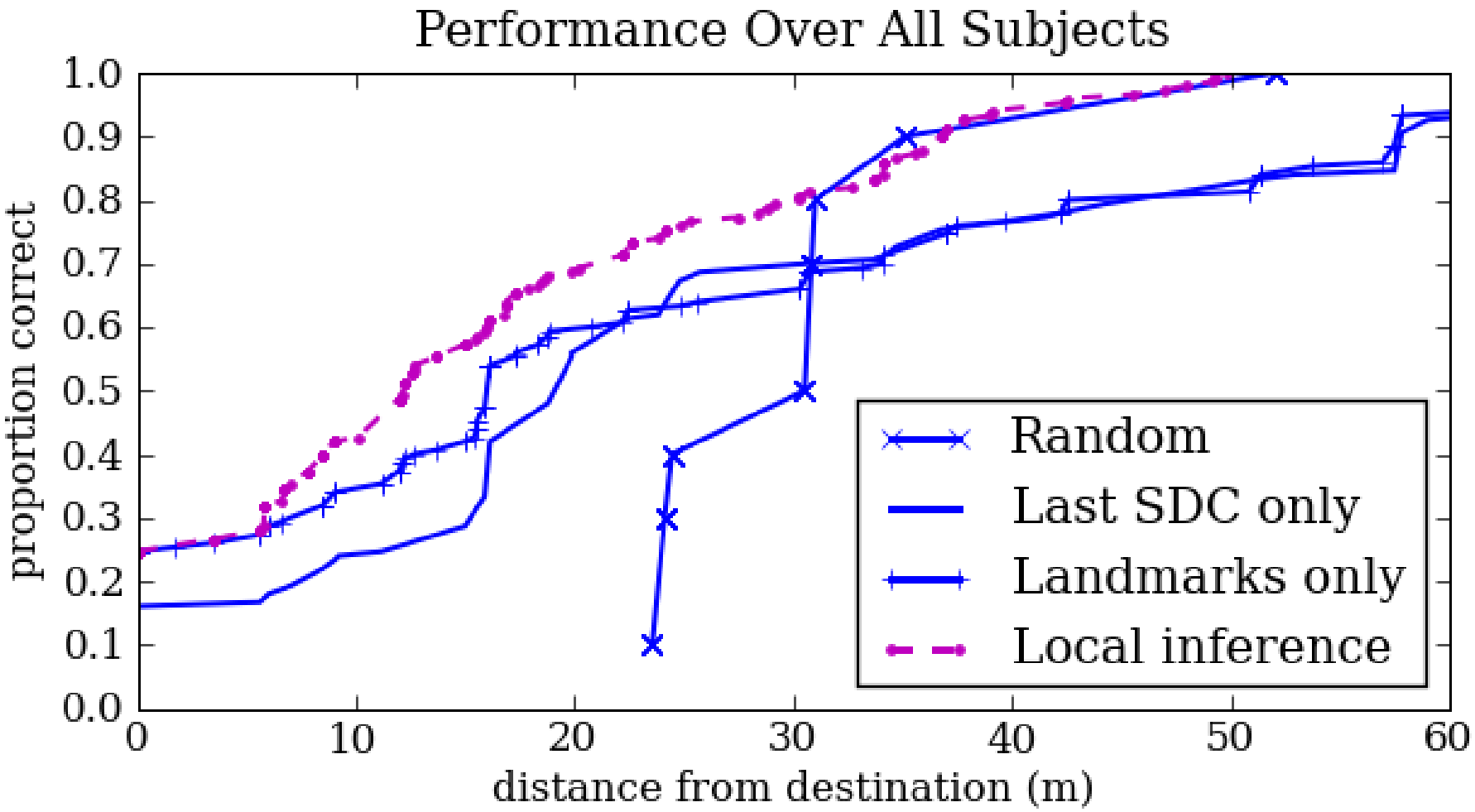
Comparison to Baselines



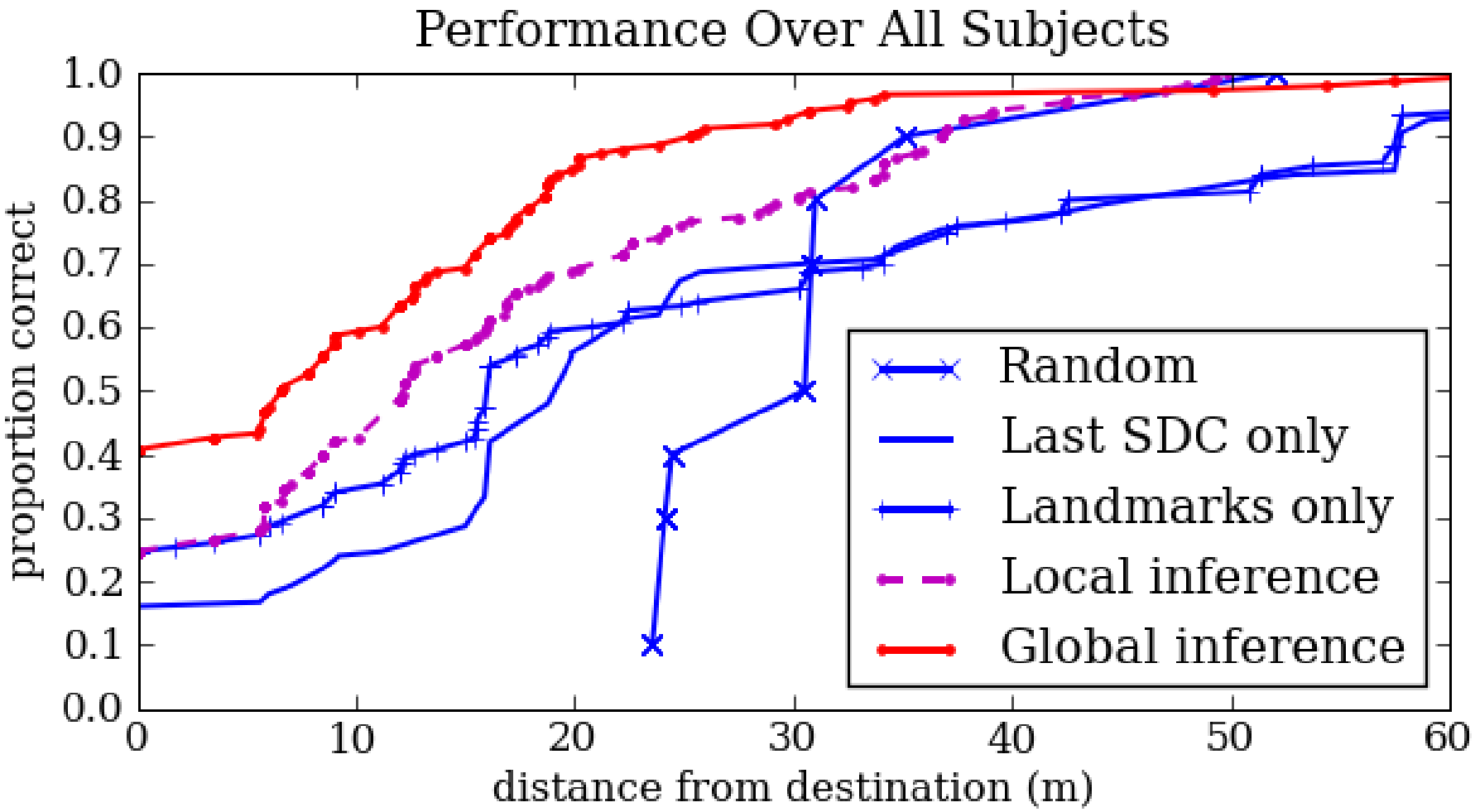
Comparison to Baselines



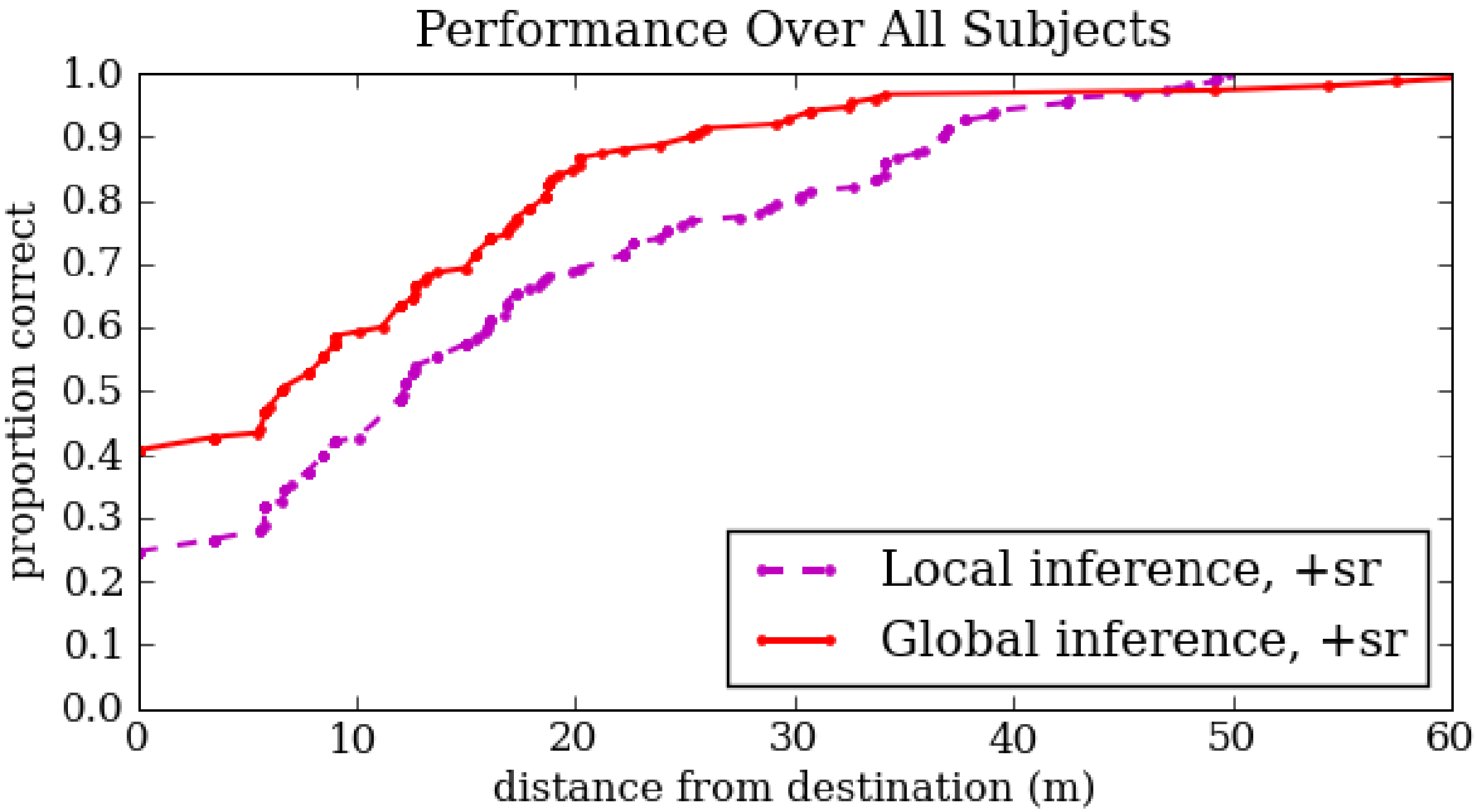
Comparison to Baselines



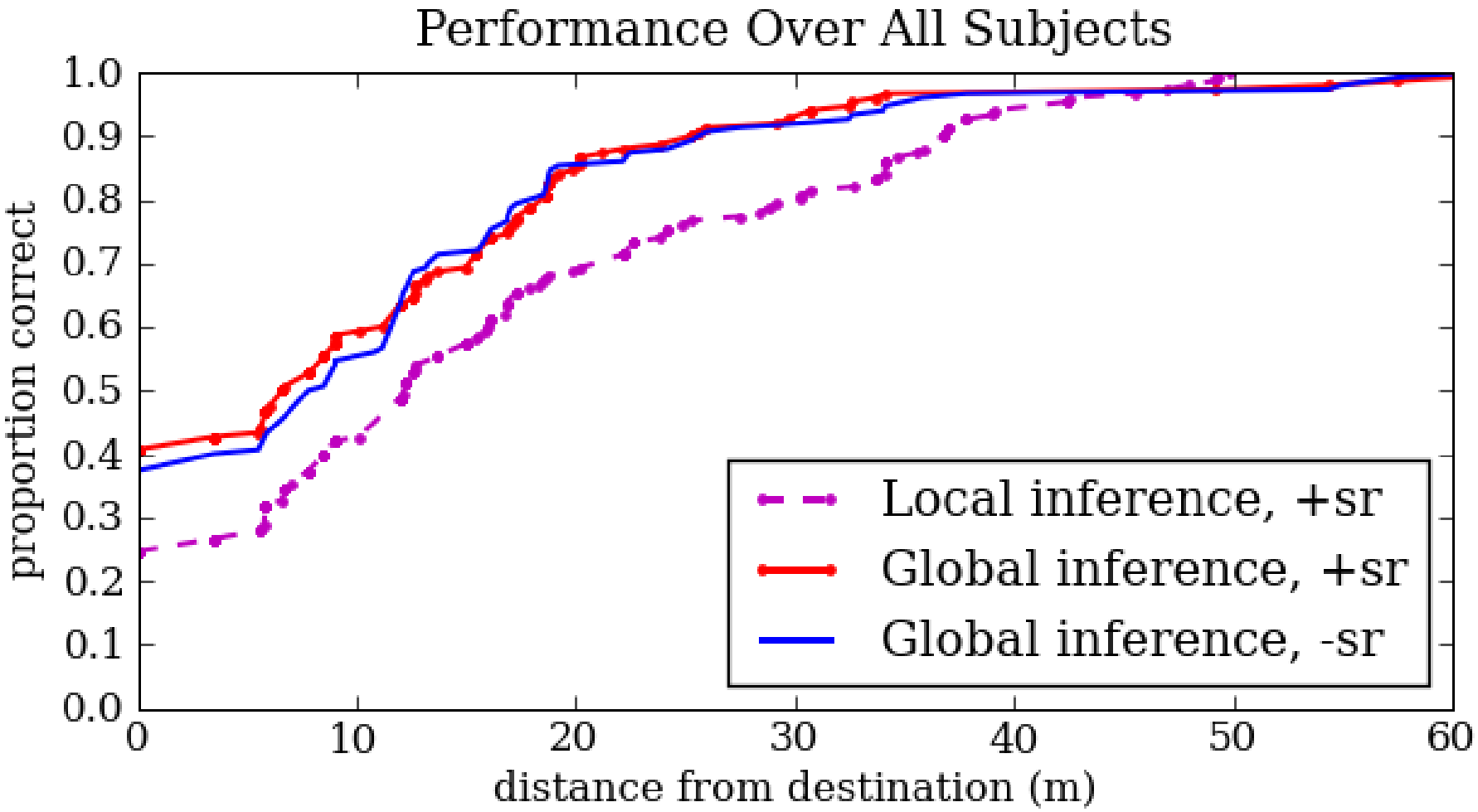
Comparison to Baselines



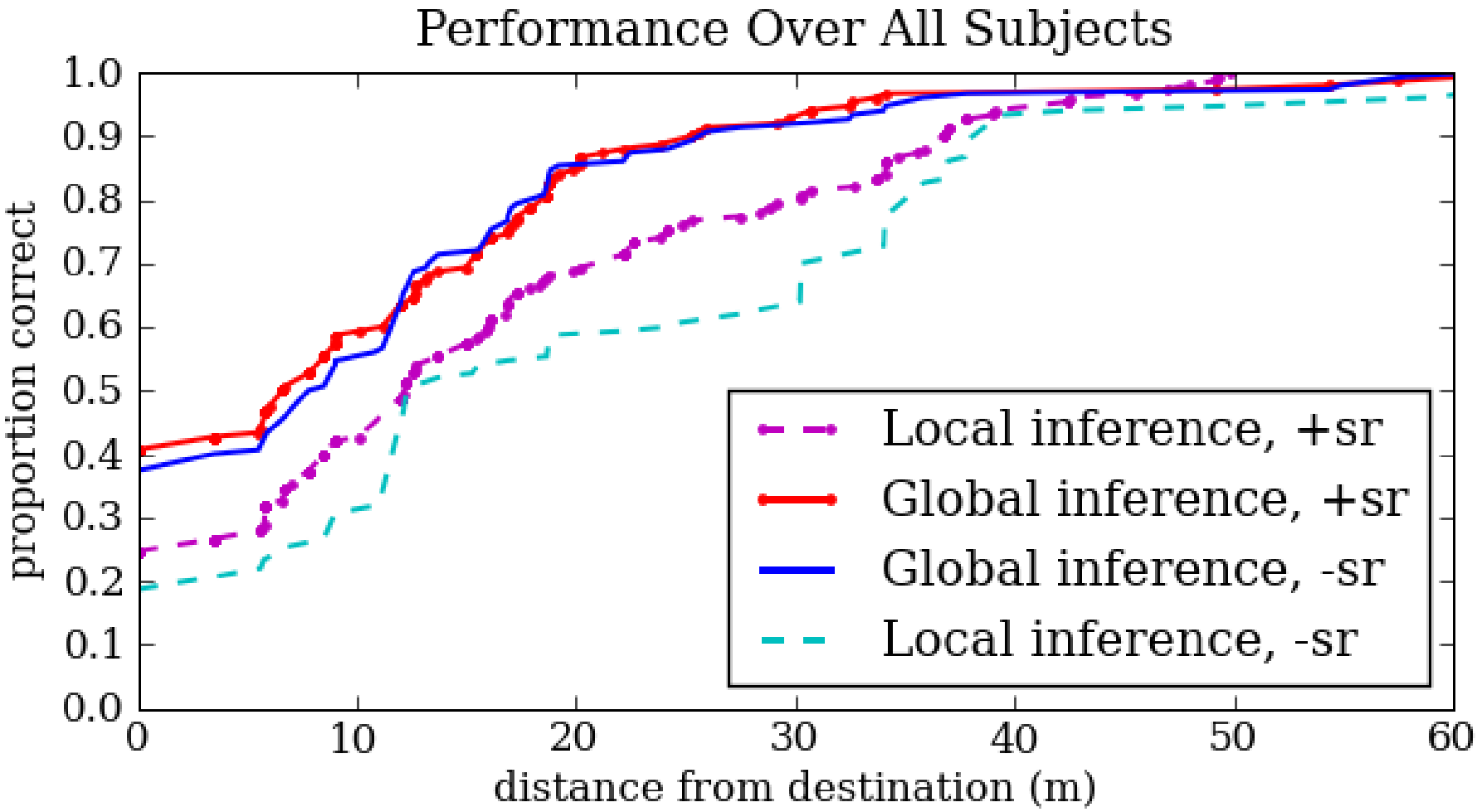
With and Without Spatial Relations



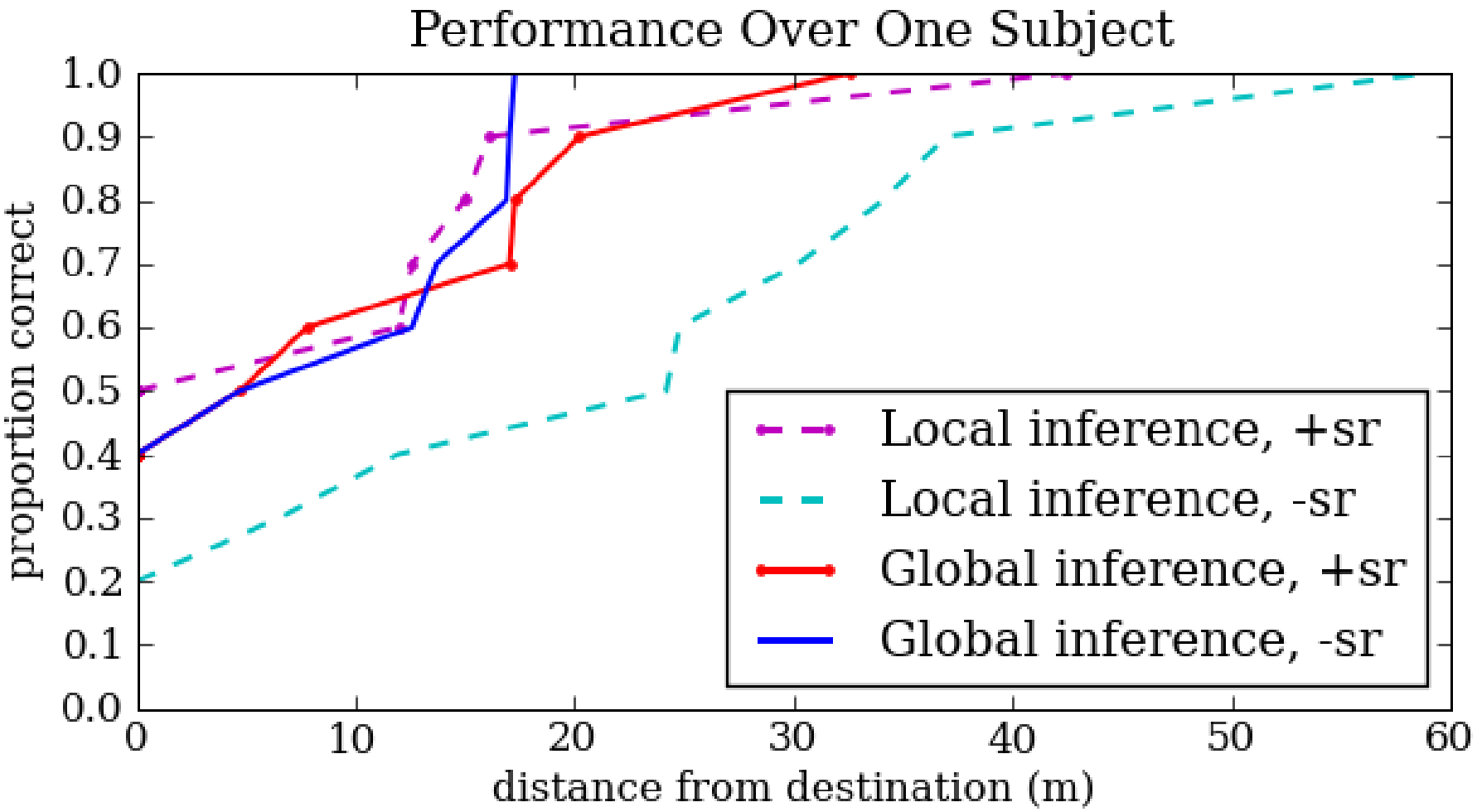
With and Without Spatial Relations



With and Without Spatial Relations



With and Without Spatial Relations

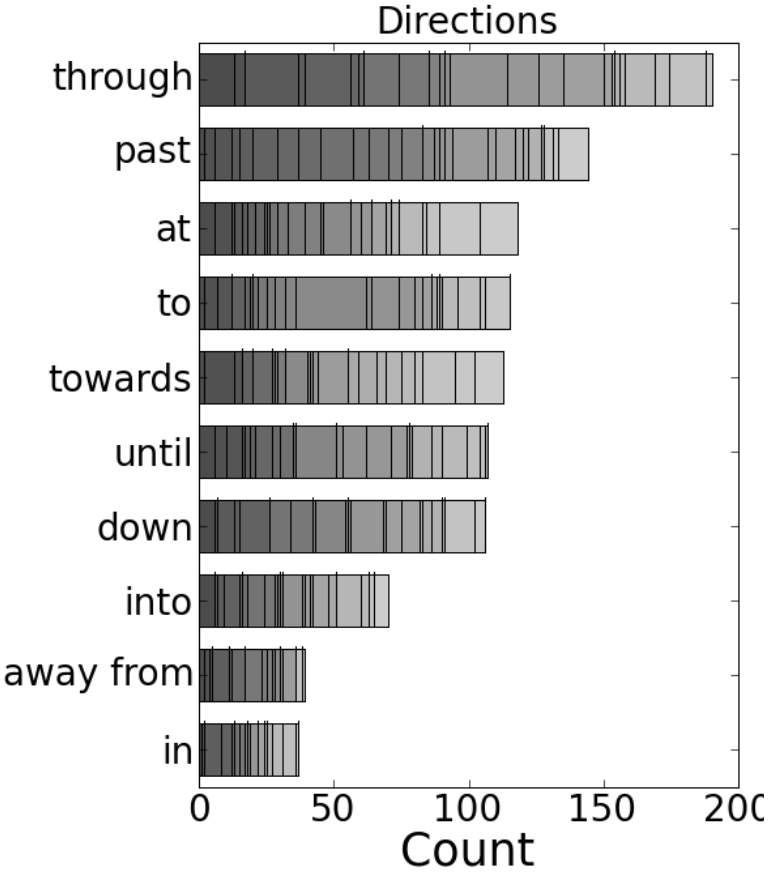


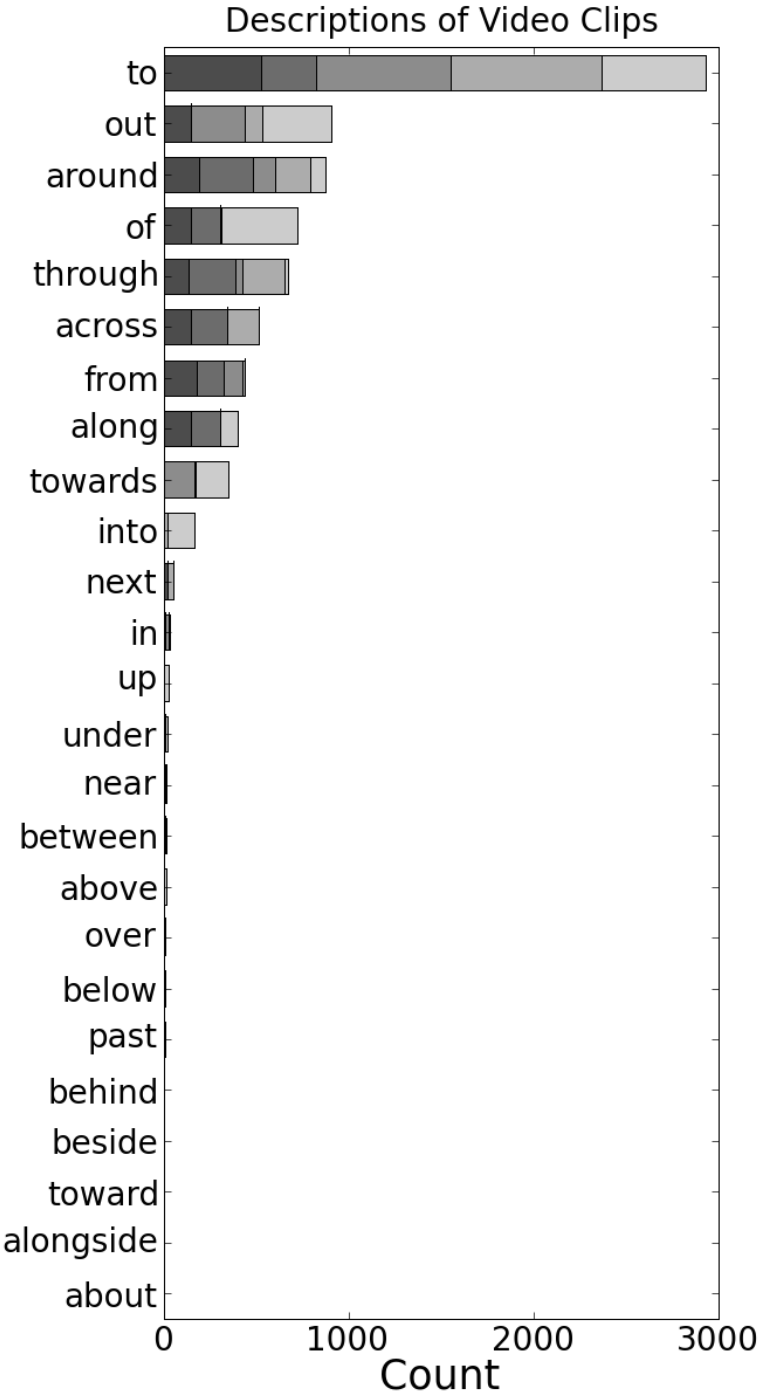
Improving Direction Understanding Results

- Polygon landmarks.
- Add more spatial relations.
- Use more complex environments.
- Changing the model.
  - Alignment.
  - Conditional Random Fields.
- Exploration.

Next Steps

- Video retrieval.
  - Collect a corpus of paragraph-length descriptions of a person's activity.
- Direction generation.
- Compare spatial prepositions across domains.
- Create a lexicon of spatial prepositions.

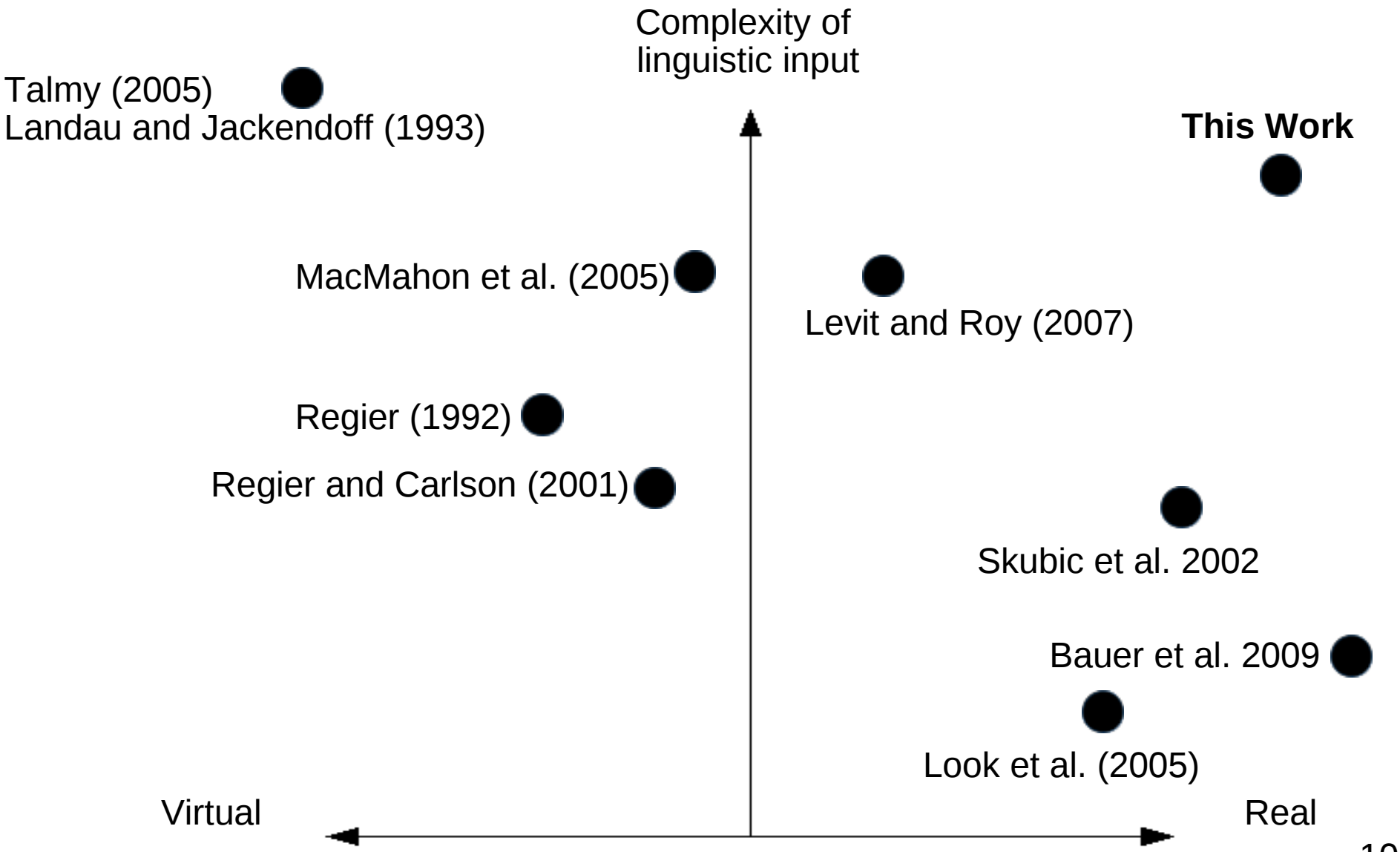

Directions
through
past
at
to
towards
until
down
into
away from
in
0
50
100
150
200
Count


Descriptions of Video Clips
to
out
around
of
through
across
from
along
towards
into
next
in
up
under
near
between
above
over
below
past
behind
beside
toward
alongside
about
0
1000
2000
3000
Count

Schedule

- December 15, 2010 – JAIR, direction understanding
- January 22, 2010 – SIGIR, video retrieval
- March 1, 2010 - TiCS on generation.
- April 1, 2010 – JAIR, lexicon of spatial prepositions.
- April 14, 2010 - thesis outline to Deb.
- May 14, 2010 – thesis draft to Deb.
- May 31, 2010 – thesis draft to committe.
- July, 2010 – thesis defense

Related Work



Related Work – Video Retrieval

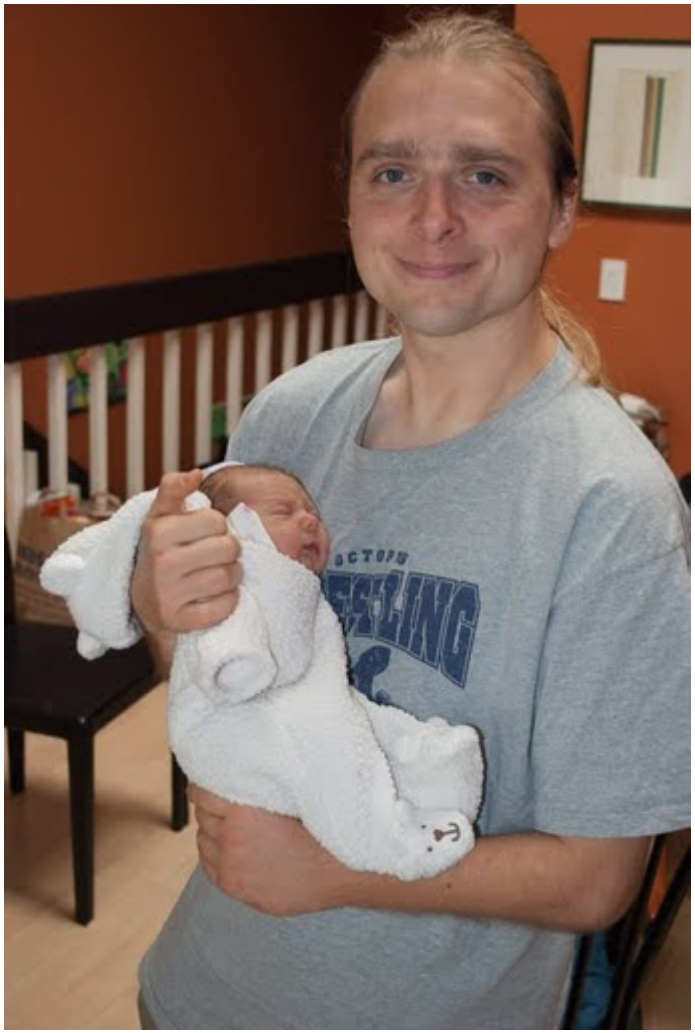
- Katz et al. (2004) - “Show me all the cars leaving the garage.”
- Ivanov and Wren (2006) – Query by example trajectory
- Fleischman et al. (2006) - “Show me people making coffee.”
- Naphade et al. (2006) - Large-Scale Concept Ontology for Multimedia

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 higher level structures.

Acknowledgements

Tom Kollar
Deb, Yuri, Boris, Cynthia,
Cogmac
ML support staff
Gremio, Dave, Carie, Kai-yuh, Lin
Ania & Seth & Moby
Grandma Patches and Grandma Nicholas
Mom, Dad, Staci & Seth, Shannon & Jon, Scott



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 higher level structures.

Collaboration

- My thesis
  - Library of features for spatial prepositions.
  - Analyze which features perform the best.
  - Comparison of meanings of spatial relations in different domains.
- Tom's thesis
  - Grounding landmarks with flickr.
  - Planning to find objects.
- Both
  - Spatial description clauses.
  - MRF model of spatial language.

Directions With Only Landmarks

the windows, the door
the elevators. one door
you an intersection a whiteboard.
the second door (sign
"Administrative Assistant").

Why is this a Media Lab thesis?

- Towards a new (or old) form of human-machine interaction.
- Multi-modal, multi-domain interface.
- Defining a new problem, with a new corpus.

Extracting Spatial Description Clauses

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

Extracting Spatial Description Clauses

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

Extracting Spatial Description Clauses

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

V: Continue to walk straight

Extracting Spatial Description Clauses

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

Extracting Spatial Description Clauses

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

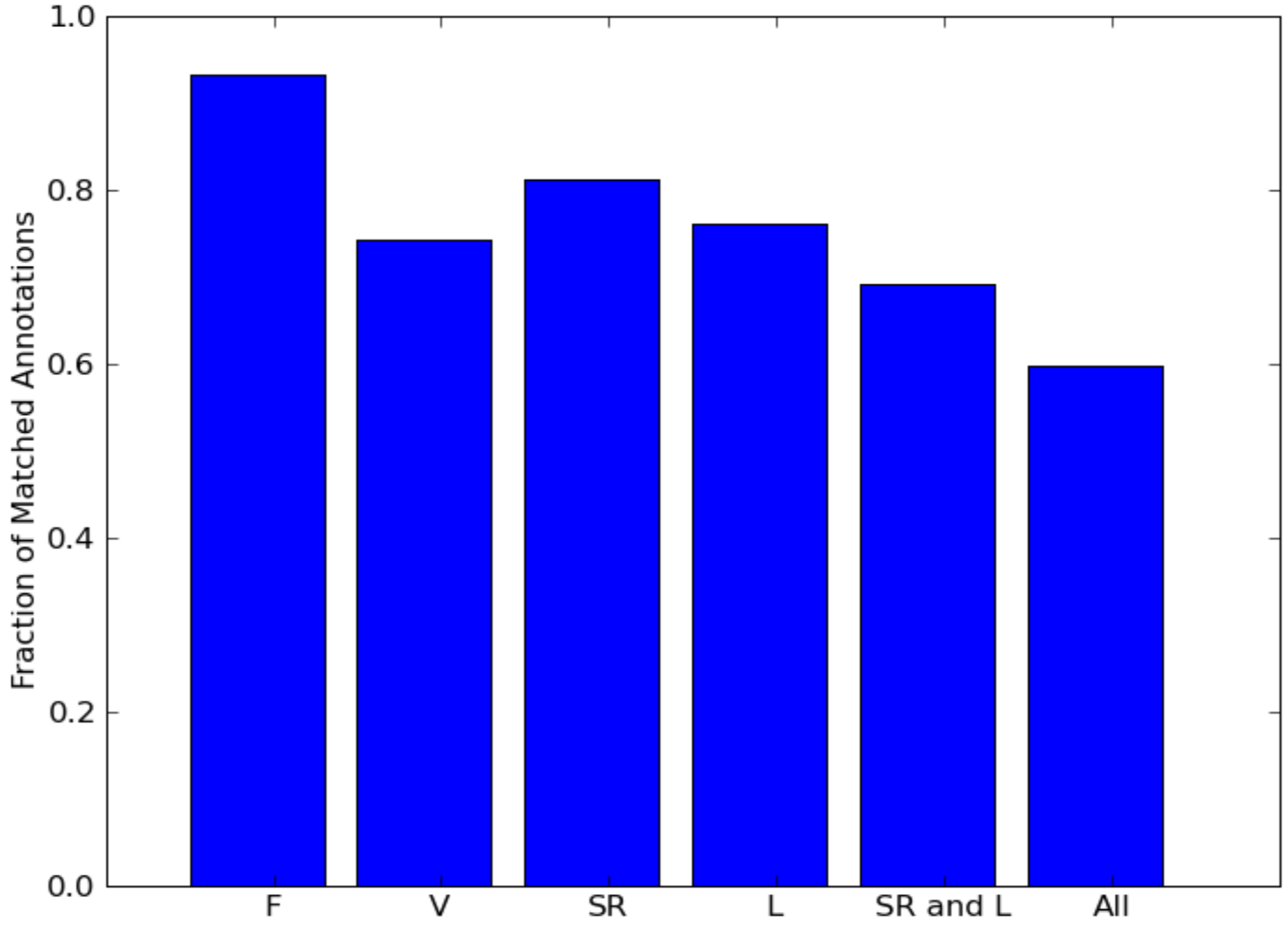
SR: until, L: you come to an intersection

Extracting Spatial Description Clauses

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

Extracting Spatial Description Clauses



Computational Linguistics

- Maron (1961)
  - First bag-of-words paper, using Naïve Bayes.
- Cleverdon (1967)
  - First corpus-based information retrieval evaluation.

Related Work

- Cognitive Semantics
  - Landau and Jackendoff (1993)
  - Talmy (2005)
- Cognitive Science
  - Regier (1992)
  - Regier and Carlson (2001) – attention vector sum for “above”
  - Carlson and Covey (2005) - “How far is near? Inferring distance from spatial descriptions.”

Related Work

- Turing (1950)
- Winograd (1971)
  - Shrdlu
- Harnad (1990)
  - Grounding