Cognitive Mapping and Planning for Visual Navigation

Saurabh Gupta
UC Berkeley

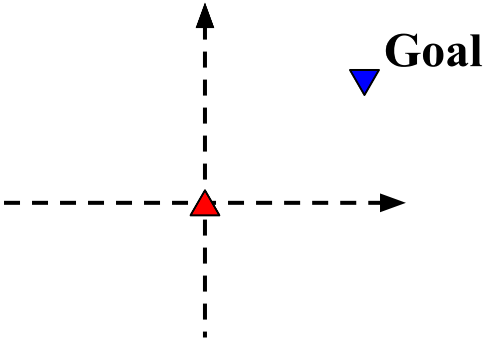
Joint work with James Davidson, Sergey Levine, Rahul Sukthankar, Jitendra Malik

Problem Statement

Robot navigation in novel environments

Robot equipped with a first person camera

Dropped into a novel environment it has not been in before.



"Go 300 feet North, 400 feet East

Or

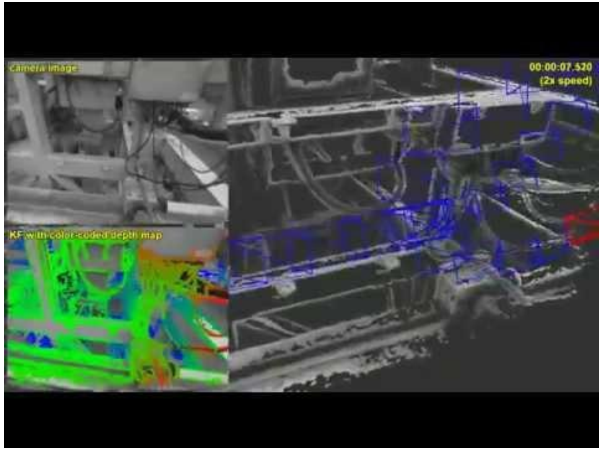
"Go Find a Chair"

Navigate in the environment

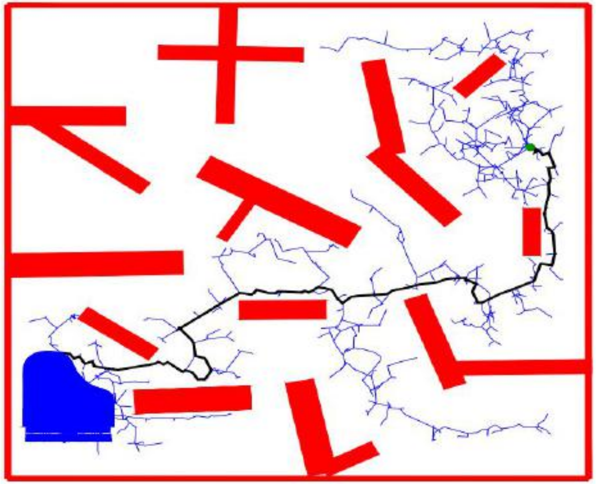
Classical Work

- **Over-complete** - Precise reconstruction of everything is not necessary
- **Incomplete**
  - Nothing is known till it is explicitly observed, fail to exploit the structure of the world.
  - Only geometry, no semantics
- Unnecessarily fragile due to separation between mapping and planning.

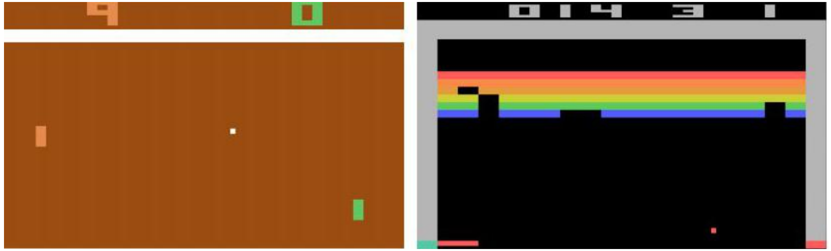
Mapping



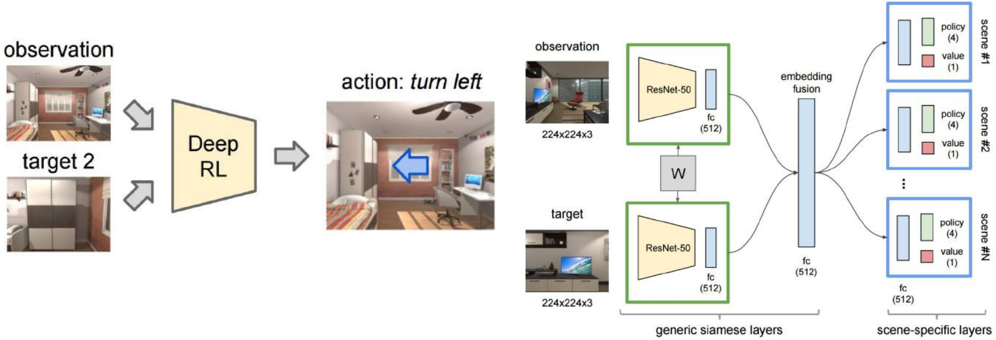
Planning



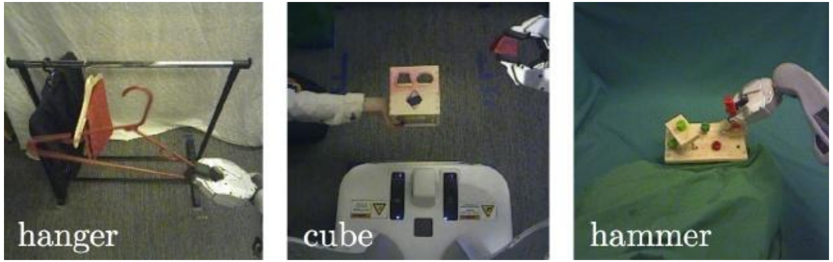
Contemporary Work



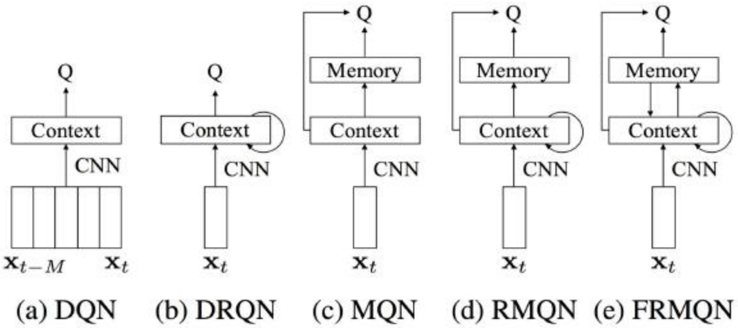
Human-level control through deep reinforcement learning, Mnih et al., Nature 2014



Target-driven Visual Navigation in Indoor Scenes using Deep Reinforcement Learning, Zhu et al., ICRA 2017

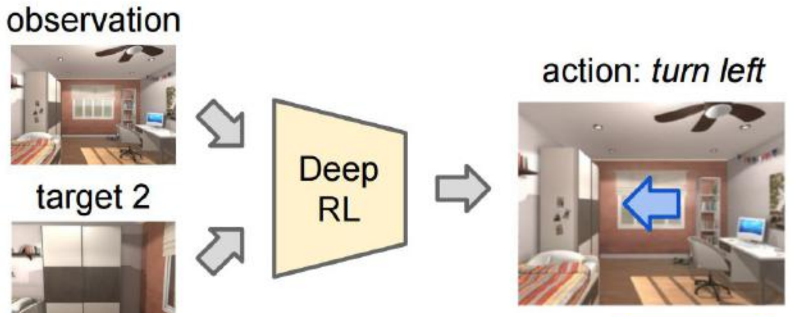


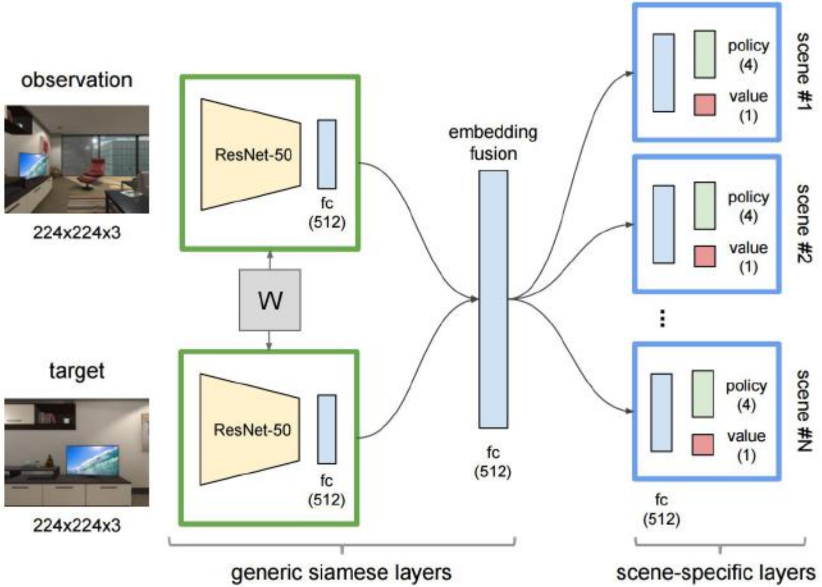
End-to-End Training of Deep Visuomotor Policies, Levine et al., JMLR 2015



Control of Memory, Active Perception, and Action in Minecraft, Oh et al., ICML 2016

Contemporary Work





Target-driven Visual Navigation in Indoor Scenes using Deep Reinforcement Learning, Zhu et al., ICRA 2017

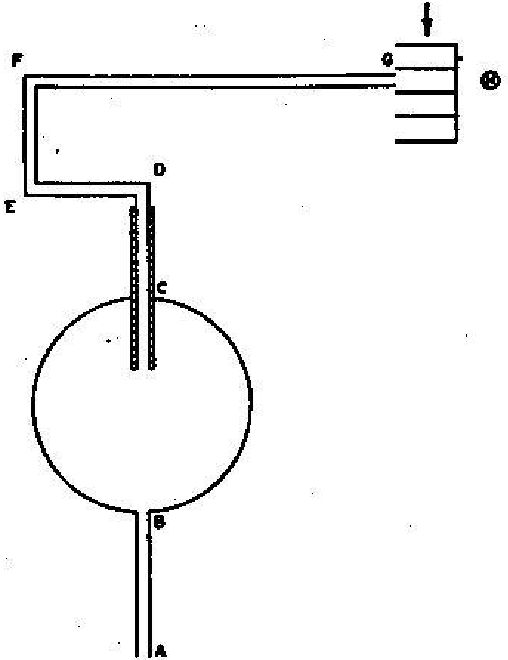
Feed forward architecture without memory.

- Agent can't systematically explore a new environment or backtrack.
- Agent needs experience with a new environment before it can start navigating successfully.

COGNITIVE MAPS IN RATS AND MEN [1]

BY EDWARD C. TOLMAN

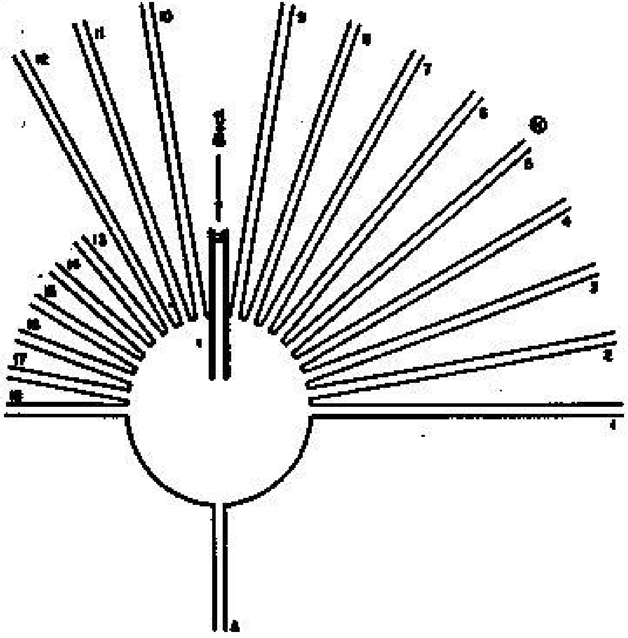
University of California



Apparatus used in preliminary training

FIG. 15

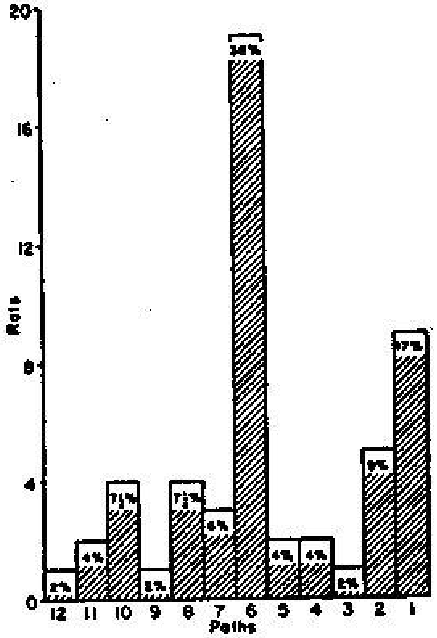
(From E. C. Tolman, B. F. Ritchie and D. Kalish, Studies in spatial learning. I. Orientation and the short-cut. *J. exp. Psychol.*, 1946, 36, p. 16.)



Apparatus used in the test trial

FIG. 16

(From E. C. Tolman, B. F. Ritchie and D. Kalish, Studies in spatial learning. I. Orientation and short-cut. *J. exp. Psychol.*, 1946, 36, p. 17.)



Numbers of rats which chose each of the paths

FIG. 17

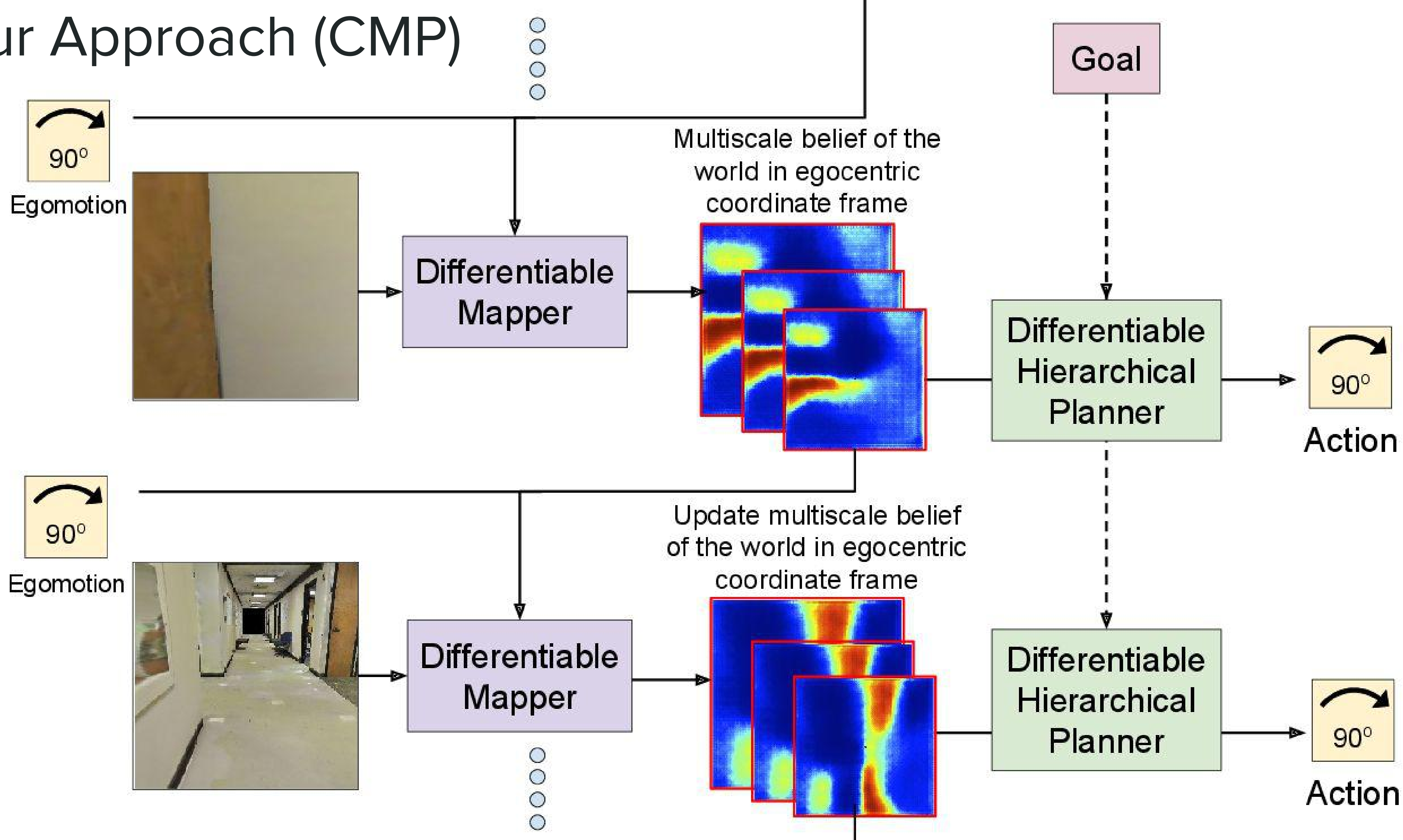
(From E. C. Tolman, B. F. Ritchie and D. Kalish, Studies in spatial learning. I. Orientation and the short-cut. *J. exp. Psychol.*, 1946, 36, p. 19.)

Our Approach

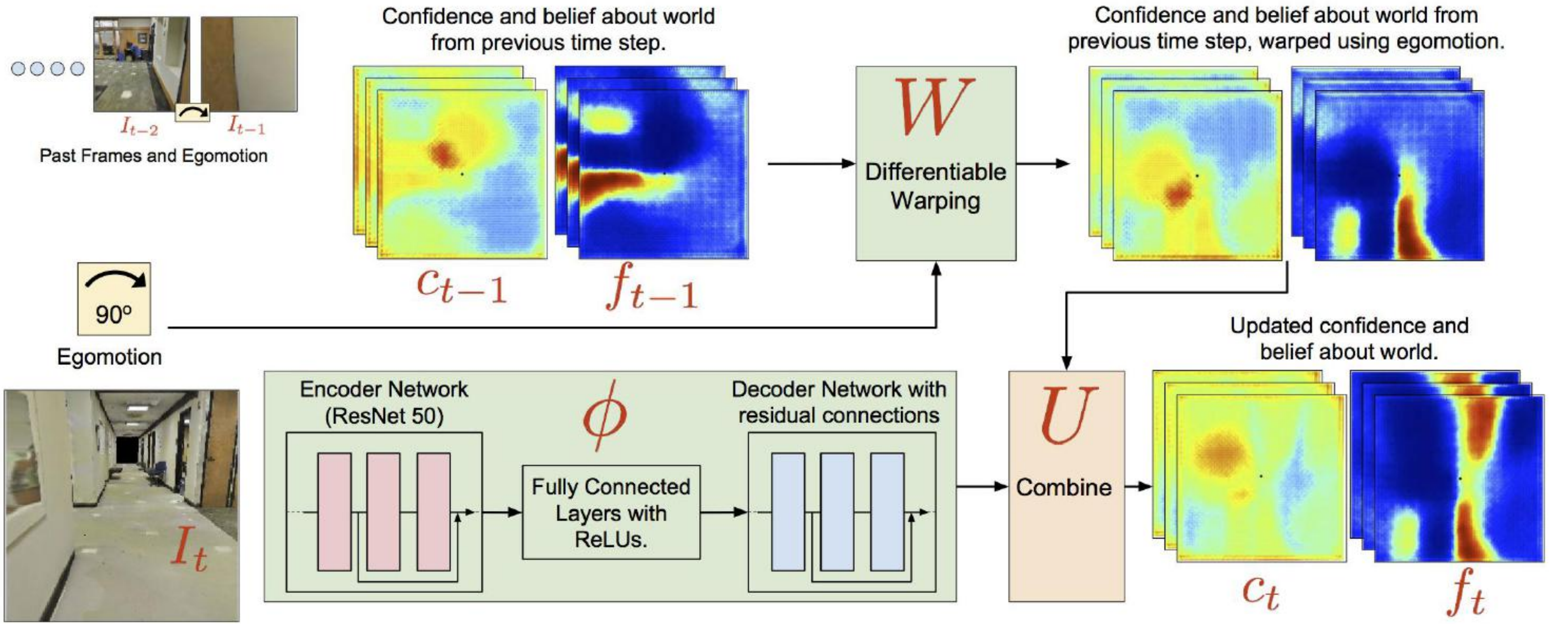
Neural network policy for visual navigation

a) Joint architecture for mapping and planning
b) Spatial memory with the ability to plan given partial observations
c) Is end-to-end trainable

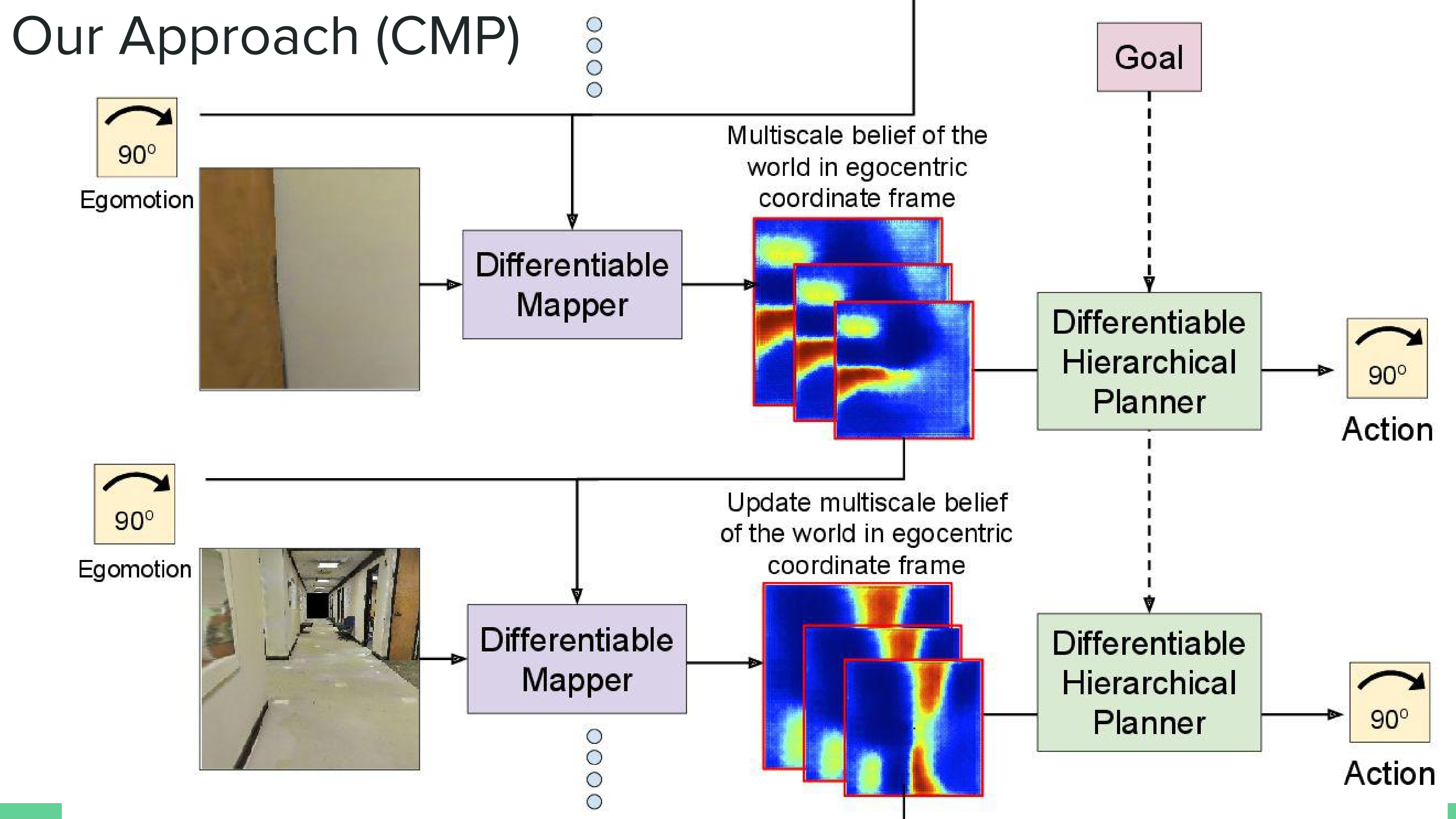
Our Approach (CMP)



Differentiable Mapper



Our Approach (CMP)



Differentiable Planner

Value Iteration Networks

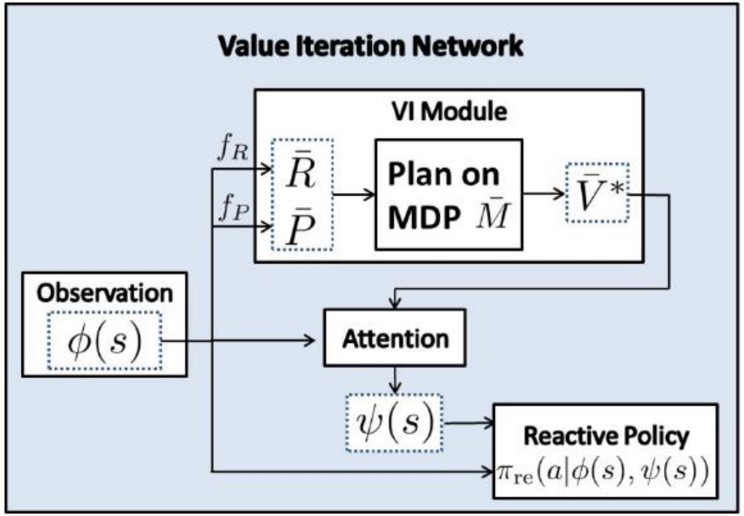
Aviv Tamar, Yi Wu, Garrett Thomas, Sergey Levine, and Pieter Abbeel

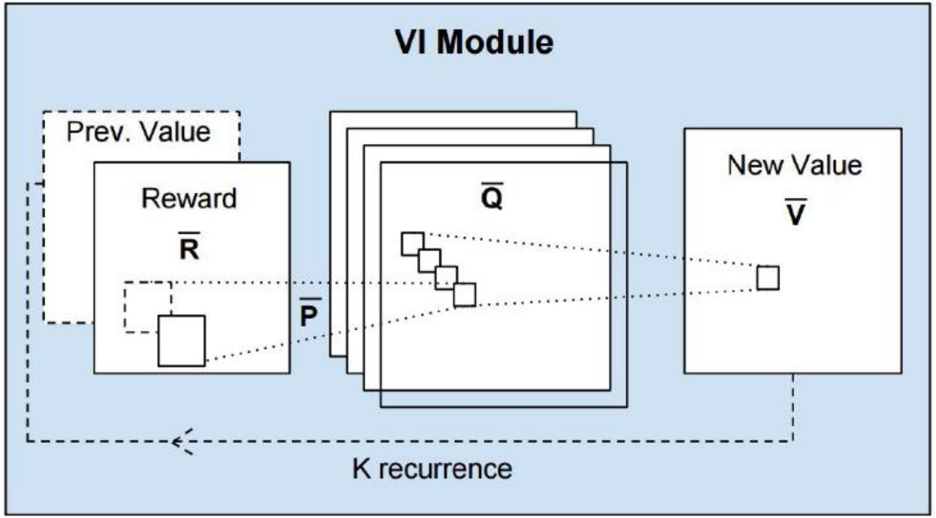
$$Q_n(s,a) = R(s,a) + \gamma \sum_{s'} P(s'|s,a) V_n(s')$$

If actions move the agent locally, then can be computed using convolutions

$$V_{n+1}(s) = \max_a Q_n(s,a) \quad \forall s$$

Max Pooling over channels





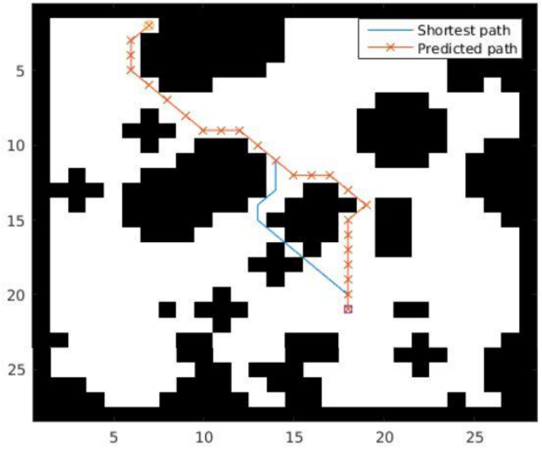
Differentiable Planner

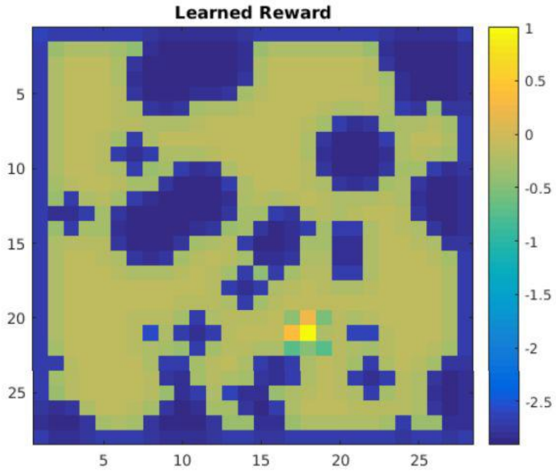
$$Q_n(s,a) = R(s,a) + \gamma \sum_{s'} P(s'|s,a) V_n(s')$$

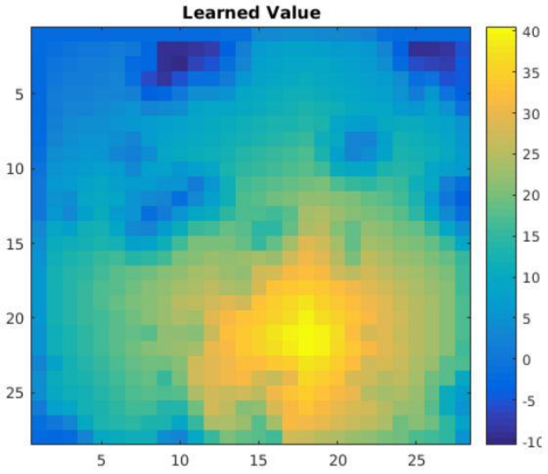
If actions move the agent locally, then can be computed using convolutions

$$V_{n+1}(s) = \max_a Q_n(s,a) \quad \forall s$$

Max Pooling over channels



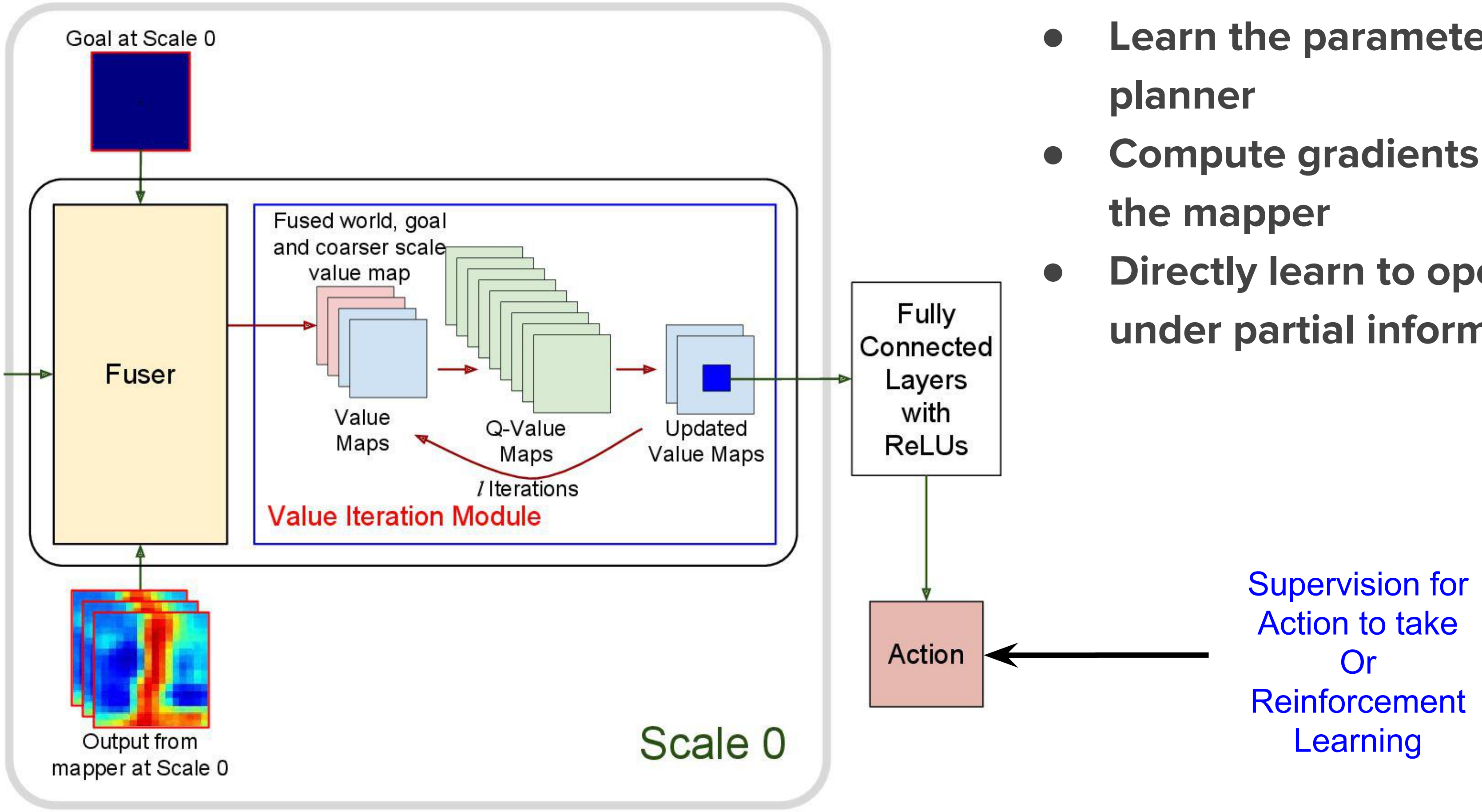




Can be trained using training data

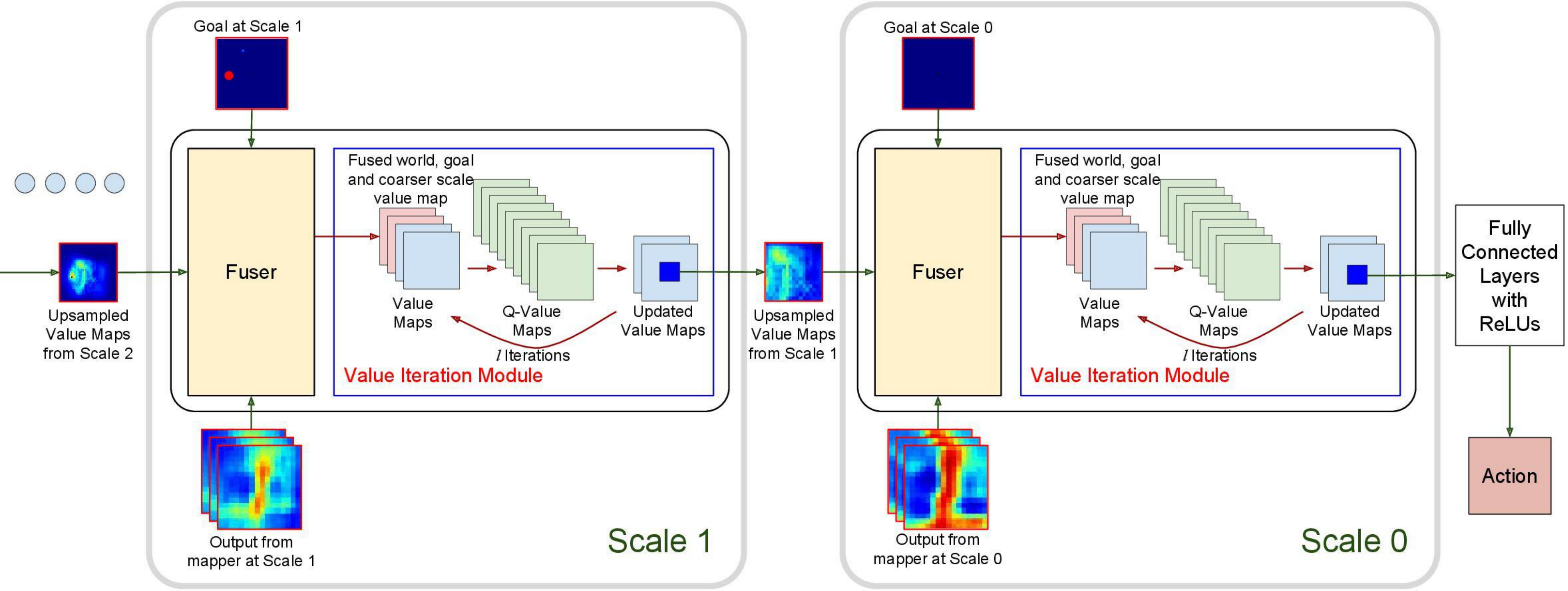
Value Iteration Networks, Aviv Tamar, Yi Wu, Garrett Thomas, Sergey Levine, and Pieter Abbeel, In NIPS 2016

Differentiable Planner

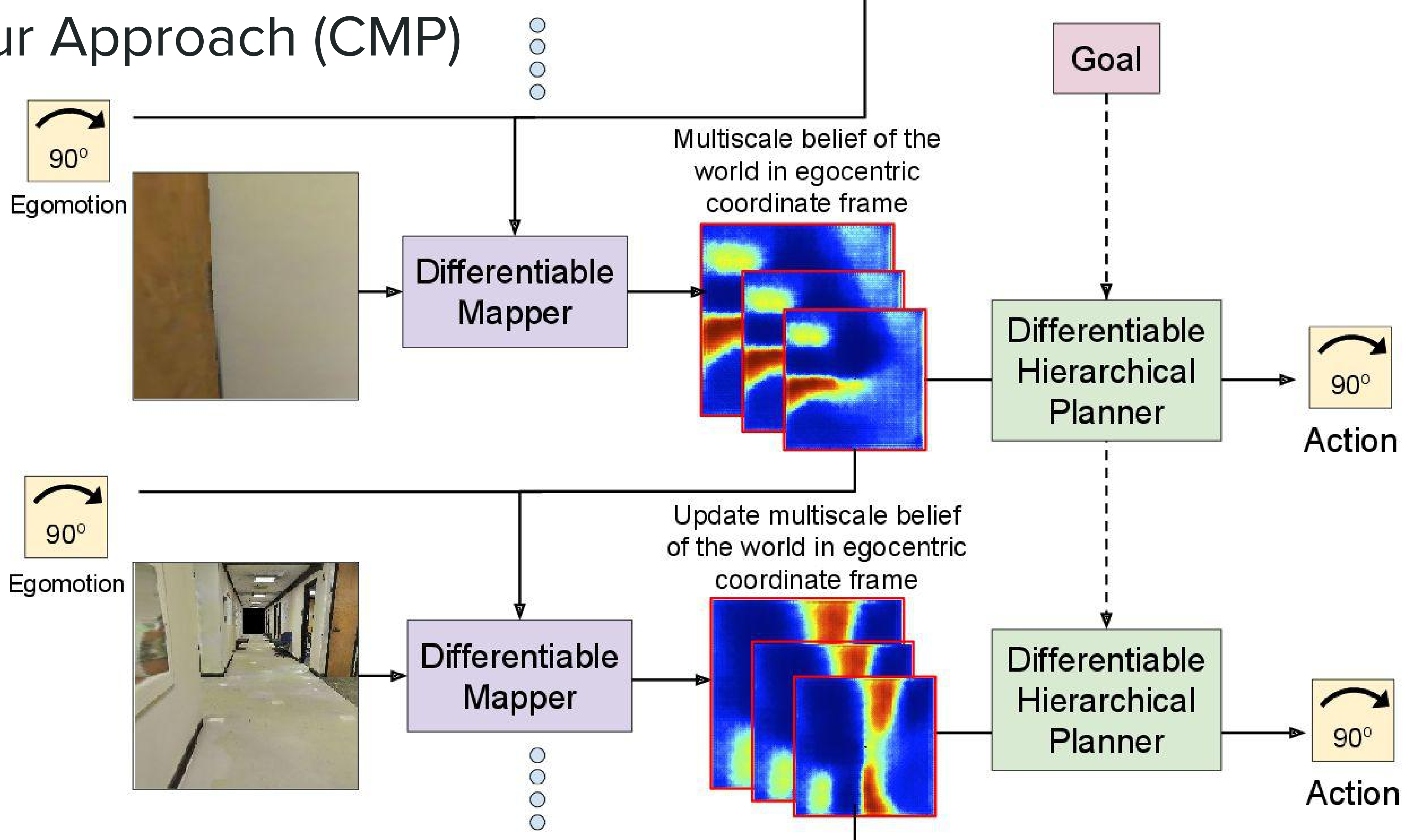


- **Learn the parameters of the planner**
- **Compute gradients to train the mapper**
- **Directly learn to operate under partial information**

Differentiable Planner (Multi Scale)



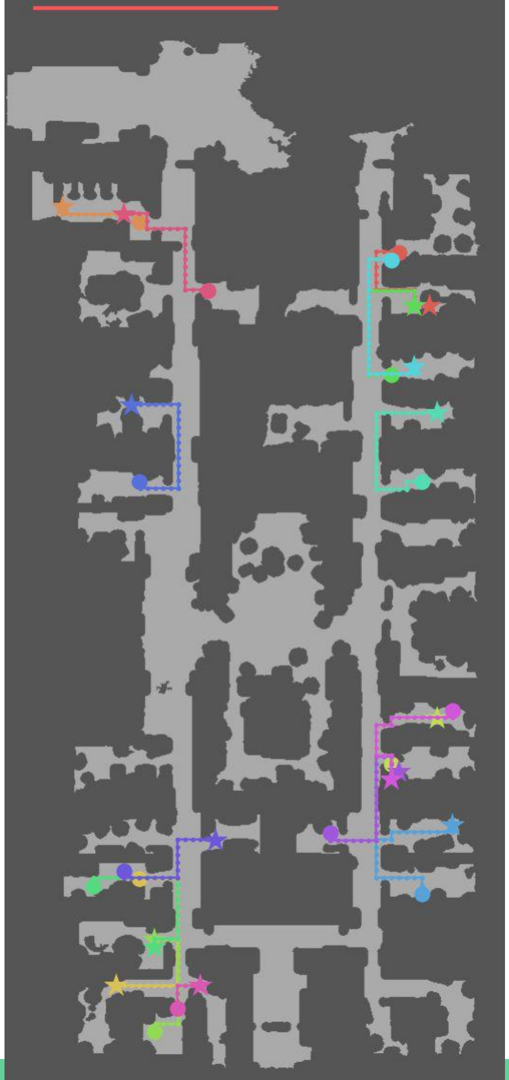
Our Approach (CMP)

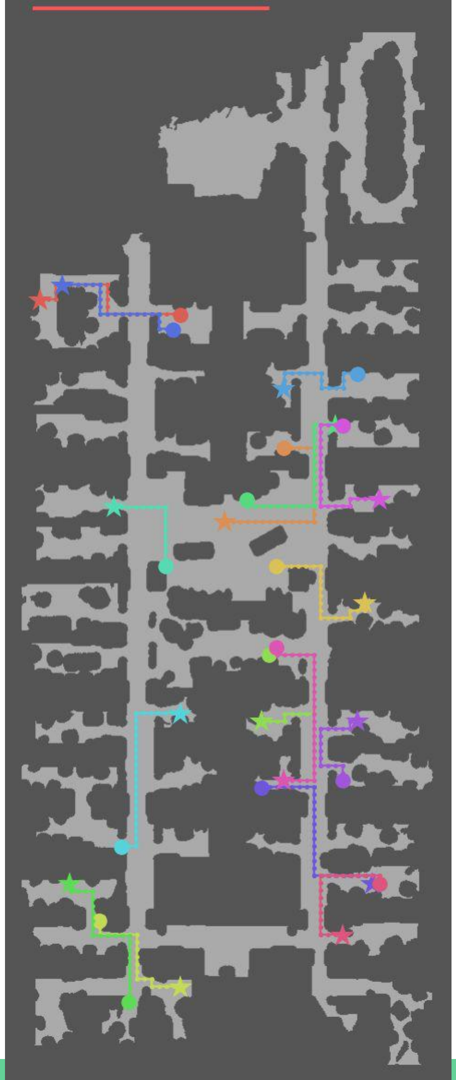


Experimental Setup

- Trained and tested in static simulated real-world environments.
- **Testing environment is different from training environments.**
- Robot:
  - Robot lives in a grid world. Motion is discrete.
  - Robot has 4 macro-actions:
    - Go Forward, Turn left, Turn right, Stay in place.
  - Robot has access to precise egomotion.
  - Robot has RGB or Depth Cameras.
- All models are trained using DAGGER.
- Geometric Task:
  - Goal is sampled to be at most 32 time steps away. Agent is run for 39 time steps.
- Semantic Task:
  - 'Go to a Chair', agent run for 39 time steps.

Stanford Building Parser Dataset

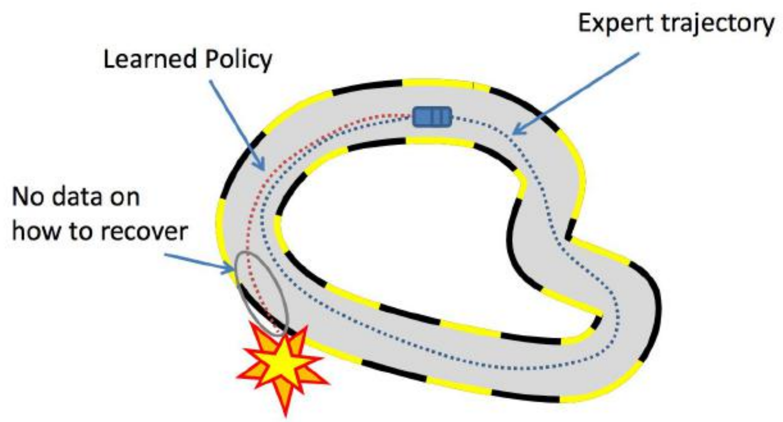


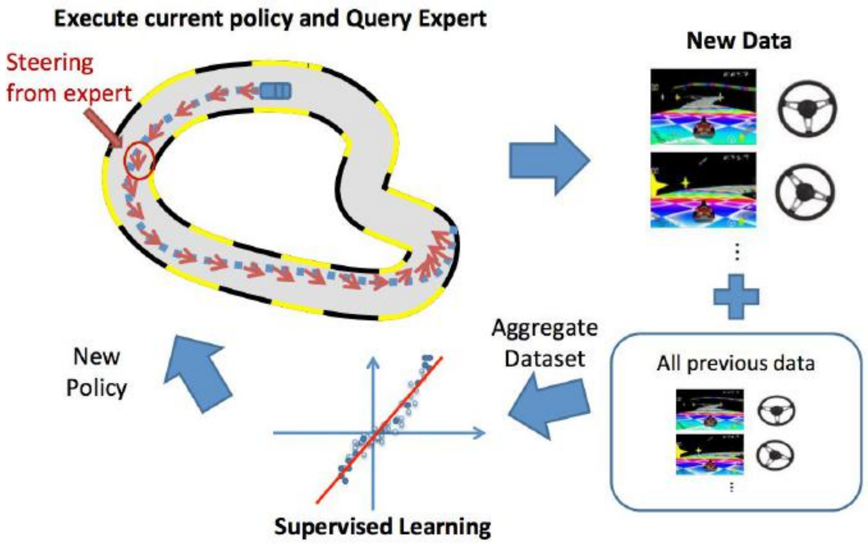


3D Semantic Parsing of Large-Scale Indoor Spaces, Armeni et al. CVPR 2016

Policy Training using DAGGER

Data Mismatch Problem

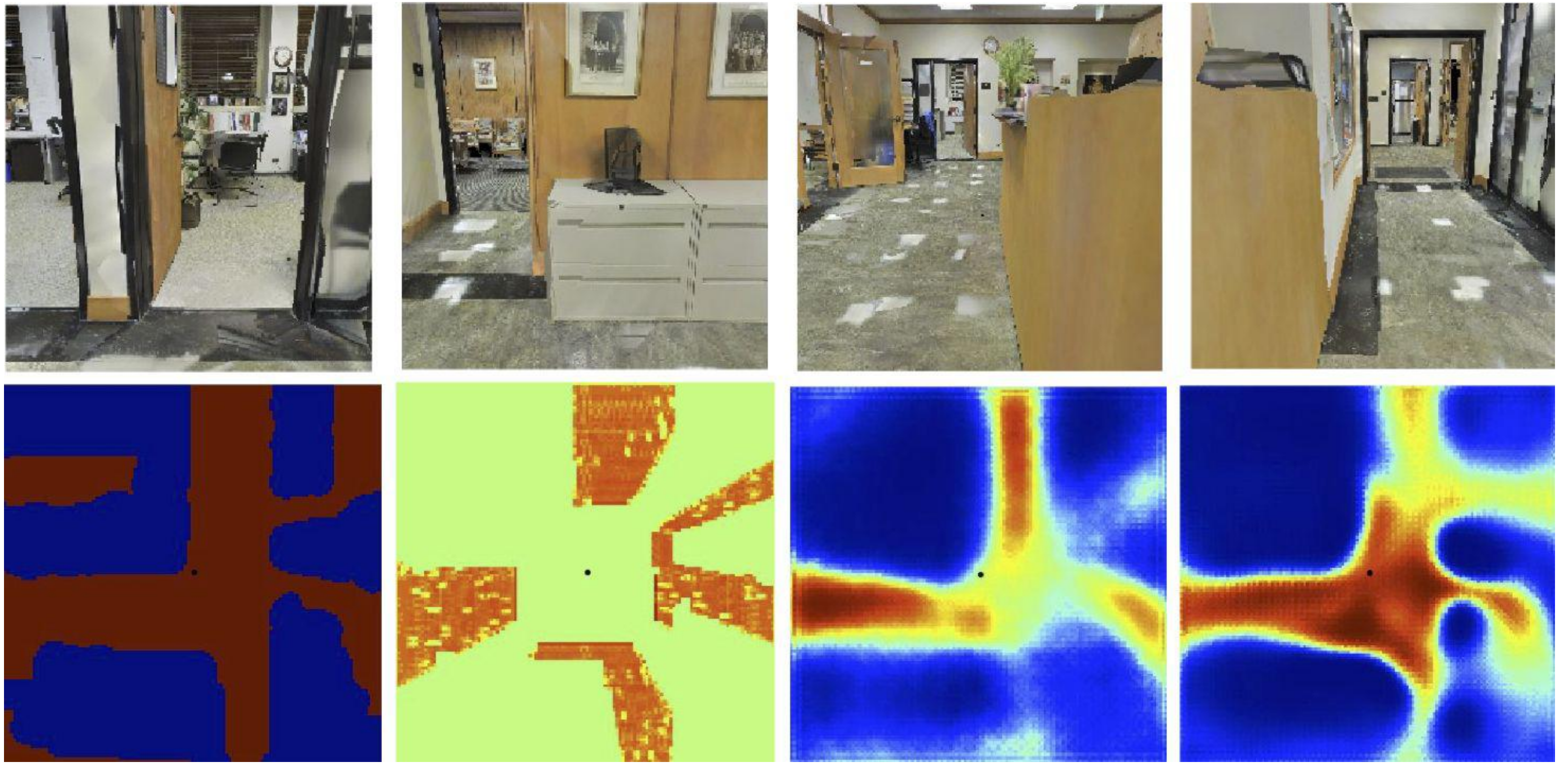




Slide: John Schulman's Lecture on Reinforcement Learning
A Reduction of Imitation Learning and Structured Prediction to No-Regret Online Learning, Ross, Gordon & Bagnell (2010).

Mapper Unit Test

Learned mapper provides more complete output



Ground Truth　Analytical Projection　RGB Pred　D Pred

Navigation Results

Methods	RGB Input			Depth Input		
	Mean Distance	75th %ile Distance	Success Rate (in %)	Mean Distance	75th %ile Distance	Success Rate (in %)
Initial	25.3	30	0.7	25.3	30	0.7
No Image	20.8	28	0.7	20.8	28	0.7
React 1	20.9	28	8.2	17.0	26	21.9
React 4	14.4	25	30.4	8.8	18	56.9
LSTM	10.3	21	53	5.9	5	71.8
Our(CMP)	**7.7**	**14**	**62.5**	**4.8**	**1**	**78.3**

Additional Comparisons

- **Larger improvements for long horizon tasks.**
- **Better Generalization: Smaller drop in performance when transferring between datasets and training on smaller amount of data.**
- **CMP performs much better on harder examples.**

	75^{th} %ile			Success Rate (in %)		
	Initial	LSTM	CMP	Initial	LSTM	CMP
Far away goal (run for 79 steps)	58	29	19.2	0.0	58.4	66.3
Far away goal (run for 159 steps)	58	19	0	0.0	69.0	78.5
Train on 5 floors	30	5	1	0.7	71.8	78.3
Train on 1 floor	30	18	10	0.7	58.9	67.9
Transfer from internal dataset	30	21	15	0.7	48.6	61.1

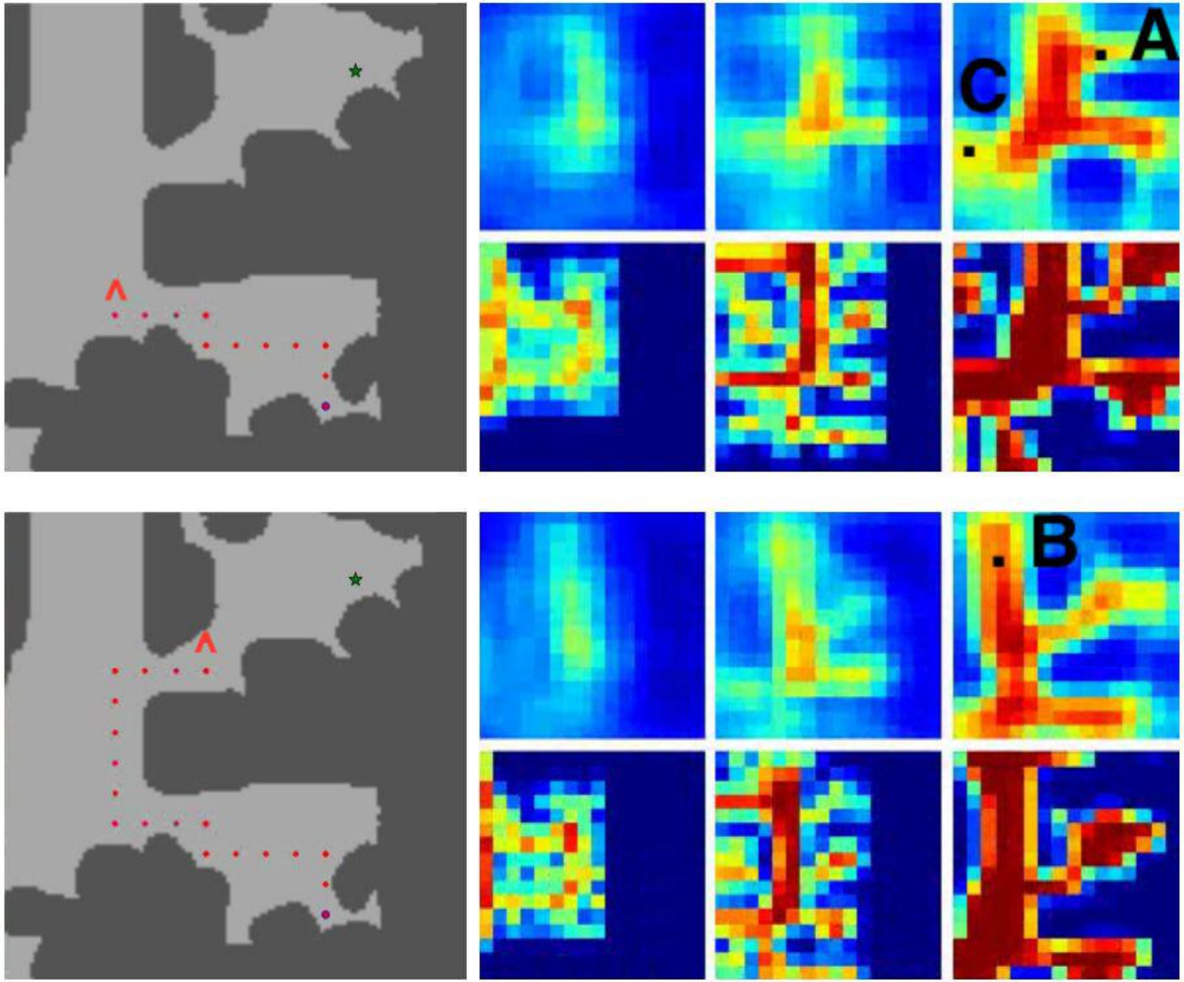
Ablations

- Single scale planner works almost as well, except for being slower
- Removing the planning module hurts performance
- Analytical mapper performs worse than a learned mapper

Method	Mean		75^{th} percentile		Success Rate (in %)	
	RGB	Depth	RGB	Depth	RGB	Depth
Geometric Task						
Initial Distance	25.3		30		0.7	
No Image LSTM	20.8		28		0.7	
Cognitive Mapper and Planner						
Full model	7.7	4.8	14	1	62.5	78.3
Single-scale planning	7.9	4.9	12	1	63.0	79.5
No planning	8.5	4.8	16	1	58.6	79.0
Single-scale planning, analytical map	-	8.0	-	14	-	62.9

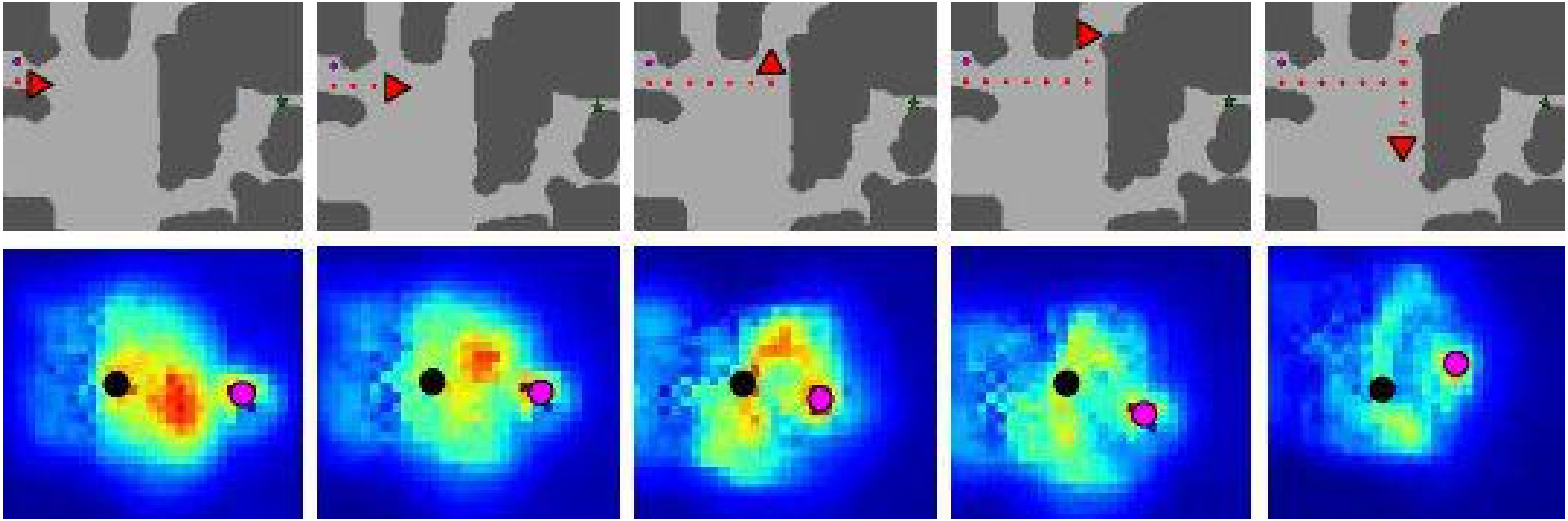
Visualizations

Reading out representation learned by mapper to predict free space.

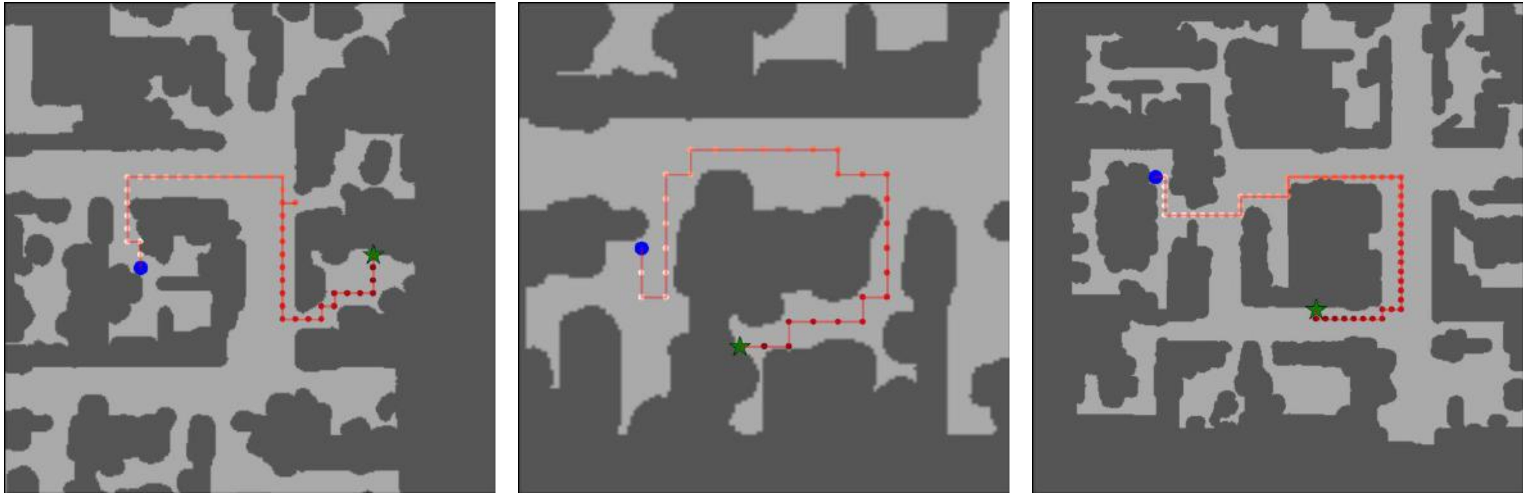


Visualizations

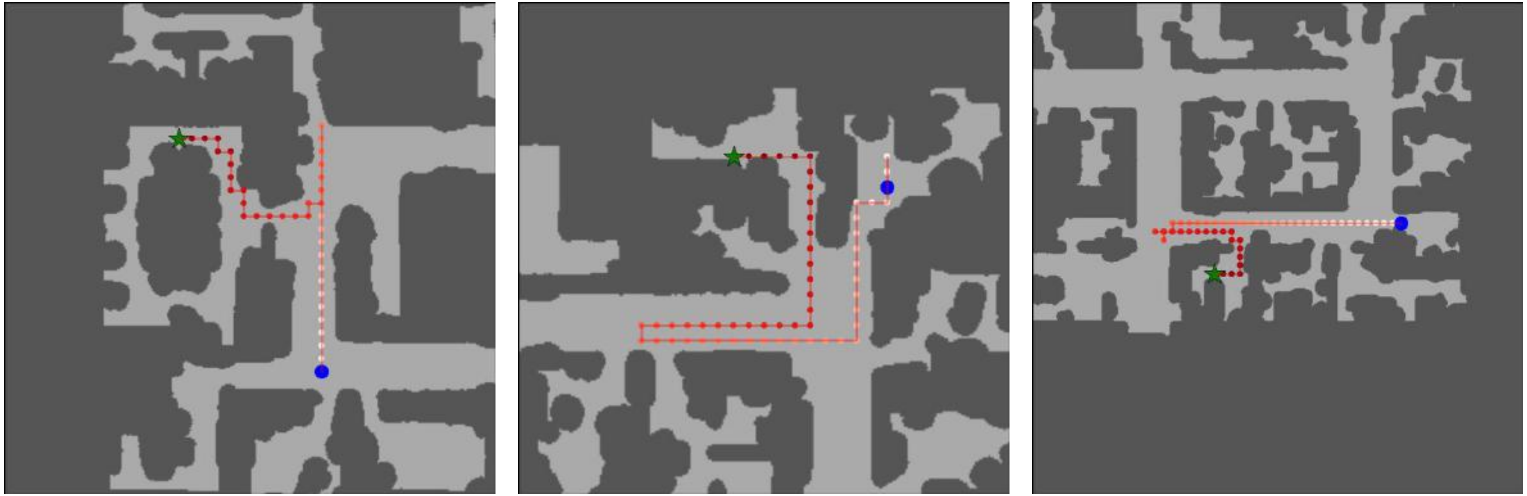
Value map at 5 snapshots from an episode



Successful Navigations



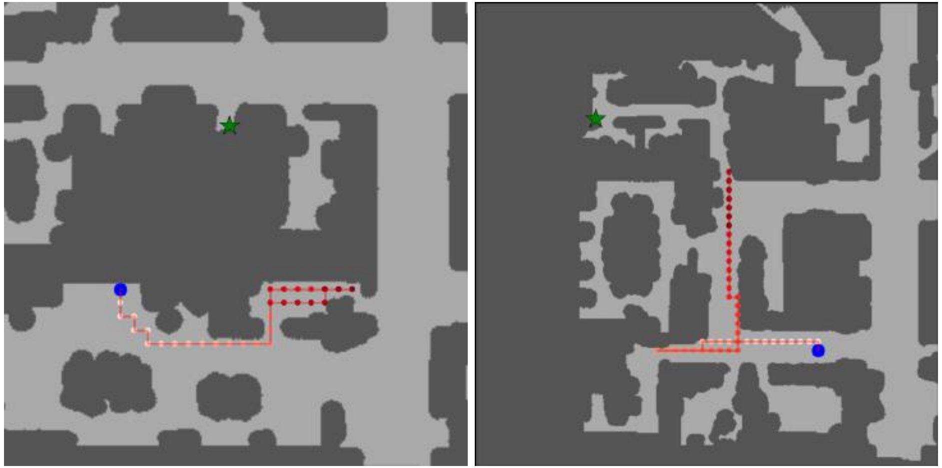
Backtracking Behaviour



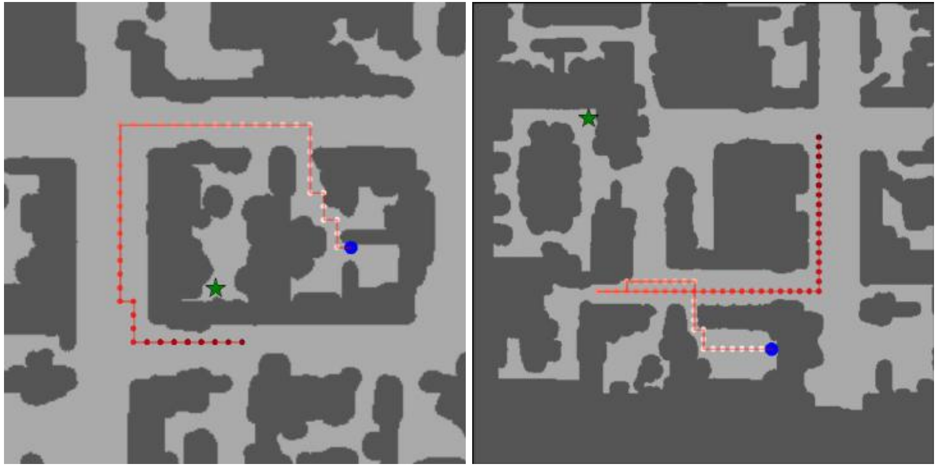
Trajectories

Failure Cases

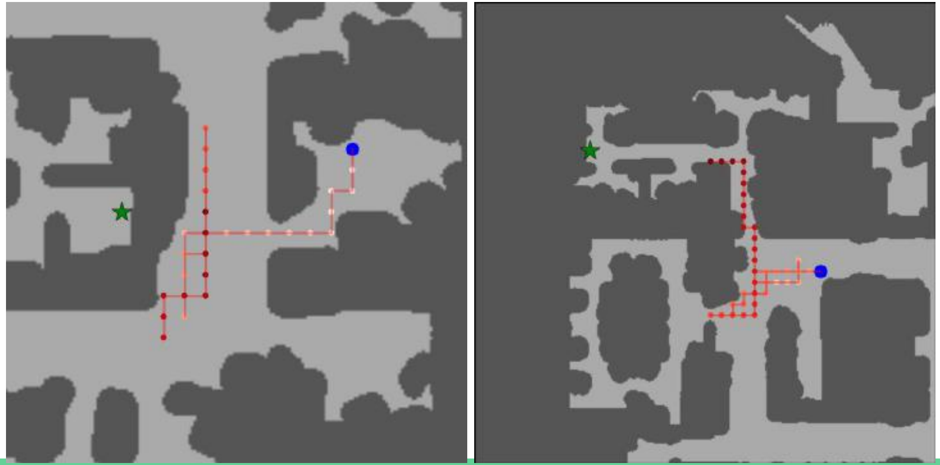
Tight spaces (quantization artifacts)



Missed pathways



Thrashing



Navigation Results (Semantic Task)

Methods	RGB Input			Depth Input		
	Mean Distance	75th %ile Distance	Success Rate (%)	Mean Distance	75th %ile Distance	Success Rate (%)
Initial	16.2	25	11.3	16.2	25	11.3
Reactive (4)	14.2	22	23.4	14.2	23	22.3
LSTM	13.5	20	23.5	13.4	23	27.2
Our(CMP)	**11.3**	**18**	**34.2**	**11.0**	**19**	**40.0**

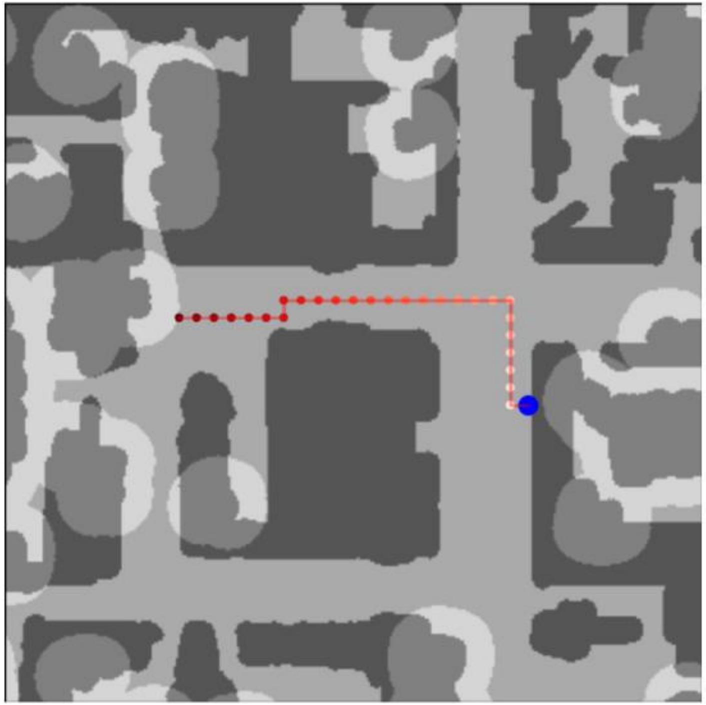
Navigation Results (Semantic Task)

Per category Results

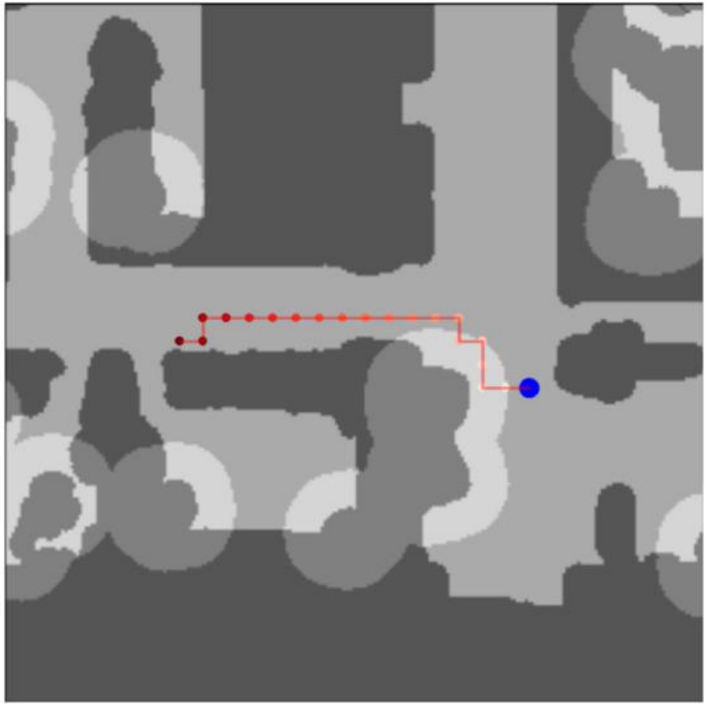
Methods	RGB Input **[Success Rate(%)]**			Depth Input **[Success Rate(%)]**		
	Chair	Door	Table	Chair	Door	Table
Initial	9.9	11.9	11.7	9.9	11.9	11.7
Reactive (4)	22.0	24.8	21.9	16.9	26.2	20.7
LSTM	17.9	26.9	23.6	23.1	28.9	28.9
Our(CMP)	**32.8**	**38.3**	**26.4**	**40.6**	**40.3**	**38.2**

Trajectory (Semantic Task)

Success Case



Failure Case




First Person View
Top corners show diagnostics (distance, agents' action) not input to agent.
Dist to Goal:
4 steps
Readout Output (Egocentric)
GT Free Space (Egocentric)
Trajectory in topview

Thank You

Backup Slides

COGNITIVE MAPS IN RATS AND MEN[1]

BY EDWARD C. TOLMAN

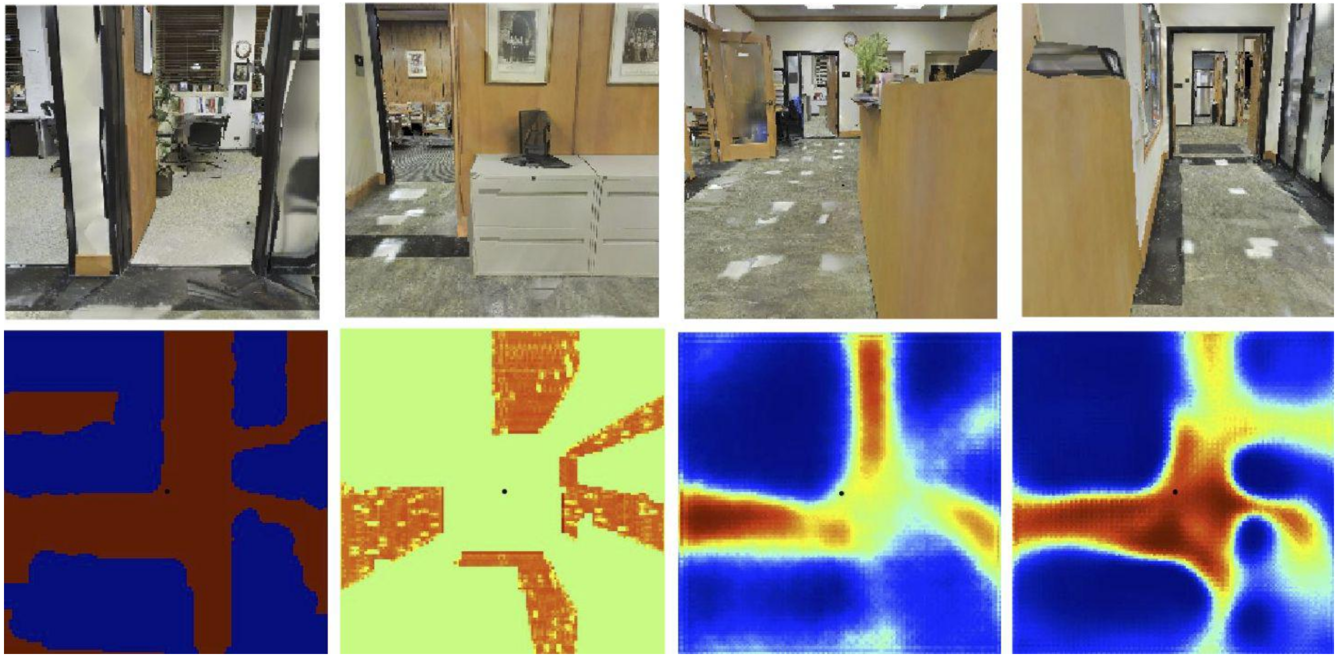
University of California

Secondly, we assert that the central office itself is far more like a map control room than it is like an old-fashioned telephone exchange. The stimuli, which

outgoing responses. Rather, the incoming impulses are usually worked over and elaborated in the central control room into a tentative, cognitive-like map of the environment. And it is this

Mapper Unit Test

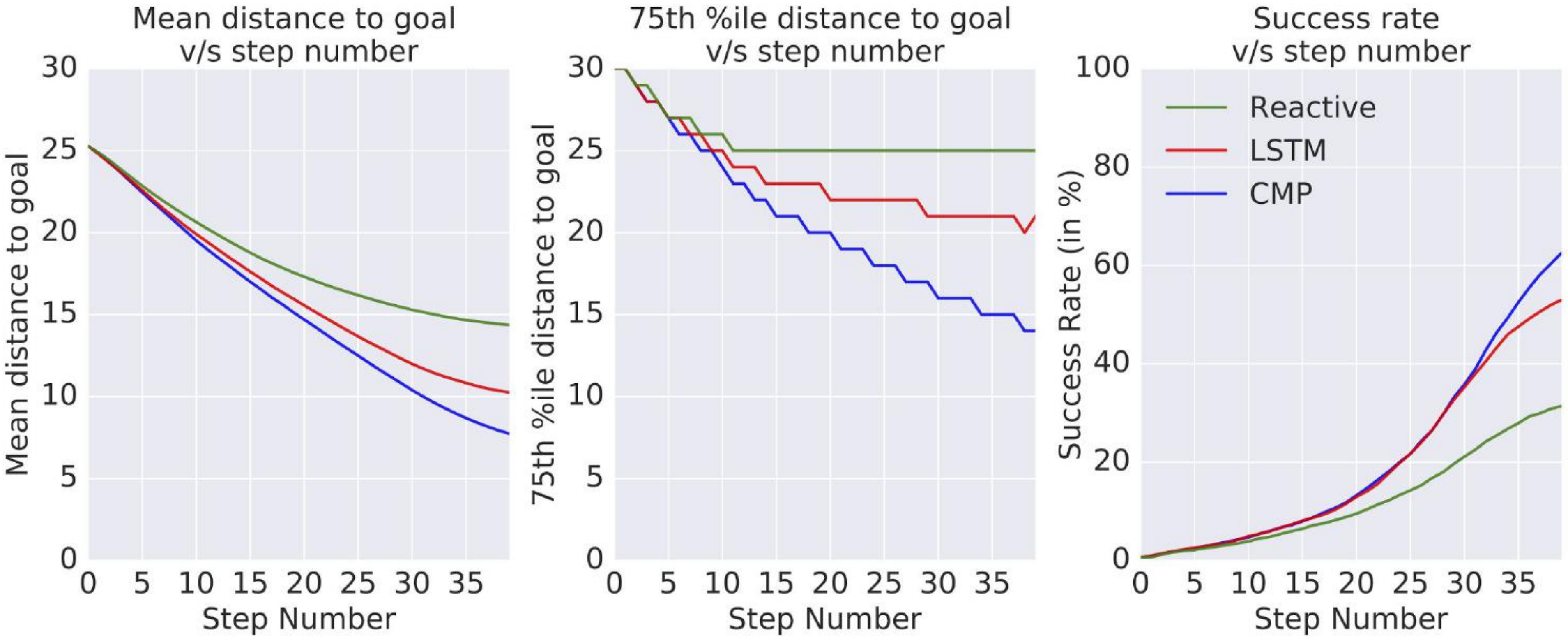
Learned mapper provides more complete output



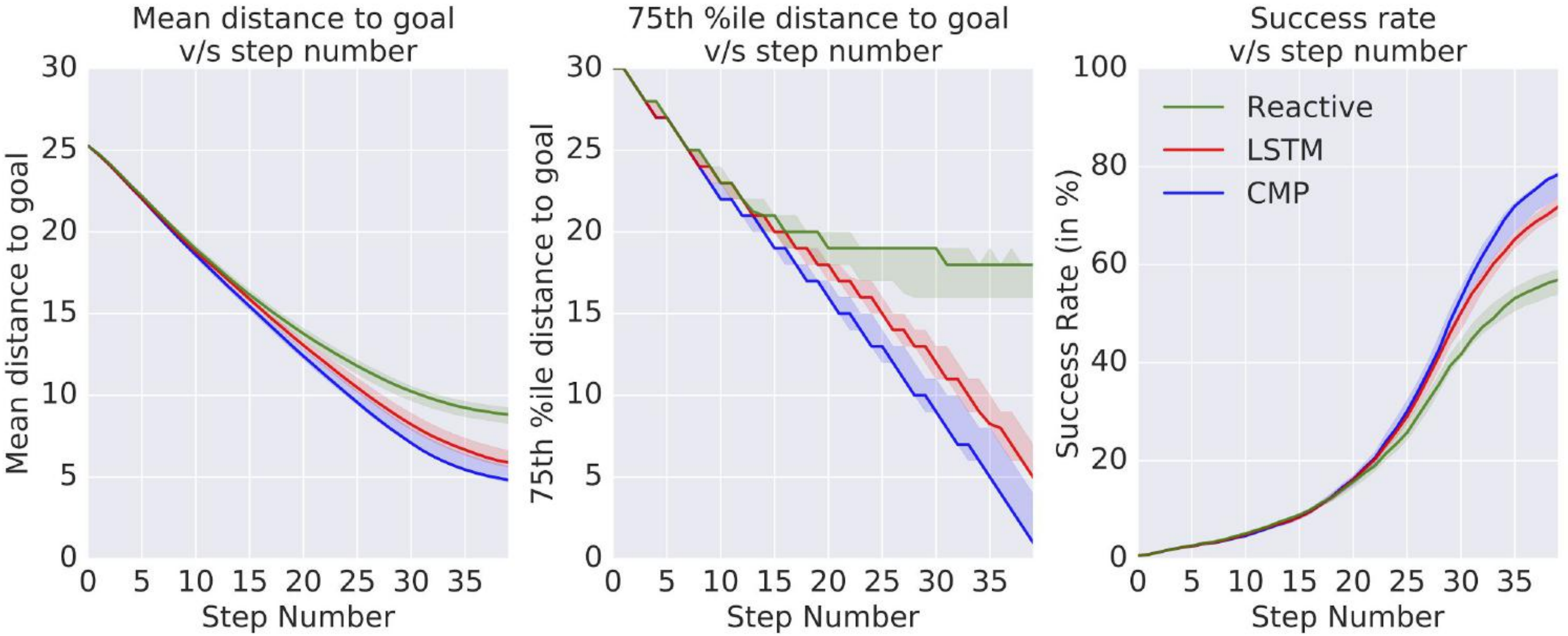
Ground Truth Analytical Projection RGB Pred D Pred

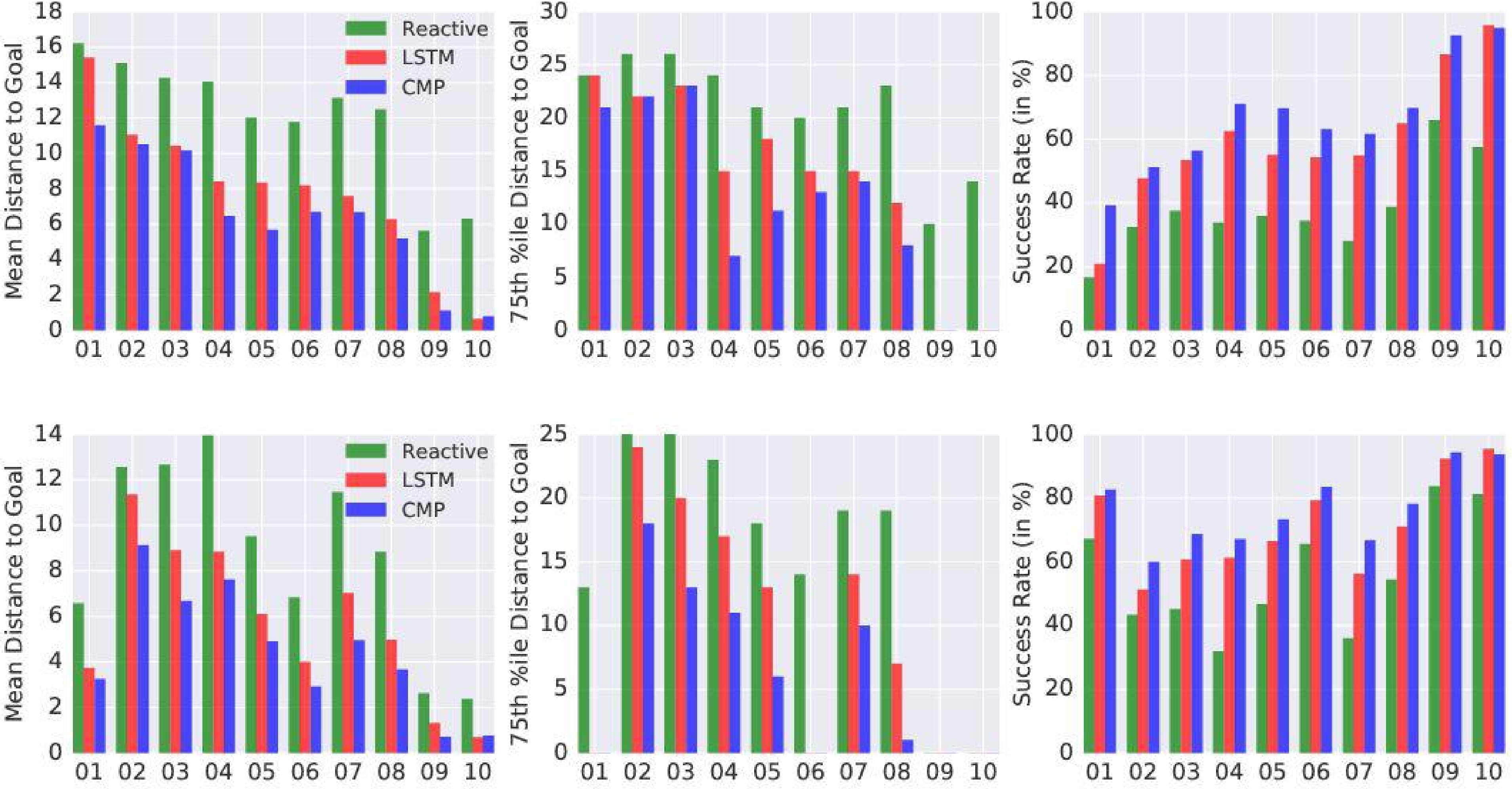
Method	Modality	CNN Architecture	Free Space Prediction AP
Analytical Projection	depth	-	56.1
Learned Mapper	RGB	ResNet-50	74.9
Learned Mapper	depth	ResNet-50 Randomly Initialized	63.4
Learned Mapper	depth	ResNet-50 Initialized using [20]	78.4

Navigation Results: RGB Images



Navigation Results: Depth Images

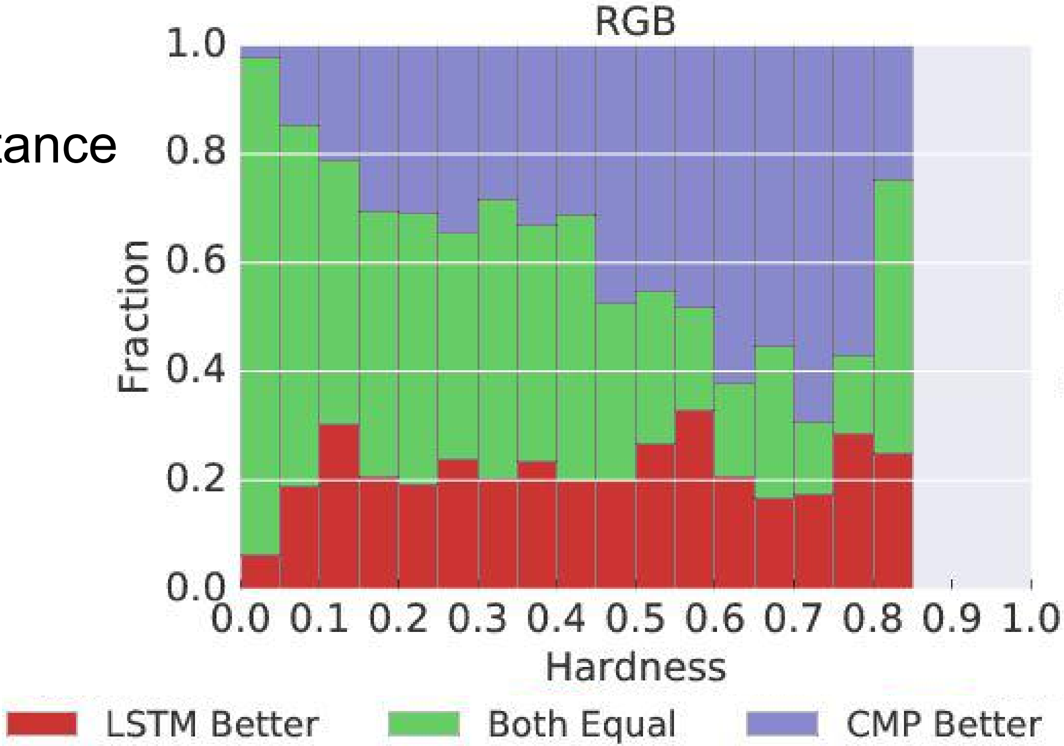



Reactive
LSTM
CMP
Mean Distance to Goal
75th %ile Distance to Goal
Success Rate (in %)
01 02 03 04 05 06 07 08 09 10
Reactive
LSTM
CMP
Mean Distance to Goal
75th %ile Distance to Goal
Success Rate (in %)
01 02 03 04 05 06 07 08 09 10

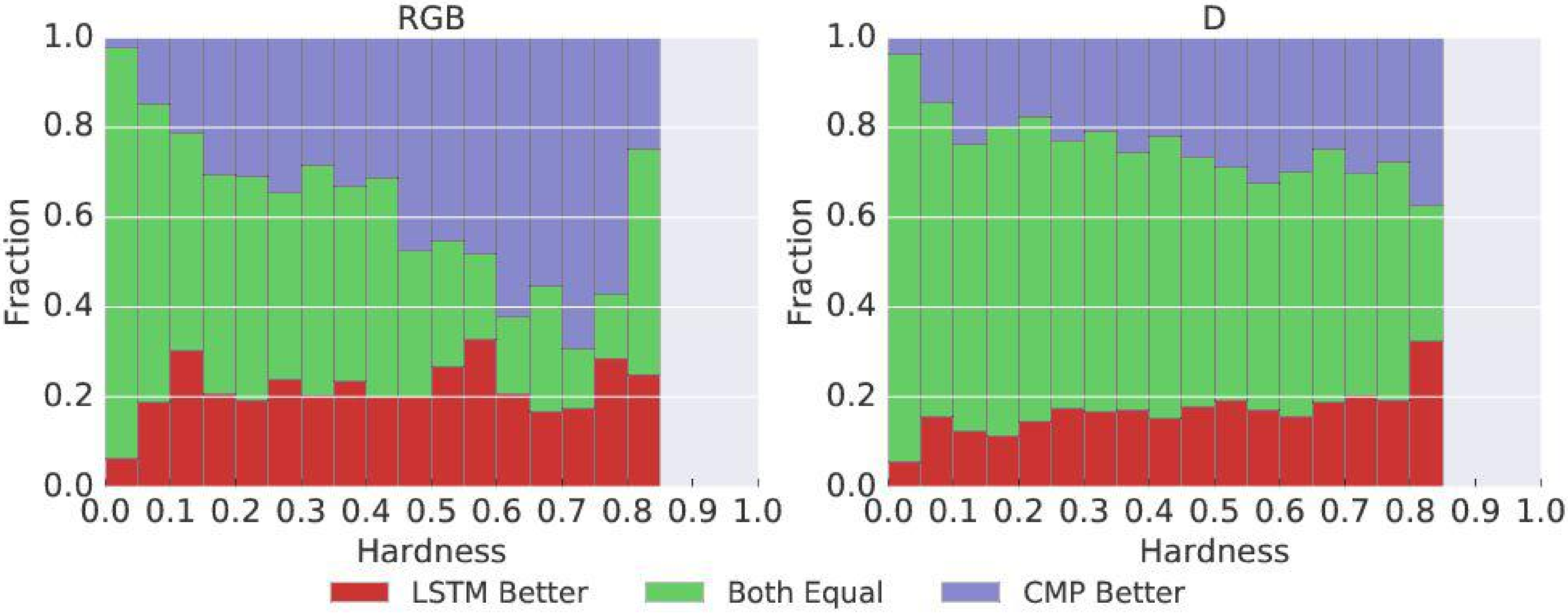
Relative performance on problems of different hardness

Hardness =

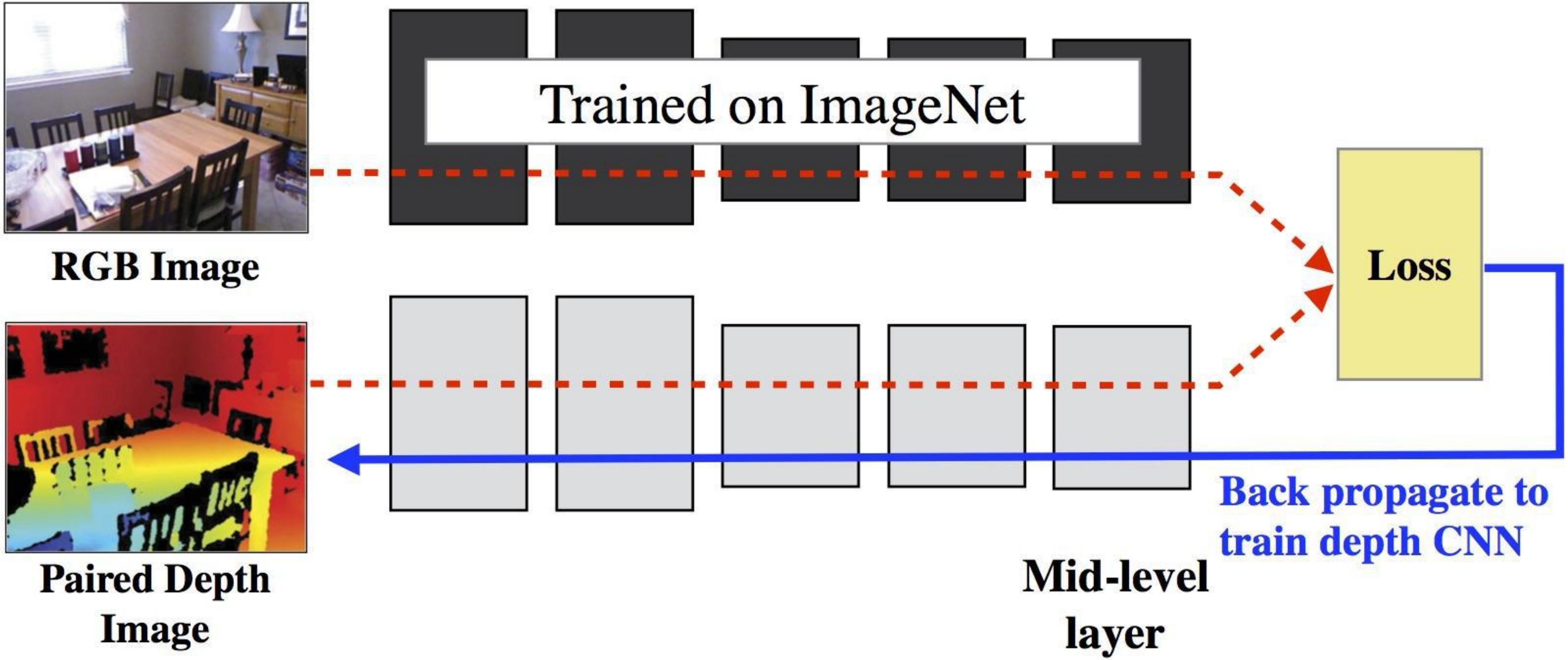
1 - heuristic distance / actual distance



Relative performance on problems of different hardness



Cross Modal Distillation for Supervision Transfer



Discussion

- Does having such structured architectures improve sample complexity when training with reinforcement learning?
- Metric representation for space v/s ‘topological’ representation for space?
- Analysis of robustness to imperfect odometry, joint training with inferred visual odometry.