Learning and Reasoning for Robot Sequential Decision Making under Uncertainty

Saeid Amiri[1], Mohammad S. Shirazi[2], and Shiqi Zhang[1]
[1]SUNY Binghamton, [2]University of Indianapolis

To appear in the proceedings of AAAI-20

Motivation

Robots need sequential decision making (SDM) algorithms to work on tasks that are not possible for individual actions. At least three AI paradigms support SDM with different strengths and weaknesses:

1. **Supervised learning**: Learning from previous interaction experiences (e.g., imitation learning).
2. **Knowledge representation and reasoning**: Reasoning with contextual knowledge.
3. **Probabilistic planning**: Active information collection for goal achievement under uncertainty.

Illustrative Example

Learning: b^{lrn}

s_0^{lrn}	s_1^{lrn}
0.00	1.00

→

Reasoning: $\hat{b}^{rsn}$

$\hat{s}_0^{rsn}$	$\hat{s}_1^{rsn}$
0.22	0.78

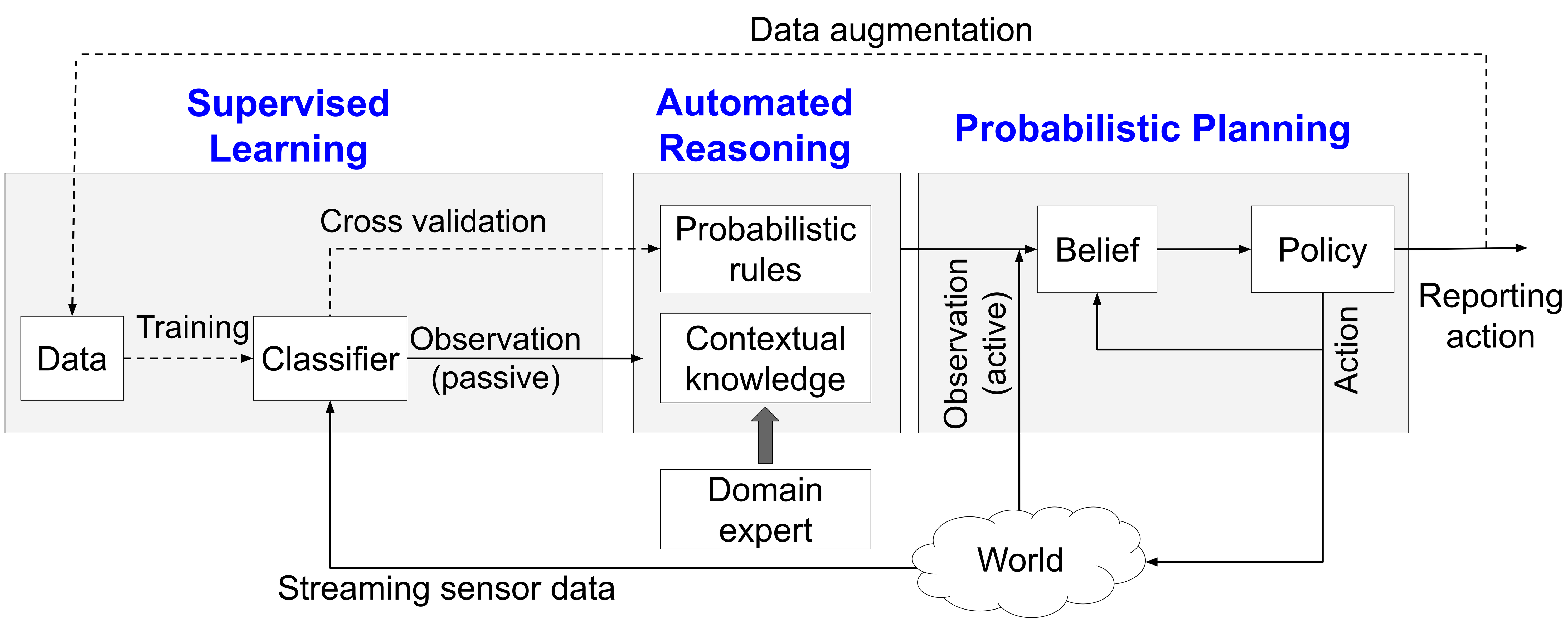
↓

Planning: b^{pln}

	s_0^{pln}	s_1^{pln}	s_2^{pln}	s_3^{pln}	s_4^{pln}	Action	Observation
1	0.22	0.78	0.00	0.00	0.00	Move forward	Positive
2	0.71	0.29	0.00	0.00	0.00	Move forward	Positive
3	0.95	0.05	0.00	0.00	0.00	Greet	Positive
4	0.98	0.02	0.00	0.00	0.00	Report human being interested	N/A
5	0.00	0.00	0.00	0.00	1.00		

LCORPP: Learning, COmmonsense Reasoning and Probabilistic Planning

- **Learner**'s output is provided to the reasoner along with classifier's cross validation.
- **Reasoner** produces a prior belief for the planner.
- **Planner** suggests actions to enable the robot to actively interact with the world, and determines what actions to take and when to terminate.



Human Intention Estimation Problem

The robot is tasked to estimate human intention of interaction with as early and accurately as possible.

LSTM-based supervised learning

- Labeled human motion trajectories
- Dataset collected in a shopping mall using 3D range sensors

Automated reasoner

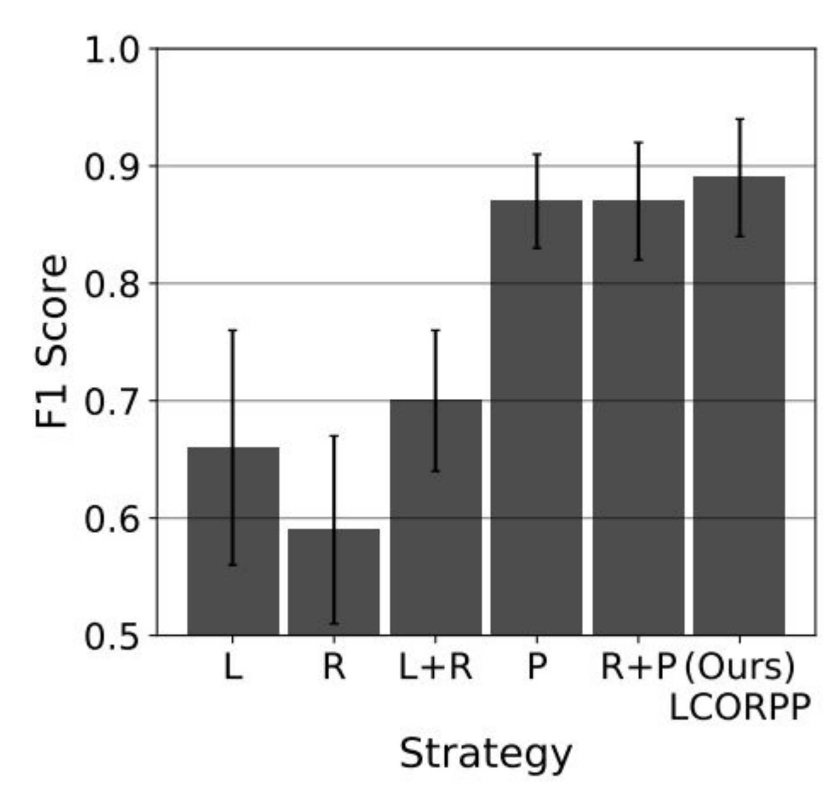
- Infers human intention based on time and location of interaction (e.g. visitors more likely to interact with robot)

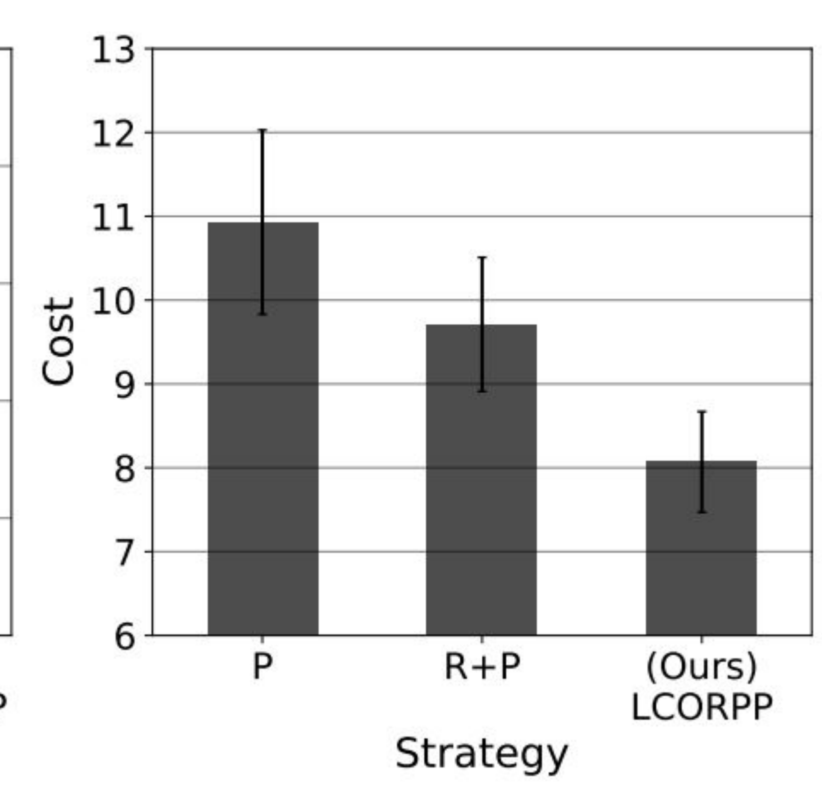
POMDP-based planner

- Motion-based actions such as "turn", and "move forward"
- Language-based actions such as "greet" (May I help you?)

Experiments

In comparison to five baselines, LCORPP produces the highest F1 score, while reducing interaction costs





References

- Zhang S, Stone P. CORPP: Commonsense Reasoning and Probabilistic Planning, as Applied to Dialog with a Mobile Robot. AAAI 2015
- Kato Y, Kanda T, Ishiguro H. May i help you?: Design of human-like polite approaching behavior. HRI 2015

This work has taken place in the Autonomous Intelligent Robotics (AIR) Group at SUNY Binghamton. AIR research is supported in part by grants from the National Science Foundation (IIS-1925044), Ford Motor Company, and SUNY Research Foundation.