HRI Reading Group

@ Instituto Superior Técnico

Spring 2019

Meeting #3 (1 Mar 2019)

Paper

R. Choudhury, G. Swarmy, D. Hadfield-Menell, and A.D. Dragan. **On the utility of model learning in HRI.** International Conference on Human-Robot Interaction (HRI), 2019. (to appear)

What is a model?

Model class? Model instance? Learn a model vs. learn model parameters? Parameters versus hyperparameters? Non-parametric models?

Model definition: A mathematical / statistical description of a system

"All models are wrong, but some are useful"

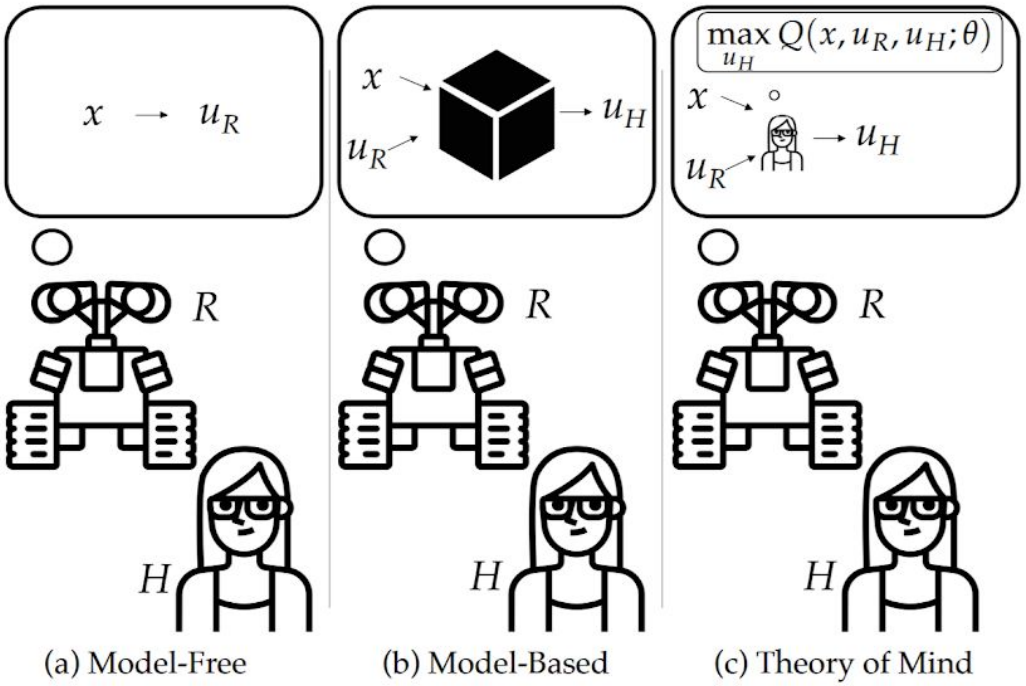
What does a model do? Predict, explain, represent, produce

Learning a model? Equivalent to “fitting through optimization”

Model-free vs. Model-based (I)RL? Known facts (?)

Model-free		Model-based	
Pros	**Cons**	**Pros**	**Cons**
Doesn't need explicit model of the world	More samples needed	Explainability	Less flexible
Doesn't on any strong assumptions about how the world behaves	Can hardly generalize / transfer	Transferability	Mathematical tool needed
			Representation more expensive

Comparing three paradigms in interactive learning



(a) Model-Free (b) Model-Based (c) Theory of Mind

Illustrated with HRI scenario: autonomous car / human driver interaction

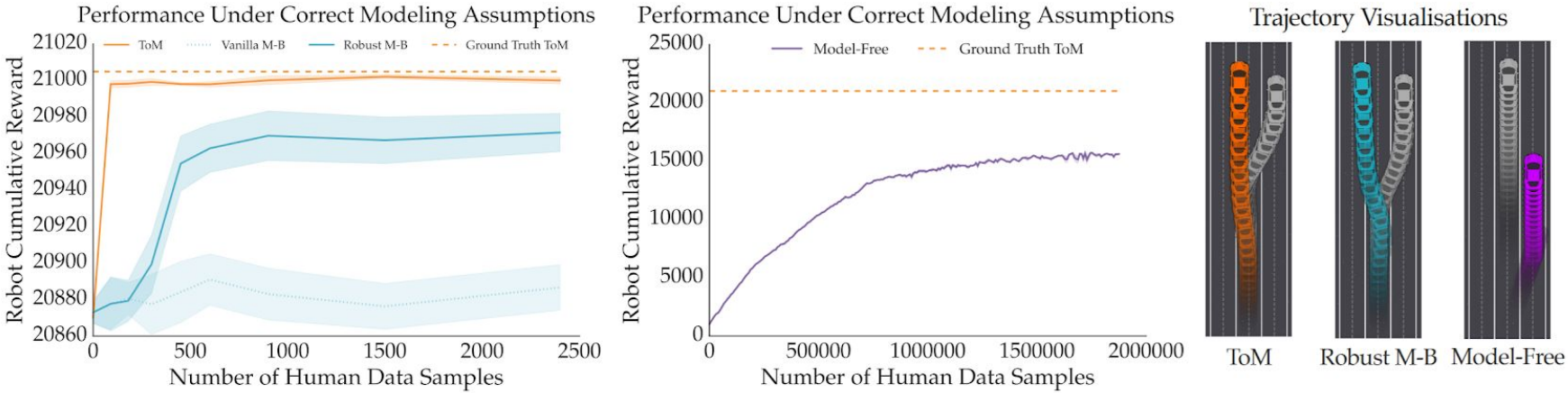


Fig. 2: The test rewards of the interaction learning algorithms on the scenario with the ground truth human simulator. The ToM learner has the smallest sample complexity and best performance, followed by the covariate-shift robust model based method. The 'vanilla' model-based method does poorly. The ToM is able to pass the human car with the least movement out of its lane, and thus obtains the highest reward.

"**Takeaways.** Overall, what we find confirms intuition: if we have a good model, learning its parameters leads to good performance compared to learning from scratch. More surprising is the poor performance of the vanilla model-based method: to get model-based methods to work, it seems like they need to be interactive. What this says is that we might not be able to use black-box models learned based on human-human interaction data: we might need human-robot interaction data, and in particular data obtained from interaction as the robot is still learning. This can be prohibitively expensive or dangerous in many scenarios."

Perturbing the assumptions

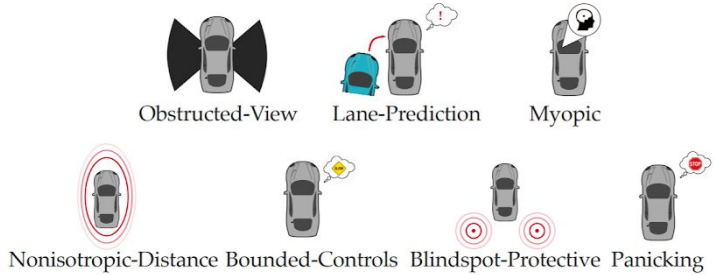


Fig. 3: The various modifications made to the ground truth human simulator. The first row corresponds to modification of planning methodology, while the first three elements of the second row correspond to changes in reward features. The last corresponds to an irrational planning heuristic humans might use while under pressure.

"**Takeaways.** We establish that there is a tipping point where ToM switches from being robust to being unable to model the human. Before this tipping point, ToM remains superior. At this tipping point, model-based eventually surpasses ToM. Past this tipping point, ToM is drastically inferior. Surprisingly, even some large inaccuracies in ToM fail to harm it enough, especially in low-data regimes." Also, ToM seems to be more transferable across the board, even on situations where its assumptions are dramatically different from reality (so different that if model-based were to be re-trained, it would vastly surpass it). This is again explained by its resiliency to covariate shift

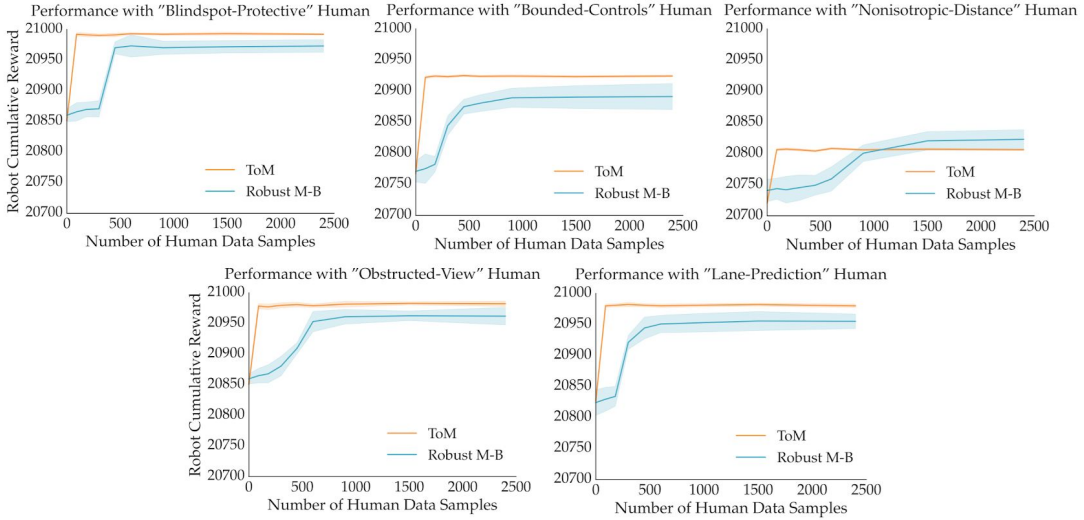


Fig. 4: ToM vs model-based on different simulators. ToM is robust to simulator modifications in some cases but is eventually surpassed by model-based when the difference between assumptions and reality is sufficiently large.

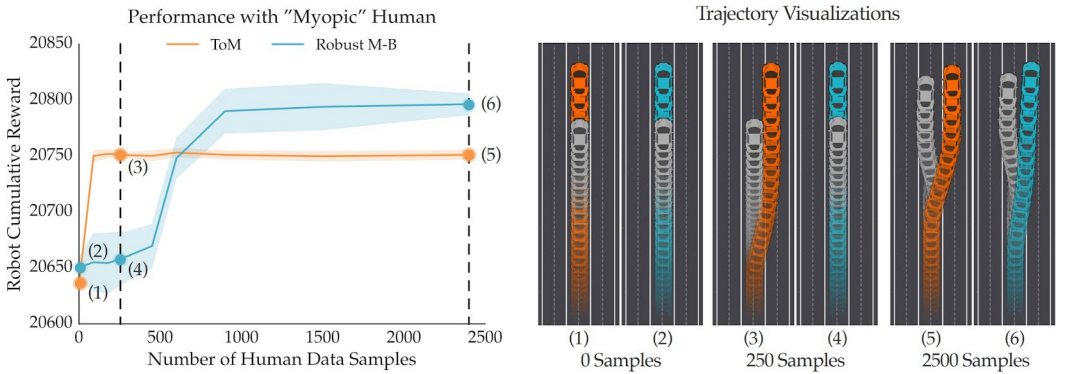


Fig. 5: The cumulative reward and taken trajectories of the interaction learning algorithms with the myopic human simulator. The ToM learner performs better in low-data regimes but the robust model-based method is able to eventually outperform the other method.

What practical conclusions can we draw from this work? What are other scenarios were those results may be relevant?

- Confirms general intuition
- Having a good model will take you a long way
- In the absence of a good model: if small disturbance, no big deal; if large disturbance: may switch to a different model-based learning mode

Group exercise

Come up with two real-world examples where you think each of the following would work best:

Model-free, Model-based (Black Box), ToM (Gray Box)

[No time to do it : Homework]

HRI Reading Group

@ Instituto Superior Técnico

Meeting #4 (8 March 2019)