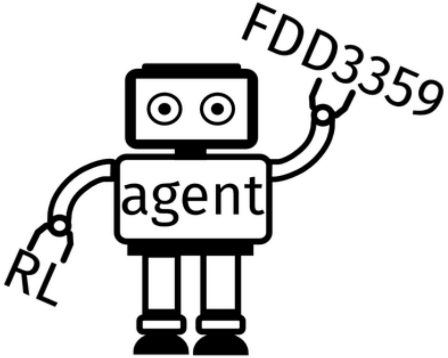
KTH ROYAL INSTITUTE OF TECHNOLOGY



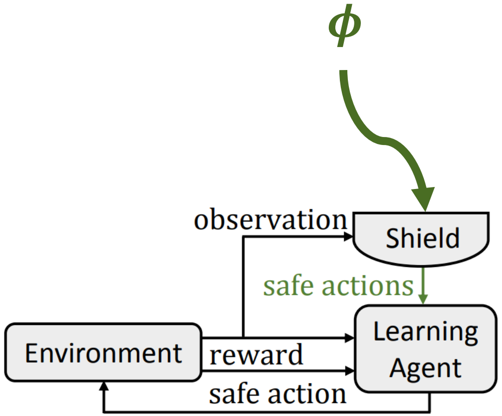
Reinforcement Learning

and Human-Robot Interaction

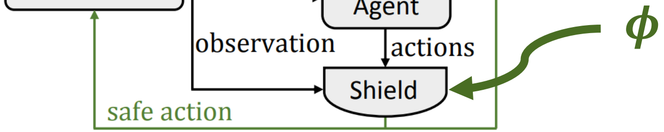
FDD3359

Alexis Linard

Safety?

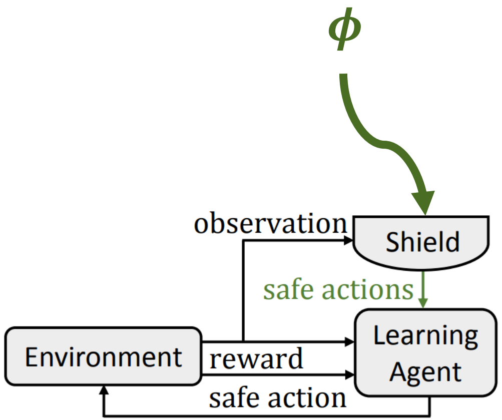


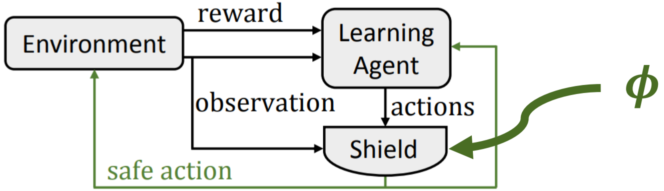
ϕ
observation
Shield
safe actions
Environment
reward
Learning Agent
safe action



Environment
reward
Learning Agent
observation
actions
ϕ
Shield
safe action

Safety?





What does that safety specification ϕ encapsulates?

How do end-users define safety?

How to include humans in RL`?

Intended Learning Outcomes

- By the end of this session, you should be able to:
  - Understand the difference between objective safety, normative safety and perceived safety
  - Know how to apply human-related notions of safety in your RL loop

- Acknowledgements:
  - Some figures taken from literature (cited along the slides)
  - Ilaria Torre's lecture on Perceived Safety (Safe Autonomy course)
  - Talks of Dorsa Sadigh and Mohamed Chetouani at the RL-CONFORM workshop

[1] Talk of Dorsa Sadigh, Stanford University, USA, at the RL-CONFORM workshop: "Bringing in the Human in the (Reinforcement) Learning Loop". https://youtu.be/Pw2bAM4ykdY
[2] Talk of Mohamed Chetouani, Sorbonne University, France, at the RL-CONFORM workshop: "Socially Interactive Learning: On the interpretation of human teaching signals". https://youtu.be/gWvZKh6suhg

HRI in RL

- Safety vs. Perceived Safety
- Bringing the human in the RL loop

HRI in RL

- **Safety vs. Perceived Safety**
- Bringing the human in the RL loop

Different types of safety?

Objective safety refers to the safety-related history in real-world situations.

Perceived safety refers to an individual’s perception of risk.

Normative safety follows best practices and standards to reduce risks.

Normative safety is the engineer’s attempt to achieve a goal, but weather or not it worked can only be assessed through objective and perceived safety.

But to increase the safety of a product, one needs to understand, measure, and eliminate perceived and objective risks.

“Objective” safety

- Objective safety relates to object related quantifications of real-world risks.
- Depending on the technical system, they can be more or less complex.

Risk	Normative safety	Objective safety
Collision avoidance	Add new sensor system	count accidents with and without that sensor
Car’s remote interface	Encrypted communication	Compare reported issues between cars with and without remote interface
Autonomous vehicle runs out of gas or electricity	Improved path planning	How many cars need towing service

Safety with humans

- Many systems are intended for human interaction / use
- Humans are complicated / unpredictable
- They don’t behave rationally / optimally
  - E.g. the Dictator Game
  - E.g. what is potentially more complex:
    autonomous car + autonomous car vs.
    autonomous car + pedestrian vs.
    pedestrian + pedestrian?

[3] Rothenbücher, D., Li, J., Sirkin, D., Mok, B., & Ju, W. (2016). Ghost driver: A field study investigating the interaction between pedestrians and driverless vehicles. In 2016 25th IEEE international symposium on robot and human interactive communication (RO-MAN), 795-802.
[4] Liu, P., Du, Y., Wang, L., & Da Young, J. (2020). Ready to bully automated vehicles on public roads?. Accident Analysis & Prevention, 137, 105457.

Perceived safety

Perceived safety relates to individual's perceived risks.

Risk	Normative safety	Perceived safety
Collision avoidance	Add new sensor system	People believe the system: No unnecessary alarm but simple tests work.
Car's remote interface	Encrypted communication	People's observation: Does the car always do exactly what's expected?
Autonomous vehicle runs out of gas or electricity	Improved path planning	Communication: Transparently show current fuel level.

Perceived safety relates to subject's assessment of risks.

Depending on the technical system, they can be more or less complex but are most often more difficult to assess than objective safety.

Human-related factors affecting perceived safety

Attitudes towards technology

Previous experience

Age

Cultural background

Gender

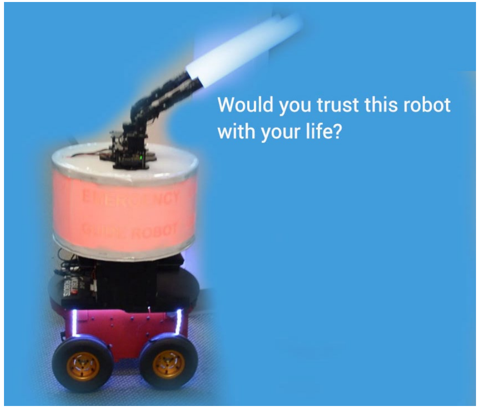
Knowledge

Propensity to trust

[5] Xu, Z., Zhang, K., Min, H., Wang, Z., Zhao, X., & Liu, P. (2018). What drives people to accept automated vehicles? Findings from a field experiment. Transportation research part C: emerging technologies, 95, 320-334.
[6] Moody, J., Bailey, N., & Zhao, J. (2020). Public perceptions of autonomous vehicle safety: An international comparison. Safety science, 121, 634-650.
[7] Nordhoff, S., Stapel, J., van Arem, B., & Happee, R. (2020). Passenger opinions of the perceived safety and interaction with automated shuttles: A test ride study with 'hidden'safety steward. Transportation research part A: policy and practice, 138, 508-524.

Related notion: Trustworthiness

- Trustworthiness involves at least two interacting members:
  - The trustor (human person)
  - The trustee (the machine/system/computer)



- **Trust**: attitude where the users are confident in that the system will help them to achieve a goal.
  - Helps to overcome risk and uncertainty
  - But how can we model/measure trust?

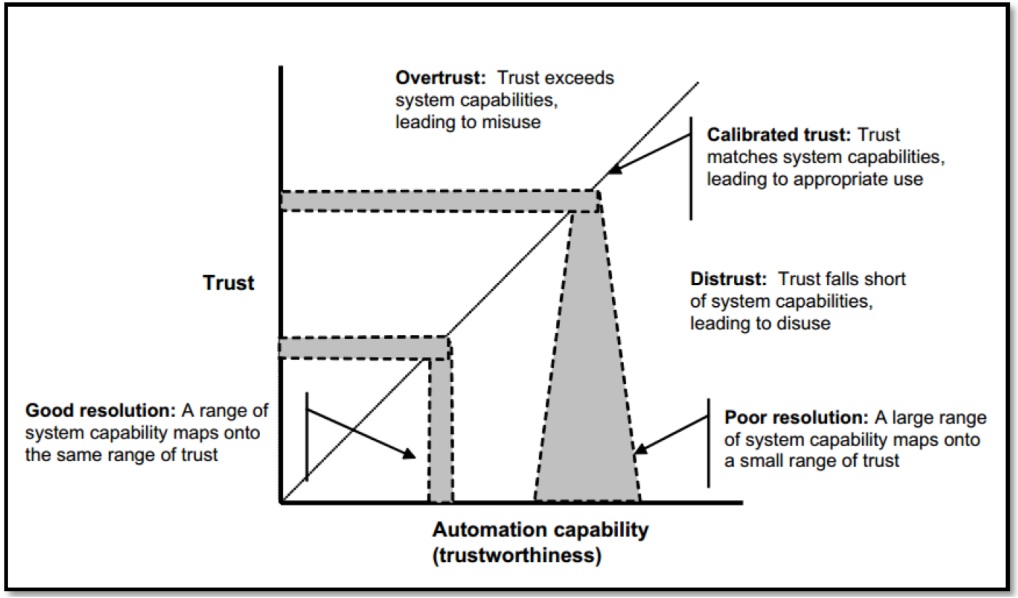
Picture: https://davidamerland.com/seo-blog/1075-humans-trust-robots-more-than-people.html

Related notion: Trustworthiness

- Calibration of trust with system's capacity
  - Overtrust
  - Distrust
  - Calibrated trust

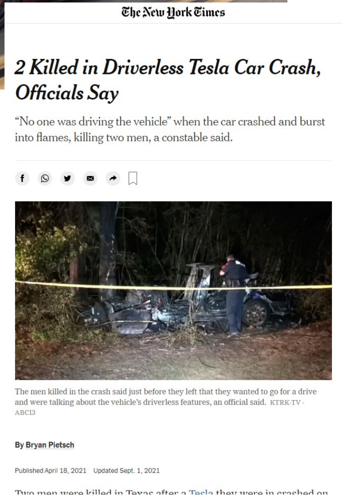
- Identify features influencing trust

- Trust scales



Picture: Schaefer, Kristin. "The perception and measurement of human-robot trust." (2013).

Related notion: Trustworthiness



The New York Times

2 Killed in Driverless Tesla Car Crash, Officials Say

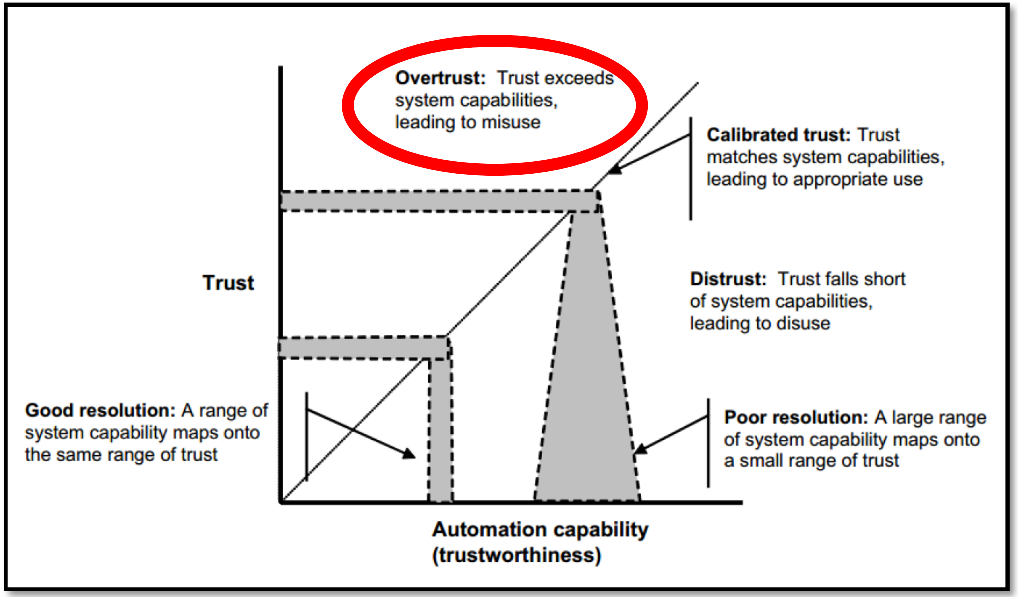
"No one was driving the vehicle" when the car crashed and burst into flames, killing two men, a constable said.

The men killed in the crash said just before they left that they wanted to go for a drive and were talking about the vehicle's driverless features, an official said. KTRK-TV - ABC13

By Bryan Pietsch

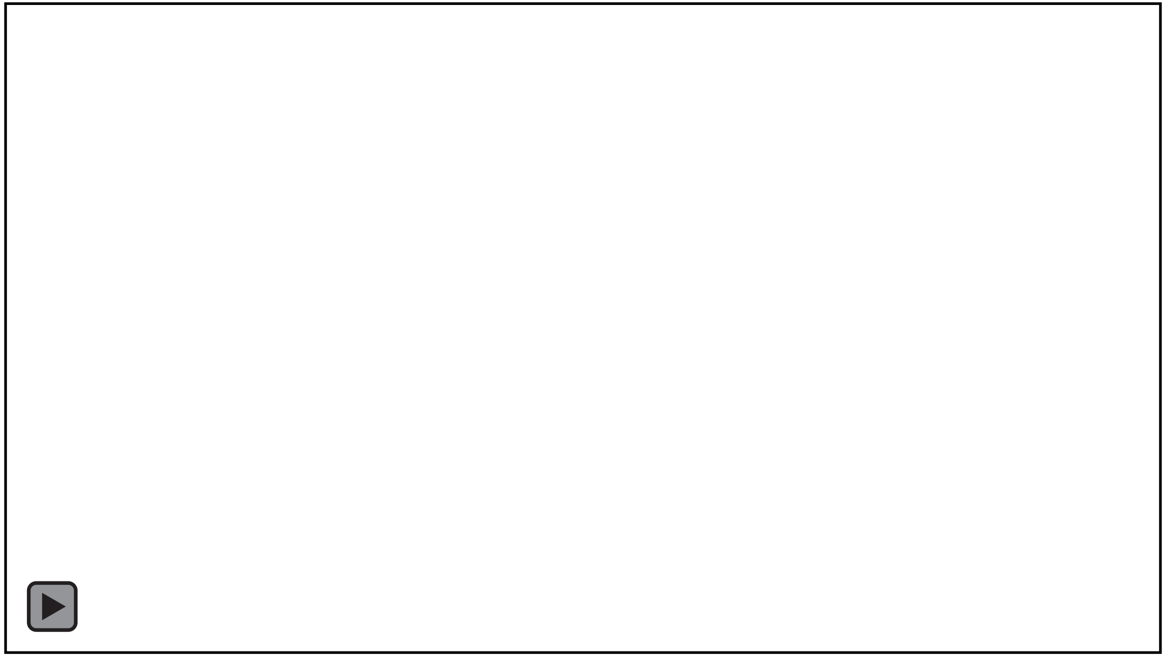
Published April 18, 2021 Updated Sept. 1, 2021

Two men were killed in Texas after a Tesla they were in crashed on



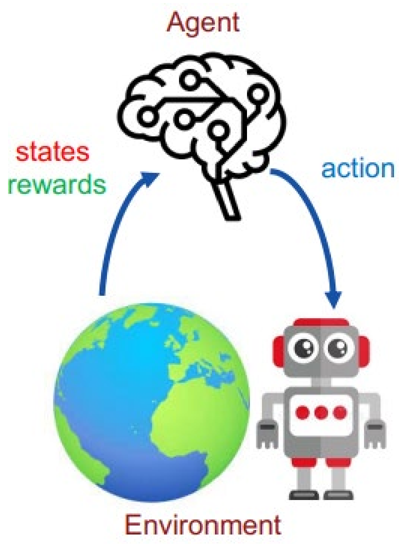
Picture: https://www.businessinsider.com/tesla-future-models-cybertruck-roadster-semi-release-date-price-2021-3?r=US&IR=T
NYT article: https://www.nytimes.com/2021/04/18/business/tesla-fatal-crash-texas.html

An example of overtrust

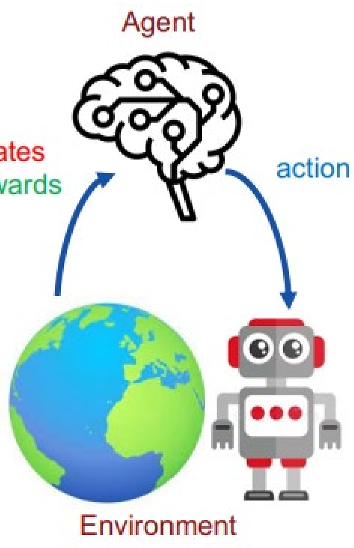


Video: Volvo auto brake system fail? https://youtu.be/_47utWAoupo

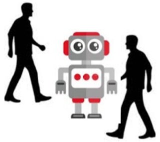
How to measure Perceived Safety/Trust?



How to measure Perceived Safety/Trust?



perceived safety



$r(s, a, s') = ?$

Explicit vs. Implicit measures

- Explicit measures relate to conscious impressions that people typically have time to reflect on
  - E.g. questionnaire: “From 1 to 7, how safe did you feel…?”
  - E.g. interview: “Why did you feel safe…?”
  - E.g. informal brainstorming: “Would you or your grandmother use this feature if we added it to the product?”
- Implicit measures refer to unconscious attitudes
  - E.g. reaction times: How quickly does someone click on the word “safe” after seeing a picture of a robot? (IAT)
  - E.g. behaviour: Are people smiling / looking uncomfortable?
  - E.g. engagement: Are people using a product’s feature at all? Are people willing to interact with the system?

Explicit vs. Implicit measures

- Keep in mind that unfortunately attitudes (survey responses) don't always correlate with behaviour!
- There is a difference between words and actions.
- People say they would 100% trust a system ≠ people actually trusting the system

[8] Greenwald, A. G.(1990). What cognitive representations underlie social attitudes? Bulletin of the Psychonomic Society, 28, 254-260.
[9] Greenwald, A. G.; McGhee, D. E. & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: the implicit association test. Journal of Personality and Social Psychology, 74, 1464.

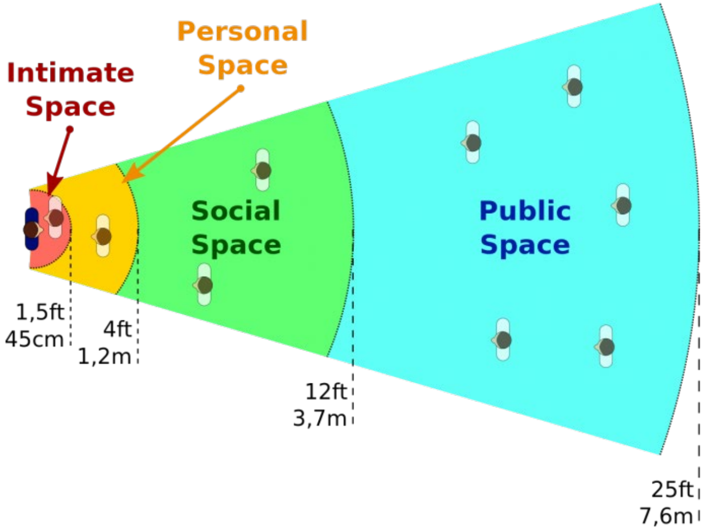
Practical tool: the CTAM

- It’s better to use a “validated” questionnaire, and not make up your own
- Car Technology Acceptance Model (CTAM)
  - Low number of participants
  - Only studied in the USA
  - Driver’s point of view
- Items reflecting performance expectancy, effort expectancy, attitudes towards technology, social influence, facilitating conditions, self efficacy, anxiety, behavioural intention to use the system, perceived safety

[10] Osswald, S., Wurhofer, D., Trösterer, S., Beck, E., & Tscheligi, M. (2012). Predicting information technology usage in the car: towards a car technology acceptance model.
In Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, 51-58.

Practical tool: Proxemics

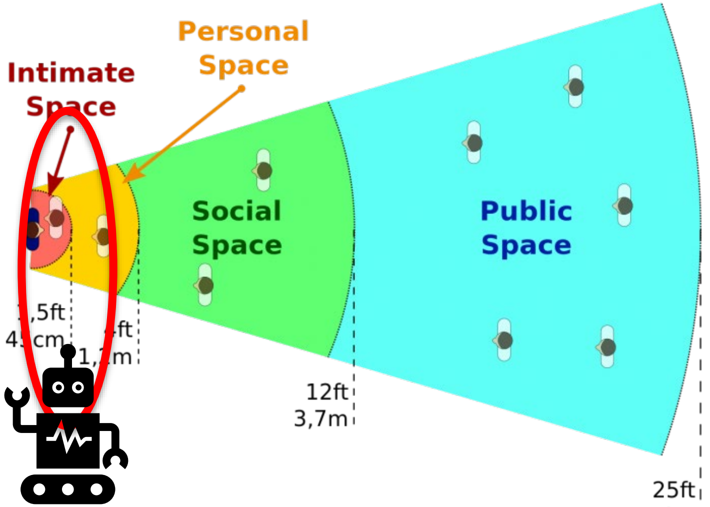
- The personal space that people maintain around themselves



[12] Hall, E. T. (1966). The hidden dimension. Vol. 609. Garden City, NY: Doubleday.
[13] Mumm, J. and Mutlu, B. (2011). Human-robot proxemics: physical and psychological distancing in human-robot interaction. In Proceedings of the 6th ACM/IEEE International conference on Human-Robot Interaction, 331–338

Practical tool: Proxemics

- Distance people keep from a robot = distance a robot should keep from a human



[14] Satake, S., Kanda, T., Glas, D. F., Imai, M., Ishiguro, H., & Hagita, N. (2009). How to approach humans? Strategies for social robots to initiate interaction. In Proceedings of the 4th ACM/IEEE international conference on Human robot interaction, 109-116.
[15] Senft, E., Satake, S., & Kanda, T. (2020). Would You Mind Me if I Pass by You? Socially-Appropriate Behaviour for an Omni-based Social Robot in Narrow Environment. In Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction, 539-547.
[16] Truong, X. T., & Ngo, T. D. (2017). "To approach humans?": A unified framework for approaching pose prediction and socially aware robot navigation. IEEE Transactions on Cognitive and Developmental Systems, 10(3), 557-572.

Practical tool: Proxemics

- But distances vary based on robot type
  - E.g. anthropomorphic / machine-like / zoomorphic

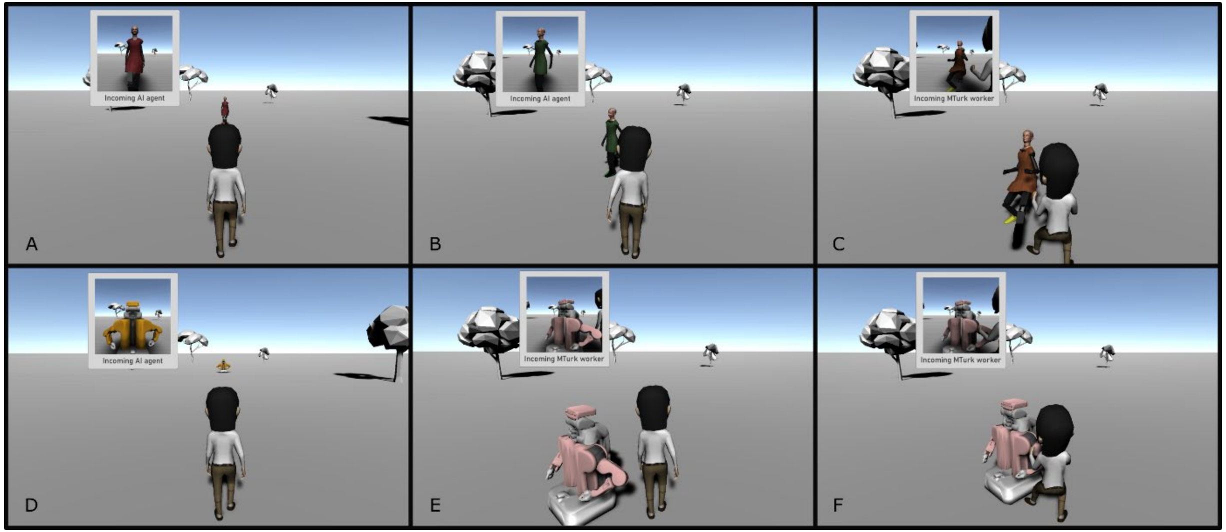
- On task type
  - E.g. communicative / cooperative

- On presence
  - E.g. physical preference / telepresence

[17] Hüttenrauch, H., Eklundh, K. S., Green, A., & Topp, E. A. (2006). Investigating spatial relationships in human-robot interaction. In 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, 5052-5059.
[18] Walters, M. L., Syrdal, D. S., Koay, K. L., Dautenhahn, K., & Te Boekhorst, R. (2008). Human approach distances to a mechanical-looking robot with different robot voice styles. In RO-MAN 2008-The 17th IEEE International Symposium on Robot and Human Interactive Communication, 707-712.

Practical tool: Game of chicken

- We used it to measure when robots should swerve

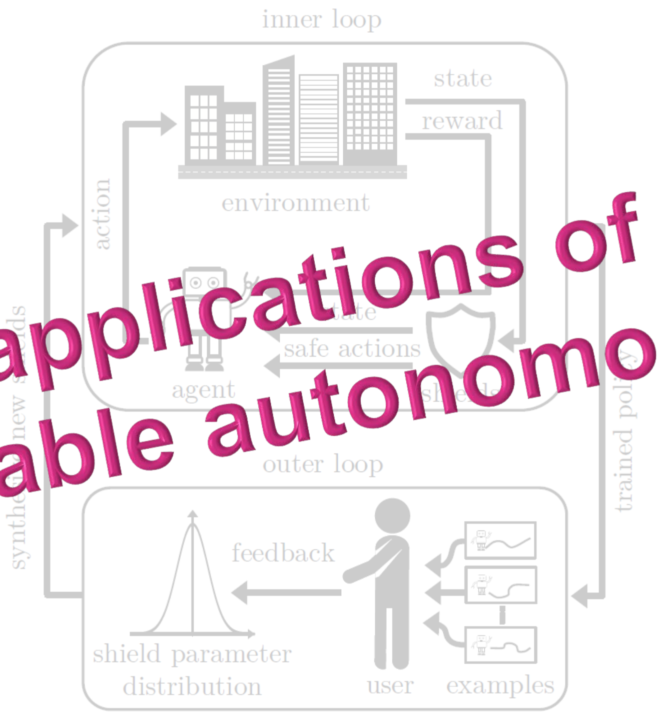


[19] Torre, I., Linard, A., Steen, A., Tumová, J. & Leite, I. (2021). Should robots chicken? How anthropomorphism and perceived autonomy influence trajectories in a game-theoretic problem. In Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (pp. 370-379).

Applications of Perceived Safety in RL

- Including perceived safety in RL:
  - In the reward function?
  - As shield?

Applications of Perceived Safety in RL

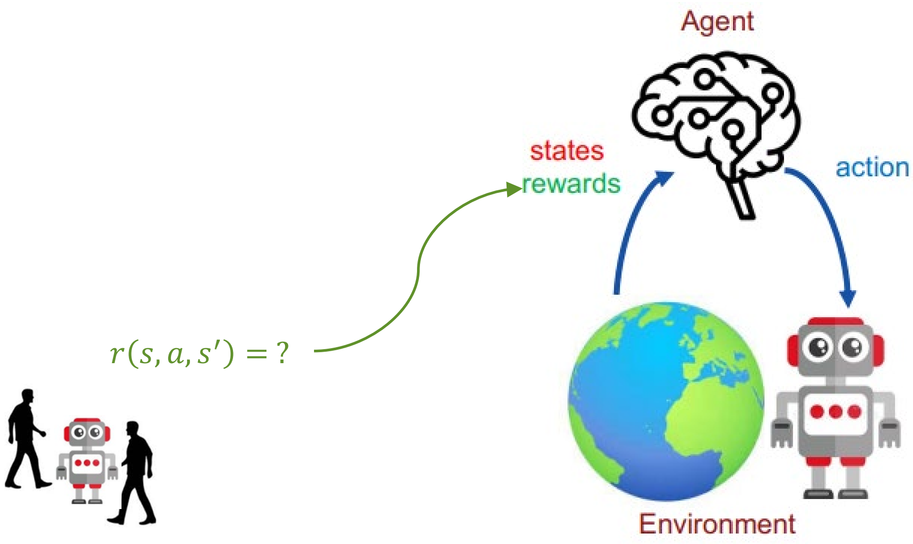


[20] Daniel Marta, Christian Pek, Gaspar Isaac, Melsión, Jana Tumova, Iolanda Leite, Human-Feedback Shield Synthesis for Perceived Safety in Deep Reinforcement Learning. IEEE Robotics and Automation Letters 7.1 (2021): 406-413.

HRI in RL

- Safety vs. Perceived Safety
- **Bringing the human in the RL loop**

Inverse Reinforcement Learning



[21] Adams, S., Cody, T., & Beling, P. A. (2022). A survey of inverse reinforcement learning. Artificial Intelligence Review, 1-40.

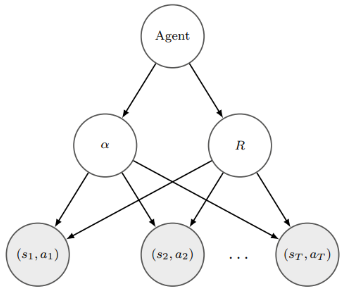
Inverse Reinforcement Learning

- max-margin methods

$$
\begin{aligned}
\max_{\alpha,t} \quad & t \\
\text{s.t.} \quad & \alpha^\intercal \mu_E \geq \alpha^\intercal \mu^{(j)} + t, \quad j = 0, ..., i-1 \\
& ||\alpha||_2 \leq 1,
\end{aligned}
$$

- Bayesian methods

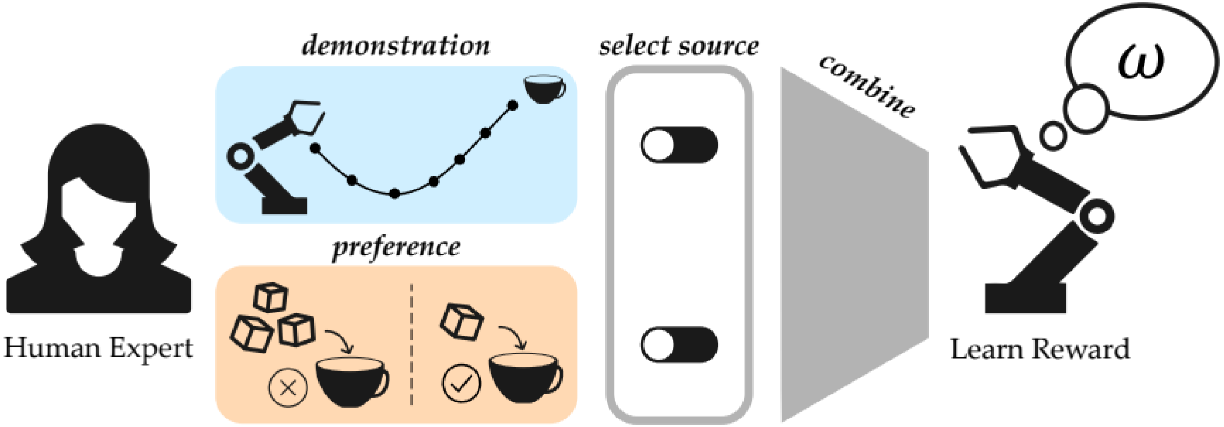
$$\mathbb{P}(R|\mathcal{T}) = \frac{1}{Z'} e^{\alpha E_R(\mathcal{T})} \mathbb{P}(R).$$



- Maximum entropy methods

[21] Adams, S., Cody, T., & Beling, P. A. (2022). A survey of inverse reinforcement learning. Artificial Intelligence Review, 1-40.

Learning from Demonstrations vs. from Preferences



[23] Bıyık, E., Losey, D. P., Palan, M., Landolfi, N. C., Shevchuk, G., & Sadigh, D. (2021). Learning reward functions from diverse sources of human feedback: Optimally integrating demonstrations and preferences. The International Journal of Robotics Research.

Learning from Human Demonstrations

[22] Palan, M., Landolfi, N. C., Shevchuk, G., & Sadigh, D. (2019). Learning reward functions by integrating human demonstrations and preferences. arXiv preprint arXiv:1906.08928.
[23] Bıyık, E., Losey, D. P., Palan, M., Landolfi, N. C., Shevchuk, G., & Sadigh, D. (2021). Learning reward functions from diverse sources of human feedback: Optimally integrating demonstrations and preferences. The International Journal of Robotics Research.

Learning from Human Demonstrations

- Demonstrations from human $\{\xi_1^D, \ldots, \xi_n^D\}$
- Assume that there exists a reward function

$$R_H(\xi) = \mathbf{w} \cdot \Phi(\xi) = \mathbf{w} \cdot \sum_{t=0}^{T} \phi(x^t, u_H^t)$$

[22] Palan, M., Landolfi, N. C., Shevchuk, G., & Sadigh, D. (2019). Learning reward functions by integrating human demonstrations and preferences. arXiv preprint arXiv:1906.08928.
[23] Bıyık, E., Losey, D. P., Palan, M., Landolfi, N. C., Shevchuk, G., & Sadigh, D. (2021). Learning reward functions from diverse sources of human feedback: Optimally integrating demonstrations and preferences. The International Journal of Robotics Research.

Learning from Human Demonstrations

- Demonstrations from human $\{\xi_1^D, \ldots, \xi_n^D\}$
- Assume that there exists a reward function

$$R_H(\xi) = \mathbf{w} \cdot \Phi(\xi) = \mathbf{w} \cdot \sum_{t=0}^{T} \phi(x^t, u_H^t)$$

- Use Bayesian IRL

$$\mathbb{P}(\xi_1^D, \ldots, \xi_n^D \mid \mathbf{w}) = \prod_{i=1}^{n} \mathbb{P}(\xi_i^D \mid \mathbf{w})$$

[22] Palan, M., Landolfi, N. C., Shevchuk, G., & Sadigh, D. (2019). Learning reward functions by integrating human demonstrations and preferences. arXiv preprint arXiv:1906.08928.
[23] Bıyık, E., Losey, D. P., Palan, M., Landolfi, N. C., Shevchuk, G., & Sadigh, D. (2021). Learning reward functions from diverse sources of human feedback: Optimally integrating demonstrations and preferences. The International Journal of Robotics Research.

Learning from Human Demonstrations

- Demonstrations from human $\{\xi_1^D, \ldots, \xi_n^D\}$
- Assume that there exists a reward function

$$R_H(\xi) = \mathbf{w} \cdot \Phi(\xi) = \mathbf{w} \cdot \sum_{t=0}^{T} \phi(x^t, u_H^t)$$

- Use Bayesian IRL

$$\mathbb{P}(\xi_1^D, \ldots, \xi_n^D \mid \mathbf{w}) = \prod_{i=1}^{n} \mathbb{P}(\xi_i^D \mid \mathbf{w})$$

- Account for human noisiness

$$\mathbb{P}(\xi^D \mid \mathbf{w}) \propto \exp(\beta^D \mathbf{w} \cdot \Phi(\xi^D))$$

[22] Palan, M., Landolfi, N. C., Shevchuk, G., & Sadigh, D. (2019). Learning reward functions by integrating human demonstrations and preferences. arXiv preprint arXiv:1906.08928.
[23] Bıyık, E., Losey, D. P., Palan, M., Landolfi, N. C., Shevchuk, G., & Sadigh, D. (2021). Learning reward functions from diverse sources of human feedback: Optimally integrating demonstrations and preferences. The International Journal of Robotics Research.

Learning from Human Demonstrations

- Demonstrations from human $\{\xi_1^D, \ldots, \xi_n^D\}$
- Assume that there exists a reward function

$$R_H(\xi) = \mathbf{w} \cdot \Phi(\xi) = \mathbf{w} \cdot \sum_{t=0}^{T} \phi(x^t, u_H^t)$$

- Use Bayesian IRL

$$\mathbb{P}(\xi_1^D, \ldots, \xi_n^D \mid \mathbf{w}) = \prod_{i=1}^{n} \mathbb{P}(\xi_i^D \mid \mathbf{w})$$

- Account for human noisiness

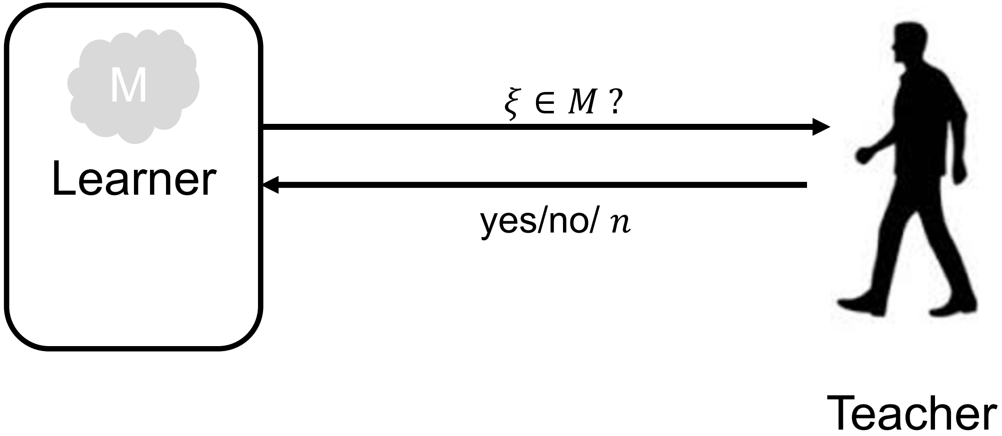
$$\mathbb{P}(\xi^D \mid \mathbf{w}) \propto \exp(\beta^D \mathbf{w} \cdot \Phi(\xi^D))$$

- Bayesian update

$$\mathbb{P}(\mathbf{w} \mid \xi_1^D, \ldots, \xi_n^D) \propto \exp\left(\beta^D \sum_{i=1}^{n} \mathbf{w} \cdot \Phi(\xi_i^D)\right)$$

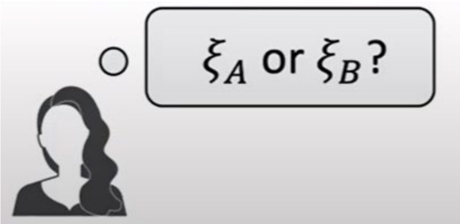
[22] Palan, M., Landolfi, N. C., Shevchuk, G., & Sadigh, D. (2019). Learning reward functions by integrating human demonstrations and preferences. arXiv preprint arXiv:1906.08928.
[23] Bıyık, E., Losey, D. P., Palan, M., Landolfi, N. C., Shevchuk, G., & Sadigh, D. (2021). Learning reward functions from diverse sources of human feedback: Optimally integrating demonstrations and preferences. The International Journal of Robotics Research.

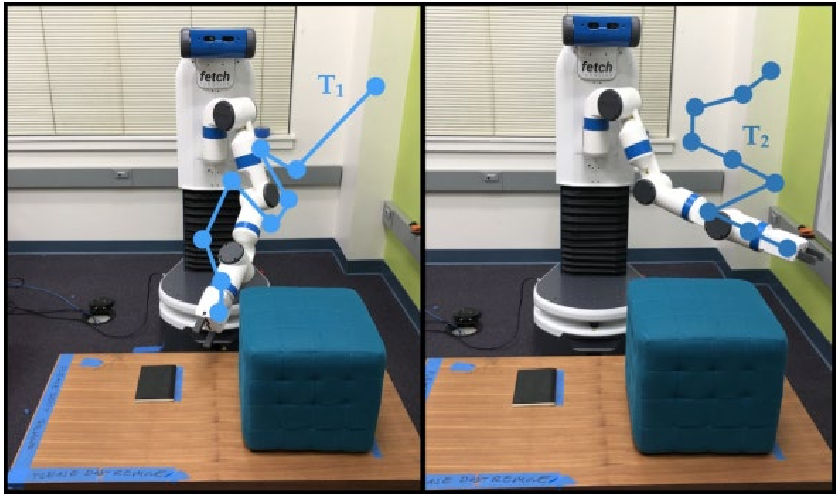
Active Learning of Reward Functions



[24] Sadigh, D., Dragan, A.D., Sastry, S. and Seshia, S.A., 2017. Active preference-based learning of reward functions.
[25] Bıyık, Erdem, et al. "Batch active learning using determinantal point processes." arXiv preprint arXiv:1906.07975 (2019).
[26] Basu, Chandrayee, et al. "Active learning of reward dynamics from hierarchical queries." 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2019.
[27] Biyik, Erdem, et al. "Active Preference-Based Gaussian Process Regression for Reward Learning." Robotics: Science and Systems. 2020.

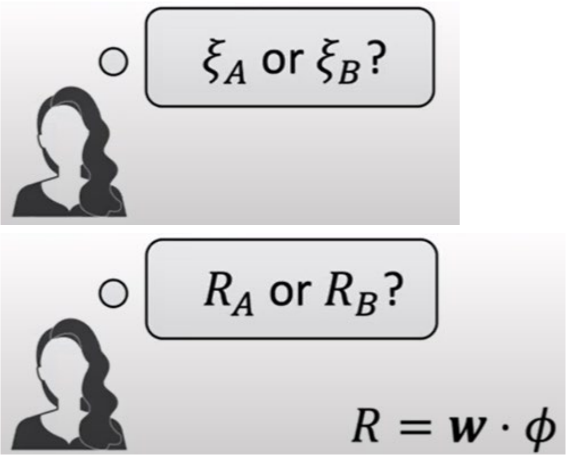
Active Learning of Reward Functions

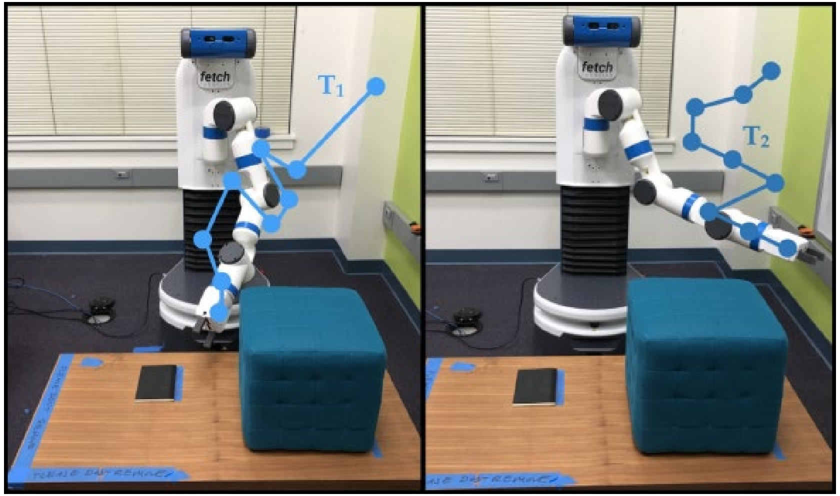




[24] Sadigh, D., Dragan, A.D., Sastry, S. and Seshia, S.A., 2017. Active preference-based learning of reward functions.
[25] Bıyık, Erdem, et al. "Batch active learning using determinantal point processes." arXiv preprint arXiv:1906.07975 (2019).
[26] Basu, Chandrayee, et al. "Active learning of reward dynamics from hierarchical queries." 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2019.
[27] Biyik, Erdem, et al. "Active Preference-Based Gaussian Process Regression for Reward Learning." Robotics: Science and Systems. 2020.

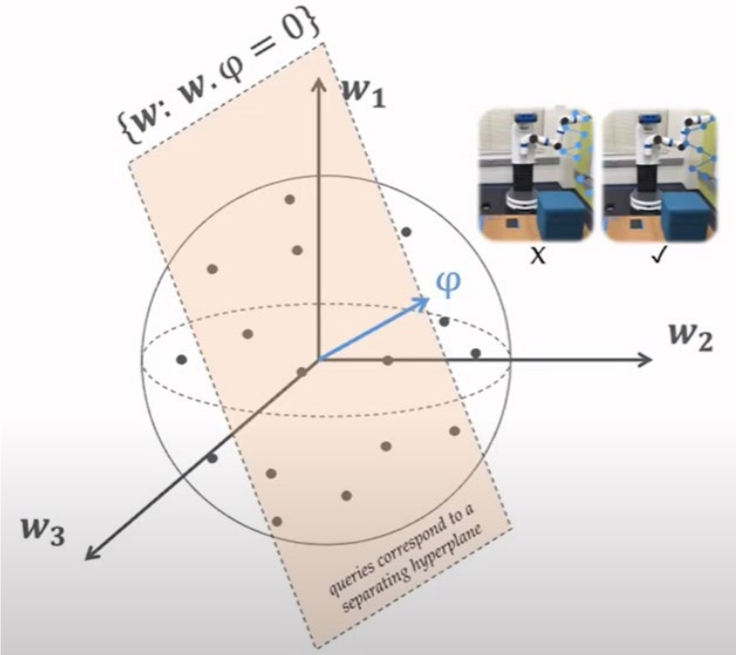
Active Learning of Reward Functions





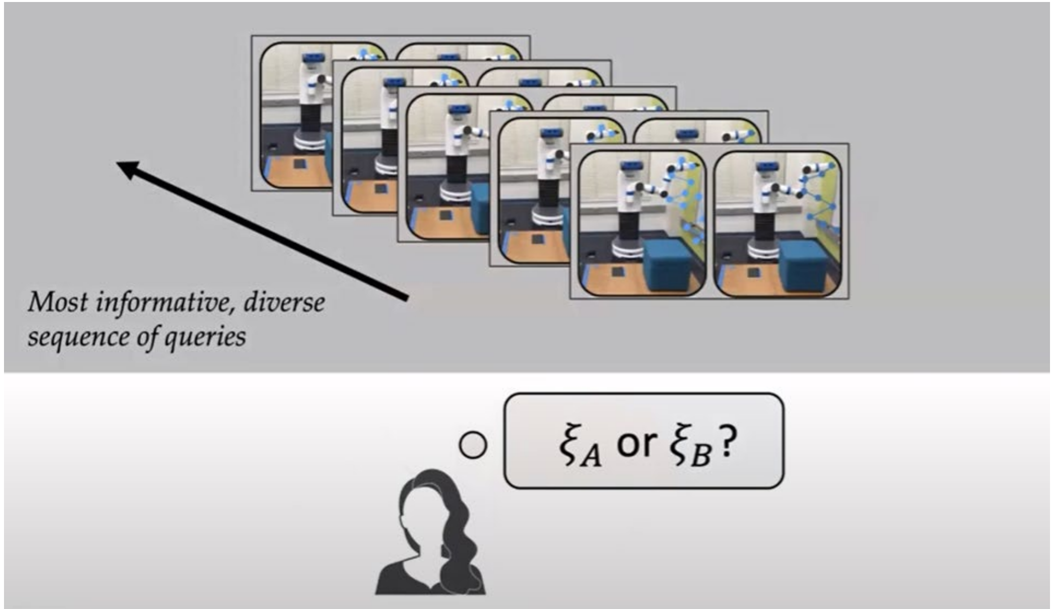
[24] Sadigh, D., Dragan, A.D., Sastry, S. and Seshia, S.A., 2017. Active preference-based learning of reward functions.
[25] Bıyık, Erdem, et al. "Batch active learning using determinantal point processes." arXiv preprint arXiv:1906.07975 (2019).
[26] Basu, Chandrayee, et al. "Active learning of reward dynamics from hierarchical queries." 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2019.
[27] Biyik, Erdem, et al. "Active Preference-Based Gaussian Process Regression for Reward Learning." Robotics: Science and Systems. 2020.

Active Learning of Reward Functions



[24] Sadigh, D., Dragan, A.D., Sastry, S. and Seshia, S.A., 2017. Active preference-based learning of reward functions.
[25] Bıyık, Erdem, et al. "Batch active learning using determinantal point processes." arXiv preprint arXiv:1906.07975 (2019).
[26] Basu, Chandrayee, et al. "Active learning of reward dynamics from hierarchical queries." 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2019.
[27] Biyik, Erdem, et al. "Active Preference-Based Gaussian Process Regression for Reward Learning." Robotics: Science and Systems. 2020.

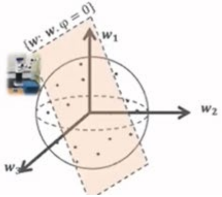
Active Learning of Reward Functions



[24] Sadigh, D., Dragan, A.D., Sastry, S. and Seshia, S.A., 2017. Active preference-based learning of reward functions.
[25] Bıyık, Erdem, et al. "Batch active learning using determinantal point processes." arXiv preprint arXiv:1906.07975 (2019).
[26] Basu, Chandrayee, et al. "Active learning of reward dynamics from hierarchical queries." 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2019.
[27] Biyik, Erdem, et al. "Active Preference-Based Gaussian Process Regression for Reward Learning." Robotics: Science and Systems. 2020.

Synthesizing Queries

Actively synthesizing queries



$$\max_{\varphi} \ \overbrace{\min\{\mathbb{E}[1 - f_{\varphi}(w)], \mathbb{E}[1 - f_{-\varphi}(w)]\}}^{\textit{minimum volume removed}}$$

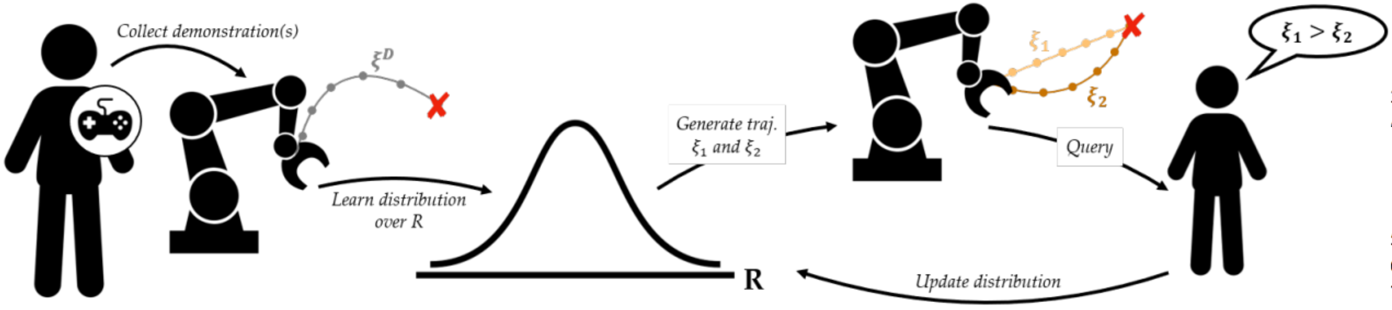
Subject to $\varphi \in \mathbb{F}$

$\mathbb{F} = \{\varphi : \varphi = \Phi(\xi_A) - \Phi(\xi_B), \xi_A, \xi_B \in \Xi\}$

Human update function $f_{\varphi}(\mathbf{w}) = \min(1, \exp(I_t \mathbf{w}^{\top} \varphi))$

[24] Sadigh, D., Dragan, A.D., Sastry, S. and Seshia, S.A., 2017. Active preference-based learning of reward functions.
[25] Bıyık, Erdem, et al. "Batch active learning using determinantal point processes." arXiv preprint arXiv:1906.07975 (2019).
[26] Basu, Chandrayee, et al. "Active learning of reward dynamics from hierarchical queries." 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2019.
[27] Biyik, Erdem, et al. "Active Preference-Based Gaussian Process Regression for Reward Learning." Robotics: Science and Systems. 2020.

Integrating Demonstrations & Preferences



Algorithm 1 DemPref with a Human-in-the-Loop

1: Collect human demonstrations: $\mathcal{D} = \{\xi_1^D, \xi_2^D, \ldots, \xi_L^D\}$

2: Initialize belief over the human's reward weights ω:

$$b^0(\omega) \propto \exp\left(\beta^D \omega \cdot \sum_{\xi^D \in \mathcal{D}} \Phi(\xi^D)\right) P(\omega)$$

3: **for** $i \leftarrow 0, 1, \ldots$ **do**

4: Choose proactive question Q_i:

$$Q_i \leftarrow \arg\max_{Q} \; I(\omega; q \mid Q, b^i) - c(Q)$$

5: **if** $I(\omega; q \mid Q_i, b^i) - c(Q_i) < 0$ **then**

6: **return** b^i

7: **end if**

8: Elicit human's answer q_i to query Q_i

9: Update belief over ω given query and response:

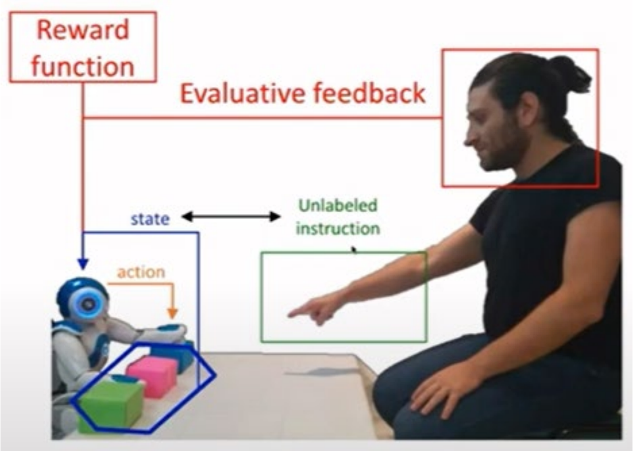
$$b^{i+1}(\omega) \propto P(q_i \mid Q_i, \omega) b^i(\omega)$$

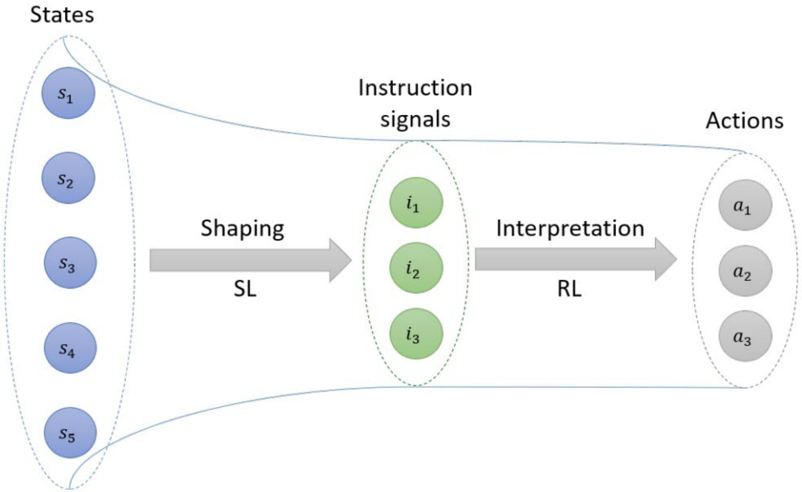
10: **end for**

[23] Bıyık, E., Losey, D. P., Palan, M., Landolfi, N. C., Shevchuk, G., & Sadigh, D. (2021). Learning reward functions from diverse sources of human feedback: Optimally integrating demonstrations and preferences. The International Journal of Robotics Research.

More on HRI and RL

- Dealing with unlabelled human instructions

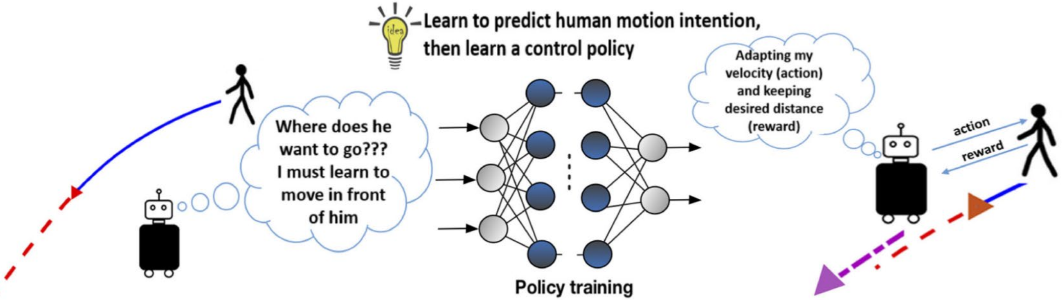


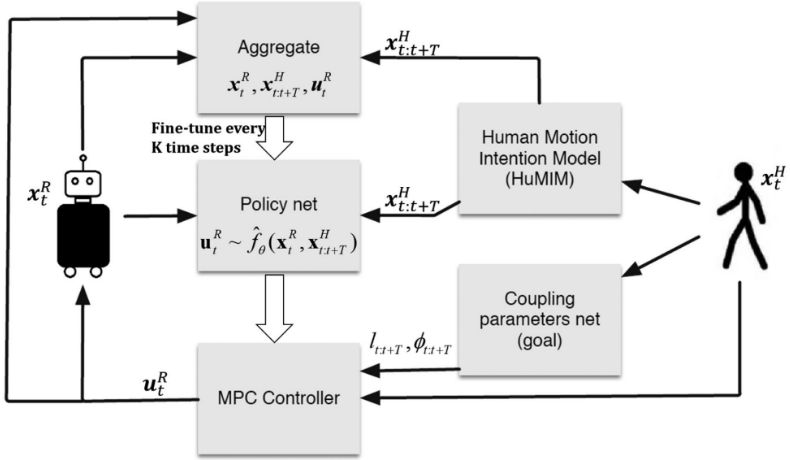


[28] Najar, A., Sigaud, O., & Chetouani, M. (2020). Interactively shaping robot behaviour with unlabeled human instructions. Autonomous Agents and Multi-Agent Systems, 34(2), 1-35.

More on HRI and RL

- Social Navigation





[29] Chalvatzaki, G., Papageorgiou, X. S., Maragos, P., & Tzafestas, C. S. (2019). Learn to adapt to human walking: A model-based reinforcement learning approach for a robotic assistant rollator. IEEE Robotics and Automation Letters, 4(4), 3774-3781.

Conclusion

We covered:

- Safety vs. Human Perceived Safety
- Learning from Human Demonstrations
- Active Learning of Reward Functions

References

[1] Talk of Dorsa Sadigh, Stanford University, USA, at the RL-CONFORM workshop:
"Bringing in the Human in the (Reinforcement) Learning Loop". https://youtu.be/Pw2bAM4ykdY

[2] Talk of Mohamed Chetouani, Sorbonne University, France, at the RL-CONFORM workshop:
"Socially Interactive Learning: On the interpretation of human teaching signals". https://youtu.be/gWvZKh6suhg

[3] Rothenbücher, D., Li, J., Sirkin, D., Mok, B., & Ju, W. (2016). Ghost driver: A field study investigating the interaction between pedestrians and driverless vehicles. In IEEE international symposium on robot and human interactive communication (RO-MAN).

[4] Liu, P., Du, Y., Wang, L., & Da Young, J. (2020). Ready to bully automated vehicles on public roads?.
Accident Analysis & Prevention, 137, 105457.

[5] Xu, Z., Zhang, K., Min, H., Wang, Z., Zhao, X., & Liu, P. (2018). What drives people to accept automated vehicles?
Findings from a field experiment. Transportation research part C: emerging technologies, 95, 320-334.

[6] Moody, J., Bailey, N., & Zhao, J. (2020). Public perceptions of autonomous vehicle safety: An international comparison.
Safety science, 121, 634-650.

[7] Nordhoff, S., Stapel, J., van Arem, B., & Happee, R. (2020). Passenger opinions of the perceived safety and interaction with automated shuttles: A test ride study with 'hidden'safety steward. Transportation research part A: policy and practice, 138, 508-524.

[8] Greenwald, A. G.(1990). What cognitive representations underlie social attitudes?
Bulletin of the Psychonomic Society, 28, 254-260.

References

[9] Greenwald, A. G.; McGhee, D. E. & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: the implicit association test. Journal of Personality and Social Psychology, 74, 1464.

[10] Osswald, S., Wurhofer, D., Trösterer, S., Beck, E., & Tscheligi, M. (2012). Predicting information technology usage in the car: towards a car technology acceptance model. In Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, 51-58.

[11] Deb, S., Strawderman, L., Carruth, D. W., DuBien, J., Smith, B., & Garrison, T. M. (2017). Development and validation of a questionnaire to assess pedestrian receptivity toward fully autonomous vehicles. Transportation research part C: emerging technologies, 84, 178-195.

[12] Hall, E. T. (1966). The hidden dimension. Vol. 609. Garden City, NY: Doubleday.

[13] Mumm, J. and Mutlu, B. (2011). Human-robot proxemics: physical and psychological distancing in human-robot interaction. In Proceedings of the 6th ACM/IEEE International conference on Human-Robot Interaction, 331–338

[14] Satake, S., Kanda, T., Glas, D. F., Imai, M., Ishiguro, H., & Hagita, N. (2009). How to approach humans? Strategies for social robots to initiate interaction. In Proceedings of the International conference on Human robot interaction, 109-116.

[15] Senft, E., Satake, S., & Kanda, T. (2020). Would You Mind Me if I Pass by You? Socially-Appropriate Behaviour for an Omni-based Social Robot in Narrow Environment. In Proceedings of International Conference on Human-Robot Interaction, 539-54

[16] Truong, X. T., & Ngo, T. D. (2017). “To approach humans?”: A unified framework for approaching pose prediction and socially aware robot navigation. IEEE Transactions on Cognitive and Developmental Systems, 10(3), 557-572.

References

[17] Hüttenrauch, H., Eklundh, K. S., Green, A., & Topp, E. A. (2006). Investigating spatial relationships in human-robot interaction. In 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, 5052-5059.

[18] Walters, M. L., Syrdal, D. S., Koay, K. L., Dautenhahn, K., & Te Boekhorst, R. (2008). Human approach distances to a mechanical-looking robot with different robot voice styles.
In RO-MAN 2008-The 17th IEEE International Symposium on Robot and Human Interactive Communication, 707-712.

[19] Torre, I., Linard, A., Steen, A., Tumová, J. & Leite, I. (2021). Should robots chicken? How anthropomorphism and perceived autonomy influence trajectories in a game-theoretic problem.
In Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (pp. 370-379).

[20] Daniel Marta, Christian Pek, Gaspar Isaac, Melsión, Jana Tumova, Iolanda Leite, Human-Feedback Shield Synthesis for Perceived Safety in Deep Reinforcement Learning. IEEE Robotics and Automation Letters 7.1 (2021): 406-413.

[21] Adams, S., Cody, T., & Beling, P. A. (2022). A survey of inverse reinforcement learning. Artificial Intelligence Review, 1-40.

[22] Palan, M., Landolfi, N. C., Shevchuk, G., & Sadigh, D. (2019). Learning reward functions by integrating human demonstrations and preferences. arXiv preprint arXiv:1906.08928.

[23] Bıyık, E., Losey, D. P., Palan, M., Landolfi, N. C., Shevchuk, G., & Sadigh, D. (2021). Learning reward functions from diverse sources of human feedback: Optimally integrating demonstrations and preferences. The International Journal of Robotics Research

[24] Sadigh, D., Dragan, A.D., Sastry, S. and Seshia, S.A., 2017. Active preference-based learning of reward functions.

References

[25] Bıyık, Erdem, et al. "Batch active learning using determinantal point processes." arXiv preprint arXiv:1906.07975 (2019).

[26] Basu, Chandrayee, et al. "Active learning of reward dynamics from hierarchical queries." 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2019.

[27] Biyik, Erdem, et al. "Active Preference-Based Gaussian Process Regression for Reward Learning." Robotics: Science and Systems. 2020.

[28] Najar, A., Sigaud, O., & Chetouani, M. (2020). Interactively shaping robot behaviour with unlabeled human instructions. Autonomous Agents and Multi-Agent Systems, 34(2), 1-35.

[29] Chalvatzaki, G., Papageorgiou, X. S., Maragos, P., & Tzafestas, C. S. (2019). Learn to adapt to human walking: A model-based reinforcement learning approach for a robotic assistant rollator. IEEE Robotics and Automation Letters, 4(4), 3774-3781.