

Research Group Data Mining and Machine Learning

July 23, 2024 // ASAI Networking Event 2024 // **Sebastian Tschitschek**



6 working groups

Database Techniques for Data Mining



Christian Böhm

Machine Learning with Graphs



Nils Kriege

Data Mining



Claudia Plant

Natural Language Processing



Benjamin Roth

Probabilistic and Interactive Machine Learning



Sebastian Tschitschek

Scalable Algorithms for Graph Mining



Yllka Velaj

Probabilistic and Interactive Machine Learning – Motivation

[C. Morrison et al., CHI'21]





Enabling efficient & seamless collaboration
with intelligent agents



Enabling efficient & seamless collaboration
with intelligent agents

Challenge: Collaboration in the face of

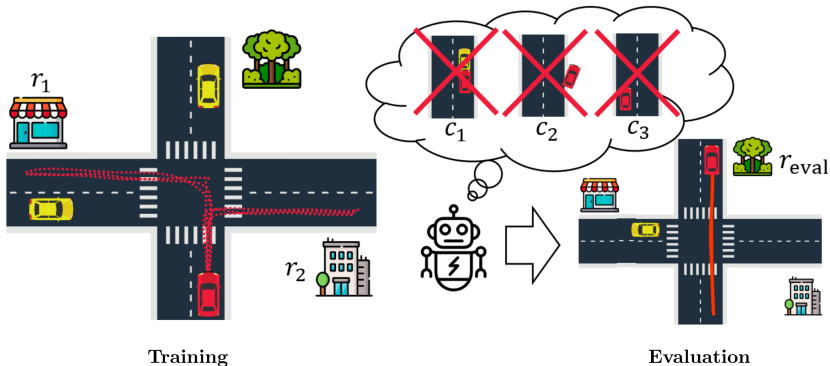
- (significant) mismatch in inputs
- non-aligned goals and constraints
- (complex) large state spaces
- constraints on sample complexity
- inaccurate mental models

3 key research "directions"

- Reinforcement Learning
 - Reward / constraint inference
 - Exploration & abstraction
- Interactive machine learning
- Probabilistic (generative) models

Learning Safety Constraints from Demonstrations with Unknown Constraints

[D. Lindner et al., AISTATS'24]

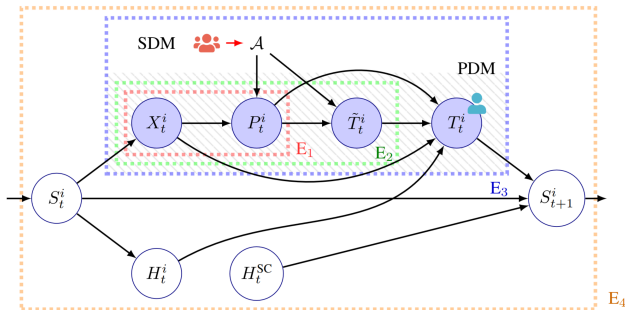


Challenging the Human-in-the-loop in ADM

[S. Tschitschek, E. Stamboliev, M. Coeckelbergh, L. Koesten, HAS@ICML'24]



Der Standard, 10. Februar 2024
Jakob Pflügl, Link



Information That Matters: Information Needs of People Affected by ADM

[T. Schmude et al., in submission]

XAI Novice Question Bank

System context

- ① ② What is the intention of deploying it?
- ① ② How does the deployment process work?
- ① ② What are the consequences after deployment?
- ② ② How is the system developed?
- ① Who is the system's intended target group?
- ① ② Who is responsible for the system's deployment?
- ② ② What do other people think about the system?
- ② ② How are moral values considered in the system?

System usage

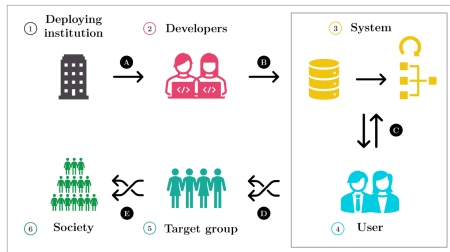
- ① ② How is the system operated?
- ① ② How will the system impact interpersonal relations?
- ① ② How is the system integrated into ex. structures?
- ① ② How can the system be misused?
- ② ② How would the system handle [this case]?
- ② What are the costs?

Data

- ③ What kind of data was the system trained on?*
- ③ What is the source of the training data?*
- ③ Are the data correct / representative / safe?

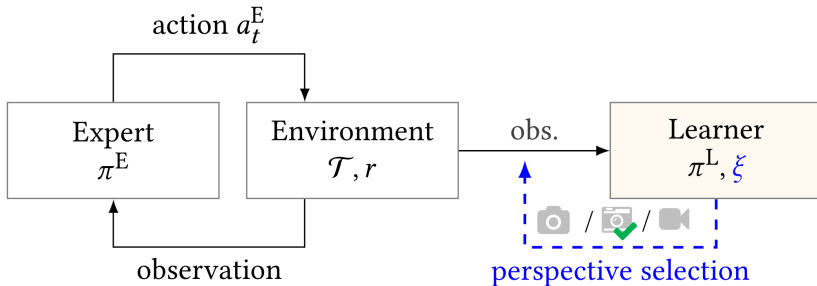
System specifications

- ③ What features does the system consider and why?*
- ③ What is the scope of the system's capability?*
- ③ What kind of output does the system give?*
- ③ How does the system learn?*
- ③ What is the system's overall logic?*
- ③ What kind of algorithm was used?*
- ③ What kind of mistakes is the system likely to make?*
- ③ What are the limitations of the system?*
- ③ How reliable are the predictions?*
- ③ What does [a machine learning concept] mean?



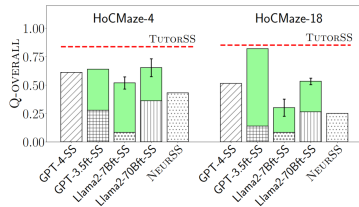
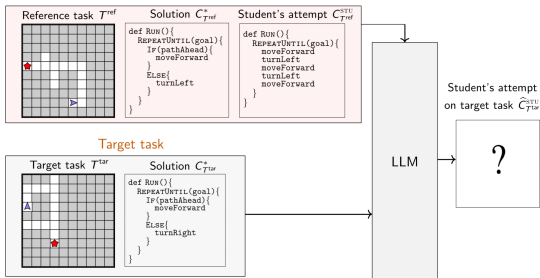
Active Third-person Imitation Learning

[T. Klein et al., in submission]



LLMs for In-Context Student Modeling

[H. Nguyen, S. Tschitschek, A. Singla, EDM'24]



References

- [Sebastian Tschiatschek](#), Eugenia Stamboliev, Mark Coeckelbergh, Laura Koesten, *Challenging the Human-in-the-loop in Algorithmic Decision-making*, Workshop on Humans, Algorithmic Decision-Making and Society @ ICML'24
- David Lindner, Xin Chen, [Sebastian Tschiatschek](#), Katja Hofmann, Andreas Krause, *Learning safety constraints from demonstrations with unknown rewards*, AISTATS'24.
- Timothee Schmude, Laura Koesten, Torsten Möller, [Sebastian Tschiatschek](#). *Information That Matters: Exploring Information Needs of People Affected by Algorithmic Decisions*. arXiv preprint arXiv:2401.13324, 2024.
- Timo Klein, Susanna Weinberger, Adish Singla, [Sebastian Tschiatschek](#), *Active Third-Person Imitation Learning*. arXiv preprint arXiv:2312.16365, 2024.
- Manh Hung Nguyen, [Sebastian Tschiatschek](#), Adish Singla, *Large Language Models for In-Context Student Modeling: Synthesizing Student's Behavior in Visual Programming*, EDM'24

References contd.

- Luis Haug, Sebastian Tschiatschek, Adish Singla, *Teaching inverse reinforcement learners via features and demonstrations*, NeurIPS'18
- Sebastian Tschiatschek, Ahana Gosh, Luis Haug, Rati Devidze, Adish Singla, *Learner-aware Teaching: Inverse Reinforcement Learning with Preferences and Constraints*, NeurIPS'19
- Cecily Morrison, Edward Cutrell, Martin Grayson, Anja Thieme, Alex Taylor, Geert Roumen, Camilla Longden, Sebastian Tschiatschek, Rita Faia Marques, Abigail Sellen, *Social Sensemaking with AI: Designing an Open-ended AI experience with a Blind Child*, CHI'21
- David Lindner, Matteo Turchetta, Sebastian Tschiatschek, Kamil Ciosek, Andreas Krause, *Information Directed Reward Learning for Reinforcement Learning*, NeurIPS'21
- Han Dongge, Sebastian Tschiatschek, *Option Transfer and SMDP Abstraction with Successor Features*, IJCAI'22
- David Linder, Sebastian Tschiatschek, Katja Hofmann, Andreas Krause *Interactively Learning Preference Constraints in Linear Bandits*, ICML'22
- Sebastian Tschiatschek, Maria Knobelsdorf, Adish Singla, *Equity and Fairness of Bayesian Knowledge Tracing*, EDM'22
- Silvia Poletti, Alberto Testolin, Sebastian Tschiatschek, *Learning Constraints From Human Stop-Feedback in Reinforcement Learning*, AAMAS'23