Jaywalk! for all I care...

Introduction: Automated driver assistance systems (ADAS) face challenges with pedestrians, especially in complex urban environments where pedestrians exhibit diverse and unpredictable behaviors. The uncertainty exhibited by these road users - in addition to the can-of-worms in collecting it (Feng et al., 2021) - is magnified in automated driving systems (ADS) designed to execute both the dynamic driving task (DDT) and strategic tasks (i.e., trip planning, route selection). However, this off-loading of the driving task requires an adequately scaled system that can handle a myriad of ambiguous pedestrian behaviors (Mitman and Ragland, 2007). Herein lies the gap between the amount of training possible to provide such systems and the training necessary to ensure *some* level of safe deployment in multimodal contexts.

Background: Reinforcement learning (RL) has gained significant popularity on this front for its flexibility to explore a state space independently, providing resiliency towards unseen events and environments. Reasonably, a lack of interpretability becomes concerning in high-risk domains, where agents act according to learned optimal policies not entirely understood (Armstrong and Mindermann, 2018; Glanois et al., 2021). Mistrust leads to distrust: before real-world applications of RL models are statistically real test-beds for them. However, creating such simulations with distribution-level accuracy is difficult for sparse events struggling with dimensionality (i.e., safety-critical driving events, Yan et al. (2023)). At scale, such is also the case for pedestrians, whom inherently and indefinitely add complexity to their environments through interaction (autocirricula, Leibo et al. (2019)). But fear not! Wang et al. (2022a,b, 2020), Wang et al. (2016, 2021), ..., have made significant efforts developing human-like driving models to exhibit socially compatible behaviors. In light, socially compatible behaviors for various kinds of road agents must be learned (i.e., vulnerable road users (pedestrians) require more sensitivity in interaction terms of space and time).

Research question: How do we walk? Due to collection costs and privacy concerns, robust (and granular) pedestrian behavior data is scarce, which limits the amount of knowledge we have about pedestrian path planning at a microenvironment scale. It may be interesting to use observed expert demonstrations to explore, for example, how pedestrians avoid re-routes due to construction (flexibility to new tasks), or what running to the bus looks like (robustness to goal specification). Johora and Müller (2021) suggests pedestrians often path plan irrationally, is it preferential? Because there are multiple expert demonstrations per participant, it would also be interesting to explore preferences among pedestrian in urban navigation.

Analysis plan: Let's explore walking. Maybe potentially with deep imitation learning and expert demonstration. These demonstrations are provided via previously collected GIS-linked walking activity from pedestrians in Seattle, WA. Of note, this dataset contains multiple traversals for ~ 300 pedestrians.

I would be remiss to not mention some recent work. What initially sparked interest into expert demonstration for intelligent planning was Rhinehart et al. (2018), who model continuous-state, discrete-time, partially-observed Markov processes in a deep imitative learning model (an exact target no longer, cars can avoid pot-holes). Vozniak et al. (2020) uses deep imitation learning to explore pedestrian behaviors in critical safety scenarios, and Trinh et al. (2020) employed a PPO-based path planning model that followed common walking conventions and human behaviours such as walking on the left side of the road and staying away from dangerous obstacles. Honorable mention to a devil's advocate, Lake et al. (2017), who kinda thinks the whole things a joke.

Next steps: Will need help scaling scope/focus. Get feedback from EV, meet to discuss scope and brainstorm implementation. Please reach out if you have significant concerns!

SIMCITY!!! (Jaywalk-v2)

1 Introduction

Automated driver assistance systems (ADAS) play a crucial role in enhancing road safety and efficiency. However, ADAS faces significant challenges when dealing with pedestrian interactions, particularly in complex urban environments. Pedestrians exhibit diverse and unpredictable behaviors, making it essential to develop robust systems that can handle this uncertainty.

1.1 Problem Statement

The uncertainty associated with pedestrian behavior is further amplified in automated driving systems (ADS). ADS must simultaneously manage the dynamic driving task (DDT) and strategic tasks (e.g., trip planning, route selection). Off-loading the driving task to an automated system requires a well-scaled solution capable of handling ambiguous pedestrian behaviors. Bridging the gap between training feasibility and safe deployment in multimodal contexts is critical.

2 Background

Reinforcement learning (RL) has gained prominence in addressing complex control problems due to its ability to explore state spaces independently. RL agents learn optimal policies through trial and error, adapting to unseen events and environments. However, in safety-critical domains like traffic safety, interpretability becomes a concern. Additionally, real-world driving scenarios involve sparse events (e.g., pedestrian interactions) with high-dimensional state spaces. Therefore creating accurate simulations for training and testing is difficult. Inverse reinforcement learning (IRL) aims to infer the underlying reward function from expert demonstrations. It helps understand why an expert behaves the way they do.

Since we are interested in a specific setting of imitation learning—the problem of learning to perform a task from expert demonstrations—in which the learner is given only samples of trajectories from the expert, is not allowed to query the expert for more data while training, and is not provided reinforcement signal of any kind. There are two main approaches suitable for this setting: behavioral cloning, which learns a policy as a supervised learning problem over state-action pairs from expert trajectories; and inverse reinforcement learning, which finds a cost function under which the expert is uniquely optimal. Behavioral cloning, while appealingly simple, only tends to succeed with large amounts of data, due to compounding error caused by covariate shift. Inverse reinforcement learning (IRL), on the other hand, learns a cost function that prioritizes entire trajectories over others, so compounding error, a problem for methods that fit single-timestep decisions, is not an issue. Unfortunately, IRL algorithms can expensive to run, often incorporating reinforcement learning methods in the process. Scaling these applications have thus become a worthy venture Ho and Ermon (2016); Finn et al. (2016). However, in IRL the learner is not taught how to act, but rather simply learns the cost function from expert training data. In comes Ho and Ermon (2016)'s GAIL, who combines a Generative Adversarial Network (GAN) with a specific form of sampling-based Maximum Entropy Inverse RL algorithm. By leveraging the adversarial training process, GAIL effectively learns from expert demonstrations while addressing the limitations of traditional IRL methods, making it well-suited for scenarios where direct reinforcement signals are absent.

Previous work that inspired interest into expert demonstration for intelligent planning was Rhinehart et al. (2018), who model continuous-state, discrete-time, partially-observed Markov processes in a deep imitative learning model (an exact target no longer, cars can avoid pot-holes). Vozniak et al. (2020) uses deep imitation learning to explore pedestrian behaviors in critical safety scenarios, and Trinh et al. (2020) employed a PPO-based path planning model that followed common walking conventions and human behaviours such as walking on the left side of the road and staying away from dangerous obstacles.

3 Methods

Data: Pedestrian data were collected as part of two previous National Institute of Health (NIH) projects designed to examine the impact of transit and transit changes on walking and physical activity (Kang et al., 2013). Both projects used adult participants living in Seattle and the surrounding areas. Participants were asked to wear electronic devices that track their movements and to complete activity diaries that detailed their travel behavior each week. Data from the electronic devices captured walking bouts, defined as continuous walking with breaks less than or equal to two minutes within a seven minute rolling window (Kang et al., 2013).

- *Travel Assessment and Community (TRAC) Study*: Walking bouts were collected in three waves (2008-2009, n=707; 2010-2011, n=599, 2012-2013, n=525). The original study goal was designed to assess walking activities given the implementation of light rail in King County, WA **?**.
- *Assessing Choices in Transportation in our Neighborhood (ACTION) Study*: Walking bouts were collected in three waves (2013-2014, n=590; 2015-2016, n=398; 2017-2018, n=382). The original study goal was designed to assess walking activities given the implementation of two bus rapid transit lines in King County, WA **?** .

Of the cleaned pedestrian data, there are 412 experts. They vary both in the amount of traversals completed, and the length of traversals. Figure 1 depicts two bar-plots demonstrating the number of traversals completed by each pedestrian (top), and the sum of distance covered by each participant. The next step would be determining if the amount of data available from these pedestrians is enough to teach an agent preferential path planning, or if the distance covered would need to confined to a single area.

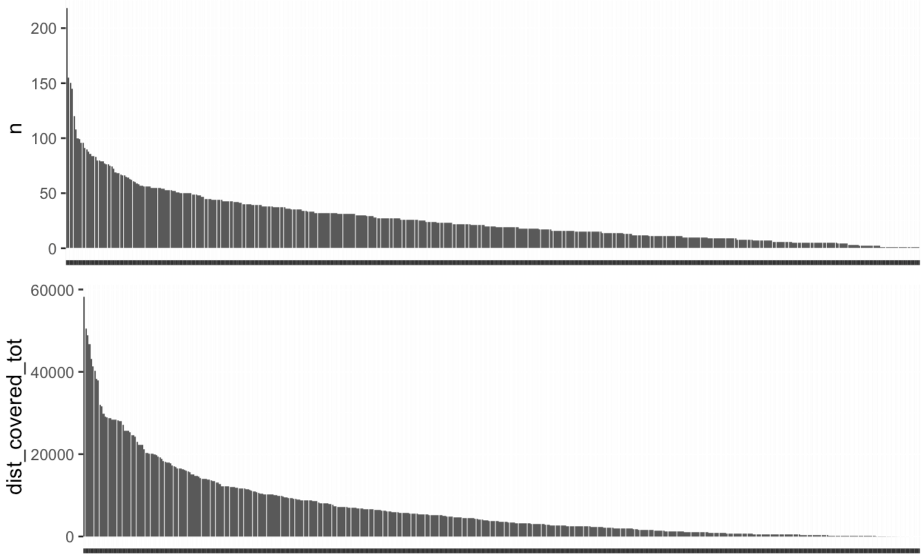

Figure 1: Top: Number of traversals completed per pedestrian. Bottom: Sum of total distance covered per pedestrian.

Figure 2b and Figure 2a visualizes both a downtown corridor and a university setting of pedestrian traversals available. Depending on the the amount of training data necessary, it may be interesting to consider the University setting, where pedestrians often have the right-of-way and urban make-up can become rather ambiguous.

(a) Cleaned pedestrian traversals in downtown Seattle (b) Cleaned pedestrian traversals in U of Washington

Figure 2: Comparison of traversal data in densely populated pedestrian areas

Proposed Approach: Generative Adversarial Imitation Learning (GAIL)
The Stablebaselines3 is a collection of RL in the PyTorch framework and will be used for this implementation. GAIL combines the power of generative adversarial networks (GANs) with imitation learning. It aims to learn a policy by imitating expert demonstrations while maintaining the flexibility of RL. Here's how it works:

- Expert Demonstrations: We collect expert demonstrations (e.g., human-driven trajectories) to guide the learning process.
- Discriminator Network (D): The discriminator distinguishes between expert trajectories and those generated by the RL agent. This network acts as a critic, differentiating between real expert trajectories and those generated by the policy network.
- Generator Network (G - policy network): The generator (policy network) learns to produce trajectories that fool the discriminator. This network aims to mimic the expert demonstrations by generating pedestrian behavior trajectories. It essentially learns a policy that maps states (pedestrian surroundings) to actions (pedestrian movement).
- Adversarial Training: GAIL optimizes both the generator and discriminator simultaneously, creating a competitive learning process.

3.1 Potential challenges:

- Mode Collapse: GAIL employs adversarial training, which can suffer from mode collapse. In the context of pedestrian behavior, this means the generated trajectories might converge to a limited set of behaviors, missing out on diverse pedestrian actions.
- Distribution Mismatch: The distribution of expert demonstrations may differ significantly from the distribution of pedestrian behavior encountered during deployment. This distribution shift can lead to poor generalization.
- High-Dimensional State Space: Pedestrian behavior modeling involves complex interactions with the environment. Representing these interactions in a high-dimensional state space can be challenging for neural networks.
- Exploration vs. Exploitation: Balancing exploration (trying out new behaviors) and exploitation (leveraging learned policies) is crucial. GAIL tends to focus on exploitation, which may hinder discovering new microenvironment pedestrian behaviors.

4 Metrics of success:

- Imitation Accuracy: Measure how closely the generated trajectories match the expert demonstrations. High accuracy indicates successful imitation.
- Policy Performance: Evaluate the RL agent's performance to learn normative pedestrian behavior (i.e., walking on the left side of the road)
- Preferential Path Planning: Johora and Müller (2021) suggests that pedestrians often path plan irrationally. Investigating whether this behavior is preferential and exploring preferences among pedestrians in urban navigation is intriguing.
- Training Efficiency: Monitor the training time and sample efficiency. Faster convergence and fewer samples needed for learning are desirable.

5 Conclusion

By integrating socially compatible behaviors, leveraging GAIL with the Stablebaselines environment, and utilizing expert demonstrations, possibly we can learn more about pedestrian decision-making in microenvironments. Ideally the proposed approach takes a step towards bridging the gap between ADS/ADAS training feasibility and real-world deployment, ensuring safer interactions in complex urban environments.

References

Armstrong, S., Mindermann, S., 2018. Occam's razor is insufficient to infer the preferences of irrational agents. Adv. Neural Inf. Process. Syst. 31. URL: https://proceedings.neurips.cc/paper/2018/file/d89a66c7c80a29b1bdbab0f2a1a94af8-Paper.pdf.

Feng, Y., Duives, D., Daamen, W., Hoogendoorn, S., 2021. Data collection methods for studying pedestrian behaviour: A systematic review. Build. Environ. 187, 107329. URL: https://www.sciencedirect.com/science/article/pii/S0360132320306983, doi:10.1016/j.buildenv.2020.107329.

Finn, C., Levine, S., Abbeel, P., 2016. Guided cost learning: Deep inverse optimal control via policy optimization URL: http://arxiv.org/abs/1603.00448, arXiv:1603.00448.

Glanois, C., Weng, P., Zimmer, M., Li, D., Yang, T., Hao, J., Liu, W., 2021. A survey on interpretable reinforcement learning URL: http://arxiv.org/abs/2112.13112, arXiv:2112.13112.

Ho, J., Ermon, S., 2016. Generative adversarial imitation learning URL: http://arxiv.org/abs/1606.03476, arXiv:1606.03476.

Johora, F.T., Müller, J.P., 2021. Modeling interactions of multimodal road users in shared spaces URL: http://arxiv.org/abs/2107.02083, arXiv:2107.02083.

Kang, B., Moudon, A.V., Hurvitz, P.M., Reichley, L., Saelens, B.E., 2013. Walking objectively measured: classifying accelerometer data with GPS and travel diaries. Med. Sci. Sports Exerc. 45, 1419–1428. URL: http://dx.doi.org/10.1249/MSS.0b013e318285f202, doi:10.1249/MSS.0b013e318285f202.

Lake, B.M., Ullman, T.D., Tenenbaum, J.B., Gershman, S.J., 2017. Building machines that learn and think like people. Behav. Brain Sci. 40, e253. URL: http://dx.doi.org/10.1017/S0140525X16001837, doi:10.1017/S0140525X16001837.

Leibo, J.Z., Hughes, E., Lanctot, M., Graepel, T., 2019. Autocurricula and the emergence of innovation from social interaction: A manifesto for Multi-Agent intelligence research URL: http://arxiv.org/abs/1903.00742, arXiv:1903.00742.

Mitman, M.F., Ragland, D.R., 2007. Crosswalk confusion: More evidence why pedestrian and driver knowledge of the vehicle code should not be assumed. Transp. Res. Rec. 2002, 55–63. URL: https://doi.org/10.3141/2002-07, doi:10.3141/2002-07.

Rhinehart, N., McAllister, R., Levine, S., 2018. Deep imitative models for flexible inference, planning, and control URL: http://arxiv.org/abs/1810.06544, arXiv:1810.06544.

Trinh, T.T., Vu, D.M., Kimura, M., 2020. A pedestrian path-planning model in accordance with obstacle's danger with reinforcement learning, in: Proceedings of the 3rd International Conference on Information Science and Systems, Association for Computing Machinery, New York, NY, USA. pp. 115–120. URL: https://doi.org/10.1145/3388176.3388187, doi:10.1145/3388176.3388187.

Vozniak, I., Klusch, M., Antakli, A., Müller, C.A., 2020. InfoSalGAIL: Visual attention-empowered imitation learning of pedestrian behavior in critical traffic scenarios. Int Jt Conf Comput Intell , 325–337URL: https://pdfs.semanticscholar.org/f10e/c4e8446dc9bc4b5d8b88b6a0b74f4207af03.pdf, doi:10.5220/0010020003250337.

Wang, L., Sun, L., Tomizuka, M., Zhan, W., 2020. Socially-Compatible behavior design of autonomous vehicles with verification on real human data URL: http://arxiv.org/abs/2010.14712, arXiv:2010.14712.

Wang, Q., Chen, J., Ma, J., 2021. Experimental study on individual level interaction between bicycle and pedestrian. J. Stat. Mech. 2021, 093403. URL: https://iopscience.iop.org/article/10.1088/1742-5468/ac1d54/meta, doi:10.1088/1742-5468/ac1d54.

Wang, S., Gwizdka, J., Chaovalitwongse, W.A., 2016. Using wireless EEG signals to assess memory workload in the n-Back task. IEEE Transactions on Human-Machine Systems 46, 424–435. URL: http://dx.doi.org/10.1109/THMS.2015.2476818, doi:10.1109/THMS.2015.2476818.

Wang, W., Wang, L., Zhang, C., Liu, C., Sun, L., 2022a. Social interactions for autonomous driving: A review and perspectives URL: http://arxiv.org/abs/2208.07541, arXiv:2208.07541.

Wang, W., Wang, L., Zhang, C., Liu, C., Sun, L., 2022b. Social interactions for autonomous driving: A review and perspectives URL: http://arxiv.org/abs/2208.07541, doi:10.1561/2300000078, arXiv:2208.07541.

Yan, X., Zou, Z., Feng, S., Zhu, H., Sun, H., Liu, H.X., 2023. Learning naturalistic driving environment with statistical realism. Nat. Commun. 14, 2037. URL: http://dx.doi.org/10.1038/s41467-023-37677-5, doi:10.1038/s41467-023-37677-5.