The perfect path

An inverse reinforcement learning approach for microenvironment pedestrian path-planning.

GRACE DOUGLAS, New York University, USA

Pedestrian path-planning behaviors are crucial in urban environments, especially with the rising demand for pedestrian-friendly spaces. However, understanding pedestrian decision-making is challenging due to the diversity and complexity of behaviors. This report proposes an inverse reinforcement learning from demonstration framework to model and understand pedestrian path-planning in informal space.

Additional Key Words and Phrases: *pedestrians, path-planning, inverse reinforcement learning, geographic information systems*

ACM Reference Format:
Grace Douglas. 2024. The perfect path: An inverse reinforcement learning approach for microenvironment pedestrian path-planning.. 1, 1 (May 2024), 11 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Pedestrian path-planning behaviors are diverse, preferential, and often irrational [Johora et al. 2017]. Additionally, efforts to understand pedestrian navigation in the microenvironment are often invasive or uncomfortable [Feng et al. 2021]. However, in the US, significant increases in urban dwelling populations have raised interest for pedestrian-friendly environments. A country historically touted as the darling of mass motorization is now seeing unprecedented rates of soft mode travel (walking, cycling).

Unlike their motorized counterpart, pedestrian behavior is also largely unregulated. This freedom creates uncertainty that often goes unnoticed by other road users, as drivers themselves can *step into their shoes* to navigate the interaction on common ground. It is only in the event of unregulated interaction, or regulation treated irregularly, that these physically mismatched road users spar for shared road space.

A far more interesting question is that of informal space, and the spillage of informal behaviors into formal space. Informal spaces are a commonly overlooked dimension of daily life. But it is in informal spaces that cultures have shaped formal societal structures. In a driving culture these norms are plentiful: the California stop, New York City's 'Honk Honk', Massachusetts' 'Rotary Rules', the Pittsburgh left, the Michigan left, and the Boston Brake are just a few that come to mind - and are technically housed under the same societal roof. Pedestrians exhibit similar structure in walking tendency, perhaps the most obvious cultural discrepancy being the correct side of a given path to walk or wait on (in a Western world the right side waits, left side walks). However, picture for a moment an empty environment, unimpeded by regulated roads and a man-made mesh of concrete, and consider a pedestrian's path from origin A to destination B. For the most part, the expected trajectory is as the crow flies, it is optimal to do so. However, in what environments is it appropriate to assume pedestrians are incessantly in search of *the perfect path*?

Author's address: Grace Douglas, grace.douglas@nyu.edu, New York University, Brooklyn, New York, USA.

XXXX-XXXX/2024/5-ART $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

This project investigates pedestrian path-planning in dense urban settings, focusing particularly on informal space hypothesized to serve leisurely walking activities. The contributions of this work can be summarized,

- Procedure for reducing naturalistic walking data from personal electronic devices
- Framework for instantiating custom (Gymnasium) environments from observed walking behavior
- ...preference learning... Comparison of reduced expert demonstrations with observed walking profiles

2 RELATED WORK

Previous studies have delved into pedestrian behaviors using methodologies like deep imitation learning and inverse reinforcement learning. However, our project takes a unique approach by leveraging expert demonstrations to model pedestrian decision-making in informal urban spaces, offering deeper insights compared to existing methodologies. [Rhinehart et al. 2018] laid the groundwork with deep imitative learning for vehicular navigation dynamics, while [Vozniak et al. 2020] and [Trinh et al. 2020] explored pedestrian behaviors through deep imitation learning and PPO-based path planning, respectively. [Fahad et al. 2018] addressed robot navigation in shared spaces with MEDIRL, [Gonon and Billard 2023] focused on navigating robots through human crowds using IRL, and [Avaei et al. 2023] integrated path and velocity preferences for robotic manipulators with incremental IRL. [Martinez-Gil et al. 2020] improved pedestrian simulations using IRL with real behavior traces. Informed by referenced work, this research examines pedestrian path-planning in informal settings, using an inverse reinforcement learning from demonstration (IRLfD) framework to understand walking preferences in unconfined space.

3 BACKGROUND

In recent years, advancements in machine learning and artificial intelligence have opened up new avenues for understanding and modeling pedestrian behavior. Traditional methods of studying pedestrian movement, such as observational studies or surveys, can suffer significantly from collection and survey biases. Reinforcement learning (RL) and inverse reinforcement learning (IRL) offer promising alternatives by allowing agents to learn optimal policies and infer underlying reward functions from expert demonstrations [Finn et al. 2016; Ho and Ermon 2016].[1][2]

RL has gained prominence for its ability to explore state space independently. RL agents learn optimal policies through trial and error, adapting to unseen events and environments. This becomes particularly interesting when applied to real-world scenarios that often involve sparse events (i.e., pedestrian interactions) with high-dimensional state spaces (i.e., informal shared space). However, crafting training environments for real-world testing is difficult. Inverse reinforcement learning (IRL) takes an altered, but rather informed approach, to infer an underlying reward function from observed expert demonstrations. One such IRL technique, Generative Adversarial Imitation Learning (GAIL), has shown promising performances by combining RL's exploration ability with supervised learning techniques in deep learning (DL). This guided exploration can minimize the computation needed to explore high dimensional state spaces. In applications of guided learning, the learner is given samples of trajectories from a expert, and is not allowed to query the expert for more information, nor is provided a reinforcement signal of any kind.

There are two main approaches suitable for the guided 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

[1] just a randomly rad repo to counter the wall of text
[2] a light-weight, high-fidelity, multi-player urban driving simulator for cross-cultural research

is uniquely optimal. Behavioral cloning, while appealingly simple, only tends to succeed with large amounts of data (see [Spencer et al. 2021] for an explanation of compounding error and covariate shift). 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 still be expensive to run because of the RL learning methods employed.

Scaling these applications have thus become a worthy venture [Finn et al. 2016; Ho and Ermon 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 IRL. 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.

4 METHODS

The proposed framework integrates Generative Adversarial Imitation Learning (GAIL) with techniques from Multivariate Functional Principle Component Analysis (MFPCA) and Maximum Entropy Inverse Reinforcement Learning (MaxEnt IRL). GAIL is used to learn optimal policies from expert demonstrations, while MFPCA helps identify common patterns in pedestrian trajectories. MaxEnt IRL is then employed to infer underlying preferences from observed demonstrations and train the model to generate realistic pedestrian trajectories. Figure 1 visualizes the approach.

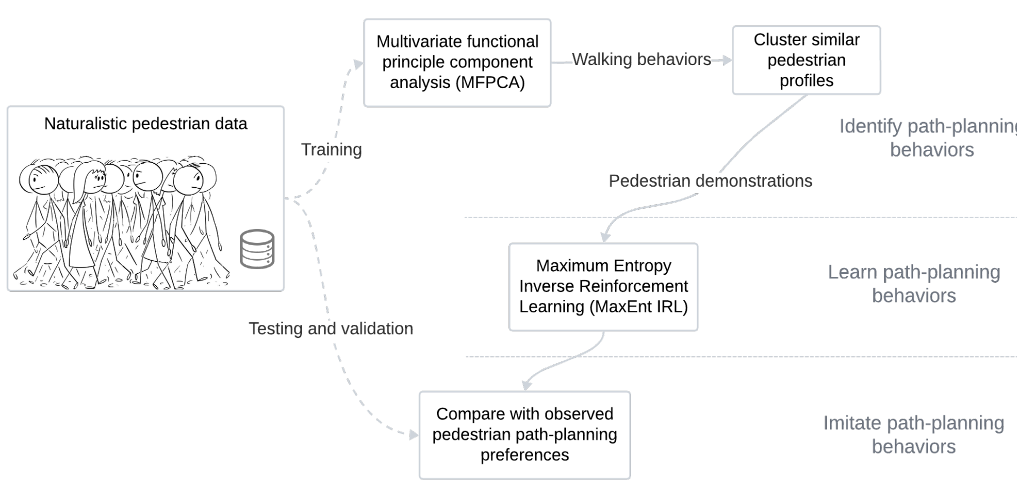

Fig. 1. Framework of proposed methodology.

4.1 Generative Adversarial Imitation Learning

GAIL, a variant of imitation learning, leverages expert demonstrations to train agents to mimic observed behaviors. By collecting expert demonstrations from real-world pedestrian activities in urban environments, a rich diversity of pedestrian behaviors is captured, including route selection, obstacle avoidance, and interaction with urban infrastructure.

The GAIL framework formulates the imitation learning problem as a min-max optimization problem based on dual representation, which can be solved by stochastic gradient algorithms [Guan et al. 2020]. The objective function of GAIL is expressed as,

$$\min_{\pi} \max_{D} \mathbb{E}_{\pi}[\log D(s,a)] + \mathbb{E}_{\pi_E}[\log(1 - D(s,a))]$$

where:

- π is the policy of the learning agent,
- π_E is the expert's policy,
- $D(s,a)$ is a discriminator function that tries to distinguish between the state-action pairs (s,a) coming from the learning agent and the expert,
- (s,a) are the state-action pairs

The expectations are taken over the state-action pairs generated by the learning agent and the expert. The learning agent aims to find a policy π that makes the discriminator D unable to distinguish its state-action pairs from those of the expert. On the other hand, the discriminator D tries to best distinguish between the state-action pairs from the learning agent and the expert.

4.2 MFPCA and MaxEnt IRL

In addition to GAIL, the proposed methodology integrates techniques from Multivariate Functional Principle Component Analysis (MFPCA) and Maximum Entropy Inverse Reinforcement Learning (MaxEnt IRL). MFPCA allows researchers to identify common patterns in pedestrian trajectories and cluster them into distinct pedestrian profiles. This enables the extraction of specific path-planning behaviors: ideally this dimensionality reduction provides more efficient input for MaxEnt IRL training. By inferring underlying preferences from observed demonstrations, MaxEnt IRL effectively trains the model to generate realistic pedestrian trajectories that align with real-world preferences. The integration of GAIL, MFPCA, and MaxEnt IRL should provide reasonable starting point to understanding pedestrian path planning in complex urban environments.

5 PROBLEM SETTING

This project considers an potentially task-less environment. Green Lake, Seattle, WA is a park located in North Seattle, and is a very popular green space with sport fields, pedestrian and cyclist facilities, a boathouse, and both paved and unpaved paths wrapped around the circumference of a lake (3.1 miles). Please refer to Figure 2 for a visualization of this microenvironment.

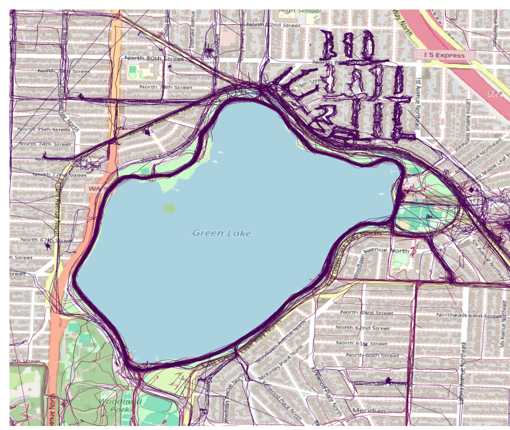

Fig. 2. Map of informal pedestrian space considered in this project. For thorough folks captioned-inclined, it may be interesting to note the most dangerous intersection in the City of Seattle (in terms of injury severity and the annual rate of road user fatality) had been previously identified as a block off from the south-end of Green Lake.

The purple lining in Figure 2 visualizes a (multi-) line-string geometry representing observed pedestrian walking traversals. Notice in the top left corner of this figure sporadic GPS traces that must be cleaned or removed as expert demonstrations are prepared for use. This kind of messingness (yes, I'll make up a derivative of missingness) illustrates a common callow facing pedestrian data collection in dense urban areas. For reference, personal GPS devices have known limitations in terms of data completeness [Krenn et al. 2011]. Missing GPS data can result from signal loss, signal dropout, variant temperatures, or low unit power.

5.1 Data Collection

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. Electronic device data was captured in 15 second epochs and used to develop a physical activity measure, *the walking bout*. Walking bouts are 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 [Saelens et al. 2014].
- *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 [Saelens et al. 2014] .

Of the cleaned pedestrian data, there are 412 experts. They vary both in the amount of traversals completed, and the length of traversals.

5.2 Data Preprocessing

Raw sensor data, including GPS coordinates and accelerometer readings, are converted into a format suitable for training. GIS-linked data is segmented by participant, and due to week-long collection periods, individuals are attributed to multiple traversals. See Figure 3 for a visualization of walking trips and distance covered on these trips by participant.

GPS data from personal devices were supplemented with velocity, acceleration, and a derived orientation measure to maintain as much granularity as possible. Brief procedural notes for computing these measures is provided.

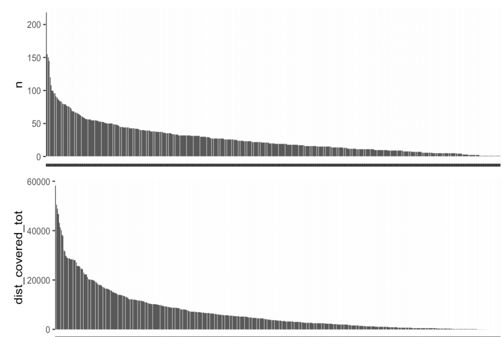

Fig. 3. A pair of histograms provide traversal counts (top) and distance covered (bottom) by participant.

5.2.1 *Consider the Chaiken.*

As Figure 2 suggests, the data from electronic devices exhibits choppy behavior due to sampling rate. The Chaikin algorithm is a simple and effective method for smoothing a jagged line into a smooth curve [Marx 2020]. It's often referred to as the *corner-cutting* algorithm. Analytically its lovely:

(1) Given vector representations of a point set $\{A, B\}$, new points C and D are calculated,
  - $C = 0.75A + 0.25B$
  - $D = 0.25A + 0.75B$

(2) Assuming $A = (A_x, A_y)$ and $B = (B_x, B_y)$, perform operations component-wise for additional points C and D,
  - $C = (0.75A_x + 0.25B_x, 0.75A_y + 0.25B_y)$
  - $D = (0.25A_x + 0.75B_x, 0.25A_y + 0.75B_y)$

(3) Continue as necessary.

This process is applied to each pair of consecutive points in the original paths, and the resulting set $\{A, B, C, D\}$ form a new, smoother path. This process can be repeated as many times as needed to achieve the desired level of smoothness. The more iterations, the smoother the resulting path will be. However, with each iteration, the resulting path also becomes less similar to the original path. A six-point Chaiken was applied to the pedestrian traversal data.

5.2.2 Position, velocity, acceleration. Positioning in meters was projected to a local coordinate reference system (CRS) (EPSG:32610 - WGS 84 / UTM zone 10N), which disregards the horizontal dimension. This is acknowledged as a limitation because the conversion from global longitude, latitude (+altitude) to a parameterised sphere requires a further zonal specification that changes depending on a region's mapping base. However, local projections do reduce compute and are advised by Quantum Geographic Information System (QGIS) when conducting microenvironment geospatial analyses [QGIS 2002].

Accelorometry provided velocity information along pedestrian traversals. However, the personal accelerometry device was equipped with a slightly lower sampling rate (10 second epochs) and was aggregated to the GPS fidelity (15 second epochs) for ease.

The choice to Chaiken in a local reference system actually benefited these computations by enhancing spatial specificity to impute velocity (and acceleration) at smaller intervals along traversals. Remember any imputed information along the smoothed line-strings was not actually observed pedestrian behavior, but rather an extrapolated extension.

5.2.3 Orientation. The *azimuth angle*, simply referred to as the *azimuth*, is a directional measure representing the horizontal angle or direction of an object with respect to a reference direction, usually north. In a spherical coordinate system, the azimuth angle is the anticlockwise angle between the positive x-axis and the projection of the vector onto the xy-plane. It is usually measured in degrees (°), in the positive range 0° to 360° or in the signed range -180° to +180°[1]. The azimuth in a spherical reference system is derived from positional changes between time steps using QGIS, an open-sourced spatial processing software that provides plug-in support for out-of-box use.

Now that the problem setting has been established, a training environment (state, action, and observation spaces) is defined in context.

6 EXPERIMENT

As a reminder, the objective of this paper is pedestrian behavior in informal space, particularly in a park where more ambiguous behavior is accepted and expected.

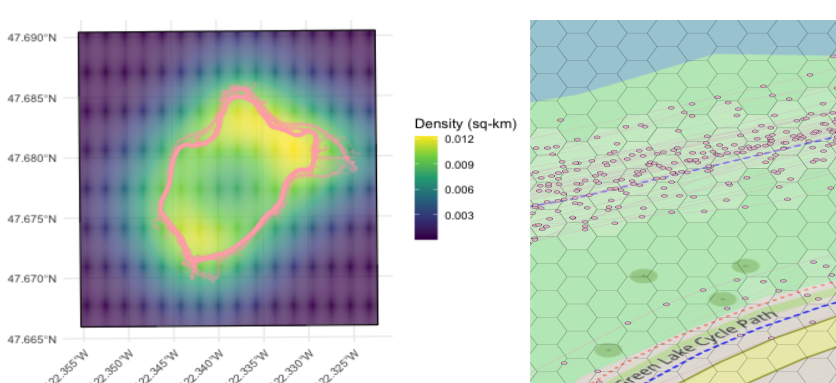

Fig. 4. Comparison of gridworld representations. Left: A 10x20 Gridworld rasterized with rewards relative to observed walking density around Green Lake, Seattle, WA. Right: Exploding Gridworld. Hexagons of 10x10 (width x height) meters are constructed and overlayed on traversals. Chaiken points are also shown, which represent observed pedestrian behavior. A (multi-) linestring layer allocates Chaiken children to their respective parent traversal.

Figure 4 depicts two interpretations of a common single-agent training environment, *GridWorld*. The environment on the left is obviously a significantly simplified representation of the target environment. This environment was used to practice environment customization to ensure no issues occurred with the geospatial data.

Stablebaselines3 is a collection of RL algorithms in the PyTorch framework used. The *imitation* library is the additional packaging to include behavioral cloning pre-training.

6.1 Expert demonstrations

The collected traversals will serve as expert demonstrations. Inverse Reinforcement Learning (IRL) requires a varying number of expert demonstrations or samples depending on the specific algorithm used. One paper proposes a more informed imitation learning reduction that utilizes the state distribution of the expert to alleviate the global exploration component of the RL subroutine, providing an exponential speedup in theory [Armstrong and Mindermann 2018]. Another paper presents an algorithm called Delphi that requires a polynomial number of expert queries and a polynomial amount of exploratory samples to recover an $\epsilon-$suboptimal policy, resulting in an exponential improvement in sample complexity [Feng et al. 2021]. Additionally, a paper introduces SEILO, an on-policy algorithm for imitation learning from observations, which requires fewer interactions with the environment to achieve expert performance compared to other state-of-the-art methods [Finn et al. 2016].

These traversals are also labeled with utility, or purpose, from travel diaries kept by participants. For example, the traversal-specific utility recorded attributes such as *work*, *leisure*, or *errands*. Origin-destination start and end times were also collected. Participant sociodemographics were also provided. These demonstrations guide the GAIL training process and serve as the ground truth for imitation learning.

6.2 Customizing the environment

Table 1 briefly describes the environment, which includes a discretized state space, and action spaces collected in hexagons as shown on the right side of Figure 4. For ease, a discretized action space is chosen to include position, velocity, acceleration, and orientation within the microenvironment. This is acknowledge as a limitation because a continuous representation is preferred, which better describes dynamic movement for realistic path-planning behaviors. The observation space encompasses the information available to the agent, including environmental cues (i.e., distance to edge of lake). Algorithm 1 shows pseudo-code for the reward function as so,

Input: state, action
Output: reward
Extract the current position;
$x \leftarrow$ state.position.x;
$y \leftarrow$ state.position.y;
Calculate the distance from the lake;
$distance\_from\_lake \leftarrow$
  calculate_distance(x, y, lake_position);
Initialize reward;
$reward \leftarrow 0$;
Check conditions for decreasing rewards;
if *action.displacement == "leave"* **then**
  $reward \leftarrow reward - 30$;
end
if *crossed_road_boundary(x, y)* **then**
  $reward \leftarrow reward - 60$;
end
if $distance\_from\_lake \geq 0.1$ *and* $distance\_from\_lake < 0.25$ **then**
  $reward \leftarrow reward - 120$;
end
Check conditions for additional penalties and rewards;
if *state.hexagon == previous_state.hexagon* **then**
  $reward \leftarrow reward - 10$;
end
if *action.direction == previous_action.direction and action.movement == previous_action.movement* **then**
  $reward \leftarrow reward + 10$;
end
return *reward*;

Algorithm 1: Reward function

7 RESULTS

As informal space was the area of interest, it was not enough to simply consider observed pedestrian activity directly mappable to popular geospatial platforms (OpenStreetMaps). Instead, the local OSM road network supplemented observed pedestrian positioning, and Djikstra's shortest path method [Lanning et al. 2014] was used to reconnect both sets of nodes (OSM + observed traversals). This was expensive to compute, and took 4/5 days to complete. In the future, a more thorough investigation of graph-based environments could improve this process efficiency [Janisch et al. 2020; Kinose and Taniguchi 2019]. Figure 5L visualizes Green Lake park after Djikstra's shortest path computation, and Figure 5R is a graphical representation of a nonspatial graph environment.

7.0.1 *State space.*

8 DISCUSSION

There did exist limitations in this work. Another limitation to this project is the discrete action space. Ideally the expert demonstrations are collected with higher fidelity: a common solution is collecting video data of expert demonstrations and using this multidimensional data to learn an optimal policy. There seems to be significant support for this approach among the imitation, stablebaselines3, and gynmasium documentation. However, video collecting surveillance data that can identify any one of three identifying human features (face, hands, ears) will face significant red tape during collection and throughout usage. Most high resolution pedestrian data is collected in controlled environments, which is not entirely conducive or reasonable for studying informal pedestrian behavior. The personal electronic devices used in this study were battery powered, which did result in significant data loss during collection. A clear gap exists here for (1) immersive (and *available*) pedestrian experimentation environments, and (b) ubiquitous technologies that don't rely on participants to conduct their own data collection.

A perhaps more significant limitation to this project is the single-agent framework. Historically pedestrian experimentation has only considered single participant environments, a shortcoming largely excused by issues with confounding factors. However, intra-urban road users rarely see such empty environments: again, a gap exists in experimentation environments that can provide this multi-player realism. I honestly do not know how the mutli-agent framework

Space	Attributes	Details
State Space	Position (x, y)	Discretized hexagons of the park area
Action Space	Discrete actions	(1) direction: six edges define available directions in each 10x10 meter hexagon edges (60°/ edge), (2) displacement: *stay*, *leave*, (3) accelerate: *accelerate*, *decelerate*, *no change*
Rewards	Reward function	(1) Stay near Green Lake: {-30, -60, -120} exponential decrease in reward for leaving the park (-30), for crossing the first road boundary surrounding lake (-60), and for being >=0.1 & <0.25 miles from lake perimeter (-120), (2) -10 for staying in the same hexagon 2x+, (3) +10 for every repeated direction, movement combination
Termination	Discrete	(1) Going in lake, (2) >= 0.25 miles from lake perimeter

Table 1. Attributes of State Space, Action Space, and Rewards

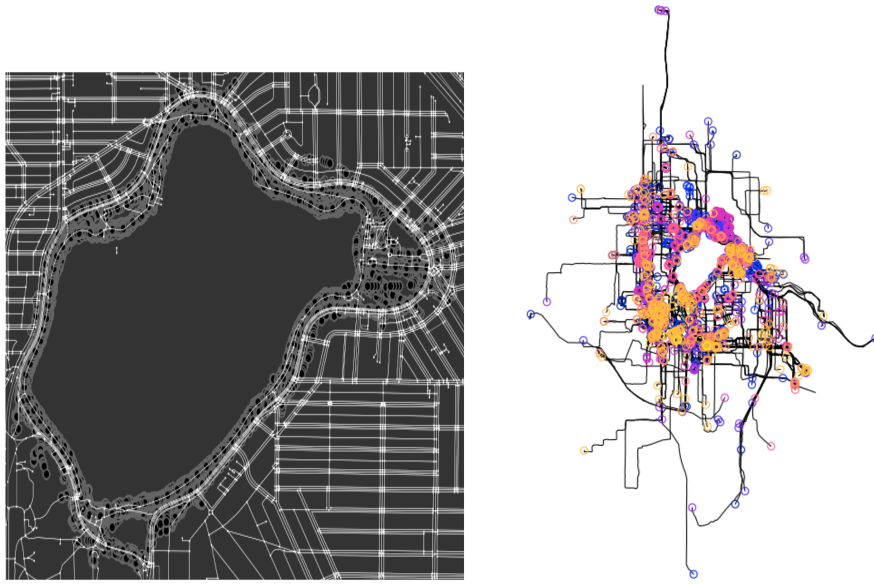

Fig. 5. Left: Open Street Map road network and path network, supplemented with observed pedestrian behaviors. Right: pure graph-based environment that does not retain spatial link.

would be formulated in the context of this study, but that is where I would go next.

8.1 Next steps, metrics of success

this section is wip. sorry :/

To evaluate the success of the learning agent in imitating pedestrian path planning behavior, the follwoing metrics could be observed.

8.1.1 Trajectory Similarity. One of the primary metrics of success is the similarity between the trajectories generated by the learning agent and the expert demonstrations. This can be quantified using various distance metrics such as the Dynamic Time Warping (DTW) distance or the Frechet distance. A lower distance indicates a higher similarity between the trajectories, suggesting that the learning agent has successfully imitated the expert's behavior.

8.1.2 Collision Avoidance. In a real-world scenario like a park, avoiding collisions with other pedestrians and obstacles is crucial. Therefore, the number of collisions encountered by the learning agent during its navigation can serve as a metric. A lower number of collisions indicates better

9 CONCLUSION

Little is known about pedestrian path-planning in the microenvironment, and there are significant obstacles facing its data collection. Inverse reinforcement learning from expert demonstrations thus becomes an attractive option for retaining human behavior when there exists scare amounts of it (....scaling...). This project developed a single-agent learning environment to explore pedestrian path-planning (fD) in the microenvironment. Moreover, the observation space in this environment is weighted with observed pedestrian activity, and the environment is supplemented with direction, speed, and orientation features from expert (technically "non-expert") observations. The chosen environment of interest is Green Lake, Seattle, WA, an park that provides informal green space for vulnerable road populations to exist amidst a densely populated Seattle backdrop.

There are also many exciting next steps for this project. The environment was finally registered with Gynasium! after undergoing some pretty significant cleaning and processing steps.

ACKNOWLEDGMENTS

To Professor Vinitsky, for the space to explore, and exploit.

REFERENCES

Stuart Armstrong and Sören Mindermann. 2018. Occam's razor is insufficient to infer the preferences of irrational agents. *Adv. Neural Inf. Process. Syst.* 31 (2018). https://proceedings.neurips.cc/paper/2018/file/d89a66c7c80a29b1bdbab0f2a1a94af8-Paper.pdf

Armin Avaei, Linda van der Spaa, Luka Peternel, and Jens Kober. 2023. An Incremental Inverse Reinforcement Learning Approach for Motion Planning with Separated Path and Velocity Preferences. *Robotics* 12, 2 (April 2023), 61. https://doi.org/10.3390/robotics12020061

M Fahad, Zhuo Chen, and Yi Guo. 2018. Learning how pedestrians navigate: A deep inverse reinforcement learning approach. *IROS* (Oct. 2018), 819–826. https://doi.org/10.1109/IROS.2018.8593438

Yan Feng, Dorine Duives, Winnie Daamen, and Serge Hoogendoorn. 2021. Data collection methods for studying pedestrian behaviour: A systematic review. *Build. Environ.* 187 (Jan. 2021), 107329. https://doi.org/10.1016/j.buildenv.2020.107329

Chelsea Finn, Sergey Levine, and Pieter Abbeel. 2016. Guided Cost Learning: Deep Inverse Optimal Control via Policy Optimization. (March 2016). arXiv:1603.00448 [cs.LG] http://arxiv.org/abs/1603.00448

David Gonon and Aude Billard. 2023. Inverse Reinforcement Learning of Pedestrian–Robot Coordination. *IEEE Robotics and Automation Letters* 8, 8 (Aug. 2023), 4815–4822. https://doi.org/10.1109/LRA.2023.3289770

Ziwei Guan, Tengyu Xu, and Yingbin Liang. 2020. When Will Generative Adversarial Imitation Learning Algorithms Attain Global Convergence. (June 2020). arXiv:2006.13506 [cs.LG] http://arxiv.org/abs/2006.13506

Jonathan Ho and Stefano Ermon. 2016. Generative Adversarial Imitation Learning. (June 2016). arXiv:1606.03476 [cs.LG] http://arxiv.org/abs/1606.03476

Jaromír Janisch, Tomáš Pevný, and Viliam Lisý. 2020. Symbolic Relational Deep Reinforcement Learning based on Graph Neural Networks and Autoregressive Policy Decomposition. (Sept. 2020). arXiv:2009.12462 [cs.LG] http://arxiv.org/abs/2009.12462

Fatema Tuj Johora, Philipp Kraus, and Jörg P Müller. 2017. Dynamic Path Planning and Movement Control in Pedestrian Simulation. (Sept. 2017). arXiv:1709.08235 [cs.MA] http://arxiv.org/abs/1709.08235

Bumjoon Kang, Anne V Moudon, Philip M Hurvitz, Lucas Reichley, and Brian E Saelens. 2013. Walking objectively measured: classifying accelerometer data with GPS and travel diaries. *Med. Sci. Sports Exerc.* 45, 7 (July 2013), 1419–1428. https://doi.org/10.1249/MSS.0b013e318285f202

Akira Kinose and Tadahiro Taniguchi. 2019. Integration of Imitation Learning using GAIL and Reinforcement Learning using Task-achievement Rewards via Probabilistic Graphical Model. (July 2019). arXiv:1907.02140 [cs.LG] http://arxiv.org/abs/1907.02140

Patricia J Krenn, Sylvia Titze, Pekka Oja, Andrew Jones, and David Ogilvie. 2011. Use of global positioning systems to study physical activity and the environment: a systematic review. *Am. J. Prev. Med.* 41, 5 (Nov. 2011), 508–515. https://doi.org/10.1016/j.amepre.2011.06.046

Daniel R Lanning, Gregory K Harrell, and Jin Wang. 2014. Dijkstra's algorithm and Google maps. In *Proceedings of the 2014 ACM Southeast Regional Conference* (Kennesaw, Georgia) *(ACM SE '14, Article 30)*. Association for Computing Machinery, New York, NY, USA, 1–3. https://doi.org/10.1145/2638404.2638494

Francisco Martinez-Gil, Miguel Lozano, Ignacio García-Fernández, Pau Romero, Dolors Serra, and Rafael Sebastián. 2020. Using Inverse Reinforcement Learning with Real Trajectories to Get More Trustworthy Pedestrian Simulations. *Sci. China Ser. A Math.* 8, 9 (Sept. 2020), 1479. https://doi.org/10.3390/math8091479

Steve Marx. 2020. Chaikin Curves: a Beautifully Simple Algorithm. https://smarx.com/posts/2020/08/chaikin-curves-a-beautifully-simple-algorithm/. https://smarx.com/posts/2020/08/chaikin-curves-a-beautifully-simple-algorithm/ Accessed: 2024-5-15.

QGIS. 2002. Working with Projections. https://docs.qgis.org/2.18/en/docs/user_manual/working_with_projections/working_with_projections.html. https://docs.qgis.org/2.18/en/docs/user_manual/working_with_projections/working_with_projections.html Accessed: 2024-5-15.

Nicholas Rhinehart, Rowan McAllister, and Sergey Levine. 2018. Deep Imitative Models for Flexible Inference, Planning, and Control. (Oct. 2018). arXiv:1810.06544 [cs.LG] http://arxiv.org/abs/1810.06544

Brian E Saelens, Anne Vernez Moudon, Bumjoon Kang, Philip M Hurvitz, and Chuan Zhou. 2014. Relation between higher physical activity and public transit use. *Am. J. Public Health* 104, 5 (May 2014), 854–859. https://doi.org/10.2105/AJPH.2013.301696

Jonathan Spencer, Sanjiban Choudhury, Arun Venkatraman, Brian Ziebart, and J Andrew Bagnell. 2021. Feedback in Imitation Learning: The Three Regimes of Covariate Shift. (Feb. 2021). arXiv:2102.02872 [cs.LG] http://arxiv.org/abs/2102.02872

Thanh-Trung Trinh, Dinh-Minh Vu, and Masaomi Kimura. 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 (Cambridge, United Kingdom) *(ICISS '20)*. Association for Computing Machinery, New York, NY, USA, 115–120. https://doi.org/10.1145/3388176.3388187

Igor Vozniak, M Klusch, A Antakli, and Christian A Müller. 2020. InfoSalGAIL: Visual attention-empowered imitation learning of pedestrian behavior in critical traffic scenarios. *Int Jt Conf Comput Intell* (2020), 325–337. https://doi.org/10.5220/0010020003250337

Received May 2024