2022 31st IEEE International Conference on Robot and
Human Interactive Communication (RO-MAN)
August 29 - September 2, 2022. Naples, Italy.

2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN) | 978-1-7281-8859-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/RO-MAN53752.2022.9900676

Observer-Aware Legibility for Social Navigation

Ada V. Taylor[1], Ellie Mamantov[2], and Henny Admoni[1]

***Abstract*— We designed an observer-aware method for creating navigation paths that simultaneously indicate a robot's goal while attempting to remain in view for a particular observer. Prior art in legible motion does not account for the limited field of view of observers, which can lead to wasted communication efforts that are unobserved by the intended audience. Our *observer-aware legibility* algorithm directly models the locations and perspectives of observers, and places legible movements where they can be easily seen. To explore the effectiveness of this technique, we performed a 300-person online user study. Users viewed first-person videos of restaurant scenes with robot waiters moving along paths optimized for different observer perspectives, along with a baseline path that did not take into account any observer's field of view. Participants were asked to report their estimate of how likely it was the robot was heading to their table versus the other goal table as it moved along each path. We found that for observers with incomplete views of the restaurant, observer-aware legibility is effective at increasing the period of time for which observers correctly infer the goal of the robot. Non-targeted observers have lower performance on paths created for other observers than themselves, which is the natural drawback of personalizing legible motion to a particular observer. We also find that an observer's relationship to the environment (e.g. what is in their field of view) has more influence on their inferences than the observer's relative position to the targeted observer, and discuss how this implies knowledge of the environment is required in order to effectively plan for multiple observers at once.**

I. Introduction

Legible robot motion is path planning in a manner that clarifies the robot's objective in order to support human interaction, allowing an observer to infer the robot's intent more confidently and quickly [2], [17]–[19]. Previous approaches to legibility focus on methods of defining highly-informative paths, but often assume total path visibility, prioritize intent-expressiveness early in the path [10], or are designed for manipulators picking up objects rather than navigation.

Our key insight is that *legible motions need to be visible to the intended observer*, otherwise their implicit message may not be successfully perceived and these added efforts will be wasted. This requires modeling the perception of the intended observer. This approach will become increasingly necessary as robots operate in larger, occluded environments and observers are unable to maintain a complete view.

In this paper, we introduce *observer-aware legibility*, which incorporates a model of the observer's field of view (FOV) into a formulation of legibility for social navigation. Observer-aware legibility can be used to create robot paths that are both easier to see and clear in their destination.

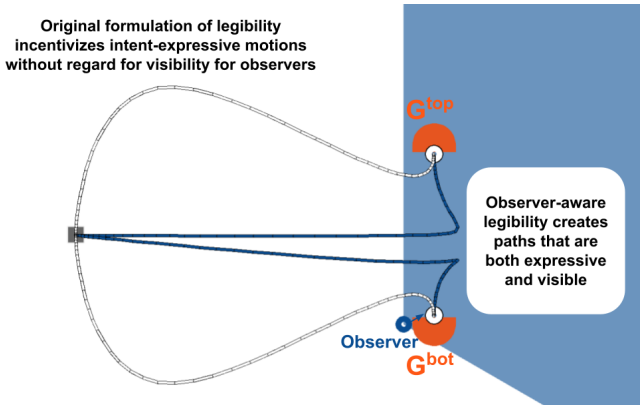


Fig. 1: *Paths of a server robot in a restaurant moving from start location on the left to one of two tables G^{top} and G^{bot}. An observer in blue is seated at G^{bot} and looking to the right, with their field of view in light blue extending towards the right. Paths in white use the original "omniscient" legibility formulation that assumes complete vision of the restaurant. Paths in blue are generated with observer-aware legibility, personalizing their movements to be both expressive and in-vision for the observer.*

We are motivated by the scenario of a robot server at a restaurant [9], [24]–[26] telegraphing its intent to visit a goal table with multiple unique observers. This social navigation scenario provides a use case where the observer fields of view are limited, pre-computable, and strongly affect their ability to infer robot intent. However, observer-aware legibility is appropriate for many other scenarios where a human and robot are collaborating, such as warehouse fulfillment, autonomous delivery, and sidewalk navigation.

Our goal is to generate approach trajectories for the robot that are more legible to a target observer, as measured by how well observers can infer the robot's goal over time and the length of the period they are ready for its arrival. To aid in a more detailed comparison of legibility, we propose three new metrics to assess observer understanding of legible paths: *envelope of readiness*, *clarity*, and *moments of confusion*.

We deployed a 300-person online user study examining how *observer-aware* paths are perceived by both the target observer and all secondary observers at the same table. We compared performance to the original "omniscient" formulation of legibility [8], [22], that does not account for observers' fields of view, and find that observer-aware

*This work is supported in part by the National Science Foundation (#IIS1755823, #DGE1745016), the Sony Group Corporation, and Sony AI.

[1]Ada Taylor and Henny Admoni are with the Carnegie Mellon Robotics Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA adat@andrew.cmu.edu, hadmoni@andrew.cmu.edu

[2]Ellie Mamantov is with Yale Computer Science, Yale University, New Haven, Connecticut, USA ellie.mamantov@yale.edu

legibility is effective for observers with an incomplete view of the scene.

However, we also recognize that personalizing movement for one observer may be confusing to secondary observers. Therefore, we also measured how observer-aware paths are perceived by non-targeted observers. We find that secondary observers have lower performance when observing paths created for others, the natural consequence of specialization.

To further understand the impact of observer-aware paths on secondary observers, we also investigated if performance for secondary observers can be predicted from just their relationship to the target observer (i.e., how much overlap there is in their fields of view). We found that the performance of secondary observers is influenced not just by their physical relationship but also by the secondary observer's own view of the scene, which indicates that knowledge of the scene is also required in order to plan for multiple observers at once.

II. Related Work

While our work is focused on a restaurant context, we believe that these metrics and findings are broadly applicable to legibility within several other navigation contexts.

Robots in Restaurants While currently limited in use, robots have already been deployed in commercial restaurant settings [26]. Among those deployed for serving tasks, the majority of their utility comes from traveling to and from tables: taking orders [25], transporting food to aid a human waiter [9], or interacting with guests [24]. An effective legible approach in this context could enable better service by dynamically informing path planning, or simply optimizing these static paths for better customer and waiter experience.

Social Navigation Work in social navigation has aimed to enable robots to be human-aware while navigating a space and improve a robot's ability to navigate through crowds of pedestrians [3]. Humans attempt to predict others' actions and motions by assessing their intentions in the given context [7]. Therefore, generating intent-expressive paths for robots may allow people and robots to more fluently navigate in the same space. For example, motion conveying a robot's plan for avoiding an oncoming person leads to both the robot and human being able to successfully predict the other's motion [20]. Studies have shown that increasing the legibility of a robot's motions decreases human effort and increases pedestrian trust in the robot [6]. These scenarios highlight both that there is a need for generalizable legibility appropriate for navigation, and that crowded navigation scenarios such as a sidewalk present important secondary observers.

Legible Motion Legible robot motion is path planning in a manner that clarifies the robot's objective in order to support human interaction. These motions are designed to allow a human observer to infer the robot's intent more confidently and quickly [2], [17]–[19]. It has been shown that path trajectories that do not match humans' expectations but convey the motion's goal or intent are more legible, and allow a human observer to infer the robot's intent swiftly and accurately [8]. Legibility has also been extended to other socially communicative domains such as pointing [12].

Prior Assumptions of Legibility Most contemporary legible motion planners assume that human observers are "omniscient"- that they are aware of all changes in the robot's configuration as it moves. While momentary occlusions or projections into a 2D-perspective have been considered, these assume a default-visible scene with only minor interruptions to vision [22]. Work to create a generalized formula for legibility found drawbacks to the early-path exaggerated movements of the original Dragan et al. formulation [5]. Deployment-specific context has also been shown to affect interpretation of legibility, which supports that legibility for navigation could present unique challenges [4], [30].

Limited Fields of View Our algorithm focuses on the limited fields of view of observers within the same scene, rather than the more general per-room models used in search and rescue [15], use of line of sight obfuscation [14], or the relative size of objects in view [27]. It does not default to the assumption the robot is visible from beginning of the path [22] and can choose a different moment to enter view.

Additional Applications The concept of combining one of many nonverbal path-based communication strategies [28] with our *observer-aware* model of when these signals can be perceived could be extended to apply to other signals such as passing direction [6], disapproval [1], hesitation [21], or confidence [13].

III. Observer-Aware Legibility Formulation

Consider a robot that is navigating an area to one of several possible goals G in $\mathcal{G}$. An observer is watching this trajectory to see which goal the robot is heading to. Our research aim is to plan a path for the robot that is maximally *legible* for the observer, i.e. gives them as much information as possible about the robot's goal for as long of a period as possible.

To formalize the problem, we parameterize the robot's path over time t as $\xi_{0\to T}$. We also assume a human inference model $P(G|\xi)$ which gives the observer's chance of predicting goal G from the specific trajectory, and use as a baseline the "omniscient" legibility $\mathcal{L}_o$ created by [8], [22] as follows:

$$\mathcal{L}_o(\xi) = \frac{\int P(G|\xi_{0\to t}) f(t) dt}{\int f(t) dt} \tag{1}$$

This equation averages across the entire path the probability assigned by an omniscient observer to the true goal given the path observed so far. Weighting function $f(t)$ controls the importance placed on this estimation at different points on the path. This is set to $f(t) = T - t$, with the goal of maximizing the clarity of goal inference as early as possible.

A. Modeling Limited Fields of View

In a navigation context, motions often take place over a much broader space than observers can see. In Eqn. 1, weighting function $f(t)$ assigned earlier points greater importance than those later in the path.

However, incentivizing early legible movements is not useful if observers cannot see these movements. Thus, we weight the importance of each point based on its *visibility*.

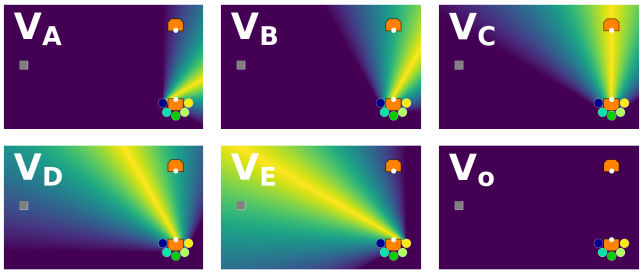


Fig. 2: *Top-down view of visibility $\mathcal{V}_i$ at all points in the restaurant for each observer $O_i = \{A, B, C, D, E\}$ seated at bottom right table G^{bot}, as well as omniscient visibility $\mathcal{V}_o$. $\mathcal{V}_A$ and $\mathcal{V}_B$ have a particularly poor view of the left side.*

We model each observer O_i as located at point p_i and gazing in direction given by unit vector u_i. Let θ_{FOV} be our selected field of view (in our case, 120°) [11]. The visibility of a point p with respect to observer O_i is defined as

$$\mathcal{V}(O, p) = \max\left(1 - \frac{|\theta(O,p)|}{\frac{1}{2}\theta_{FOV}}, 0\right) \tag{2}$$

where $\theta(O, p)$ indicates the angle between gaze vector u_i and the vector from p_i to p,

$$\theta(O, p) = \arccos\left(\frac{p - p_i}{|p - p_i|} \cdot u_i\right) \tag{3}$$

Intuitively, points along the path which are outside the observer's field of view will have a visibility of zero. Once inside the viewable range, visibility linearly increases as an object moves closer to the center of an observer's field of vision, with a score of one at the center. This enables a smooth transition between in vision and out of vision behavior. We can see how this works for every point in the restaurant from each of our observers in Figure 2.

Now, unseen movements will no longer contribute to legibility, and movements central in vision are preferred.

To incorporate this into assessment of legibility, we weight the importance of points by their visibility $\mathcal{V}(O, \xi_t)$. This weighting is no longer time-dependent. We normalize by the maximum possible visibility for the path, which is T for 100% visibility. This guarantees $\mathcal{L} \mapsto [0, 1]$.

Therefore we define observer-aware legibility as:

$$\mathcal{L}_{oa}(O, \xi) = \frac{1}{T}\int P_{oa}(G|\xi_t) \cdot \mathcal{V}(O, \xi_t)dt \tag{4}$$

B. Impact of Limited Visibility on Likeliness of a Goal

In a navigation scenario with limited visibility, we cannot assume that observers are able see prior parts of the path. Including this history in our model of inference is particularly harmful for observers who can only see late moments in the path: out-of-sight history is incorrectly given credit for impacting their inference. We instead only assume that observers have knowledge of the static start and candidate goal locations, not the entire path so far.

We define a local version of the human inference model for the path so far, $P_{oa}(G^*|\xi_{S\to Q})$, that depends only on the point Q along the path where the robot is currently observed. The probability of a goal based on the current point is computed from the efficiency of arriving at that point on the way to goal G compared to to other goals in set $\mathcal{G}$. Given that $\hat{\xi}_{A\to B}$ represents the shortest path from A to B, C is the cost of a path, and S is the start location, this is defined to be:

$$P_{oa}(G|\xi_{S\to Q}) \propto \frac{\exp\left(-C(\hat{\xi}_{S\to Q}) - C(\hat{\xi}_{Q\to G})\right)}{\exp\left(-C(\hat{\xi}_{S\to G})\right)} \tag{5}$$

This equation is similar to the likeliness of a goal in [8], but considers the relative efficiency of arriving at a point from the start, rather than along a particular path. This heuristic does not require observers have total knowledge of the path.

This incidentally makes our formulation compatible with classic path planning algorithms such as A^*, no longer violating the Markov property by relying on knowledge of the path so far, nor therefore requiring an additional dimension for search.

IV. Experimental Design

Our overall aim is to investigate if our modifications to observer-aware legibility improve the ability of observers to infer the goal of each path in time for service. We anticipate observer-aware legibility will enhance performance for the targeted viewpoint, but that performance will fall off for other increasingly dissimilar viewpoints.

A. Experimental Scenario

To investigate the effect on limited fields of view on legibility, we created a restaurant scenario with a single choice: "is the robot approaching me, or the other table?"

To gain a better understanding of the relationship between viewpoint and path, we will be presenting a variety of paths, and viewing these paths from the locations of five observers O_A, O_B, O_C, O_D, and O_E, with orientations 30°, 60°, 90°, 120°, and 150°, all seated around the same circular table seen in Fig 1 labeled G^{bot}. The other table is referred to as G^{top}. All observers are able to see both goal locations, to avoid requiring that users remember out of sight goal locations.

Independent Variables We will vary the following:

Goal: If the robot is approaching G^{bot} or G^{top}.

Observer: Which of the five observers O_A, O_B, O_C, O_D, O_E is viewing the scene.

Path: The path ξ observed. ξ_A, ξ_B, ξ_C, ξ_D, and ξ_E each maximize observer-aware legibility for their matching observers, and baseline ξ_o maximizes $\mathcal{L}_o$.

Dependent Variables Our dependent variables are *clarity*, *envelope of readiness*, and *moments of confusion*. These are determined from user estimates of the likeliness that the robot is approaching each goal, with definitions in section IV-C.

This comes out to 60 stimuli shown to each participant, and is within-subjects.

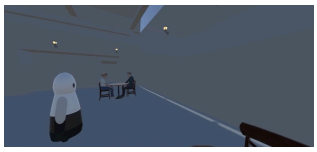


Fig. 3: O_A *simulator view.*

Fig. 4: O_E *simulator view.*

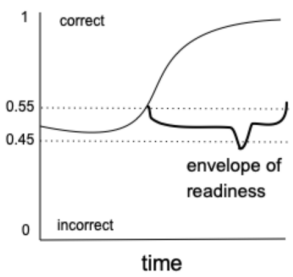


Fig. 5: *Envelope of Readiness: length of the final period where the observer is correct.*

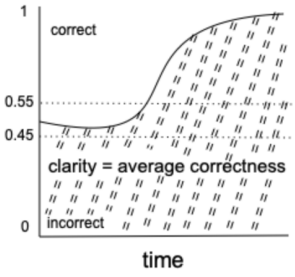


Fig. 6: *Clarity: Average correctness of observer's guess over the path.*

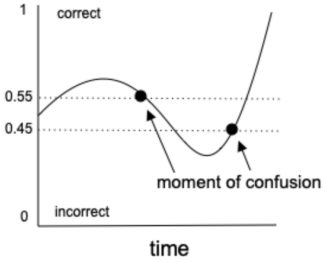


Fig. 7: *Moments of confusion: times the observer re-enters the unsure zone.*

B. Deployment

We have developed a 3D restaurant simulator seen in Fig. 3 and 4. The robot waiter was a scaled up Kuri [16], and we played a video of each of our 2D paths in this 3D environment, captured from the perspective of each observer.

For each stimuli, participants were asked to report on a slider their confidence that the robot was approaching G^{top} or G^{bot}. To ensure participants continuously reported this, the video only played while participants held the slider control.

We deployed the study online, which consisted of a tutorial introducing participants to the restaurant's layout and the mechanics of the slider, the 60 trials, and a final questionnaire. Paths were deployed in a randomized order.

Ecological Validity Our simulation environment was intended to be as realistic as possible, and included both a restaurant interior and animations of other patrons eating. We used a locked view with $\theta_{FOV} = 120°$, intended to account for human field of view including peripheral vision [11]. In the wild, an observer could move their head, but the goal of this study is to first examine legibility from a single viewpoint. Future applications can investigate dynamic views, but that will first require understanding static viewpoints.

C. Metrics

We introduce new metrics for legibility that take advantage of the continuous nature of our data collection: *envelope of readiness*, *clarity*, and *moments of confusion*. They are applicable to any dataset with multiple datapoints of user certainty over the path, and expand on state of the art [29].

These metrics target the goals of intent-expressiveness: we want observers to have a high overall understanding of the robot's goals, and be ready when the robot arrives.

Envelope of Readiness (EoR) We define "*envelope of readiness*" as the final percent of a path an observer is certain of the robot's goal without wavering. This is the period for which an observer is correct and ready for robot arrival.

To quantify this, we divide slider responses into three discretized zones: *unsure*, *correct*, and *incorrect*. Based on our pilot studies, *unsure* is defined as +/- 5% from the middle of the slider. *Correct* and *incorrect* are on corresponding sides of this boundary. Our *envelope of readiness* is the length of the final period for which the observer is *correct* (Fig 5).

top	O_A	O_B	O_C	O_D	O_E	O_o
ξ_o	25%	43%	68%	100%	100%	**100%**
ξ_A	**44%**	50%	61%	100%	100%	100%
ξ_B	44%	**50%**	63%	100%	100%	100%
ξ_C	38%	51%	**66%**	100%	100%	100%
ξ_D	25%	45%	70%	**100%**	100%	100%
ξ_E	26%	39%	65%	100%	**100%**	100%

bot	O_A	O_B	O_C	O_D	O_E	O_o
ξ_o	25%	29%	32%	41%	100%	**100%**
ξ_A	**43%**	47%	57%	100%	100%	100%
ξ_B	44%	**50%**	59%	100%	100%	100%
ξ_C	43%	48%	**59%**	100%	100%	100%
ξ_D	36%	43%	57%	**100%**	100%	100%
ξ_E	28%	35%	49%	100%	**100%**	100%

TABLE I: *Percent of time visible for each path and observer. Bolded values indicate the observer targeted by each path.*

This metric improves on state of the art by decreasing the value of paths that are confusing in the middle [29], and verifies if early guesses remain accurate. By counting from the end of the path instead of the beginning, this metric accounts for consistent accurate inference through the path's critical final approach. We expect observers to be correct by the end of paths when seeing the robot arrive.

Clarity We define "clarity" as the overall confidence observers have of the correct goal over the path. We assess this by mapping 100% correct confidence to one, and incorrect to zero. This value is averaged over the path (Fig 6).

Moments of Confusion (MoC) We pinpoint confusing times during paths by defining "moments of confusion" as when a participant re-enters *unsure* after exiting that zone (Fig 7). These timestamps can be used to localize confusing sections of the path, and multiple can occur within a path.

D. Path Selection

Our paths were selected via a sampling approach, rather than direct optimization. We generated a series of paths from the start to each of the goals with the constraint that they had a final approach angle that is perpendicular to the approach table with no collisions. This mimics the behavior of an actual waiter, and provides a final approach that all viewers can see. These paths were segmented to reflect the physics of a real-world wheeled robot executing the path.

For each observer O_i we select the path ξ_i with the highest observer-aware legibility for that observer. These paths can be viewed in Figure 8.

To gain intuition for how observer-aware legibility creates an improved amount of time in sight, Table I shows the percent of time that each path is within-view for each observer. Note that ξ_o is not visible for much of O_D, but

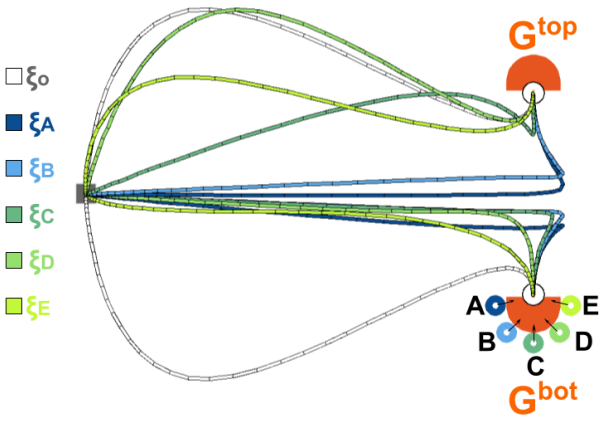


Fig. 8: *The complete set of paths tested, with paths ξ_i corresponding to observer O_i and colored accordingly. White paths ξ_o are created using omniscient legibility $\mathcal{L}_o$.*

that observer can see it at the beginning of the path.

E. Hypotheses

Our hypotheses seek to first determine the value of observer-aware legibility, then investigate how these paths are perceived by non-targeted secondary observers.

H1: Observer-aware legibility improves an observer's ability to see the path, and thus overall legibility.

H2: Performance falls off for non-targeted observers. An observer's neighbors will perform worse than their personalized observer-aware baseline when looking at paths made for a different observer.

H3: Neighbors with an equal offset from the target observer will not have the same performance; the observer with a better view of the area the robot approaches from will perform better.

V. Results

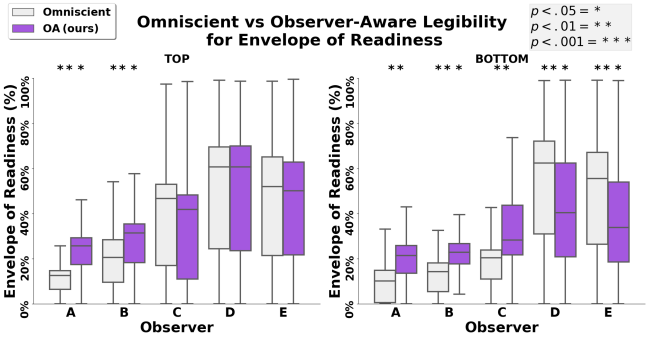
We recruited 300 participants (126 Female, 162 Male, 9 Nonbinary, 3 Other) through Prolific [23]. The age of participants ranged from 18 to 49 (M=24.3, SD=5.14). Our research was approved by CMU's institutional review board, and participants were paid US$5 for the 20-minute survey. Under 1% of individual trials did not successfully log; users without complete sets of trials for a condition were excluded from corresponding repeated measures (RM) ANOVAs.

For each hypothesis, separate statistical tests were performed for independent sets ξ^{top} and ξ^{bot}. For all statistically significant initial results, we conducted a post hoc using a paired-samples t-test with Bonferroni correction.

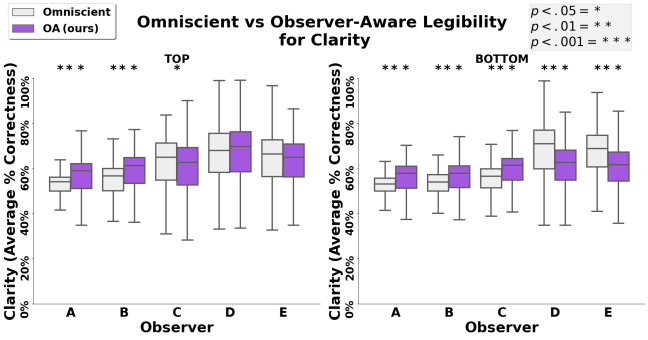
A. H1: Omniscient vs Observer-Aware Legibility

For each goal location, we conducted a two-way repeated measures ANOVA between observers and the two path personalizations (PP): omniscient (ξ_o) and *observer-aware* legibility targeted at each observer.

EoR. Among ξ^{top}, both observer ($F(4, 1104) = 150.2$, $p < .001$) and PP ($F(1, 276) = 4.5$, $p < .05$) had a main effect on EoR. There was an interaction effect between observer and PP ($F(4, 1104) = 10.4$, $p < .001$). Among ξ^{bot}, only observer had a main effect ($F(4, 1112) = 197.1$, $p < .001$), not PP. There was an interaction effect between observer and PP ($F(4, 1112) = 49.3$, $p < .001$). Post hoc test results are in Fig 9a.



(a) *Envelope of Readiness.*



(b) *Clarity.*

Fig. 9: *Results for H1: Observer-aware legibility improved both EoR and clarity performance for O_A, O_B, and O_C^{bot}, but was not effective for $O_{D,E}^{bot}$ with more complete views. Clarity was also lower for O_C^{top}, a relatively complete view.*

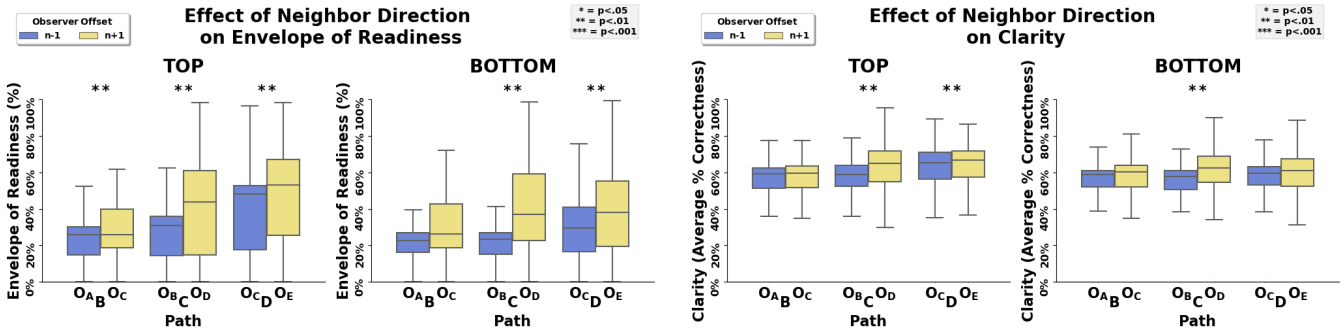
Clarity. Among ξ^{top} only observer ($F(4, 1104) = 76.1$, $p < .001$) had a main effect on clarity, not PP. There was also an interaction effect between observer and PP ($F(4, 1104) = 9.2$, $p < .001$). Among ξ_{bot}, only observer had a main effect ($F(4, 1112) = 99.5$, $p < .001$), not PP. There was an interaction effect between observer and PP ($F(4, 1112) = 27.7$, $p < .001$). Post hocs are in Fig 9b.

MoC. Among ξ^{top} no main effect or interaction effect for PP was found. Among ξ^{bot}, observer ($F(4, 1112) = 5.4$, $p < .001$) and PP ($F(1, 278) = 19.4$, $p < .001$) had main effects, and there was an interaction effect between observer and PP ($F(4, 1112) = 4.6$, $p < .001$). Among ξ^{bot}, post hoc tests showed that ξ_o had fewer moments of confusion than three observer-aware paths: O_A (.34 vs .59, $p < .05$), O_B (0.38 to 0.81, $p < .001$), and O_E (0.18 to 0.31, $p < .05$).

B. H2: Non-Targeted Observers.

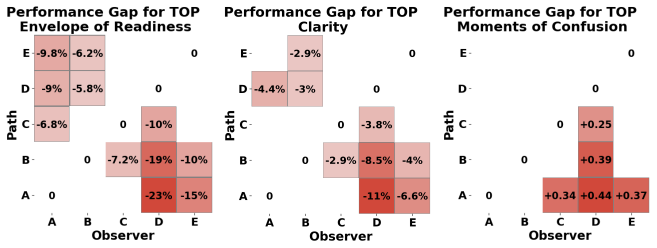
Our hypothesis is only about interaction effects, so we only report these results of the RM ANOVA.

EoR. Among ξ^{top}, there was an interaction effect between observer and path target ($F(16, 4160) = 28.1$, $p < .001$). Among ξ^{bot}, there was an interaction effect between observer



(a) *Envelope of Readiness* (b) *Clarity*

Fig. 10: *Results for H3: Performance for observers to the left and right of the observer that the path is personalized for. Despite both observers having the same offset from the target, the right observer consistently outperforms the left, likely due to having more of the scene within their FOV. Paths for a central observer are not equally effective for both neighbors.*



(a) *EoR* (b) *Clarity* (c) *MoC*

Fig. 11: *Results for H2: Difference in performance for secondary observer on paths targeted at a different observer. Values along the diagonal are by definition zero. Intuitively, paths designed for other observers are less effective than those targeted for the viewing observer.*

and path target ($F(16, 4112) = 3.4$, $p < .001$), but post hoc tests did not show statistical significance.

Clarity. Among ξ^{top}, there was an interaction effect between observer and path target ($F(16, 4160) = 21.9$, $p < .001$). Among ξ^{bot} there was an interaction effect between observer and path target ($F(16, 4112) = 2.3$, $p < .001$), but posthocs revealed no results of relevance.

MoC. Among ξ^{top} and metric MoC there was an interaction effect between observer and path target ($F(16, 4160) = 3.2$, $p < .001$). Among ξ^{bot} no interaction effect was found.

For all significant interaction effects found, G^{top} for each metric, significant post hoc test results are visualized in Figure 11.

C. H3: Neighbor Offset Direction

To assess H3, we examine observers with a neighbor offset on either side: $O_{B,C,D}$. For each index, we examine the impact on performance for a secondary observer offset to either direction, left and right. For B this is O_A and O_C, for C this is O_B and O_D, and for D this is O_C and O_E.

EoR. Among ξ^{top} both index ($F(2, 568) = 76.8$, $p < .001$) and offset ($F(1, 284) = 66.9$, $p < .001$) had a main effect on EoR. There was also an interaction effect between index and offset ($F(4, 1104) = 10.4$, $p < .001$). Among ξ^{bot}, both index ($F(2, 566) = 21.8$, $p < .001$) and offset ($F(1, 283) = 101.8$, $p < .001$), and there was an interaction effect between index and offset ($F(2, 566) = 13.3$, $p < .001$). Post hoc test results can be seen in Fig 10a.

Clarity. Among ξ^{top} both index ($F(2, 568) = 33.7$, $p < .001$) and offset ($F(1, 284) = 23.2$, $p < .001$) had a main effect on clarity. There was also an interaction effect between index and offset ($F(2, 568) = 4.2$, $p < .05$). Among ξ^{bot} both index ($F(2, 566) = 6.3$, $p < .01$) and offset ($F(1, 283) = 38.6$, $p < .001$) had a main effect, and there was an interaction effect between index and offset ($F(2, 566) = 6.1$, $p < .01$). Post hoc results are in Fig 10b.

MoC. Among ξ^{top}, there was an interaction effect between index and offset ($F(2, 568) = 3.1$, $p =< .001$). No significant results were found among ξ^{bot}, nor in post hoc tests for ξ^{top}.

D. Takeaways

H1 is partially accepted. *Observer-aware legibility* is effective for observers with a limited view of the scene. For observers with a complete view of the scene, observer-aware legibility was not as effective.

H2 is partially accepted. Performance for non-targeted observers was often lower than for their own on G^{top}.

H3 is accepted. Neighbors to the right have consistently better performance than those to the left, despite both having the same offset in angle and location from the target observer.

E. Qualitative Feedback

Table II presents a summary of the qualitative feedback gathered in the post-study questionnaire. Participants frequently commented on the robot's gaze direction, explaining that the robot's inability to make eye contact with guests at its intended table was a source of confusion. Participants said that it would be easier to understand the robot's behavior if "... the face turn[ed] in the direction of the table to which the robot [would] go" and if it had "facial expressions and eye contact". Another participant noted that when "the robot was looking in another way, [it] made it a cold and impersonal experience, even when it stopped at my own table".

If you were a server, how would you make your path clear?	#	How did you determine which table the robot was approaching?	#	Hard to understand about robot motion	#
Eye contact or gaze	156	Direction of movement	158	Sudden changes in direction	89
Direct/straight path	88				
Nonverbal indication other than eye contact	14	Eyes or where it was looking	92	The robot's eye-gaze or heading	41
Verbal indication	11	Robot's distance from goal tables (path efficiency)	63		
Follow lines on floor	7				

TABLE II: *Frequency of participants' comments. Totals do not sum to 300 due to multiple comments.*

Similarly, participants reported being surprised by the robot turning, especially in combination with confusion surrounding the robot's gaze. This was exemplified by comments such as "[it was surprising] when [the robot] was headed to a table and suddenly turned to another" and "[it was surprising] when [the robot would] look at me when it was to go to the other table". This supports the idea that heading is a primary signal for inference in social navigation, rather than distance from table.

VI. Discussion

Overall, we found that observer-aware legibility was most effective for disadvantaged viewpoints such as O_A and O_B and O_C^{bot}. For better viewpoints such as O_D^{bot}, O_E^{bot}, and O_C^{top}, observer-aware legibility was less effective.

Secondary observers found paths targeted toward others less clear than the path personalized to them. This indicates that we are correctly tailoring paths to each observer, but also that personalizing a path to one observer means it is unlikely to be equally effective for others.

We also found that performance for secondary observers is not just dependent on their relationship to the target observer, and requires understanding their view of the environment.

A. Audience-Aware Legibility with Multiple Observers

While our formulation of observer-aware legibility allows us to model how multiple individual observers with limited views may perceive the same path, developing an *audience-aware* algorithm for creating paths that are maximally legibile for multiple observers is more complex than targeting a single observer. Positive H2 results indicate observer-aware planning for a single observer does not generalize to all other potential observers. Positive H3 results imply we cannot create a multiple-observer formulation without considering observers' relationships to the scene, not just each other. Thus, a policy for modeling multiple viewpoints at once must incorporate more than just observers' relative locations.

To adapt to multiple viewpoints, our legibility formulation needs to consider multiple fields of view, but it is not clear how to combine them. Summing strong viewpoints with weak ones can lead to paths that neglect weak viewpoints. On the other hand, overvaluing the most restricted viewpoint can lead to degraded performance for observers with wide views. As a result, audience-aware legibility may need to decide when to prioritize certain observers over others.

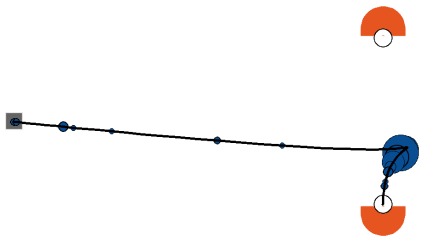


Fig. 12: *Moments of confusion for all O_A viewing ξ_A, count denoted by dot size. The cluster of MoC after the sharp turn may mean this change of heading is misleading despite Eqn. 5 indicating a high likelihood of G^{bot} throughout.*

B. Heading-Based Goal Inference

The original omniscient and new observer-aware legibility formulations both calculate the human inference model $P(G|\xi)$ based on path efficiency. However, in social navigation, we propose that a formulation of $P(G|\xi)$ based on heading may be more effective. Small deviations from the centerline between the goals can strongly impact Eqn. 5, but are often unclear to observers. This is supported by comments in V-E. This may also explain why despite performance improving for targeted observers, MoC also increased. We suggest that the sharp turn of these paths led to users feeling the robot was going to pass them despite it always being closer to G^{bot} than G^{top}, as seen in Fig. 12.

C. Impact of Preferring Centrality Within FOV

When comparing ξ_E^{top} and ξ_E^{bot} in Fig. 8, it seems surprising that these paths are so different from each other and ξ_o. Observer-aware legibility (Eqn. 4) is calculated by maximizing centrality within the field of view, $\mathcal{V}$, combined with likelihood of goal, $P(G|\xi)$, over the entire path. In this case, $P(G|\xi)$ pushes paths away from the centerline between the goals, while $\mathcal{V}$ incentivizes moving directly to the center of the observer's sight. In the ξ_E^{top} case these work together, but for ξ_E^{bot} these are in opposition and lead to a less exaggerated path.

This suggests for complete viewpoints it might be more effective to change visibility to a binary variable indicating if a point is in sight. The drawback is this can create an abrupt change in curvature when reaching in-vision areas, which may lead to confusion or mistrust from observers.

A model of goal inference based on heading as described in VI-B could also prevent this issue, as ambiguous headings below the centerline would no longer be considered clear.

D. Additional Applications

Although this work focused on creating a very visible legible robot, we could imagine inverting this to disincentivize being in view. An unobtrusive "butler" or "ninja" robot could plan paths which avoid visibility completely, or only appear when their clarity is sufficiently unambiguous. One could also imagine a "town crier" styling built on these

equations where the robot's path is planned so as to engage the attention of all the observers in an area, then proceed to a most-visible area to deliver a performance or message.

While our simulation environment is most effective at testing the informational aspects of these paths, an in-person study would be able to characterize the visceral and emotional reactions observers have to these different interaction styles. Opinions may change if observers perform a distractor task such as eating, rather than solely focusing on the robot.

Now that $P(G|\xi_t)$ is uncoupled from path history, it can be assessed for static locations in the restaurant. This can enable us to assess a restaurant layout and fields of view within it for how potentially informationally dense or ambiguous they are for observers, which could inform design for traffic, or placement of screens to constrain observer focus.

VII. Contributions

In this paper, we developed an algorithm for observer-aware legibility that is appropriate for limited viewpoints, and we assessed it using novel metrics relevant to legibility within social navigation contexts such as *envelope of readiness*, *clarity*, and *moments of confusion*. We also created a method for gathering continuous data about user inferences throughout a robot's path. We performed a 300-person user study of observer-aware legibility studying the experience of different targeted observers in a restaurant context. Our study verified that observer-aware legibility is most effective for observers with limited perspectives, but may have drawbacks for observers with more complete viewpoints of the scene. We showed that assessing the relative performance of different observers must take into account their relationships to the environment, which means *audience-aware* legibility must consider both observers and their environment.

References

[1] Santosh Balajee Banisetty and Tom Williams. Implicit communication through social distancing: Can social navigation communicate social norms? In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, pages 499–504, 2021.

[2] Michael Beetz, Freek Stulp, Piotr Esden-Tempski, Andreas Fedrizzi, Ulrich Klank, Ingo Kresse, Alexis Maldonado, and Federico Ruiz. Generality and legibility in mobile manipulation. *Autonomous Robots*, 28(1):21, 2010.

[3] Aniket Bera, Tanmay Randhavane, Rohan Prinja, and Dinesh Manocha. Sociosense: Robot navigation amongst pedestrians with social and psychological constraints. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 7018–7025. IEEE, 2017.

[4] Christopher Bodden, Daniel Rakita, Bilge Mutlu, and Michael Gleicher. Evaluating intent-expressive robot arm motion. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 658–663. IEEE, 2016.

[5] Christopher Bodden, Daniel Rakita, Bilge Mutlu, and Michael Gleicher. A flexible optimization-based method for synthesizing intent-expressive robot arm motion. *The International Journal of Robotics Research*, 37(11):1376–1394, 2018.

[6] Yuhang Che, Allison M Okamura, and Dorsa Sadigh. Efficient and trustworthy social navigation via explicit and implicit robot–human communication. *IEEE Transactions on Robotics*, 36(3):692–707, 2020.

[7] Gergely Csibra and György Gergely. 'obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans. *Acta psychologica*, 124(1):60–78, 2007.

[8] Anca D Dragan, Kenton CT Lee, and Siddhartha S Srinivasa. Legibility and predictability of robot motion. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 301–308. IEEE, 2013.

[9] Pantelis Elinas, Jesse Hoey, Darrell Lahey, Jefferson D Montgomery, Don Murray, Stephen Se, and James J Little. Waiting with josé, a vision-based mobile robot. In *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No. 02CH37292)*, volume 4, pages 3698–3705. IEEE, 2002.

[10] Jaime F Fisac, Chang Liu, Jessica B Hamrick, Shankar Sastry, J Karl Hedrick, Thomas L Griffiths, and Anca D Dragan. Generating plans that predict themselves. In *Algorithmic Foundations of Robotics XII*, pages 144–159. Springer, Cham, 2020.

[11] Riad I Hammoud. *Passive eye monitoring: Algorithms, applications and experiments*. Springer Science & Business Media, 2008.

[12] Rachel M Holladay, Anca D Dragan, and Siddhartha S Srinivasa. Legible robot pointing. In *The 23rd IEEE International Symposium on robot and human interactive communication*, pages 217–223. IEEE, 2014.

[13] Julian Hough and David Schlangen. It's not what you do, it's how you do it: Grounding uncertainty for a simple robot. In *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI*, pages 274–282. IEEE, 2017.

[14] Anagha Kulkarni, Siddharth Srivastava, and Subbarao Kambhampati. A unified framework for planning in adversarial and cooperative environments. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 2479–2487, 2019.

[15] Anagha Kulkarni, Siddharth Srivastava, and Subbarao Kambhampati. Planning for proactive assistance in environments with partial observability. *arXiv preprint arXiv:2105.00525*, 2021.

[16] KuriRobot. Kurirobot/kuri-documentation: Documentation for kuri the adorable home robot.

[17] Christina Lichtenthäler and Alexandra Kirsch. Towards legible robot navigation-how to increase the intend expressiveness of robot navigation behavior. 2013.

[18] Christina Lichtenthäler and Alexandra Kirsch. Goal-predictability vs. trajectory-predictability: which legibility factor counts. In *Proceedings of the 2014 acm/ieee international conference on human-robot interaction*, pages 228–229, 2014.

[19] Christina Lichtenthäler and Alexandra Kirsch. Legibility of robot behavior: a literature review. 2016.

[20] Christoforos I Mavrogiannis, Wil B Thomason, and Ross A Knepper. Social momentum: A framework for legible navigation in dynamic multi-agent environments. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pages 361–369, 2018.

[21] AJung Moon, Boyd Panton, HFM Van der Loos, and EA Croft. Using hesitation gestures for safe and ethical human-robot interaction. In *Proceedings of the ICRA*, pages 11–13, 2010.

[22] Stefanos Nikolaidis, Anca Dragan, and Siddhartha Srinivasa. Based legibility optimization. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 271–278. IEEE, 2016.

[23] Stefan Palan and Christian Schitter. Prolific. ac—a subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17:22–27, 2018.

[24] Ronald PA Petrick and Mary Ellen Foster. What would you like to drink? recognising and planning with social states in a robot bartender domain. In *CogRob@ AAAI*, 2012.

[25] Sakari Pieskä, Markus Liuska, Juhana Jauhiainen, Antti Auno, and Delitaz Oy. Intelligent restaurant system smartmenu. In *2013 IEEE 4th International Conference on Cognitive Infocommunications (CogInfoCom)*, pages 625–630. IEEE, 2013.

[26] Sakari Pieska, Mika Luimula, Juhana Jauhiainen, and Van Spiz. Social service robots in wellness and restaurant applications. *Journal of Communication and Computer*, 10(1):116–123, 2013.

[27] Jakob Reinhardt and Klaus Bengler. Design of a hesitant movement gesture for mobile robots. *PloS one*, 16(3):e0249081, 2021.

[28] Shane Saunderson and Goldie Nejat. How robots influence humans: A survey of nonverbal communication in social human–robot interaction. *International Journal of Social Robotics*, 11(4):575–608, 2019.

[29] Sebastian Wallkotter, Mohamed Chetouani, and Ginevra Castellano. A new approach to evaluating legibility: Comparing legibility frameworks using framework-independent robot motion trajectories. *arXiv preprint arXiv:2201.05765*, 2022.

[30] Min Zhao, Rahul Shome, Isaac Yochelson, Kostas Bekris, and Eileen Kowler. An experimental study for identifying features of legible manipulator paths. In *Experimental robotics*, pages 639–653. Springer, 2016.