Legible Robot Pointing

Rachel M. Holladay[1] Anca D. Dragan[2] Siddhartha S. Srinivasa[2]

***Abstract*— Good communication is critical to seamless human-robot interaction. Among numerous communication channels, here we focus on gestures, and in particular on spacial deixis: pointing at objects in the environment in order to reference them. We propose a mathematical model that enables robots to generate pointing configurations that make the goal object as clear as possible — pointing configurations that are *legible*. We study the implications of legibility on pointing, e.g. that the robot will sometimes need to trade off efficiency for the sake of clarity. Finally, we test how well our model works in practice in a series of user studies, showing that the resulting pointing configurations make the goal object easier to infer for novice users.**

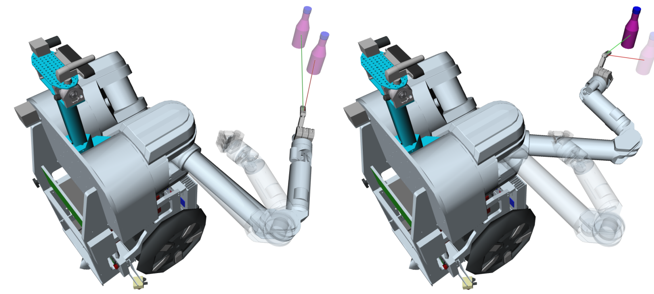


Fig. 1: Left: An efficient pointing configuration that fails to clearly convey to an observer that the goal is the further bottle. Right: Its less efficient, but *more legible* counterpart, which makes the goal clear.

I. Introduction

In a world of increasing robot functionality, communication has become key to successful human-robot interactions. Communication can entail explicit verbal statements [1–3], or nonverbal cues through motion [4–6], gaze [7, 8], or gestures [9–12].

Among many communicative actions, here we focus on spacial deixis — on producing *pointing* gestures. Regardless of language and culture, we rely on pointing to refer to objects in daily interactions [13], be it at the grocery store, during a meeting, or at a restaurant.

Imagine pointing at one of the objects on a table. This pointing configuration has to accomplish two tasks: (1) it has to convey to the observer that you are pointing at the goal object, and (2) it has to convey that you are *not* pointing at any other object.

Myopically deciding on a pointing configuration that ignores this second task can lead to the situation in Fig.1(left), where even though the robot's pointer is directly aligned with the further bottle, it is unclear to an observer which of two objects is the goal. It is the second task, of not conveying other goals, that ensures the clarity – or *legibility* – of the pointing gesture.

With this work, we introduce a mathematical model that enables robots to point legibly, producing pointing configurations like the one in Fig.1(right). Although this configuration is not as efficient, requiring the robot to move further from its starting configuration, it is more legible: it sacrifices efficiency to make the goal object clear.

The problem of generating pointing configurations has been studied in robotics as an inverse kinematics problem of aligning an axis with a target point [14], or a visually-guided alignment task [10]. Here, we explicitly focus on finding an axis that will make the target object most clear, analogously to work on legible motion [5, 6, 9, 15–17] or handovers [18].

Legible *pointing* has been a focus in the computer graphics community [19]. There, it is possible to go beyond the physical constraints of robots and augment a character's configuration with virtual features, such as extending the character's arm to the goal object [20], or visually highlighting the object [21]. Here, we focus on legible pointing while constrained by the physical world.

Our work proposes a mathematical model for generating legible pointing configurations for robots, studies its implications for the way a robot points, and puts it to the test in a series of three studies.

A Mathematical Model for Legible Pointing: If the robot in Fig.1 had a laser ray going out of its index finger and landing on the target object, then both configurations would be perfectly legible. In reality though, there is ambiguity about the pointing direction. We do not have the accuracy of laser pointers — not in forming pointing configurations, and definitely not in observing them. What we have is more akin to a torch light, shooting rays in a range of directions.

Starting with such a ray model, we create a cost function that balances two expectations: (1) that the pointing configuration will point directly at the target and not be occluded by obstacles, and (2) that the pointing configuration is efficient, in that it stays close to the robot's starting configuration.

Such a model leads to the configuration in Fig.1(left), where the central rays emerging from the robot's pointer do hit the goal object, and are not occluded by the obstacle. Yet, the configuration is illegible: this model focuses on the first task, of conveying the goal object, but fails to capture the second, of not conveying other goals.

To address this issue, we use this ray-based cost function to derive a legibility reward function that explicitly accounts for the probability that the observer will assign to the other potential goal objects in the scene. The best pointing configuration is the one that maximizes the probability of the

[1] Rachel Holladay is with the School of Computer Science, Carnegie Mellon University, rmh@andrew.cmu.edu

[2] Anca Dragan and Siddhartha Srinivasa are with the Robotics Institute, Carnegie Mellon University, {adragan,siddh}@cs.cmu.edu

goal object, at the same time minimizing the joint probability of other goals.

Implications: Next, we discuss the implications of legibility on how the robot points, for both the position of the pointer, as well as the orientation. A main implication is that efficiency has no bearing on legibility: the robot will spare no expense in order to make the goal object as clear as possible, including getting closer to it, or even (counter-intuitively) exaggerating the orientation of its pointer away from other objects.

User Study Evaluation: We run a series of three user studies to test the legibility pointing model. Our results do support that users can more easily identify the robot's intent, but they also suggest that the observer's viewpoint plays a crucial role in this process. Legibility optimization never worsens how clear a pointing configuration is, but the extent of the improvement varies with the viewpoint. Thus, incorporating the observer's viewpoint into the legibility model is a main area of future improvement.

Further along, we are excited to explore our legible pointing model's applicability further, including to other channels. In particular, we believe legible *gaze* will share a lot of similarities with legible pointing, especially for robots with neck articulations with can move their head to a position that enables them to more clearly indicate where they are looking. Finally, we look forward to exploring how these channels could be combined to make human-robot interactions and collaborations more seamless.

II. Pointing as Cost Optimization

We begin by modeling pointing as the minimum of a cost function based on rays that shoot out from the pointer and intersect the goal objects, or get blocked by obstacles in the scene.

Formally, the robot is in a starting configuration, $S \in \mathcal{Q}$, and needs to point at the goal object G within a set of objects $\mathcal{G}$. The robot must find a pointing configuration $P \in \mathcal{Q}$. We model finding this pointer as an optimization problem.

The natural human end effector shape when pointing is to close all but the index finger [22, 23], which serves as the pointer. We assume the robot's end effector is in some equivalent shape, as in Fig.1. Let $\phi(P)$ denote the transform of the pointer when the robot is in configuration P.

We expect a good pointing configuration to satisfy the following trivial properties: (1) the pointer must be directly oriented towards the goal object; (2) there should be no obstacles in between the pointer and the goal object.

We design a cost function for pointing such that the minima satisfy these properties, and deviating from them is more and more expensive. To this end, we propose a *ray model* as in Fig.2, where ray vectors r shoot out from the pointer. Rays that do not contact the goal object are assigned no weight. Rays that contact the goal object are assigned a higher weight when they are more aligned with the pointer $\phi(P)$:

$$R_G(P) = \frac{\int \delta(P, r, G) w(r) dr}{\int w(r) dr} \tag{1}$$

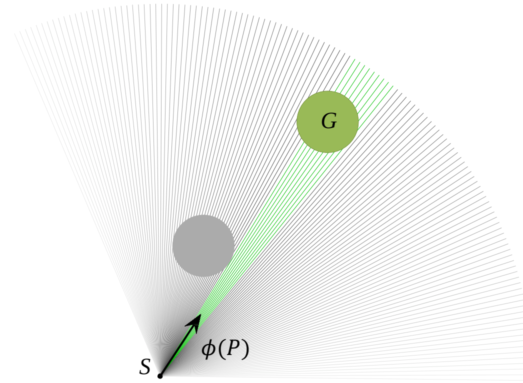


Fig. 2: The ray model only takes into account rays that hit the object, weighing them more when they are more aligned with the pointer.

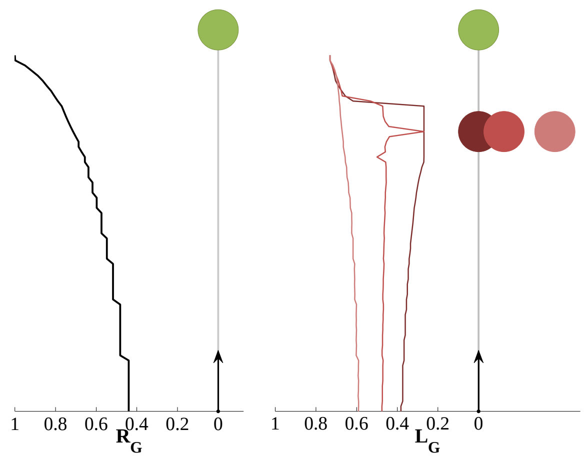


Fig. 3: Left: R_G as a function of the distance to the goal object. Right: L_G as a function of the distance, with different other object positions in the way.

with w increasing with the dot product between the pointer and the ray, and δ a function evaluating to 1 when the ray at angle r intersects the goal object, and 0 otherwise.

However, simply accounting for the ray intersections does not tell the whole story. As the pointer $\phi(P)$ moves closer the the goal G, more rays intersect and therefore R_G increases – see Fig.3(left). This would imply that the best pointing position would be to be as close to the object as possible to the point of touching it.

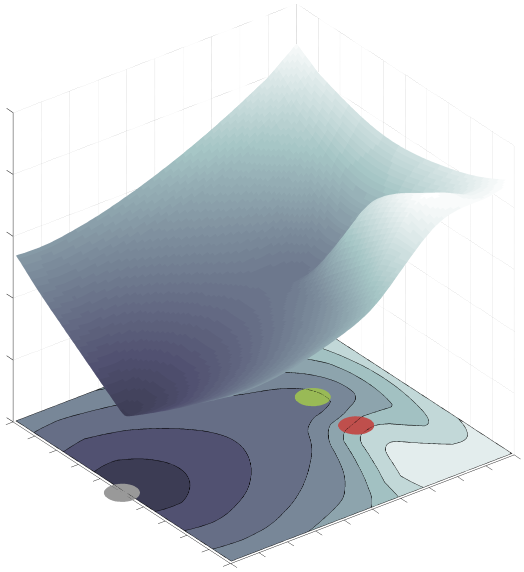
In contrast, humans observing agents tend to apply the principle of rational action, expecting them to take efficient actions to achieve their goals [24]. In the case of pointing, this implies we expect robots to not deviate too much from their starting configuration. Thus, we model the cost of a pointing configuration as the trade-off between a high reward R_G and moving the minimal distance from the start:

$$C_G(P) = (1 - R_G(P)) + \frac{\lambda}{M} ||S - P||^2 \tag{2}$$

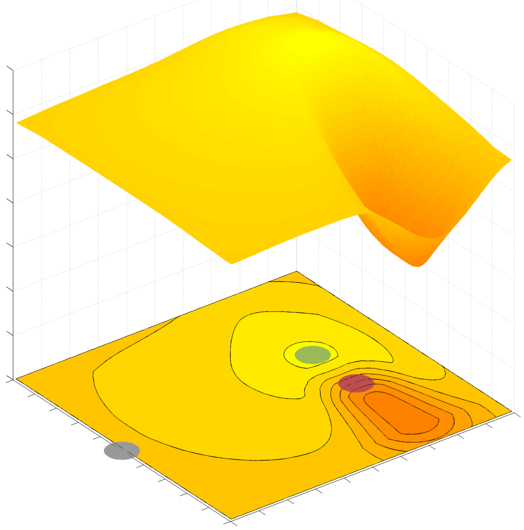
with

$$M = \max_{p \in \mathcal{Q}} ||S - p||^2 \tag{3}$$

Fig.4a plots this cost for all positions in a 2D grid, assuming the direction of the pointer is aligned with the goal object (in green). There is a large increase in cost around the other



(a) Cost C_G surface



(b) Legibility L_G surface

Fig. 4: Surface plots for the cost model, and for the corresponding probability of the goal object.

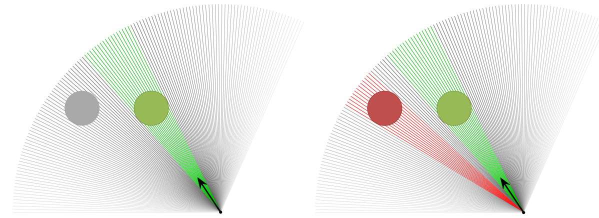
object (in red), because this objects starts blocking the rays when the pointer is in those positions.

III. Legible Pointing

Optimizing $C_G(P)$ only accounts for the goal object and any obstacles that might be in the way. However, it fails to account for other candidate goal objects in the scene, as indicated in Fig.5.

As a result, the expected pointer could create an ambiguous situation if another object is sufficiently aligned with the goal: pointing to the goal may also seem to point at this other object.

Therefore, in order to generate a pointer that clearly points at the goal *we need to take into account other objects*. For a given pointing configuration, we form a probability distribution over candidate objects and search across it to determine the pointing configuration that makes the actual goal the most probable.



(a) Cost (b) Legibility

Fig. 5: While the ray model only treats other objects in the scene as obstacles, the legibility model directly accounts for the probability that an observer will assign to them given a pointing configuration.

We start with a probability distribution over pointing configurations given an object. Based on our cost function C_G and the principle of maximum entropy [25], we define a probability distribution analogous to [6]:

$$P(P|G) \propto e^{-C_G(P)} \tag{4}$$

Using Bayes' Rule we can then compute the probability of an object given a pointing configuration:

$$P(G|P) \propto P(P|G)P(G) \tag{5}$$

We compute a legibility reward by normalizing this formula over all candidate objects in the scene, $\mathcal{G}$:

$$L_G(P) = P(G|P) = \frac{e^{-C_G(P)}}{\sum_{g\in\mathcal{G}} e^{-C_g(P)}} \tag{6}$$

Fig.4 shows a comparison between this probability (L_G) and the cost C_G.

Fig.3(right) plots L_G as a function of distance to the goal for different locations of another object in the scene. In the middle location, the object is tangent to the trajectory of the pointer as it moves towards the goal: at that point, all rays are blocked by this other object, leading to a sudden drop in the probability.

The most legible pointer maximizes the probability of the correct goal:

$$P^* = \max_{p\in\mathcal{Q}} P(G|p) = \max_{p\in\mathcal{Q}} L_G(p) \tag{7}$$

To find P^* in high-dimensional spaces, we can perform gradient ascent on the legibility reward L_G:

$$p \leftarrow p + \alpha\nabla L_G \tag{8}$$

with

$$\nabla L_G = \frac{\sum_{g\in\mathcal{G}} e^{-(C_G(p)+C_g(p))}[\nabla C_g(p) - \nabla C_G(p)]}{[\sum_{g\in\mathcal{G}} e^{-C_g(p)}]^2} \tag{9}$$

The gradient has an intuitive interpretation. For each other object $g \in \mathcal{G}$, we update the pointing configuration such that the cost increases for that object (by following the positive gradient direction $\nabla C_g(p)$), and the cost of G decreases (by following the negative gradient direction $-\nabla C_G(p)$.

Since the ray model is inherently non-analytical and thus requires numerical differentiation, making the gradient computation for C_g more efficient in high-dimensional spaces is still an area of future research, which we discuss in Sec. VI.

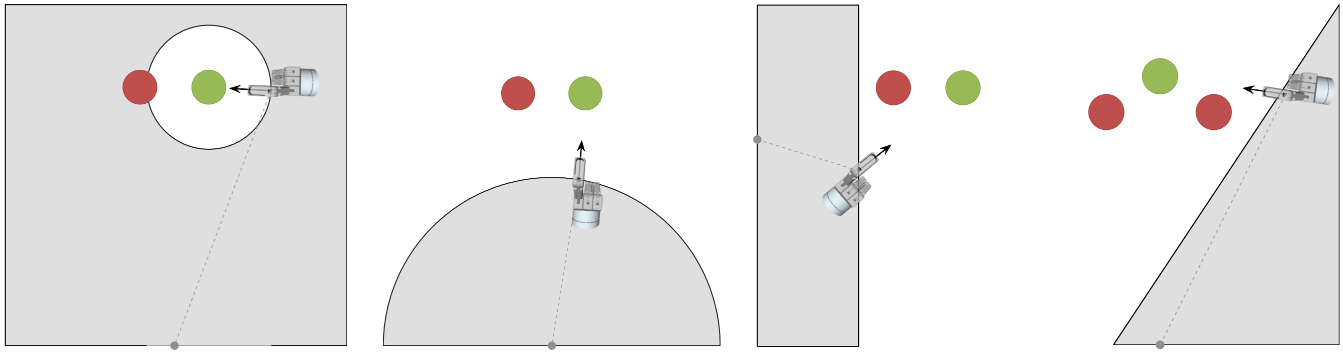


Fig. 6: The legible pointer to the green object for different constraints.

IV. Implications

Optimizing for legibility, as opposed to for C_G, has several implications.

A main implication is that the distance from the starting configuration S becomes inconsequential: $P(G|P)$ does not depend on the distance to S.

Proof: $P(G|P) = \frac{e^{-C_G(P)}}{\sum_{g\in\mathcal{G}} e^{-C_g(P)}} \Rightarrow P(G|P) = \frac{e^{-(1-R_G(P))+\frac{\lambda}{M}||S-P||^2}}{\sum_{g\in\mathcal{G}} e^{-(1-R_g(P))+\frac{\lambda}{M}||S-P||^2}} \Rightarrow P(G|P) = \frac{e^{-(1-R_G(P))}}{\sum_{g\in\mathcal{G}} e^{-(1-R_g(P))}}$ ■

This happens because the purpose of legibility is to find the absolute clearest pointing configuration, even if that requires more effort: legibility will spare no expense in making the goal object clear.

As a result, the optimal pointing configuration is different from the optimum with respect to C_G. And because legibility incorporates the probability of the other potential goals in the scene, the resulting pointing configuration is also different from simply using the ray model only, R_G – both in position and orientation (we detail this is Sec. IV-C and Sec. IV-C).

As with motion, legibility can result in too much inefficiency (analogous to over-exaggeration), and the optimization for legibility can be constrained to a trust region [6].

A. Pointer Position Under Different Constraints

Fig.6 illustrates the the most legible pointer given that the position is constrained to the shaded area, and the orientation is constrained to direct towards the goal object in green.

The first situation is a constraint arising when the pointer is not allowed to touch the goal object: often times when we point, we avoid contact with the object even when it is within reach. In this case, the robot is pointing from the right, because from this position all the rays that would hit the red object are blocked by the goal object: the legible pointer uses the goal object to block all ray vectors that would hit the other object.

The second situation corresponds to a reachability constraint: the robot cannot travel past a certain distance from the starting configuration S. The third is a constraint akin to a "glass wall": the robot has to point from one side of the wall to the next, and the red object blocks the green, causing the robot to have to move significantly to increase legibility. Finally, the last situation is a multiple objects one with the same type of constraint.

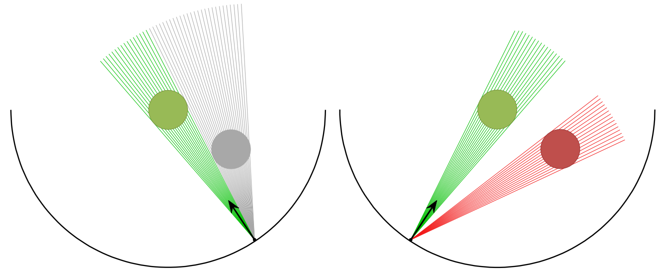


Fig. 7: Legibility is different from the ray model because it accounts for the probability that will be assigned to the other objects. In this example, both pointers are equally good according to the ray model, because the other object does not occlude either pointer. However, the pointer in right the right image is more legible. We put this to the test in practice in our last experiment.

B. Difference Between Legibility and Ray Model in Position

Even though legibility does not account for distance, it is different from solely using the ray model R_G, because it accounts for other candidate goal objects. We create an illustrative example in Fig.7, where we constrain the position of the pointer to a fixed distance to the goal object.

The figure shows two different pointers. They both have the same ray value R_G, because in both cases the other object does not block any rays that would normally hit the goal.

However, the pointer in the left image is much less legible because it does not account for the probability an observer would assign to the other object. In contrast, the pointer on the right is the result for optimizing L_G, and makes the intended goal much more clear.

C. Pointer Orientation

The orientation of the pointer is also different for legibility. Depending on the weighting function w for the rays, instead of pointing directly at the target as both R_G and C_G would suggest, a legible pointer can angle away from the other objects in order to decrease their probability relative to the goal object. We test this orientation exaggeration strategy in one of our experiments in the next section.

V. Experimental Results

We test our pointing model in a series of experiments with HERB [26], a bi-manual mobile manipulator with two Barrett WAMs that moves on a Segway RMP.

(a) Cost (View 1) (b) Legibility (View 1) (c) Cost (View 2) (d) Legibility (View 2)

Fig. 8: The four experimental conditions for our main study, which manipulates legibility and observer viewpoint.

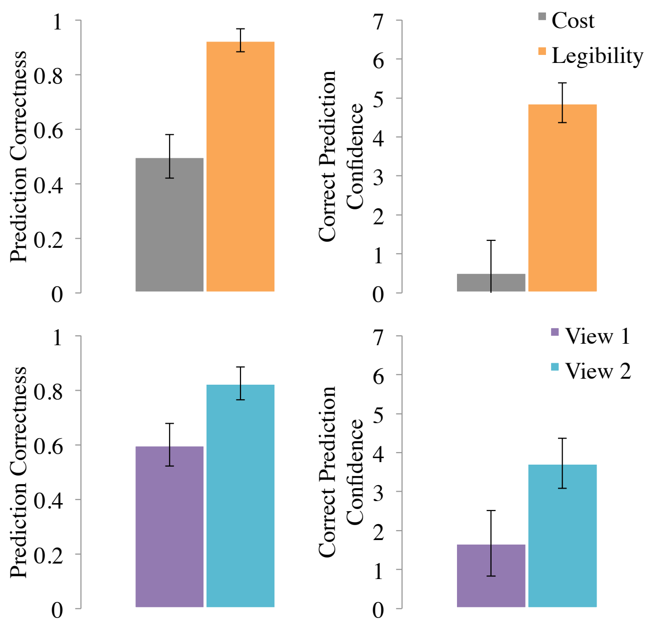


Fig. 9: Effects of legibility (top) and viewpoint (bottom) on correctness of predictions (left), and correct prediction confidence (right).

A. Main Study: Cost vs. Legibility

Our main study compares how clearly a pointing configuration conveys its goal object for the cost and legibility optimizations, testing our model's prediction that maximizing legibility will be more effective than minimizing cost.

Manipulated Factors: We manipulate *legibility* – whether the pointing is generated by minimizing the cost C_G from (2) or by maximizing the legibility score L_G from (6). For efficiency, we perform the optimization in a restricted space of pointers, where the pointer is constrained to point directly at the goal object (we explore effects of orientation exaggeration in a side study), and the optimization over position happens in the 2D plane, constrained by the robot's arm reachability.

We also manipulate the *viewpoint*. The point of view of the observer can change the perception of geometry. To control for this potential confound, we test two different opposite view points, one from the right of the robot and the other from the left.

We use a factorial design, leading to a total of four conditions, shown in Fig.8.

Dependent Measures: We measure how clearly the pointing configuration expresses its goal object (as opposed to other objects in the scene).

We show the participants (an image of) the robot pointing, and ask them to 1) select which of the two objects on the table the robot is pointing at (the objects are labeled in the images) — we use this to measure *prediction correctness*, and 2) rate their confidence on a 7-point Likert scale — we use this to measure *correct prediction confidence* by computing a score equal to the confidence for correct predictions, and equal to the negative of the confidence for incorrect prediction (i.e. we penalize being confidently wrong).

We also ask participants to rate how *expected* or natural the robot's pointing configuration is, on a 7-point Likert scale, since the cost minimization was designed to better match the expectation of efficiency, while the legibility optimization was designed to be more clear about which object is conveyed.

Hypotheses:

H1. *Legibility positively affects prediction correctness and correct prediction confidence.*

H2. *Legibility negatively affects expectedness.*

Subject Allocation: We opted for a between-subjects design in order to avoid biasing the participants. This is especially important because all conditions have the same target object, and seeing one pointer affects the prior over what the robot is pointing at.

We recruited 20 participants per condition (leading to a total of 80 participants) using Amazon's Mechanical Turk. We imposed two selection criteria for the participants: a high acceptance rate on their previous work to avoid participants who are not carefully considering the task, and a US location to avoid language barriers.

Analysis: In line with our first hypothesis, a logistic regression on *prediction correctness* with *legibility* and *viewpoint* as factors revealed a significant main effect for *legibility* (Wald $\chi^2(1, 80) = 12.68$, $p < .001$): *legible pointing was indeed more legible* (or clear) than minimizing the pointing cost. The *viewpoint* factor was marginal ($\chi^2(1, 80) = 2.86$, $p = .09$), with the first viewpoint leading to worse predictions.

With *correct prediction confidence*, the differences were all the more clear. A factorial ANOVA also showed a significant main effect for *legibility* ($F(1, 76) = 21.86$, $p < .0001$), and also one for *viewpoint* ($F(1, 76) = 4.85$, $p = 0.03$). The interaction effect was only marginal ($F(1, 76) = 64.8$, $p = .057$).

Fig.9 plots the two measures for each factor. We see that legibility increase both measures, but increases the confidence score more, and that it has a larger effect than the viewpoint. Our data also revealed that *legibility optimization is less susceptible to viewpoint changes than cost optimization*: for the legible pointing, the mean difference between viewpoints for confidence is only 0.25, compared to 3.85 for the cost minimization.

Looking at the rating for how expected or natural the pointing configuration is, we found that the second hypothesis was only supported for one of the easier viewpoints (view 2). An ANOVA revealed only a significant interaction effect ($F(1,76) = 12.8$, $p = .028$), with the Tukey HSD post-hoc analysis showing that for the second viewpoint (which led to large differences for the cost minimization configuration), the cost minimization configuration was significantly more expected than the legible configuration ($p = .0446$).

This was not true for the first viewpoint, where the cost minimization rating was much lower than for the first viewpoint, *despite the actual configurations being identical*. This shows the importance of viewpoints: an expected/natural configuration from one viewpoint can seem unnatural from a different viewpoint. Our conjecture is that this happens because certain viewpoints deem the cost minimization output too unclear.

B. Side Studies

We ran two follow-up studies exploring smaller, more subtle differences: one focused on testing the effectiveness of orientation exaggeration (the implication of legibility from Sec. IV-C), and the other on the difference between legibility and directly using the ray model, from Sec. IV-B.

Because these differences are more subtle, we opted for within-subjects designs, which would enable participants to make direct comparisons. As a result of this, we changed our dependent measures: we explicitly told participants which object the robot is pointing at, and asked how clearly each pointing configuration indicates that object.

Orientation Exaggeration: While our main study was about making pointing legible by altering position, here we alter the angle at which the robot is pointing, and explore orientation exaggeration.

Based on the results of our main study, we select the viewpoint that puts the cost minimization at an advantage, making it more clear. We then produce a pointing configuration that exaggerates the orientation slightly, and show participants the two images side by side for a comparison.

Our hypothesis is that the legible pointing configuration will be rated as more clearly expressing the goal object.

We recruited 20 additional participants for this study, and each participant rated both configurations. Our analysis revealed that the legible configuration was rated higher, but not significantly so ($t(19) = 1.23$, $p = .23$).

As a follow-up, we re-ran the study, but with the robot pointing at the other object (the left instead of the right, as in Fig.10): while this viewpoint is advantageous when the robot is pointing to the right object, it makes pointing to the left object less clear.

Indeed, here we saw a significant difference in the clarity ratings from participants, with the legible (orientation exaggeration) configuration received a significantly higher rating

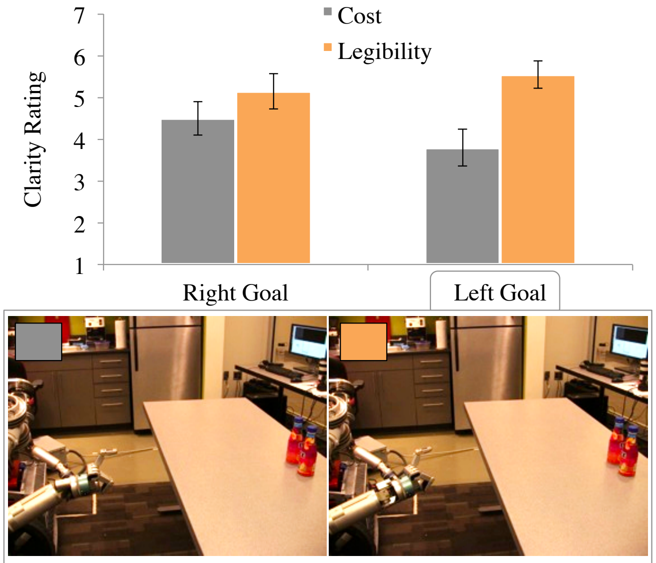


Fig. 10: Legibility by orientation exaggeration. The pictures show the direct way of pointing at the left bottle (left), and the exaggerated way (right).

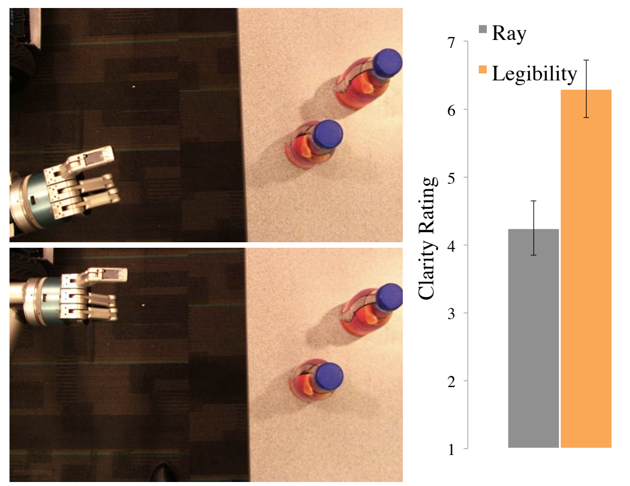


Fig. 11: The difference between the ray model (for which both configurations are identical, since there are no occlusions and the pointer is at the same distance from the goal), and the legibility model, which reasons about the probability assigned to the other object.

($t(19) = 2.68$, $p = .0147$).

Again, we see that legibility tends to help in both cases, but that here the viewpoint seems to be even more important: legibility can help when the viewpoint is favorable, but it has larger advantages with viewpoints that make the direction of pointing more ambiguous. By exaggerating the orientation, the robot can add that extra degree of clarity.

Legibility Is Different From the Ray Model: Our last study explores the difference between the ray model, which reasons about distance from the object and occlusions by obstacles, and the legibility model, which explicitly reasons about the probability that a pointing configuration could be interpreted as pointing to a different object.

In particular, we explore the difference between the two pointers from Fig.7, replicated on the robot in Fig.11: the two configurations are identical from a ray model perspective, but the bottom one is more legible. The distance from the goal object is the same for both pointers.

Our hypothesis is that the second configuration (the legible one) will be rated as more clearly expressing the goal object.

We recruited 20 additional participants to do a side-by-side comparison of the two pointers. A paired t-test showed that the legible pointer was indeed rated significantly higher ($t(19) = 3.82$, $p = .0011$).

This result emphasizes the importance of explicitly reasoning about the other objects in the scene that the observer could potentially interpret as the goal of the pointer — the importance of legibility.

VI. Discussion

We proposed a model for legible robot pointing, analyzed its implications for the way the robot points, and evaluated it in a series of three user studies. We found that optimizing for legibility does make the goal object of the pointer more clear in practice.

Like any user studies, ours too suffer from certain limitations. Key among them is the use of images as opposed to in-person views of the robot: this was a logistical necessity for the between-subjects design, but future work should follow up with a smaller pool of in-person users, and include the full gesture from the starting configuration to the pointing one. However, even using images provided insight into the utility of legibility and the biases introduced by changes in the viewpoint.

A main direction of improvement in the theory, as revealed by our studies, is incorporating the viewpoint directly into the legibility model. For this, the robot needs a model of how the viewpoint should change the weighting of virtual rays emitted by its pointer.

A second direction of improvement is a faster numerical approximation of the gradient of C_G, which remains difficult to compute in high-dimensional spaces.

We are excited to explore these directions, along with the applicability of our model beyond pointing, to legibility of robot gaze: much like with pointing, gaze has direction uncertainty, and can benefit from a more carefully chosen position and orientation of the head.

Acknowledgements

This material is based upon work supported by NSF-IIS-0916557, NSF-EEC-0540865, ONR-YIP 2012, and the Intel PhD Fellowship. We thank the members of the Personal Robotics Lab for helpful discussion and advice.

References

[1] A. Vogel, C. Potts, and D. Jurafsky, "Implicatures and nested beliefs in approximate Decentralized-POMDPs," in *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, (Sofia, Bulgaria), Association for Computational Linguistics, August 2013.

[2] N. D. Goodman and A. Stuhlmüller, "Knowledge and implicature: Modeling language understanding as social cognition," *Topics in Cognitive Science*, vol. 5, no. 1, pp. 173–184, 2013.

[3] R. Deits, S. Tellex, P. Thaker, D. Simeonov, T. Kollar, and N. Roy, "Clarifying commands with information-theoretic human-robot dialog," *Journal of Human-Robot Interaction*, 2013.

[4] C. Breazeal, C. D. Kidd, A. L. Thomaz, G. Hoffman, and M. Berlin, "Effects of nonverbal communication on efficiency and robustness in human-robot teamwork," in *Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on*, pp. 708–713, IEEE, 2005.

[5] L. Takayama, D. Dooley, and W. Ju, "Expressing thought: improving robot readability with animation principles," in *Proceedings of the 6th international conference on Human-robot interaction*, pp. 69–76, ACM, 2011.

[6] A. Dragan and S. Srinivasa, "Generating legible motion," in *Robotics: Science and Systems*, 2013.

[7] B. Mutlu, J. Forlizzi, and J. Hodgins, "A storytelling robot: Modeling and evaluation of human-like gaze behavior," in *Humanoid Robots, 2006 6th IEEE-RAS International Conference on*, pp. 518–523, IEEE, 2006.

[8] H. Admoni, C. Bank, J. Tan, M. Toneva, and B. Scassellati, "Robot gaze does not reflexively cue human attention," in *Proceedings of the 33rd Annual Conference of the Cognitive Science Society, Boston, MA, USA*, pp. 1983–1988, 2011.

[9] M. J. Gielniak and A. L. Thomaz, "Generating anticipation in robot motion," in *RO-MAN, 2011 IEEE*, pp. 449–454, IEEE, 2011.

[10] P. Maes, M. Mataric, J. Meyer, J. Pollack, and S. Wilson, "Self-taught visually-guided pointing for a humanoid robot,"

[11] E. Sato, T. Yamaguchi, and F. Harashima, "Natural interface using pointing behavior for human–robot gestural interaction," *Industrial Electronics, IEEE Transactions on*, vol. 54, no. 2, pp. 1105–1112, 2007.

[12] A. Sauppé and B. Mutlu, "Robot deictics: How gesture and context shape referential communication," 2014.

[13] S. Kita, *Pointing: Where language, culture, and cognition meet*. Psychology Press, 2003.

[14] D. Tolani, A. Goswami, and N. I. Badler, "Real-time inverse kinematics techniques for anthropomorphic limbs," *Graphical models*, vol. 62, no. 5, pp. 353–388, 2000.

[15] A. D. Dragan, K. C. Lee, and S. S. Srinivasa, "Legibility and predictability of robot motion," in *Human-Robot Interaction (HRI), 2013 8th ACM/IEEE International Conference on*, pp. 301–308, IEEE, 2013.

[16] M. Beetz, F. Stulp, P. Esden-Tempski, A. Fedrizzi, U. Klank, I. Kresse, A. Maldonado, and F. Ruiz, "Generality and legibility in mobile manipulation," *Autonomous Robots*, vol. 28, pp. 21–44, 2010.

[17] R. Alami, A. Albu-Schaeffer, A. Bicchi, R. Bischoff, R. Chatila, A. D. Luca, A. D. Santis, G. Giralt, J. Guiochet, G. Hirzinger, F. Ingrand, V. Lippiello, R. Mattone, D. Powell, S. Sen, B. Siciliano, G. Tonietti, and L. Villani, "Safe and dependable physical human-robot interaction in anthropic domains: State of the art and challenges," in *IROS Workshop on pHRI*, 2006.

[18] J. Mainprice, E. A. Sisbot, T. Siméon, and R. Alami, "Planning safe and legible hand-over motions for human-robot interaction," in *IARP Workshop on Technical Challenges for Dependable Robots in Human Environments*, vol. 2, p. 7, 2010.

[19] N. Wong and C. Gutwin, "Where are you pointing?: the accuracy of deictic pointing in cves," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1029–1038, ACM, 2010.

[20] M. Fraser, S. Benford, J. Hindmarsh, and C. Heath, "Supporting awareness and interaction through collaborative virtual interfaces," in *Proceedings of the 12th annual ACM symposium on User interface software and technology*, pp. 27–36, ACM, 1999.

[21] J. Hindmarsh, M. Fraser, C. Heath, S. Benford, and C. Greenhalgh, "Fragmented interaction: establishing mutual orientation in virtual environments," in *Proceedings of the 1998 ACM conference on Computer supported cooperative work*, pp. 217–226, ACM, 1998.

[22] M. Gokturk and J. L. Sibert, "An analysis of the index finger as a pointing device," in *CHI'99 Extended Abstracts on Human Factors in Computing Systems*, pp. 286–287, ACM, 1999.

[23] R. Cipolla and N. J. Hollinghurst, "Human-robot interface by pointing with uncalibrated stereo vision," *Image and Vision Computing*, vol. 14, no. 3, pp. 171–178, 1996.

[24] G. Gergely, Z. Nadasdy, G. Csibra, and S. Biro, "Taking the intentional stance at 12 months of age," *Cognition*, vol. 56, no. 2, pp. 165 – 193, 1995.

[25] B. D. Ziebart, A. Maas, J. A. Bagnell, and A. Dey, "Maximum entropy inverse reinforcement learning," in *AAAI*, 2008.

[26] S. S. Srinivasa, D. Berenson, M. Cakmak, A. Collet, M. R. Dogar, A. D. Dragan, R. A. Knepper, T. Niemueller, K. Strabala, M. Vande Weghe, *et al.*, "Herb 2.0: Lessons learned from developing a mobile manipulator for the home," *Proceedings of the IEEE*, vol. 100, no. 8, pp. 2410–2428, 2012.