Knowledge-based Reasoning for Navigation in Public Spaces

Roya Salek Shahrezaie
Socially Assistive Robotics Group
University of Nevada, Reno
Reno, Nevada, USA
Email: rsalek@nevada.unr.edu

Lotey Betta Manalo
University of Nevada, Reno
Reno, Nevada, USA
Email: bmanalo@unr.edu

David Feil-Seifer
Socially Assistive Robotics Group
University of Nevada, Reno
Reno, Nevada, USA
Email: dave@cse.unr.edu

***Abstract*—Robots' autonomous navigation in public spaces and their social awareness suited to the environmental context is an active investigation in HRI. In this paper, we are presenting a methodology to achieve this goal. While most navigation models focus on objects, context, or human presence in the scene, we will incorporate all three to perceive the environment more accurately. Other than scene perception, the other important aspect of socially aware navigation is the social norms associated with the context. To do so, we have included interviews with museum visitors, volunteers, and staff to gather information about museums and convert the text data to social rules. This effort is currently in progress, we present a framework for future study and analysis of this problem.**

I. Introduction

As we study the use of robotics in public spaces for role-specific functions, such as tour guide or gallery educator, we also want to investigate how to make the social navigation behaviors of those robotics appropriate for these spaces. We have developed a Socially-Aware Navigation (SAN) system to optimize for both navigation performance as well as social performance. This works by using a non-linear optimization over a set of objectives, which can include navigation objectives like "minimize the time to goal," as well as social objectives, like "don't walk too closely to other people" [3]. We have since augmented that approach to by selecting these objectives based on detected features of the environment [1], resulting in changing navigation behavior based on what type of environment it is in (i.e., hallway vs. art gallery). Further enhancements utilize a navigation language ontology to select navigation objectives based on detected objects and people and their relationship to navigation rules in an established knowledge base. [7]

Ongoing research regarding stakeholders of a museum space (docents, tour participants, bystanders, and custodial/maintenance staff) and the general museum experience has highlighted the importance of understanding the art museum environment prior to integrating social robotics. Art museum stakeholders have expressed unanimous interest in the ability to nurture an individualized museum experience. Due to the COVID-19 pandemic, museums have considered integrating social robotics to uphold gallery tours while maintaining social distancing guidelines. There is also a nationwide movement towards promoting Diversity, Equity, Inclusion and Accessibility (DEIA, known alternatively by IDEA and DEAI) in multiple educational institutions. While research participants expressed concern of introducing Socially Assistive Robotics (SAR) into their spaces prematurely, there is an overwhelming consensus that robots can eventually close gaps in limitations such as language barriers and cultural biases. [6]

This poses the need for continued research on understanding human-human interactions in various museum spaces. Once appropriate movements and interactions are identified to acknowledge tour participants, other visitors in the space, and potential collaborative interactions between museum employees and robots, social norms can be created as guidelines for navigating the museum environment. Given that the social norms of various spaces can differ based on a significant number of factors as well as internal cultures, it would be advantageous to be able to navigate spaces based on the established social norms of those places. Some more universal norms might include, "don't walk between a patron and a piece of art they are viewing/near." However, other norms might be space-specific such as, "it is respectful to move more slowly in the room cataloging a traumatic historical event than in the main space." In this case, it would be useful that such information is available for query in an ontology specific to space. Such an ontology could be automatically generated from transcripts of conversations related to the use of space in order to learn such social norms from the people utilizing the space themselves.

II. Background

Robots sharing workspace with people needs to consider human comfort and safety for their long-term acceptance in public places such as hospitals and factory floors. Traditional navigation of mobile robots aims to find the short path to reach its goal without considering if such a performance-oriented path is optimal in terms of social objectives such as human comfort and safety. Recently, researchers incorporated social costs in mobile robot path planning to keep appropriate interaction distance [5], avoiding personal space, avoiding passing behind a person [9], avoiding activity spaces and waiting in a line [3] to approach humans, and having a preferred passing side. There has been a growing interest in knowledge-based methods and their application in robotics, such as socially aware navigation. Semantic awareness has presented new frontiers in robot navigation, enabling more powerful abstraction tools in representing information [2]. A location-based mobile service was developed and evaluated to study an indoor navigation service [10], which helped people navigate physical difficulties. This work uses navigation context to enhance navigation behavior similar to our work, but our application is autonomous robot navigation instead of an online service for people with disabilities. In a similar work [8], the authors propose a knowledge engine that learns and shares knowledge representations for robots to complete various responsibilities.

III. Study Aim

This section describes our design considerations and modeling concepts, together with a context-aware scenario in the gallery to be used to illustrate our context model. Our high-level goals for this work include:

Aim 1: Interview corpus of socially appropriate and inappropriate behaviors and movements in particular museum spaces. This corpus will be used to develop a SAN person-aware ontology that might be specific to various museum spaces. In particular, we will be looking for rules that might be specific to galleries or artworks in order to create behavior for a robot that should be role- or exhibit-specific. This will be used to develop an evaluation plan for particular spaces for candidate robot behavior.

Aim 2: Delivering a data-driven knowledge base representation (ontology) can serve as a socially aware navigation approach capable of making inferences about the behavior necessary in a given environment to observe the social rules specific to space. This ontology will be automatically updated from new images and the corpus described above, see Figure 1. This ontology will enhance the robot's decision-making related to navigation behavior specific to a space when completed.

Aim 3: A socially-aware planner that leverages the above ontology to select appropriate behaviors in museum spaces for appropriate navigation. This system will be evaluated with visitors to the two museums participating in the project and the recently-validated Perception of Social Intelligence (PSI) instrument to assess the navigation performance of the system.

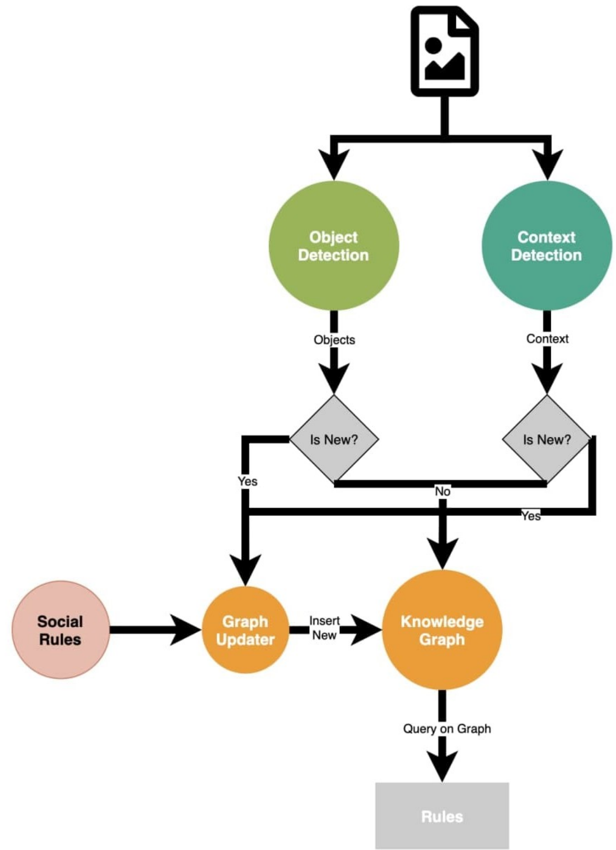

Fig. 1. The ontology update and execution illustration, showing the fellow from input media to detecting context and objects/human in the environment and reasoning on social rules.

IV. Methods

A. Interviews to Establish Social Rules of Museum Spaces

Typically, navigation behavior is established based on a hand-coding of rules for given spaces. These can include appropriate speeds, social distances to maintain with people, how to orient a robot with regard to an object in the environment. There is also the consideration of when to engage a human and how to effectively disengage if the human sends cues of discomfort during the interaction. While such manual coding of rules can be effective, it is likely the case that specific spaces might have specific social rules. Even within a space, different rooms or times might result in different desired navigation behavior, necessitating different navigation rules.

As a preliminary step for realizing autonomous robot behavior in a museum space, we will interview stakeholders of diverse museum spaces to establish what customs exist in the museum environment that a robot might need to be aware of. Interviews would be rooted in open-ended questions to allow for participants to elaborate on why certain social interactions make them feel more or less annoyed or uncomfortable. Candidate questions to facilitate discussion-based responses would include, but are not limited to:

- When it comes to people moving around your museum environment, what social norms might be at play in your exhibits?
- What would make you feel welcomed into a tour guide's space?
- How would you prefer to be treated or acknowledged?
- People leave the tour for various reasons; what speculations can you make or experiences can you share that cause a participant to leave a tour?

A portion of the interview would be focused on human-human interactions, with participants rating their level of discomfort in various scenarios. Then we will transition into human-robot interactions for the same scenarios to understand the participants' comfort with humans in their personal space and gauge their comfort and acceptance of robots in a public space. Possible scale-based questions include, but are not be limited to:

- Rate your level of annoyance if you are viewing artwork and someone walks up and stands right behind you.
- Rate your level of annoyance if you are viewing artwork and someone walks up and tries to engage in a conversation with you.
- Rate your level of annoyance if you are viewing artwork, and a robot tries to engage with you.
- Rate your level of annoyance if you are walking through a gallery, and a robot leading a tour of people intercept your path.

Stakeholders of the museum space acknowledge that their personal bubble relies heavily on the artistic environment, their sight-line of artwork, or their experience in a gallery. Because of this, interviewees will be shown photos and videos of a robot in a gallery or moving in a museum space and other crowded social settings to allow them to visualize a potential robot encounter. These videos will show a robot giving a tour, presenting at an exhibit, and showing people a directory-like map on a screen showing locations of exhibits. We will then ask follow-up questions regarding introducing social robotics into the museum environment and potential benefits they may foresee in such a shift towards blending art and technology on a social level.

Transcribed interviews will then be cleaned and used to create a text corpus for a few different museum spaces detailing social behavior that is appropriate and inappropriate for space. Since these interviews can be conducted remotely, we hope to interview museum stakeholders at museums outside of the Reno area in order to get a more diverse text dataset. Once identifying information is removed, the corpus will be made publicly available for others to use as well.

Expected Outcomes: We expect to find a number of social rules that are consistent between spaces (i.e., appropriate social distances from people). However, we also expect to find a number of social criteria that differ from one space to another. These distinctions can be used to develop space-specific movement behavior.

Evaluation: We will conduct a qualitative evaluation of the interview data to see if the data appear complete. In particular, we want to do a deep examination of the rules of given museum space for scenarios that the robot is expected to operate in (e.g., mobile directory mode, tour-guide mode, and presentation mode). In this case, we will start to develop space-specific navigation criteria for the follow-up study. If no such differentiation among spaces is found, we will conduct follow-up interviews to try to examine space-related behavior in more detail.

B. Knowledge Base Representation

This is the core of our method consist of a few modules where possible information coming from 1. *Context Classification*, 2. *Object/Person Detection*, and 3. *Ontology* for either update knowledge graph or execute related social rule, see figure 1.

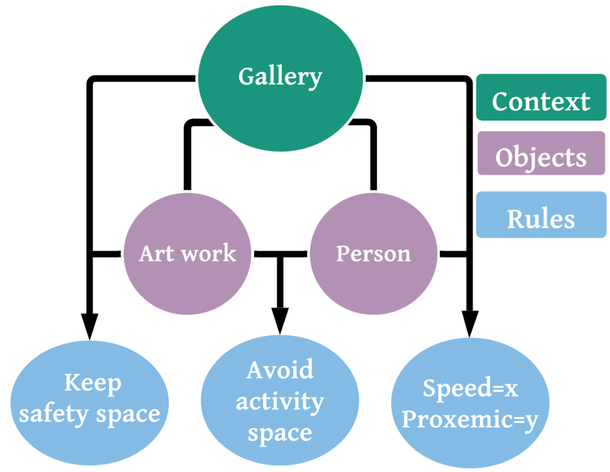

Fig. 2. An illustration of ontology holding rule dependency, showing the relationship between context, objects, and rules. As an example, in an empty gallery, the agent does not observe activity space unless a person and objects like artwork are detected.

1- A context classifier, a trained CNN that distinguishes between environment contexts such as gallery rooms, hallways, and other areas.

2- Understanding the general context category alone is not enough to extract related rules as the objects within the context and the interactions between them play a vital role. For example, in an empty gallery, the agent does not observe activity space unless a person and objects like artwork are detected. YOLOv3 [14] is used to detect objects and persons because more specific rules can be added to a context on top of general contextual social rules by detecting objects and people in the environment.

3- Ontologies are referred to the shared understanding of domains, which are often conceived as a set of concepts, relations, functions, axioms, and instances [12]. An ontology can formalize high-level representations of knowledge of various concepts. We use OWL language to build and expand a knowledge graph of concepts and relationships between them. We use context, concepts, and social rules models associated with them to form an ontology in our approach.
We will query the knowledge base to extract applicable social rules associated with the context with the output label from the context classifier and the objects detected. We will use the objects detected in the environment and their physical relationships to other detected objects, people, and features of the environment to get the associated social rules. We use the SPARQL language to retrieve and manipulate data stored in Resource Description Framework (RDF) format.

Figure 3 on the right presents sample results on images of three categories: an art gallery, hallway, and the vending machine. Image (a) is an art gallery context with a person in it. Our system extracted the rule “do not get too close to the artwork”, “Respects peoples’ personal space”, and “Respects activity zones” as social rules for this situation. Similarly, image (b) is a vending machine (ATMs) situation where the robot would apply general rules like “wait in line”; however, when people are using the ATMs, other specific social rules like respecting privacy and “respecting peoples’ personal space” should be considered. To illustrate this, consider images (c) and (d) both are hallway context; however, one of the images is just a hallway without any people; in this case, the general rule of “stay on the right side” is applicable. In the other case, a hallway context with people in it, a specific rule of “respecting people’s personal space” is also extracted by our system along with the general rule of “stay on the right”, as shown in image (c) and (d).

Expected Outcomes: Our ontology will be able to identify social rules associated with context and objects and people in the context. A sample of related rules based on the observed context and detected objects are presented in the above figure and description.

Evaluation: Several criteria can be used to analyze semantic measures. Some can be studied theoretically, while others require empirical analyses. Among the criteria that are the most frequently considered evaluating semantic measures, we use the accuracy and precision of executed rules, mathematical properties and semantics, and characterization regarding technical details.

Fig. 3. Sample of results on images of three categories: an art gallery, hallway, and the vending machine. To illustrate the importance of human presence, consider images (c) and (d) both are hallway context; however, one of the images is just a hallway without any people; in this case, the general rule of stay on the right side is applicable.

C. Socially-Aware Planner

Most of the recent planners involve social path planning for a single context. We propose a language-based augmentation to our socially-aware navigation planner that can plan and execute trajectories for an autonomously sensed interaction context that adheres to social norms [3]. We use a multi-objective optimization tool called the Pareto Concavity Elimination Transformation (PaCcET) to capture the nonlinear human navigation behavior [13]. We use autonomously sensed distance-based features that capture the social norms and associated social costs for a given trajectory point towards the goal. Rather than use a finely-tuned linear combination of these costs, we use PaCcET to select an optimized future trajectory point, associated with a non-linear combination of the costs. Our approach augments the navigation stack of the Robot Operating System (ROS) utilizing optimization tools. This optimization-based approach optimizes multiple navigation cardinal objectives for a sensed setting to achieve the goals.

We will use the ontology described above to select the multiple cardinal objectives, which should be optimized and guide the navigation behavior of a robot. For example, in larger spaces, we will ensure that we do not get between people and artwork, while in hallways, we will prioritize respecting social norms with regard to staying on the right side of the hallway. In other spaces, we will select behavior derived from the interview dataset described above to derive space- and context-specific behavior.

Expected Outcomes: We expect that socially acceptable navigation in simulation and later in the real world with a wide range of social scenarios will respect considerations specific to the robot’s environment (i.e., places of interest, such as paintings on a wall). We expect that this observation of social norms will make the robot seem to bystanders as more socially intelligent as measured by the perception of social intelligence scale[4] and using social performance metrics.

Evaluation: To evaluate the results of our socially aware planner, we consider the perception of social intelligence (PSI) [11] and consider other social metrics, such as: comfort; sociability; naturalness; safety; legibility; predictability; fluency; overall efficiency; and acceptance. Some of these measures (safety, efficiency, predictability) can be measured directly by the performance of the robot. Other measures will be measured via quantitative survey after participating in an interaction with the robot directly or observing robot behavior in a museum space.

V. Discussion/Future Work

In this research project, we focus on the process of robot decision-making while navigating in public spaces in the presence of humans. For different spaces, the agent has to consider particular social norms. Our model consists of a few modules. Context Classification, Object/Person Detection, Knowledge graph (ontology), and Reasoner. Our ongoing efforts include gathering interviews with museum experts and analyzing text data, building a broader knowledge base using a more extensive dataset such as MIT Indoor Scenes dataset, and finally integrating our knowledge base with an optimization-based social navigation planner and validating our proposed method on a real-world robot.

Acknowledgments

The authors would also like to acknowledge the financial support of this work by the National Science Foundation (NSF, #IIS-1757929, #IIS-1719027)

References

[1] S. B. Banisetty. *Multi-Context Socially-Aware Navigation Using Non-Linear Optimization*. PhD thesis, 2020.

[2] J. Crespo, J. C. Castillo, O. M. Mozos, and R. Barber. Semantic information for robot navigation: A survey. *Applied Sciences*, 10(2):497, 2020.

[3] S. Forer, S. B. Banisetty, L. Yliniemi, M. Nicolescu, and D. Feil-Seifer. Socially-aware navigation using non-linear multi-objective optimization. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1–9. IEEE, 2018.

[4] A. Honour, S. B. Banisetty, and D. Feil-Seifer. Perceived social intelligence as evaluation of socially navigation. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, pages 519–523, 2021.

[5] R. Kirby, R. Simmons, and J. Forlizzi. Companion: A constraint-optimizing method for person-acceptable navigation. In *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication*, pages 607–612. IEEE, 2009.

[6] L. B. Manalo and C. Lynch. Integrating social robotics into the art museum environment: Guidelines for facilitating a meaningful museum experience, 07 2021.

[7] R. Salek Shahrezaie, S. B. Banisetty, M. Mohammadi, and D. Feil-Seifer. Towards deep reasoning on social rules for socially aware navigation. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, pages 515–518, 2021.

[8] A. Saxena, A. Jain, O. Sener, A. Jami, D. K. Misra, and H. S. Koppula. Robobrain: Large-scale knowledge engine for robots. *arXiv preprint arXiv:1412.0691*, 2014.

[9] E. A. Sisbot, L. F. Marin-Urias, R. Alami, and T. Simeon. A human aware mobile robot motion planner. *IEEE Transactions on Robotics*, 23(5):874–883, 2007.

[10] V. Tsetsos, C. Anagnostopoulos, P. Kikiras, and S. Hadjiefthymiades. Semantically enriched navigation for indoor environments. *International Journal of Web and Grid Services*, 2(4):453–478, 2006.

[11] R. E. Walker and J. M. Foley. Social intelligence: Its history and measurement. *Psychological Reports*, 33(3): 839–864, 1973. doi: 10.2466/pr0.1973.33.3.839. URL https://doi.org/10.2466/pr0.1973.33.3.839.

[12] X. H. Wang, D. Q. Zhang, T. Gu, and H. K. Pung. Ontology based context modeling and reasoning using owl. In *IEEE annual conference on pervasive computing and communications workshops, 2004. Proceedings of the second*, pages 18–22. Ieee, 2004.

[13] L. Yliniemi and K. Tumer. Paccet: An objective space transformation to iteratively convexify the pareto front. In *Asia-Pacific Conference on Simulated Evolution and Learning*, pages 204–215. Springer, 2014.

[14] L. Zhao and S. Li. Object detection algorithm based on improved yolov3. *Electronics*, 9(3):537, 2020.