Social Navigation: Context is King
A Play in Three Acts

David V. Lu[1] and William D. Smart[2]

[1] Washington University in St. Louis, St. Louis, MO 63130, USA,
davidlu@wustl.edu,
[2] Oregon State University, Corvallis, OR 97331, USA

Abstract. While many approaches exist to giving context and structure to interactions between humans and robots, one that has gained increasing support in recent years is to view the interactions as theatre. This has the benefits of being strongly communicative in nature as well as giving us the capacity to view the interaction as a complete action that changes over time in a structured way. In this paper, we examine robot navigation in this capacity and discuss how theatrical frameworks can guide the behavior.

1 Introduction

When some people think about how robots move and behave, the characteristics that come to mind are not always flattering, especially given certain media portrayals. Their moves are jerky and have sharp accelerations, as though they were dancing "The Robot." They are unresponsive or slow to respond to the environment around them. They act with very narrow scope, and anything outside of that scope "does not compute." This collection of attributes has come to be represented by the term "robotic" even in non-robotic scenarios. These qualities can be reinforced by the actual behavior of certain types of robots. If the only goal is efficiency, the result is robots programmed to perform specific tasks in efficient ways, resulting in many of the qualities that people have come to expect.

What such work lacks is an appreciation for *context*, i.e. the circumstances of the robot's current situation. The easiest solution is often to only consider the contexts that directly lead to an efficient solution, e.g. what objects are in the way, where is the goal located. With the increasing use of proper human-robot interaction techniques, additional contexts are added to the optimization, such as the effects on nearby people and their impressions of the robot. Taking all contexts into account is likely intractable. To circumvent that, we turn to a multidisciplinary approach to help us sort through a number of different contexts. We use principles derived from the theatre to motivate our work in human-robot interaction. In theatre, all of the physical actions on stage are motivated by contexts, whether in relation to characters' objectives or to the larger structure of the narrative arc.

In this paper we aim to examine the problem of social robot navigation (see Kruse et al. [2] for an excellent survey) and apply the principles of theatre (first discussed in Lu and Smart [3]). In particular, we will show how the navigation task can be broken up to mirror traditional dramatic structures, namely, the three-act play.

2 Three Act Structure

One particular context that is not taken into account in navigation is the robot's progress on its task. Consider the navigation task of moving from one location to another. In order for a robot to navigate in dynamic environments with uncertain elements, navigation algorithms are written to allow the robot to continuously re-plan and change its high-level behavior (the global path) almost instantly. This creates relatively robust behavior, but does not take into account how people viewing the robot's behavior may interpret it. This behavior implements an implicit Markov assumption, in that the robot only takes into account its current state, rather than also including its progress through the action.

One way to explore this context is to confer every physical action the robot performs with a narrative arc, specifically that of the archetypal three act structure. The first act is the exposition, in which the characters and setting are established. Second is the rising action, the main course of action in which the protagonist faces multiple obstacles as they move toward their goal. Then finally, the protagonist will arrive at a point where there is only one possible scenario: the climax in which we see the protagonist either achieve or fail to achieve their goal.

To properly use this structure with a navigation task, we must first define what is meant by "goal". For navigation, it may refer to simply the goal pose of the robot, or the goal pose with some constraints on how to get there (no collisions, minimal path length). In a theatrical context, the term for goal is objective, the motivation behind every action the character does. In a play, if a character crosses the stage, it might be to move away from someone to make someone feel isolated. These objectives are always posed in relation to others and not in isolation. Hence, the robot's objective cannot just be to move from place to place, but to move from place to place in relation to others around it. (Note, in this discussion we will use "objective" to refer to the motivation, and "goal" to refer to the desired pose.)

In the exposition, we need to establish the robot's "character," i.e. what it is capable of and likely to do. In most scenarios, it is impossible and impracticable to endow the robot with as much character as traditional dramatic characters like Austen's Mr. Darcy. Instead, the aim is merely to introduce static qualities of the robot that will be present throughout the action as a way of providing information to help people predict what the robot will do. This could mean exploring the different modalities of the robot (i.e. the different ways the robot can move/act). Establishing this is important, even if the modalities are not functionally necessary, so that if/when the robot employs these behaviors later,

they do not come as a surprise to the audience. Furthermore, establishing the type of movements the robot will perform can also be beneficial. Consider the difference between a robot that starts moving in a straight line to its goal, and one that moves more erratically. An observer may think the latter may need more attention or that the former is more deserving of trust. Not only does introducing these qualities early on have the benefit of helping predict future behavior, but also molds an observer's vital first impression.

The first act is also where the robot will begin to move toward its goal, which may require some preparation. The start of the action must be done in a way that is consistent with the robot's objective. For many robot navigators, the objective is simply to move towards the goal. However, in social navigation, the objective includes moving toward the goal in a way that does not disturb the people around it or cause them to be uncomfortable. For some large robots, the simple act of them starting to move their bulk toward the goal can be unquieting. One way around that is to use the additional modalities of the robot besides its mobile base to indicate that the robot is about to move. This could entail moving the head around to ensure the area is clear or a slight raise in the torso to indicate imminent action. This sort of anticipatory gesture is also suggested by Van Breemen [7] and Mead and Matarić [4].

During the middle act, i.e. the bulk of the movement toward the goal, the robot's objective must be to move to the goal, deal with unforeseen obstacles it encounters along its planned path, and to make people aware of those activities in a way that makes them continue to be comfortable. Importantly, the robot should react to the obstacles during this middle act in a way appropriate to the context of the action as a whole. The robot should not stop completely and act as though it were planning a brand new motion from the beginning again. It should react in a way that indicates that it is still pursuing the same goal while taking into account new information about the obstacles. Similarly, the scale of the reaction to unforeseen obstacles needs to be adjusted based on when it happens. One would not expect a robot to react the same way to a change in plans at the beginning of an action than at the end when it is almost at its goal.

The relationship between the robot and the people in the environment is centered around the idea of legibility, i.e. making the robot's actions clear and readable[1, 5]. Legibility is particularly important in this middle act for ensuring a smooth transition between the robot's initial goal and the ultimate outcome of that goal, since illegible behavior could be read as not acting toward that goal. Certain people in the environment require extra consideration since the robot is actively moving in the same space as them. As a result, the robot must take into account contexts related to them, their objectives, and the robot's relationship to those objectives. A person should be treated differently than other mere obstacles, in that they are often mobile and have personal space which it is better not to enter. The person may be particularly sociable and want explicit interaction with the robot. On the other hand, if the person's sole objective is to get to their destination as quickly as possible, then the robot may need to adjust its behaviors for that. If the robot's only objective is its own

goal, then it can ignore the other person's objective and move in a way that my detrimental to their goal. However, if the robot's objective does include other people's objectives, then it should move in a way that enables both to complete their objectives.

In the final act, it becomes clear whether the robot has achieved its goal or not. If a robot just stops, it is difficult to determine whether the robot is actually at its goal position or whether there is a problem, especially if the goal is unknown to the people observing the robot. [3] Thus, adding cues like looking at the goal or changing the robot pose in some manner will help indicate the final outcome, making its behavior more legible. One particularly useful approach that has not seen common-place usage in navigation tasks is including success and failure animations to explicitly mark the outcome of the action[6].

3 Discussion

Thinking of the navigation tasks as a three act play as the first step toward a context-sensitive implementation of behavior can help to alleviate the perception of robots as robotic (as defined in the introduction). Robots should ease into the initial motion, to avoid the perception of jerky motion. They should react appropriately when encountering unforeseen obstacles during the middle of their actions. Finally, they should demonstrate their success or failure in order to acknowledge the additional context of the entire action that has come before. We argue that adding in these additional layers of context will make robot navigation behavior more legible and more naturally understood by people.

Additional contextual data will improve behavior in specific contexts. This approach contrasts with the usual aim of creating universally applicable behavior. Instead of pursuing behaviors that work adequate in most situations, we should attempt to create interactions which work particularly well in given situations with given contexts. The failure to do so will lead to human robot interactions that are both homogeneous and mediocre, a phenomenon we term, the heat death of robotics.

Acknowledgements

Many thanks to Annamaria Pileggi for her insights on all things theatrical and to Doug Dooley for his animated ideas.

References

[1] Anca Dragan and Siddhartha Srinivasa. Generating legible motion. In *Robotics: Science and Systems*, June 2013.

[3] This is especially true while testing new iterations of navigation algorithms. Spoken from experience.

[2] Thibault Kruse, Amit Kumar Pandey, Rachid Alami, and Alexandra Kirsch. Human-aware robot navigation: A survey. *Robotics and Autonomous Systems*, 2013.

[3] David V. Lu and William D. Smart. Human-Robot Interactions as Theatre. In *RO-MAN 2011*, pages 473–478. IEEE, 2011. ISBN 978-1-4577-1572-3.

[4] Ross Mead and Maja J Matarić. Automated caricature of robot expressions in socially assistive human-robot interaction. In *The 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI2010) Workshop on What Do Collaborations with the Arts Have to Say about HRI*, 2010.

[5] 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.

[6] 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*, pages 69–76. ACM, 2011.

[7] AJN Van Breemen. Bringing robots to life: Applying principles of animation to robots. In *Proceedings of Shaping Human-Robot Interaction workshop held at CHI*. Citeseer, 2004.