Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16)

Action Selection Methods for Multi-Agent Navigation in Crowded Environments

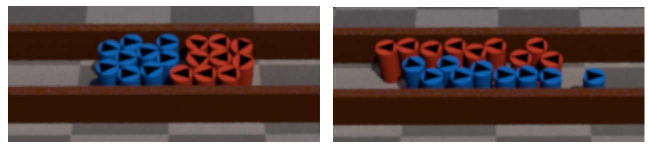
Julio Godoy
Department of Computer Science and Engineering
University of Minnesota
200 Union Street SE, Minneapolis, MN 55414

Abstract

In real-time multi-agent navigation, agents need to move towards their goal positions while adapting their paths to avoid potential collisions with other agents and static obstacles. Existing methods compute motions that are optimal locally but do not account for the motions of the other agents, producing inefficient global motions especially when many agents move in a crowded space. In my thesis work, each agent has only a limited sensing range and uses online action selection techniques to dynamically adapt its motion to the local conditions. Experimental results obtained in simulation under different conditions show that the agents reach their destinations faster and use motions that minimize their overall energy consumption.

Real-time navigation of multiple agents in crowded environments has important applications in many domains such as swarm robotics, planning for evacuation, and traffic engineering. This problem is challenging because agents have conflicting constraints. On one hand, they need to reach their goals as soon as possible while avoiding collisions with each other and the static obstacles present in the environment. On the other hand, due to the presence of many agents and the real-time constraints, agents need to compute their motions quickly (every 0.1s), independently of each other and in a decentralized manner instead of planning in a joint configuration space.

A recently introduced decentralized technique for real-time multi-agent navigation, the Optimal Reciprocal Collision Avoidance (ORCA) framework [van den Berg *et al.*, 2011] guarantees collision-free motion for the agents. Although ORCA generates locally efficient motion for each agent, the overall behavior of the agents can be far from efficient; actions that are locally optimal for one agent are not necessarily optimal for the entire group of agents. Consider, for example, the two groups of agents in Figure 1a that try to move past each other in a narrow hallway. The agents navigate using ORCA, which guarantees collision-free motion, but still end up getting stuck in congestion. In contrast, I seek to develop navigation methods that encourage the agents to adapt their motions to their surroundings, for example, by accounting for their neighbors' intended velocity during their local motion planning. By doing this, the global motion of all agents becomes more efficient, and they reach their goals faster. Figure 1b exemplifies this type of coordinated motion.



(a) (b)

Figure 1: Two groups of 9 agents each move to the opposite side of a narrow corridor. a) ORCA agents get stuck in the middle. b) Agents using my coordination approach (C-Nav) coordinate their motion and reach their goals faster.

My thesis research focuses on applying *online* action selection methods for planning, learning and coordination that can be completely distributed and require no communication or only limited one-way communication among the agents. Online approaches are more suitable than offline approaches for dynamic environments, as agents must be able to quickly adapt their behaviors to changes in their surroundings. In addition, the computational complexity of centralized offline learning methods becomes prohibitively high as the number of agents increases. Because of this, current literature on multi-agent learning focuses on scenarios with only a few agents and sparse interactions.

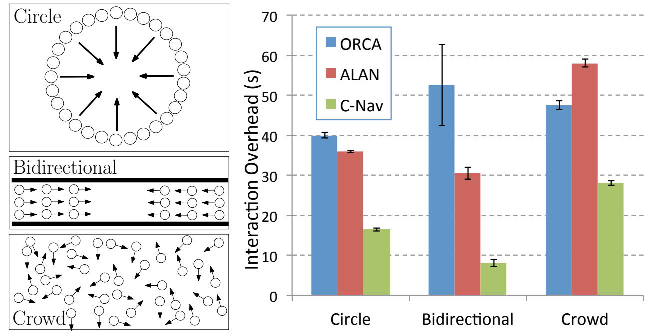
Contributions. I have proposed three methods for improving the global motion of the agents: a planning-based, a learning-based and a coordination-based method. In all of these methods, I increased the number of motions an agent can do and added new ways for agents to decide how to move. This enables them to produce different motions.

On the planning side, I proposed an anytime local approach to plan the motions of the agents in a decentralized manner, by adapting the Hindsight optimization technique in a progressive manner. I called this method Progressive Hindsight Optimization (PHOP) [Godoy *et al.*, 2014]. With PHOP, each agent simulates possible plans of actions for a given time horizon, and after assessing each one of these plans, it evaluates in 'hindsight' the quality of the first action of the plan. Each plan consists of a sequence of motion primitives. By simulating what will happen when choosing each one of the actions available, the agent can project the consequences of each choice over a time horizon and make less myopic choices. The process is repeated after each action to account for changes in the motions of the other agents. With PHOP, agents are able to predict regions in the environment where their motion is more constrained, allowing them to act accordingly (for example, by completely avoiding paths going through that region). Results of comparing PHOP with ORCA indicate lower travel time for the agents using PHOP.

I also proposed a novel framework for incorporating online learning methods in multi-agent navigation, the ALAN framework (Adaptive Learning for Agent Navigation). I formulated the problem of selecting the best motion at each time as an action selection problem in a multi-armed bandit setting. In this formulation, the challenge is to carefully balance action exploration and exploitation. ALAN uses an action selection method inspired by the principles of two well-known action selection techniques, ϵ-greedy and Upper Confidence Bounds (UCB). Agents using ALAN exploit the best action in a greedy fashion and perform biased exploration using a version of UCB more suited to non-stationary domains. Further, ALAN introduces game-theoretic elements, considering the local context of an agent to strategically adapt the amount of exploration performed [Godoy *et al.*, 2015]. Combined with a reward function that considers goal-oriented motion and neighborhood interactions, ALAN allows agents to adapt their motion to their local conditions (i.e., move back or sideways when goal-oriented motion is constrained). Agents using ALAN take advantage of pure goal-oriented motion when they are able to, and perform biased exploration when that motion is constrained. This indirectly improves the global efficiency of the motions of all agents, allowing them to reach their destinations faster while outperforming other action selection techniques in different environments.

Although both PHOP and ALAN produce time-efficient motions for all agents, they do not scale appropriately to highly-dense environments, where the agents' movement is very constrained. I proposed a method for an agent to select its velocity in a way that benefits not only itself but also its neighboring agents. I implemented this idea in C-Nav (short for Coordinated Navigation), a distributed approach to improve the global motion of a set of agents in crowded environments by implicitly coordinating their local motions [Godoy *et al.*, 2016]. This coordination is achieved by allowing each agent to learn the likely motion of its nearby neighbors. C-Nav requires no bidirectional communication, and as such, it can scale well to hundreds of agents. Agents using C-Nav choose velocities that help their nearby agents to move to their goals, mimicking the way humans move in congested environments and effectively improving the time-efficiency of the entire crowd.

Finally, to compare the performance of my proposed approaches with existing navigation algorithms, I proposed a new metric called *Interaction Overhead* which measures the time that agents spend in interactions with other agents. An *Interaction Overhead* value of 0 represents the minimum travel time for the agents in each environment, and is also the best possible result that any approach can achieve. Figure 2 shows the results (in *Interaction Overhead*) of comparing two of my proposed approaches, ALAN and C-Nav, with ORCA in three example environments. Here, ALAN agents reach their goals faster than ORCA in two of the three environments, while agents using C-Nav consistently outperform both ORCA and ALAN in all three cases.



(a) Environments (b) Results

Figure 2: a) Example environments used to evaluate my work. b) *Interaction Overhead* comparison between ALAN, C-Nav and ORCA, in the three environments (lower is better).

Future Work. A key assumption in the approaches I described is that the agents are homogeneous (i.e., they use the same navigation algorithm), and only differ in their initial and goal positions. As robotic agents are deployed in the real world, they will need to interact with humans and other types of robots, whose motion might not be accurately predicted. In the next few months, I will focus on techniques that allow an agent to navigate through environments populated by heterogeneous agents whose type is initially unknown. In these settings, the agent needs to learn, through inference, the model used by its neighbors in order to compute a safe and efficient path among them. I will present some initial results on this topic in the upcoming IJCAI conference.

Acknowledgments: This research was partially supported by the University of Minnesota Informatics Institute, and a CONICYT/Fulbright fellowship.

References

[Godoy *et al.*, 2014] Julio Godoy, Ioannis Karamouzas, Stephen J. Guy, and Maria Gini. Anytime navigation with progressive hindsight optimization. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2014.

[Godoy *et al.*, 2015] Julio Godoy, Ioannis Karamouzas, Stephen J. Guy, and Maria Gini. Adaptive learning for multi-agent navigation. In *Proc. Int. Conf. on Autonomous Agents and Multi-Agent Systems*, pages 1577–1585, 2015.

[Godoy *et al.*, 2016] Julio Godoy, Ioannis Karamouzas, Stephen J. Guy, and Maria Gini. Implicit coordination in crowded multi-agent navigation. In *Proc. AAAI Conference on Artificial Intelligence*, 2016.

[van den Berg *et al.*, 2011] Jur van den Berg, Stephen J. Guy, Ming Lin, and Dinesh Manocha. Reciprocal n-body collision avoidance. In *Proc. International Symposium of Robotics Research*, pages 3–19. Springer, 2011.