Apprenticeship Scheduling for Human-Robot Teams

Matthew C. Gombolay
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, Massachusetts 02139
gombolay@csail.mit.edu

Introduction

Resource optimization and scheduling is a costly, challenging problem that affects almost every aspect of our lives. One example that affects each of us is health care: Poor systems design and scheduling of resources can lead to higher rates of patient noncompliance and burnout of health care providers, as highlighted by the Institute of Medicine (Brandenburg et al. 2015). In aerospace manufacturing, every minute re-scheduling in response to dynamic disruptions in the build process of a Boeing 747 can cost up to $100,000. The military is also highly invested in the effective use of resources. In missile defense, for example, operators must solve a challenging weapon-to-target problem, balancing the cost of expendable, defensive weapons while hedging against uncertainty in adversaries' tactics.

Researchers in artificial intelligence (AI) planning and scheduling strive to develop algorithms to improve resource allocation. However, there are two primary challenges. First, optimal task allocation and sequencing with upper and lower-bound temporal constraints (i.e., deadlines and wait constraints) is NP-Hard (Bertsimas and Weismantel 2005). Approximation techniques for scheduling exist and typically rely on the algorithm designer crafting heuristics based on domain expertise to decompose or structure the scheduling problem and prioritize the manner in which resources are allocated and tasks are sequenced (Tang and Parker 2005; Jones, Dias, and Stentz 2011). The second problem is this aforementioned reliance on crafting clever heuristics based on domain knowledge. Manually capturing domain knowledge within a scheduling algorithm remains a challenging process and leaves much to be desired (Ryan et al. 2013).

The aim of my thesis is to develop an autonomous system that 1) learns the heuristics and implicit rules-of-thumb developed by domain experts from years of experience, 2) embeds and leverages this knowledge within a scalable resource optimization framework, and 3) provides decision support in a way that engages users and benefits them in their decision-making process. By intelligently leveraging the ability of humans to learn heuristics and the speed of modern computation, we can improve the ability to coordinate resources in these time and safety-critical domains.

The Role of the Human In-The-Loop

The first step in designing an effective autonomous system to coordinate a teams' activities is understanding the respective roles of the human and robotic (or virtual) team members within the decision-making loop. I have been conducting a series of human-subject experiments in which a team of two human team members, the subject and one confederate, and one robot team member work together to complete a set of tasks in an experimental setup analogous to a manufacturing environment. I studied how decision-making authority over task allocation decisions should be shared between the team members (Gombolay et al. 2014) and how the team's composition, mixed human-robot or strictly human-human, affected the team's fluency depending on how decision-making authority was shared (Gombolay et al. 2015). I mentored three undergraduate students to assist me in conducting the experiments. We found that increasing the robot's authority over task allocation decisions decreased the time to schedule the team and the time to execute and improved the humans' perception of their robotic counterpart. Furthermore, subjects' perception of the team member responsible for scheduling improved more significantly when that team member was a robot as opposed to another human. These studies motivate the development of autonomous systems for scheduling of human and human-robot teams.

Scalable Methods for Scheduling

The next step in developing an autonomous resource optimization mechanism is creating scalable computational models and techniques. In the key areas of interest, such as manufacturing, healthcare, and military operations, one must consider the interplay of tight upper- and lowerbound temporal constraints, spatial constraints, and heterogeneous agent and resource capabilities. Approaches in prior work typically attempt to decompose techniques and heuristics to reduce the computational search space. However, the best solvers are only able to scale to approximately 5 agents and 50 tasks within a solution time of approximately thirty minutes (Korsah, Stentz, and Dias 2013). In prior work, I developed a fast, near-optimal sequencing technique for task sets with complex dependencies (Gombolay, Wilcox, and Shah 2013). The key to the success of the approach is an analytical schedulability test I developed (Gombolay and Shah 2015; Gombolay, Wilcox, and Shah 2013) that provides an

informative upperbound on the temporal and physical resources required to execute a set of tasks with upper and lowerbound temporal constraints. I worked with a colleague to integrate my sequence technique within a task allocation algorithm to dynamically reschedule human-robot teams. A demonstration of our algorithm, which we call Tercio, can be found at `http://youtu.be/_qb2_jJID5c`.

Learning Implicit Constraints and Goals for Team Coordination

I have motivated the advantages of autonomous scheduling algorithms in team coordination. The challenge then remains of how we can learn the heuristics and rules-of-thumb of human domain experts to automatically schedule processes. I have personally seen human domain experts who are able to effectively manage large teams in practice. I leveraged such domain knowledge to understand how to develop the computational methods used within Tercio. However, manually encoding experts' domain knowledge in an autonomous framework is difficult in practice and leaves much to be desired. As such, we need to develop a technique that scales beyond the one-expert, one-apprentice model.

In my thesis, I propose a technique, which I call "apprenticeship scheduling", to capture this domain knowledge in the form of a scheduling policy and scale the power of the expert beyond a single-expert, single-trainee apprenticeship model. My approach efficiently utilizes domain-expert demonstrations without the need to train within an environment emulator. Rather than explicitly modeling a reward function and relying on dynamic programming (Zheng, Liu, and Ni 2015) or constraint solvers, which become computationally intractable for large-scale problems of interest, my approach uses action-driven learning to extract the strategies of domain experts. The key to my approach is using pairwise comparisons between the actions taken (e.g., schedule agent a to complete task τ_i at time t) and the actions not taken (e.g., unscheduled tasks at time t) to learn relevant model parameters and heuristics demonstrated by the training examples. I have submitted to the AAAI-16 main technical track a paper which presents and validates my approach using both a synthetic data set of solutions for a variety of scheduling problems and a real-world data set of demonstrations from human experts solving a resource optimization problem.

However, there are a number of items I must consider in future work. First, domain experts may be able to manage resources equally well but do so in different ways. Prior work has reled on learning one model for each operator to account for inter-demonstrator disagreement (Sammut et al. 1992). In future work, I will investigate new computational methods for reconciling the difference between operators in a way that does not dismiss disagreement as "noise" so that the key attributes that make each operator proficient are captured. Second, creating features to minimally describe the state space is often a manual process. Convolution neural networks have been shown to generate compact, visual features for dynamic programming; however, feature learning for scheduling remains an open challenge. In future work, I want to investigate computational methods that can capture the features for scheduling problems.

Attending AAAI, the premier conference in artificial intelligence, would help me to connect with the top researchers in the field to better understand the latest work in developing artificial intellgience mechanisms for planning, scheduling, and control of autonomous systems. Immersing myself within the AAAI community by sharing ideas and developing collaborations would be invaluable for me to realize the full potential of an apprenticeship scheduling approach. Eventually, the collaborations I would develop at AAAI may grow into long-term projects that improve resource management in manufacturing, healthcare, defense, and so many other domains that affect every area of our lives.

References

Bertsimas, D., and Weismantel, R. 2005. *Optimization over Integers*. Belmont: Dynamic Ideas.

Brandenburg, L.; Gabow, P.; Steele, G.; Toussaint, J.; and Tyson, B. J. 2015. Innovation and best practices in health care scheduling. Technical report.

Gombolay, M. C., and Shah, J. A. 2015. Schedulability analysis of task sets with upper- and lower-bound temporal constraints. *Journal of Aerospace Information Systems* 11(12):821–841.

Gombolay, M. C.; Gutierrez, R. A.; Sturla, G. F.; and Shah, J. A. 2014. Decision-making authority, team efficiency and human worker satisfaction in mixed human-robot teams. In *Proceedings of the Robots: Science and Systems*.

Gombolay, M. C.; Gutierrez, R. A.; Clarke, S. G.; Sturla, G. F.; and Shah, J. A. 2015. Decision-making authority, team efficiency and human worker satisfaction in mixed human-robot teams. *Autonomous Robots* 39(3):293–312.

Gombolay, M. C.; Wilcox, R. J.; and Shah, J. A. 2013. Fast scheduling of multi-robot teams with temporospatial constrints. In *Proceedings of the Robots: Science and Systems*.

Jones, E.; Dias, M.; and Stentz, A. 2011. Time-extended multi-robot coordination for domains with intra-path constraints. *Autonomous Robots* 30(1):41–56.

Korsah, G. A.; Stentz, A.; and Dias, M. B. 2013. A comprehensive taxonomy for multi-robot task allocation. *The International Journal of Robotics Research* 32(12):1495–1512.

Ryan, J. C.; Banerjee, A. G.; Cummings, M. L.; and Roy, N. 2013. Comparing the performance of expert user heuristics and an integer linear program in aircraft carrier deck operations. *IEEE Transaction on Cybernetics* (9).

Sammut, C.; Hurst, S.; Kedzier, D.; and Michie, D. 1992. Learning to fly. In *Proceedings of the Ninth International Conference on Machine Learning*, 385–393.

Tang, F., and Parker, L. E. 2005. ASyMTRe: Automated synthesis of multi-robot task solutions through software reconfiguration. In *Proceedings of the International Conference on Robotics and Automation*.

Zheng, J.; Liu, S.; and Ni, L. 2015. Robust bayesian inverse reinforcement learning with sparse behavior noise. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2198–2205.