Efficiency and Safety in Autonomous Vehicles through Planning with Uncertainty (Thesis Summary)

Zachary N. Sunberg*
Aeronautics and Astronautics Dept.
Stanford University
Stanford, CA 94305

Introduction

Autonomous vehicles are quickly becoming an important part of human society for transportation, monitoring, agriculture, and other applications. As autonomous vehicles and robots of all types become ubiquitous, they will interact more closely with humans, and it will become increasingly important for them to act efficiently in terms of time, energy, and other resources while also maintaining safety. However, there is a fundamental tradeoff between safety and efficiency because safety constraints prohibit some actions that are efficient.

A key to maintaining safety without sacrificing efficiency is dealing with uncertainty properly. Improper treatment can result in control systems that are too aggressive or too conservative, but with proper treatment, autonomous vehicles can be assertive when it is appropriate and careful in dangerous situations. The Markov decision process (MDP) and partially observable Markov decision process (POMDP) frameworks provide a systematic way to plan in the presence of uncertainty. However, there are still challenges to using POMDP and MDP approaches in the real world including scaling solution techniques and maintaining safety gaurantees.

My research contributes to this goal of safety and efficiency by analyzing the effects of safety systems in several domains, quantifying the advantages of reasoning about uncertainty, and proposing a new algorithm that is capable of solving problems with continuous state, action, and observation spaces, a qualitative increase in capability over the previous leading solution techniques.

Planning with internal states in autonomous driving

Autonomous driving is one of the most important current applications of artificial intelligence. One challenge is that it will involve extensive interaction with human drivers. Human drivers routinely negotiate with one another, for example, when merging and navigating intersections.

* zsunberg@stanford.edu

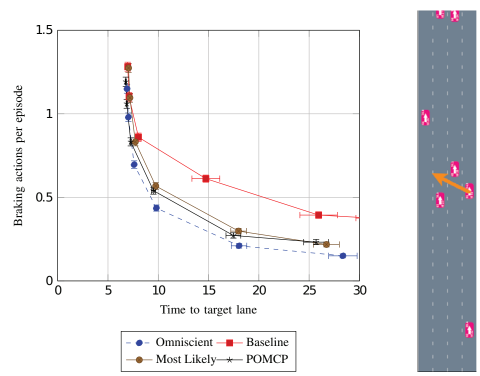
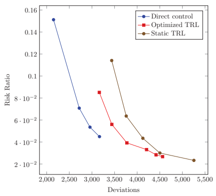


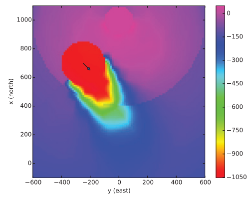
Figure 1: Pareto performance frontiers of different planning approaches in a lane-changing task.

The thesis will include an approach for planning multiple lane changes on a crowded freeway by modeling the problem as a POMDP (Sunberg, Ho, and Kochenderfer 2017). In such a POMDP, the sensing systems of the autonomous vehicle are able to measure all of the physical states of the other cars perfectly. The partially observable states of the system are the *internal states* of the other drivers, for example their aggressiveness, attentiveness, and intentions. In the first round of study, the experiments only considered static internal states related to driving style. Specifically, human drivers behave according to the intelligent driver model (IDM), and the IDM parameters are the latent internal states.

In order to gauge the importance of planning with internal states, I established a baseline and an upper performance bound by planning using two MDPs. The baseline is determined by solving a naive MDP model in which all vehicles have the same normal IDM parameters. The upper bound is established by solving an omniscient MDP in which all parameters are know exactly. The results in Fig. 1 show that inferring traffic participants' models online by solving the POMDP formulation can increase both safety and efficiency nearly to the upper bound. All problems were solved with variants of Monte Carlo tree search augmented with a safety constraint that prevents crashes. Research between now and my defense will study dynamic intentions as hidden states.

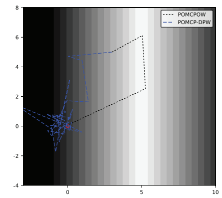


(a) Pareto optimal performance of several approaches.

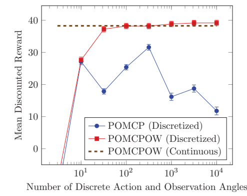


(b) Value function slice for optimized TRL approach.

Figure 2: Collision avoidance results.



(a) Typical behavior of POMCP-DPW and POMCPOW on a light-dark problem. The goal is to localize in the light region, then target the origin.



(b) Performance of POMCPOW on a continuous tag problem compared with POMCP planning with a discretized model.

Figure 3: Sample POMCPOW simulation results.

The price of trustworthiness in aerial collision avoidance

The thesis will also include an investigation of the trade-off between trustworthiness and performance in the context of aerial collision avoidance (Sunberg, Kochenderfer, and Pavone 2017). Some approaches to aerial collision avoidance consist of a set of rules that can be verified or analytically proven to prevent collision given constraints on the behavior of other aircraft. Such a system, which we refer to as trusted resolution logic (TRL) is trusted to provide safety, but the cost of certification can be very large. In some cases, TRL contains parameters with a range of certified values that can be tuned to change performance.

In order to analyze the effect of the TRL on performance, the research considers three approaches:

1. Using the TRL with static parameters.
2. Using the TRL with parameters that are dynamically optimized by solving an MDP.
3. Solving an MDP without obeying the TRL.

The difference between (1) and (3) gives an estimate of the *price of trustworthiness*, that is, the loss in performance caused by following the TRL. In the horizontal aircraft collision avoidance problem that we analyzed, the price of trustworthiness is significant (see Fig. 2a), but using an offline MDP solution to optimize the TRL parameters for each state significantly reduces the price.

POMCPOW: An online algorithm for POMDPs with continuous state, action, and observation spaces

The most general and wide-reaching contribution of my thesis is the proposal of a new online POMDP algorithm that is capable of solving continuous POMDPs (Sunberg and Kochenderfer 2018). The current most popular online algorithm, POMCP (Silver and Veness 2010), will fail on continuous action and observation spaces because the trees grow too wide and will not extend beyond one level deep.

I first investigated using double progressive widening (DPW) to mitigate this issue, but found that this introduces another issue. Specifically, I proved analytically that, when applied to continuous observations, POMCP with DPW will converge to the QMDP solution, which is suboptimal for cases where information gathering is important.

The new algorithm that I proposed, partially observable Monte Carlo planning with observation widening (POMCPOW) uses weighted particle filtering on top of DPW to correctly handle continuous problems. So far, I have demonstrated POMCPOW's effectiveness in simulation (see Fig. 3), and before my thesis defence, I hope to prove its consistency analytically by using results from Auger, Couetoux, and Teytaud (2013) and investigate its performance in simulation more thoroughly on more problems and against other solution methods.

References

Auger, D.; Couetoux, A.; and Teytaud, O. 2013. Continuous upper confidence trees with polynomial exploration–consistency. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, 194–209. Springer.

Silver, D., and Veness, J. 2010. Monte-Carlo planning in large POMDPs. In *Advances in Neural Information Processing Systems (NIPS)*.

Sunberg, Z. N., and Kochenderfer, M. J. 2018. POMCPOW: an online algorithm for POMDPs with continuous state, action, and observation spaces. In *International Conference on Automated Planning and Scheduling (ICAPS)*. Under Review.

Sunberg, Z. N.; Ho, C. J.; and Kochenderfer, Mykel, J. 2017. The value of inferring the internal state of traffic participants for autonomous freeway driving. In *American Control Conference (ACC)*.

Sunberg, Z.; Kochenderfer, M. J.; and Pavone, M. 2017. Optimized and trusted collision avoidance for unmanned aerial vehicles using approximate dynamic programming. In *IEEE International Conference on Robotics and Automation (ICRA)*.