Autonomous Air Traffic Control for Non-Towered Airports

Zouhair Mahboubi
PhD Candidate
Aeronautics and Astronautics
Stanford University
Stanford, California, USA

Abstract

Half of all reported near mid-air collisions involve at least one general aviation aircraft, and the majority of them occur in the vicinity of an airport. This research frames the problem of traffic collision prevention in the vicinity of non-towered airports as a partially-observable Markov decision process (POMDP). The solution to this problem will help enable an autonomous air traffic control system that is non-intrusive, ground-based and with no additional requirements to participating aircraft except for radio communication. We present initial results from recent work under the assumption of full observability with simulation of aircraft in the traffic pattern. We then outline extensions and future work.

Introduction

Mid-air collisions are a major safety concern in aviation, which is why the Traffic Alert and Collision Avoidance System is mandated on all large aircraft. Unfortunately, few general aviation (GA) aircraft are equipped with the system due to its high cost and weight.

In its 2000 report, the Aircraft Owners and Pilots Association highlighted that most mid-air collisions happen in the vicinity of airports (AOPA Air Safety Foundation 2000). The majority of them occurred in the pattern of non-towered airports. It is more difficult to get accurate statistics for near mid-air collision incidents (Kochenderfer, Griffith, and Olszta 2010), but nearly half of those reported through NASA's Aviation Safety Reporting System were between two GA aircraft (Kunzi and Hansman 2011).

With almost 12,000 airports without towers in the US compared to only 400 with towers, the issue of mid-air collision at non-towered airports is worth addressing. Previous research have explored various solutions to collision avoidance for GA aircraft (Diefes and Deering 1996). However, the majority of suggested solutions focus on on-board sensors and systems for detection, alert, and resolution. However, GA pilots are cost sensitive and their aircraft have little power and weight margins for additional payloads, which makes new on-board systems unlikely to be adopted (AOPA, EAA, and others 2015).

Recent research recognized the need for a low cost solution. A ground-based surveillance at small airports concept was investigated by Campbell (2014) and demonstrated promising performance. Their goal is to deploy the sensor at towered airports and high density non-towered airports. Alternatively, a radio direction finder can be used to determine the bearing to a transmitting aircraft, and its position can be triangulated if two or more antennas are available. Additionally, such a system would allow the aircraft to be identifiable for the purposes of communication.

In this dissertation abstract, we suggest the use of the decision making under uncertainty framework to the problem of traffic collision prevention between GA aircraft. We hypothesize that an autonomous air traffic control system (auto-ATC) can be modeled using a partially observable Markov decision process (POMDP). For initial work, we assume full observability and efficiently solve the problem as an MDP for a reasonable number of aircraft and states. We further propose the use of structured continuous-time Markov decision processes (CTMDP) to improve the model.

We envision the auto-ATC at a non-towered airport to be advisory in nature. The pilot would still be responsible for maintaining separation, but the system would provide advice to participating aircraft with the aim of reducing incidents by issuing high-level recommendations. The aircraft would be tracked through ground-based sensors in the immediate vicinity of an airport, and would be expected to fly according to a model that can be influenced through the recommendations transmitted over radio. Such a system could also be useful at busy towered airports as an assistive technology to ATC controllers. The idea is inspired by the work of Nikoleris et al. (2014) where 4D trajectories from an expert repertoire are uplinked to aircraft operating within an Advanced Airspace Concept, as well as Gunawardana (2012), where a module capable of autonomously communicating within the existing ATC framework is proposed.

In the following sections, we outline how such a system can be modeled using a POMDP. We then present results of a proof-of-concept model, including 3D simulations of aircraft in the pattern. Finally, we present planned extensions to our work with the aim of improving the model's accuracy. Note that part of this work is being published in the upcoming Air Traffic Management Seminar (Mahboubi and Kochenderfer 2015a), and the suggested future work on CTMDP was submitted for peer review at the Conference on Decision and Control (Mahboubi and Kochenderfer 2015b).

Problem Formulation

In this section, we introduce partially observable Markov decision processes (POMDPs) and explain how they can be used to model the problem of autonomous collision avoidance for aircraft operating at a non-towered airport.

POMDPs model sequential problems where decisions need to be made under uncertainty. They have been successfully used in the context of aircraft collision avoidance by Billingsley, Kochenderfer, and Chryssanthacopoulos (2012) as well as Holland, Kochenderfer, and Olson (2013). In this section, we briefly introduce the concept of POMDPs and explain how they can be applied to developing the decision-making component of the auto-ATC system. Because our goal is to investigate this new concept, we make a few simplifying assumptions in our modeling of the problem.

A POMDP is defined by a state space $\mathcal{S}$, action space $\mathcal{A}$, transition function T, reward function R, set of observations Ω, and observation model O (Kochenderfer 2015). If the current state of the process is $s \in \mathcal{S}$ and we execute action $a \in \mathcal{A}$, then the next state will be $s' \in \mathcal{S}$ with probability $T(s' \mid s, a)$. The reward associated with executing a from s is given by $R(s, a)$, and the observation o available to the system is distributed according to $O(o|s)$.

The goal in a POMDP is to select actions in a way that maximizes the expected discounted reward:

$$\mathrm{E}\left[\sum_{k=0}^{\infty} \gamma^k r_k\right] \tag{1}$$

where r_k is the reward received in step k and $\gamma \in [0, 1)$ is a discount factor which decays the value of rewards received in the future. A decision making policy that is dependent on the current state is denoted π, and the action recommended at state s is denoted $\pi(s)$.

State Space The state space $\mathcal{S}$ is composed of a discrete set of states specifying the locations of K aircraft in the traffic pattern. A particular state s is represented by a tuple $(s^{(1)}, \ldots, s^{(K)})$, where $s^{(i)} \in \{\ell_1, \ldots, \ell_n\}$ represents the location of the ith aircraft. In this formulation, there are a set of $n = 27$ possible locations, e.g., Taxi ($\ell_1 =$ T) and Runway ($\ell_2 =$ R). Hence, if there are K aircraft, then $|\mathcal{S}| = n^K$. The set of possible states that can immediately follow a particular location ℓ_i is denoted $\mathcal{N}(\ell_i)$. Figure 1 shows a representation of the various pattern locations along with possible transitions between them.

Action Space The action space $\mathcal{A}$ is composed of a discrete set of actions specifying a particular aircraft and a location. Action $a = (a_i, a_l)$ involves commanding aircraft $a_i \in \{1, \ldots, K\}$ to transition immediately to location a_l. If no aircraft is to be addressed, we use $a = (0, \varnothing)$. The valid set of actions depends on the current state. We denote the set of actions available from state s as $A(s)$. For example, if $s = (\mathrm{R}, \mathrm{U2})$, then $A(s) = \{(0, \varnothing), (1, \mathrm{T}), (1, \mathrm{U1}), (2, \mathrm{LX2}), (2, \mathrm{RX2})\}$. In this formulation, it is not possible to request an aircraft to depart the pattern.

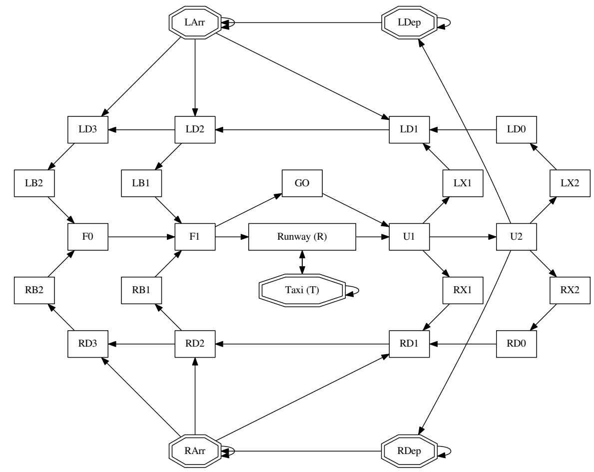

Figure 1: Aircraft states in the pattern.

Transition Function The transition function specifies the probability of transitioning to some next state given the current sate and action taken. The probabilities governing the aircraft transitions are independent from each other:

$$T(s' \mid s, a) = \prod_{i=1}^{K} T\left(s'^{(i)} \mid s^{(i)}, \begin{cases} a_l & \text{if } a_i = i \\ \varnothing & \text{otherwise} \end{cases}\right), \tag{2}$$

The transition model assumes the following:

- If an aircraft is not being addressed by an action, all of the possible next states are equally likely with probability $1/|\mathcal{N}(s^{(i)})|$.
- If the aircraft is being addressed, the commanded state a_l is selected with probability α (a cooperation factor), while other possible states are selected with probability $(1 - \alpha)/(|\mathcal{N}(s^{(i)})| - 1)$
- A pilot will fly the pattern without detailed instructions from the tower, but will generally take the runway only when instructed by the system: $T(\mathrm{R} \mid \mathrm{T}, \varnothing) = 1 - \alpha$.

Reward Function The reward function is designed to increase aircraft separation while minimizing intervention. We make the following assumptions:

- The rewards are additive over states and actions, i.e. $R(s, a) = R(s) + R(a)$.
- Two aircraft occupying the same state result in a cost C_c, unless they are in one of the states considered safe (Taxi, Departures, and Arrivals). Each aircraft in the T state incurs a cost $C_t < C_c$ to avoid the system grounding all aircraft.
- There is a cost C_a for issuing an advisory. We define the advisory-cost ratio $\beta = C_a/C_c$.

The reward function can be written by defining $\tilde{s} = (s^{(i)} \in s : s^{(i)} \notin \{\mathrm{T}, \mathrm{LDep}, \mathrm{RDep}, \mathrm{LArr}, \mathrm{Rarr}\})$:

$$\begin{aligned} R(s,a) = &- C_c(|\tilde{s}| - |\mathrm{unique}(\tilde{s})|) \\ &- C_t|(s^{(i)} \in s : s^{(i)} = \mathrm{T})| \\ &- \begin{cases} 0 & \text{if } a = (0, \varnothing) \\ C_a & \text{otherwise} \end{cases} \end{aligned} \tag{3}$$

where we use the list comprehension notation $(x \in X : F(x))$ to mean the list of all elements in X that satisfy the logical expression $F(x)$. The notation $|y|$ denotes the number of elements in the list y.

Observation Model

The observations $o \in \Omega$ are the north-east-down coordinates of the aircraft relative to a reference point at the airport. These observations are obtained from a ground-based sensor, and are related to the aircraft state using the observation model $O(o \mid s)$. The details of this model are not yet developed and are part of future work.

Solution Approach

For the purposes of this abstract, we simplify the problem by assuming full-observability, and therefore solve the POMDP as an MDP. There are different approaches for finding the optimal policy for an MDP, and we chose to use a dynamic programming algorithm known as Gauss-Seidel value iteration (Kochenderfer 2015) due to its simplicity. We begin by assigning 0 to all states of the value function U. We then iterate through all the states, updating U as we go along according to

$$U(s) \leftarrow \max_{a \in A(s)} \left[R(s,a) + \gamma \sum_{s' \in \mathcal{S}} T(s' \mid s, a) U(s') \right] \tag{4}$$

Gauss-Seidel value iteration sweeps over the states repeatedly until convergence. Once converged, an optimal policy π^* can be extracted from U as follows:

$$\pi^*(s) = \arg\max_{a \in A(s)} \left[R(s,a) + \gamma \sum_{s' \in \mathcal{S}} T(s' \mid s, a) U(s') \right] \tag{5}$$

Value iteration is an efficient way to solve MDPs since it is polynomial in the size of the state-space $|\mathcal{S}|$. However, $|\mathcal{S}|$ itself grows exponentially with the number of aircraft. For example, with four aircraft, there are over 500,000 states and 100 actions. Fortunately, we can take advantage of the transitions being sparse, and can compute an optimal solution in six minutes using a single thread on a 1.9 GHz Intel i7 CPU.

3D Simulation Framework

The MDP model presented in the previous section is not an accurate representation of the real world. In practice, aircraft in the pattern do not all move to the next states at the same time, and the amount of time they spend in each leg of the pattern depends on several factors. Therefore, we developed a higher fidelity 3D aircraft model that would capture some of these factors. We describe the details and modeling assumptions we made to construct this model.

Each aircraft is parametrized with the following states:

- $x = [x_N, x_E, x_D]^\top$, aircraft position in north-east-down world coordinates,
- V, aircraft airspeed, which is assumed to be along the aircraft longitudinal axis, and
- ψ, aircraft heading in world coordinates.

Additionally, we assume that the pilot perfectly regulates the aircraft roll ϕ and its glide path angle γ. The resulting equations of motion are:

$$\dot{x} = V \begin{bmatrix} \cos\psi \\ \sin\psi \\ -\sin\gamma \end{bmatrix}, \dot{\psi} = \frac{g \tan\phi}{V} \tag{6}$$

The equations are integrated using the Euler method. This model does not account for any wind effects, assumes that the pilot is maintaining coordinated flight, and neglects any dynamics associated with achieving the necessary roll and glide path angles. These are reasonable assumptions for this model as we are not concerned with the details of the flight dynamics, but rather with the motion of the aircraft in the 3D world coordinates.

In order to make the aircraft fly around the pattern, we implemented a two-layer logic controller to emulate a pilot:

- *Navigate.* The pilot flies towards a waypoint by setting a desired bearing ψ_d and altitude h_d.
- *Aviate.* The pilot regulates the aircraft at h_d by commanding γ and steers towards ψ_d by commanding ϕ.

We define target spatial waypoints for each of the locations in the pattern shown in Fig. 1 with the runway fixed at $(0, 0, 0)$. The simulation assumes that the pilot flies from their current waypoint to the next one, and whenever the aircraft is close to the destination waypoint, the pilot chooses the next leg they will be flying according to the MDP model introduced previously. A random position error with Gaussian distribution is added to the east-north-down coordinates of the destination waypoint. A sample of the resulting trajectories can be seen in Fig. 2.

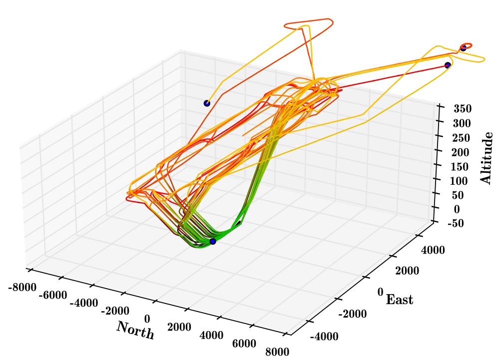

Figure 2: 3D simulation of four aircraft in the pattern.

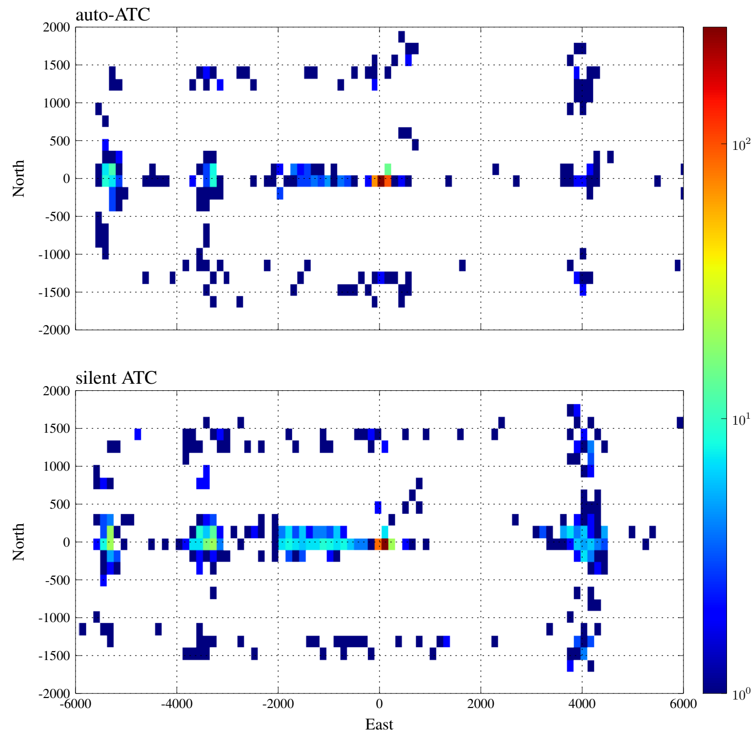

Figure 3: Position distribution of NMAC events.

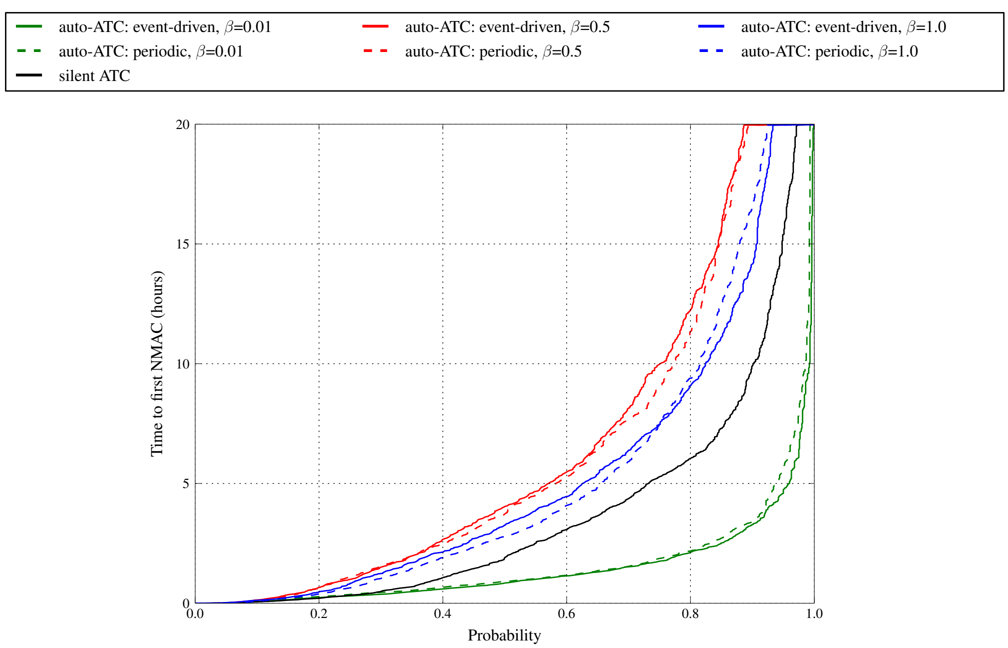

Figure 4: Inverse CDF for NMAC events.

Preliminary Results

We evaluated the policy by running Monte Carlo 3D simulations. For each policy, a total of 1000 cases are initialized with four aircraft in at random locations in the pattern and with random airspeeds. The states are then simulated forward until either a near mid-air collision (NMAC) event occurs or 20 hours is elapsed.

We compared the performance of the policy obtained from the MDP formulation to a silent policy which issues no advisories. A consequence of this silent policy is that the aircraft eventually end in the taxi state. Hence, the silent policy has an artificial advantage over the auto-ATC policy in long-running simulations because collisions are not counted in the taxi state.

Figure 3 shows the location of all the NMACs as a heat map on the pattern for both the event driven auto-ATC with $\beta = 0.5$ and the silent ATC. The majority of events occur when the aircraft are turning from base to final and on the runway. The other hotspots are other convergence points such as when aircraft are arriving in the pattern or cutting the base turn when following an upwind aircraft. These observations are consistent with the analysis of actual NMAC events (Kunzi and Hansman 2011) and is not surprising given our modeling assumptions. The NMACs between these hotspots (e.g., on the downwind leg) are due to faster aircraft overtaking slower aircraft. Because the simulation horizon is relatively long, both auto-ATC and silent ATC have roughly the same number of NMAC events (900 and 1000 respectively). However, while the number of events are similar and the 2D distribution of the events look similar, there is a difference in how long it takes for each event to occur.

Figure 4 shows the inverse cumulative density function (CDF) of the time to first NMAC for different policies. Although it performed well on the MDP transition model, the $\beta = 0.01$ auto-ATC policy performs poorly in the 3D simulation and is worse than the silent ATC. However, $\beta = 0.5$ and $\beta = 1.0$ policies perform better than the silent ATC. This is counter to what we observed in the MDP simulations where lower β values (i.e., more verbose policies) lead to less collisions. The difference in performance between the MDP model and the 3D simulation can be explained by the fact that in the MDP, the aircraft transition to next state at the same time, whereas in the 3D simulation, the amount of time they spend in each leg of the pattern depends on several factors (airspeed, position noise, etc.)

Future work

For this dissertation abstract, we suggested a concept for an autonomous ATC that could help reduce the risk of air traffic collisions in the vicinity of non-towered airports. We showed how the system can be posed as a POMDP by defining the states, actions, transitions, rewards and observations.

Although these preliminary results are promising, this first formulation makes many simplifying assumptions. As part of this abstract, we outline future work and planned extensions to increase the accuracy of the model, such as accounting for continuous-time and partial observability.

Accounting for transition times

In the real life, the transitions do not all occur at the same time as modeled by the MDP and the actions need to be taken at potentially non-uniform time steps. This could be accounted for by modeling the problem as a semi-Markov decision process, where the transition time from one state to the next follows a probability density function (Puterman 2005). The traditional semi-MDP formulation assumes that the holding time is the same for all states. However, in our model, the holding time varies depending on the state. Additionally, the MDP is structured since the state is a Cartesian product of each aircraft's location in the pattern. This makes it difficult to use the traditional formulations for semi-MDPs. Instead, we are investigating the use of Continuous-Time Bayesian networks (CTBN) to formulate the dynamics of the problem (Nodelman, Shelton, and Koller 2002). The decision making can then be posed as a structured continuous-time Markov decision process (CTMDP) (Kan and Shelton 2008).

We give a quick outline of how the auto-ATC can be formulated as a CTMDP. A CTBN describes a stochastic process which progresses over continuous time. It uses the concept of intensity matrices Q to define the transition from one state to anot her given an action. An intensity matrix can be factored into a state duration matrix M and a transition probability matrix P as:

$$Q = M(T - I) \tag{7}$$

In a CTMPD, each action $a \in \mathcal{A}$ is associated with an intensity matrix $Q_a = M_a(T_a - I)$. The transition probability matrix T_a for each aircraft can be obtained from the transition function $T(s'^{(i)} \mid s^{(i)}, a)$. The duration matrix M_a contains the rate parameters for the holding time in each state which is assumed to be distributed according to an exponential time function $f(t) = \lambda e^{-\lambda t}$. In this manner, we can define the transition process for each aircraft conditional on whether it received an ATC advisory or not. The transition process for the entire system can then be constructed from the individual Q for each aircraft. This is done by using the Kronecker product as outlined in (Shelton and Ciardo 2014) to obtain the full-joint transition intensity matrix.

Kan and Shelton (2008) presented a method of solving CTMDP using a linear program. Under assumptions of factorisable reward function, an approximate solution can be computed in linear time. We are currently working on adapting our problem formulation to this framework (Mahboubi and Kochenderfer 2015b).

Model uncertainty

A major assumption in this work is that the positions of the aircraft are exactly known by the system. In practice, we will need to estimate the location of each aircraft from the ground sensors. Doing so would require an observation model and application of Bayes' rule to track a belief state over the aircraft positions. The problem can then be reformulated as a partially observable Markov decision process (POMDP) (Kaelbling, Littman, and Cassandra 1998).

When combined with the CTMDP formulation, this leads to a PO-CTMDP, or a continuous-time hidden MDP (CTHMDP). A literature review indicates that CTHMDPs have received little attention by the research community (Yanjie, Baoqun, and Hongsheng 2005).

The parameters for the 3D simulation (airspeeds, turning radius, controller gains, etc.) were chosen using engineering judgment. These parameters, along with transition probabilities $T(s' \mid s, a)$ and observation model $O(o \mid s)$ for the POMDP, could be derived from data collected of aircraft in the pattern.

Extending the states

The number of states in the pattern and available actions should be extended. For example, if one aircraft is on left base and another is on right base, they are on a collision course but there are no actions in our formulation to prevent collision. A possible extension would be to add commands for S turns. In addition, our model has the taxi state acting as a sink. To enhance realism, it is necessary to incorporate a better model of aircraft behavior when transitioning between the runway and taxi states.

Practical implications

We envision the system issuing commands over a Common Traffic Advisory Frequency. Hence, there needs to be a way to identify the aircraft in the pattern. One way to achieve this is to refer to aircraft by their position, transponder code, or their call-sign inferred through speech recognition (Gunawardana 2012) combined with a VHF direction finder. Additionally, a practical implementation would require the ability to handle special cases such as varying number of aircraft in the pattern, aircraft overflying the runway above the traffic pattern altitude, and change of runway direction due to shifting wind.

References

AOPA Air Safety Foundation. 2000. Nall Report: Accident Trends and Factors for 1999. *AOPA Air Safety Foundation*.

AOPA; EAA; et al. 2015. Open Letter to Federal Aviation Administration.

Billingsley, T. B.; Kochenderfer, M. J.; and Chryssanthacopoulos, J. P. 2012. Collision Avoidance for General Aviation. *IEEE Aerospace and Electronic Systems Magazine* 27(7):4–12.

Campbell, S. D. 2014. Small Airport Surveillance Sensor. In *MIT Lincoln Laboratory ATC Workshop*.

Diefes, D., and Deering, R. 1996. Concept Design for a Low Cost Cockpit Display/Collision Avoidance System for General Aviation Aircraft. In *IEEE Position, Location and Navigation Symposium (PLANS)*, 352–359.

Gunawardana, S. 2012. *A Rule-Based Dialog Management System for Integration of Unmanned Aerial Systems into the National Airspace System*. Ph.D. Dissertation, Stanford.

Holland, J. E.; Kochenderfer, M. J.; and Olson, W. A. 2013. Optimizing the Next Generation Collision Avoidance System for Safe, Suitable, and Acceptable Operational Performance. *Air Traffic Control Quarterly* 21(3):275–297.

Kaelbling, L. P.; Littman, M. L.; and Cassandra, A. R. 1998. Planning and Acting in Partially Observable Stochastic Domains. *Artificial Intelligence* 101(1–2):99–134.

Kan, K. F., and Shelton, C. R. 2008. Solving Structured Continuous-Time Markov Decision Processes. In *International Symposium on Artificial Intelligence and Mathematics (ISAIM)*.

Kochenderfer, M. J.; Griffith, J. D.; and Olszta, J. E. 2010. On Estimating Mid-Air Collision Risk. In *AIAA Aviation Technology, Integration, and Operations Conference*.

Kochenderfer, M. J. 2015. *Decision Making Under Uncertainty: Theory and Application*. Cambridge, MA: MIT Press.

Kunzi, F., and Hansman, R. J. 2011. Mid-Air Collision Risk And Areas Of High Benefit For Traffic Alerting. In *AIAA Aviation Technology, Integration, and Operations Conference (ATIO)*.

Mahboubi, Z., and Kochenderfer, M. J. 2015a. Autonomous Air Traffic Control for Non-Towered Airports. In *USA/Europe Air Traffic Management Research and Development Seminar*.

Mahboubi, Z., and Kochenderfer, M. J. 2015b. Continuous Time Autonomous Air Traffic Control for Non-Towered Airports. In *IEEE Conference on Decision and Control (CDC)*. Preprint.

Nikoleris, T.; Erzberger, H.; Paielli, R. A.; and Chu, Y.-C. 2014. Autonomous System for Air Traffic Control in Terminal Airspace. In *AIAA Aviation Technology, Integration, and Operations Conference (ATIO)*.

Nodelman, U.; Shelton, C. R.; and Koller, D. 2002. Continuous time Bayesian Networks. In *Conference on Uncertainty in Artificial Intelligence (UAI)*, 378–387.

Puterman, M. L. 2005. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Hoboken, NJ: Wiley.

Shelton, C. R., and Ciardo, G. 2014. Tutorial on Structured Continuous-Time Markov Processes. *Journal of Artificial Intelligence Research* 51:725–778.

Yanjie, L.; Baoqun, Y.; and Hongsheng, X. 2005. The Policy Gradient Estimation of Continuous-Time Hidden Markov Decision Processes. In *IEEE International Conference on Information Acquisition*. IEEE.