Specification

Erwan

February 27, 2017

Hereby we present the specifications for the constitution of a benchmark simulator thanks to which we will implement, test and compare funny algorithms in the context of Active Exploration of Unknown Environment for Unmanned Aerial Vehicle.

Scenario

We put ourselves in the context of the control of a glider drone flying in the open air. The environment has several characteristics namely:

- it presents some convective atmospheric phenomenon allowing thermal soaring flight of the agent;

- it is uncertain – the agent only knows a few things about its environment (detailed further);

- it is non-stationary – the location of the thermals and more generally the wind field vary over time, increasing the uncertainty about the environment.

For the sake of illustration, a simplified 2D overview of the problem can be found in figure 1.
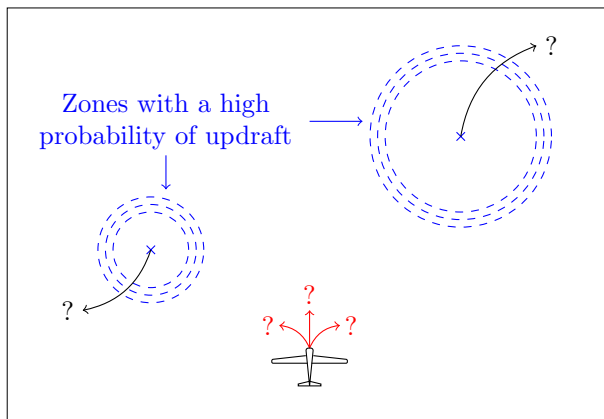

Figure 1: Simplified 2D overview of the problem

In this particular environment, we identified some missions that the agent has to fulfill. Table 2 gives a non-exhaustive list of the considered tasks. Each time we provided a short description of

the mission and tried to identify the main difficulties. Additionally, some extension to a motorized glider or any other aircraft like the very widespread quadcopter could be considered, as well as changes of the environment with the inclusion of obstacles for instance. Indeed, the algorithms to be tested aim at wider applications that we would like to try in the future. For now, we only consider the Thermal Soaring A and B missions.

Name	Label	Description	Difficulties
Thermal Soaring	A	The glider has to maximize its autonomy (altitude + velocity) in a convective situation i.e. inside a thermal. The latter is non-stationary, it can move and fade away.	Find the optimal trajectory maximizing the autonomy gain
	B	The glider has to maximize its autonomy in a field of thermals whose positions are unknown.	Exploration-Exploitation trade-off within the updraft map
Thermal Soaring + Travel	C	The glider starts from a point A and has to go to a point B which is far away. In order to achieve this goal it needs to gain autonomy with Thermal Soaring.	Exploration-Exploitation trade-off within the updraft map + deal with potentially antagonist missions
	D	Same as mission-C but with a list of points (A, B, C, ...). The list can be either ordered or not, meaning that the agent has to find its own path optimally.	Difficulties like in mission-C + establishment of a traveling plan
Thermal Soaring + Exploration	E	The glider has to fly by every places of a zone e.g. exploration mission. Obviously is must gain autonomy for the goal to be reachable.	Exploration-Exploitation trade-off within the updraft map + finding a shortest path trajectory for the exploration of the zone + deal with potentially antagonist missions
Thermal Soaring + Surveillance	F	The glider has to keep a zone under surveillance. An external player tries to reach a goal within this zone.	Have an optimal policy in the context of a 2-player game with a disequilibrium in the information access (the drone does not know the opponent's position and movement but the opposite is false)

Figure 2: Table of the considered missions

Simulator

A Reinforcement Learning (RL) problem can be modeled as an agent-environment interface. As described in figure 3, the environment is composed with an atmospheric model and an aircraft model together providing the agent with a next state s' and a reward r given an action a. Additionally, one can separate the absolute state space (state of the whole system) from the observable state space (available for the agent), dividing the notion of state space in two parts. In this section, we detail these seven components.
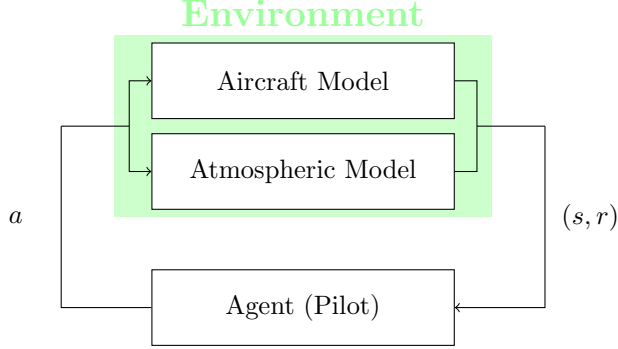

Figure 3: Overall architecture

Atmospheric model; we seek for a realistic model of the atmospheric phenomenon including the convective effects a.k.a. thermals. State of the art thermal models have been developed at ISAE based on the work of Allen. This model could be interfaced with any simulator and seems to be a good basis to start with. In addition, a global wind direction (constant wind field) and some noisy effects could be enough for our needs.

Aircraft model; we used a 1 point dynamic model of the aircraft in former work about the same subject [Beeler et al., 2003]. This simple model allows us to recreate realistic behavior of a glider drone and can be reused in a first time. However, some model instabilities have been spotted, particularly when it is about setting a maximum threshold in the range of the attitude angles. Thus, we can either decide to fix these issues in order to keep on working with it or chose another model. It could be interesting to use a benchmark model such that those provided by the open source flight simulator JSBSim. Indeed, the use made of the latter in the scientific community allows for more visibility.

Absolute state space; it includes the whole information about the state of the aircraft and the environment. Detailing a specific format is not necessary here and will depend on technical considerations relative to the code.

Observable state space; it includes what the agent sees i.e. the information harvested by the sensors. Here we try to establish an exhaustive list:

- (x, y, z) the position in the 3D space, practically acquired by a GPS system and in our case

we can simply access its real value used for the environment's transition;

- (V, V_{air}, V_{wind}) respectively the velocity of the aircraft in the earth frame acquired with a GPS; the velocity of the wind in the aircraft frame acquired with a pitot tube; and the velocity of the wind in the earth frame which can be obtained via the following formula:

$$V = V_{wind} - V_{air}$$

- (θ, ϕ, ψ) respectively pitch angle, roll angle and yaw angle, representing the attitude of the aircraft in the earth frame;

- (α, β) respectively angle of attack and side slip angle, the first represents the angle between the direction of the airflow and the aircraft's reference line (e.g. chord line), the second is the angle between the velocity vector and the aircraft plane of symmetry;

- (γ, χ) respectively the flight path angle representing the angle between V and the earth horizontal; and the course heading representing the angle between V and the vector indicating the north in the earth frame. We have the following formula:

$$\gamma = \theta - \alpha$$
$$\chi = \psi + \beta$$

- the rate of the aforementioned quantities, especially the climbing rate $\dot{z}$ measuring the vertical velocity of the aircraft in the earth frame;

- Vision-based data: here is a tricky one, we intend to use vision based data giving some clue about the position of the thermals (clouds for instance). The difficulty is that it has to be uncertain in a realistic way. Some clouds indicate the presence of a thermal but some thermals do not have clouds and vice versa. We propose to simulate the effect by randomly decide on the presence of a cloud on top of a thermal. Additionally some clouds can be randomly added to the environment. The input given to the glider could be the position of clouds in the vision field.

Reward; the reward is essentially built with the data of the absolute state space accordingly to the chosen mission.

Action space; several level of actions can be defined, from the lower to the higher we can have the following: the actuators control; the attitude of the aircraft described in the state space; and the trajectory of the aircraft. From a spatial point of view i.e. if we characterize the environment spatially (for instance if there is a cloud at position (x_c, y_c) how can I get there) it can be convenient to consider the action space as a trajectory space. However more practically with respect to what we control on board of a UAV the attitude of the aircraft seems more realistic.

Agent; The agent is the pilot i.e. the policy realizing the processing of the harvested data and the deterministic or probabilistic mapping between state and action. No definition is provided here.

The aforementioned observable state space is fully detailed, but it is not relevant to consider its entirety for every missions. We separated the relevant features of the state space with respect to the missions presented in table 2 in table 4. Similarly the action space is not always the same and details are also provided in this table.

Mission Label	Relevant state components	Relevant action components
A	$V_{earth}, \dot{z}$	θ, ϕ, β
B, C, D, E, F	$V_{earth}, \dot{z}, x, y, z, Vision$	θ, ϕ, β or $\dot{x}, \dot{y}, \dot{z}$

Figure 4: Table of the relevant state and action components for each mission presented in table 2

References

[Beeler et al., 2003] Beeler, S. C., Moerder, D. D., and Cox, D. E. (2003). A flight dynamics model for a small glider in ambient winds.