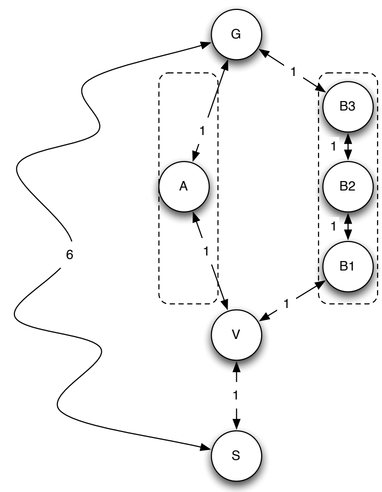
Field exam — Autonomy

2010

A robotic vehicle must move from a starting point S to a goal point G in minimum time. There is a safe "highway" path from S to G, shown to the left in the figure below, which is known to require 6 time units[1]. The vehicle also has the option to travel to a vantage point V on the way to G; traveling from S to V (and vice-versa) requires 1 time unit. Continuing from V towards G, the robotic vehicle can take one of two possible "shortcuts," A and B, requiring 2 and 4 time units, respectively, to travel from V to the goal G. See the figure below.



However, the shortcut paths are patrolled by a hostile UAV, which is programmed to move among the two shortcut paths, according to a fixed, but unknown, control policy. From the vantage point V, the (friendly) robotic vehicle can observe where the (hostile) UAV is at any time, while remaining undetected. The hostile UAV is able to move in one time unit from one shortcut path to the other. If at any time the friendly vehicle and the hostile UAV are on the same shortcut path, the former will have to stop and hide in the foliage for a time unit, to avoid being spotted and destroyed. The vehicle always has the option of stopping for one or more time units at any location.

Discuss how you would determine the optimal policy for the robotic vehicle and the optimal cost in the following cases. Please make sure to clearly describe the approach, it is not necessary to carry out all the calculations.

1. The UAV control policy is as follows: the UAV patrols path A for 2 time units, then path B for 4 time units, and repeats this pattern. The initial state is such that the UAV has just moved to path A.
2. The UAV control policy is as above, but the initial state of the UAV is not known.
3. The UAV control policy is the following:
   - if the UAV is patrolling path A at a given time step, with probability $1/2$ will remain on the same path at the next time step, and with probability $1/2$ will move to path B.
   - if the UAV is patrolling path B at a given time step, with probability $3/4$ will remain on the same path at the next time step, and with probability $1/4$ will move to path A.

[1] For simplicity, you can assume that time is discrete.

Solution, part 1

Let us indicate with $g_{i,j}(t)$, $i, j \in M = \{S, V, A, B_1, B_2, B_3, G\}$, the time needed to move directly from node i to node j. If nodes i and j are not connected at time t, we set $g_{i,j}(t) = \infty$. Based on the problem data, we have that

- $g_{i,i}(t) = 1$, for all $i, j \in M$, and $t \in \mathbb{N}$.
- $g_{S,G}(t) = 6$, for all $t \in \mathbb{N}$.
- $g_{S,V}(t) = g_{V,S}(t) = g_{V,A}(t) = g_{V,B}(t) = 1$ for all $t \in \mathbb{N}$.
- $g_{A,G}(t) = g_{A,V}(t) = 1$ if $t \mod 6 \in \{0, 1\}$, $g_{A,G}(t) = g_{A,V}(t) = \infty$ otherwise.
- ... similar formulas for the other transitions.

This is a shortest path problem with time-varying costs, which can be solved using dynamic programming. For example, define $f_i(t)$, $i \in M$, as the length of the shortest path starting from node i at time t. It is desired to compute $f_S(0)$, and the path achieving the corresponding shortest path. Note that since we know that $f_S(0) \leq g_{S,G}(0) = 6$, we can limit our analysis to $t \in \{0, 6\}$. Furthermore, given the periodicity of the UAV trajectory, $g_{i,j}(t) = g_{i,j}(t + 6)$, for all i, j, t.

The problem can be solved using dynamic programming, i.e., relying on the principle of optimality and Bellman's equation, which takes the form

$$f_i(t) = \min_{j \in M} \left[g_{i,j}(t) + f_j(t + g_{i,j}(t)) \right], \quad i \neq G,$$

$$f_G(t) = 0.$$

For example, one can fill the following table using a Dijkstra-like algorithm (i.e., starting from the "G" (goal) line, filled with zeroes):

$f_i(t)$	0	1	2	3	4	5
S	3	3	3	3	5	4
V	3	2	2	2	2	4
A	3	2	1	1	1	1
B_1	3	3	7	6	5	4
B_2	2	6	6	5	4	3
B_3	1	1	5	4	3	2
G	0	0	0	0	0	0
UAV	A	A	B	B	B	B

From the table, we get $f_S(0) = 3$, and the optimal path is (S, V, A, G).

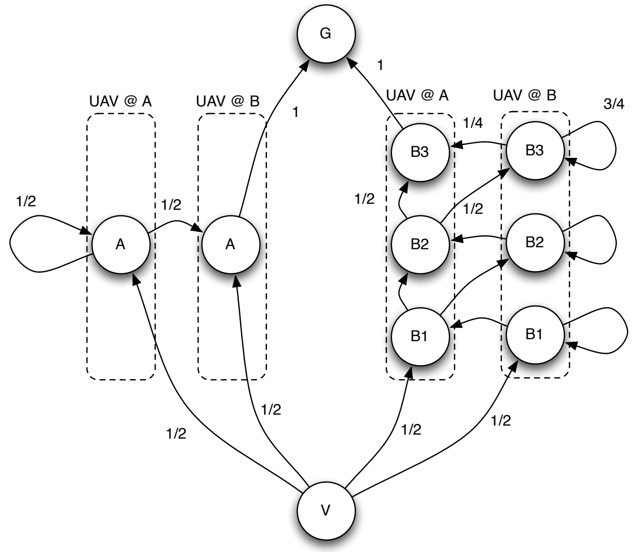
Solution, part 2

Given the table, we find that the optimal policy requires to go through V and A, and does **not** depend on the observation made at V. In the worst case, the cost could be 5 time units, which is better than the "safe" past of cost 6.

Solution, part 3

To simplify the problem, it is convenient to consider the subproblem of computing the optimal cost-to-go from point V, in the two cases in which the UAV is observed to be patrolling either path.

Each of these cases is a stochastic shortest path problem, and can be solved, e.g., using value or policy iteration. See the figure below for a sketch of a transition system modeling one of the sub problems. Each transition cost is equal to one; labels indicate transition probabilities. As a starting value function (or policy), one could use, e.g., that derived from part 1.



Assuming the optimal cost of the physical/information states $(V|\text{UAV@}A)$ and $(V|\text{UAV@}B)$ is known, the optimal cost of the starting point V can be obtained by computing the probability of the two events (UAV observed at A or B). This can be done recognizing that the UAV behavior is in fact a Markov chain, with transition matrix:

$$P = \begin{bmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{bmatrix}$$

The stationary distribution of the Markov chain can be obtained as the (normalized) left eigenvector of the matrix P associated with the unit eigenvalue, i.e.,

$$\Pr[\text{UAV@}A] = 1/3, \qquad \Pr[\text{UAV@}B] = 2/3.$$

Summarizing, we will get that

$$J(V) = \frac{1}{3} J(V|\text{UAV@}A) + \frac{2}{3} J(V|\text{UAV@}B)$$