Efficient Single-Agent Capture of a Moving Target

Martin Freeman

AA228

Problem Statement

- Pursuer seeks to capture a non-adversarial target:
  - Want to do this efficiently (minimum time) or the target may escape
  - The pursuer has only an initial idea of the target location
  - While target is non-adversarial, the belief of the target position with disperse over time-steps
  - Both the pursuer and the target are active (i.e they are constantly changing rooms)
  - Pursuer and target move one room at a time, can only see what is in their current room

- Given initial idea of the target location, how can we capture the target in the *minimum expected time?*

Belief state of target pos: $b(t)$

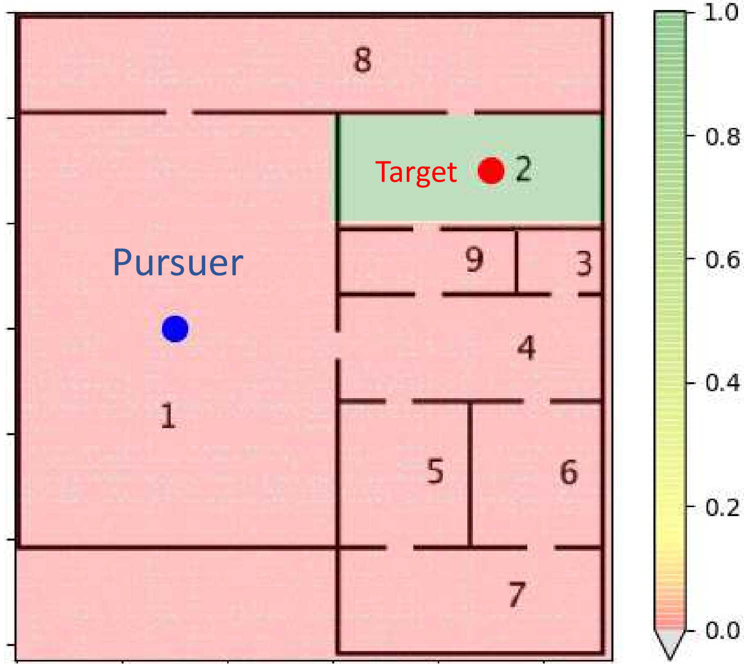


Fig. 1: Pursuer, Target in House Environment

Relevant Literature

- Efficient Multi-Robot Search for a Moving Target (Hollinger et. Al, 2009)
  - Efficient Search Path Planning (ESPP) approach uses finite-horizon planning to solve the problem for multiple searchers
  - Yielded results comparable to Heuristic Search Value Iteration (HSVI) in the POMDP formulation with a single searcher
    - Despite this, becomes quickly intractable with larger state spaces and more than one pursuer
  - ESPP approach can thus lead to accurate, tractable solutions to problems up until a certain depth (i.e finite horizon)
  - Lower bounds achieved for finite-horizon path enumeration on the ESPP problem:

$$F(A^{FH}) \geq F(A^{OPT}) - R\gamma^{d+1}$$

  - That is, the returned trajectory reward is greater than or equal to the optimal trajectory reward minus the discounted capture reward

Approach

- <u>Introduce problem constraints:</u>
    - Pursuer has a *deterministic* motion model
    - Target has a *probabilistic* motion model, **D**, known to pursuer
    - Pursuer cannot see through doorways
    - Pursuer observes the current room at every time-step
    - Discretize the world into an undirected graph (Fig. 2)
- <u>Develop a model:</u>
    - Include belief dispersion and capture mechanics in belief vector:
        - $b_{1:N}(t)$ are the probabilities for the target being at each node

$$b(t) = [b_0(t), \ldots, b_N(t)]$$

    - Disperse belief according to **D** every time-step, incorporate observation of pursuer's current room with a node-dependent matrix $\mathbf{C}_{s(t)}$, based on pursuer node s(t)
    - Evolve the belief vector as:

$$b(t+1) = b(t)DC_{s(t)}$$

    - Evolve the state as a function of the control path, U

$$S(t) = [s(t), b_0(t), \ldots, b_N(t)] \qquad S(t+1) = f(S(t), U(t+1)).$$

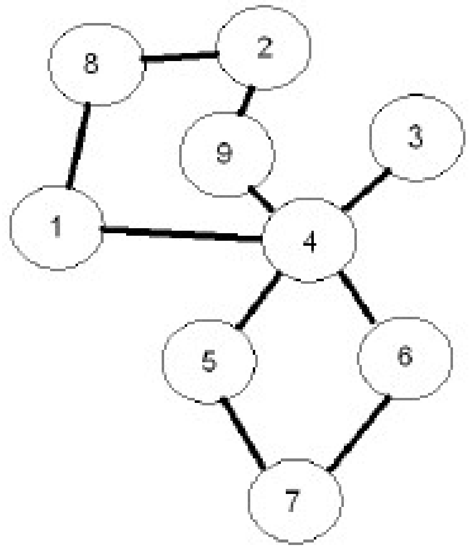


Fig 2. Graph Representation Fig. 1

Approach

- <u>Determine the d-length pursuer path that maximizes reward:</u>
  - Express the discounted reward (objective):

$$J(U(1), \ldots, U(d)) = \sum_{t=0}^{d} \gamma^t P(s\ (t) = e(t))$$

  - Enumerate all d-length paths for pursuer, target
  - Apply the discounted reward for each pursuer path based on ALL d-length target paths and their probabilities:

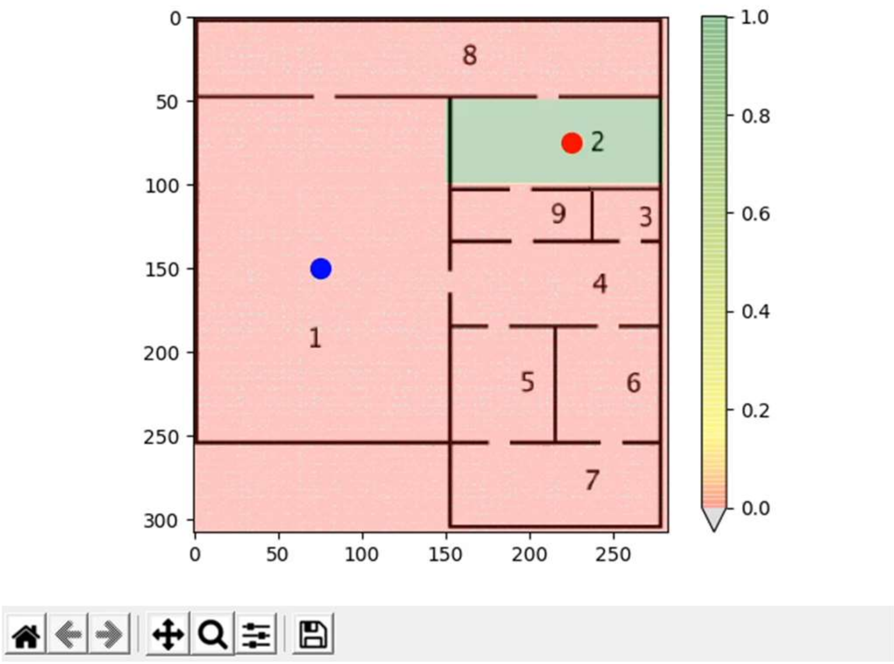
$$J(A) = \sum_{Y \in \Psi} P(Y) F_Y(A)$$

  - Choose the pursuer path(s) that thus lead to the *highest probability of quickly capturing the target*

$$A^{FH} = \underset{A}{\operatorname{argmax}}(J(A))$$

Experiment

- Evaluate for agent position 1, vary target position and depth
- Benchmark:
  - Compare *ESPP finite horizon path* against the average of **100** *random pursuer paths*
  - Generate **1000** target paths sampled from target path distribution P(Y)
  - Compare average discounted reward and capture rate for both approaches over the target paths
  - Parameters: R=1, discount factor gamma = 0.95



Results

- Compare the average discounted reward and capture rate:
    - As expected, higher average reward and capture rates for the ESPP path
- Very reasonable runtime (for d=8, maximum runtime was 6 seconds), but quickly becomes intractable with increased depth

- Finite-horizon ESPP can be a helpful tool for problems with a small state space and limited time (horizon) to maximize reward