Autonomous Area Coverage with Unmanned Aerial Vehicles

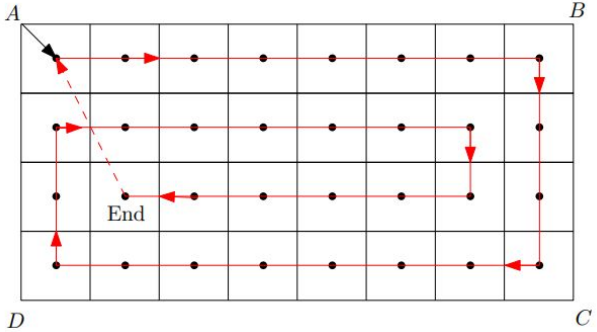
With Prof. Esra Kadioglu
Summer 2020

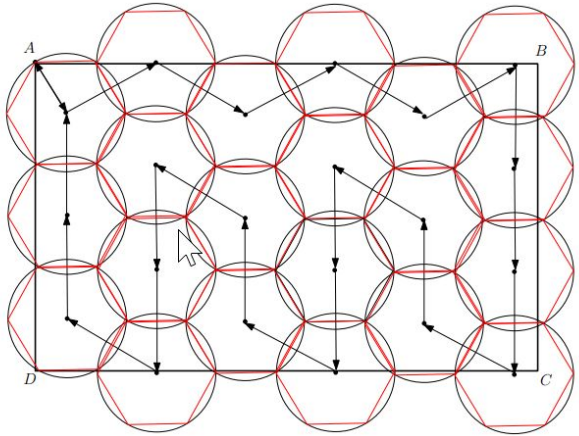
Elisabeth Landgren, Frank Zhang, Charles Zhang, and Aaron Gould

What we are going to talk about

- What is Area Coverage?
- What is Reinforcement Learning?
- How Have we Applied Them?
- What Are Our Next Steps?

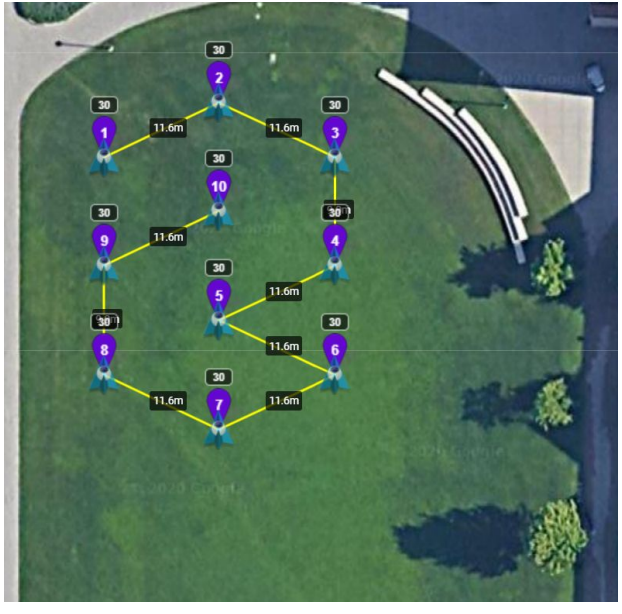
What is Area Coverage?



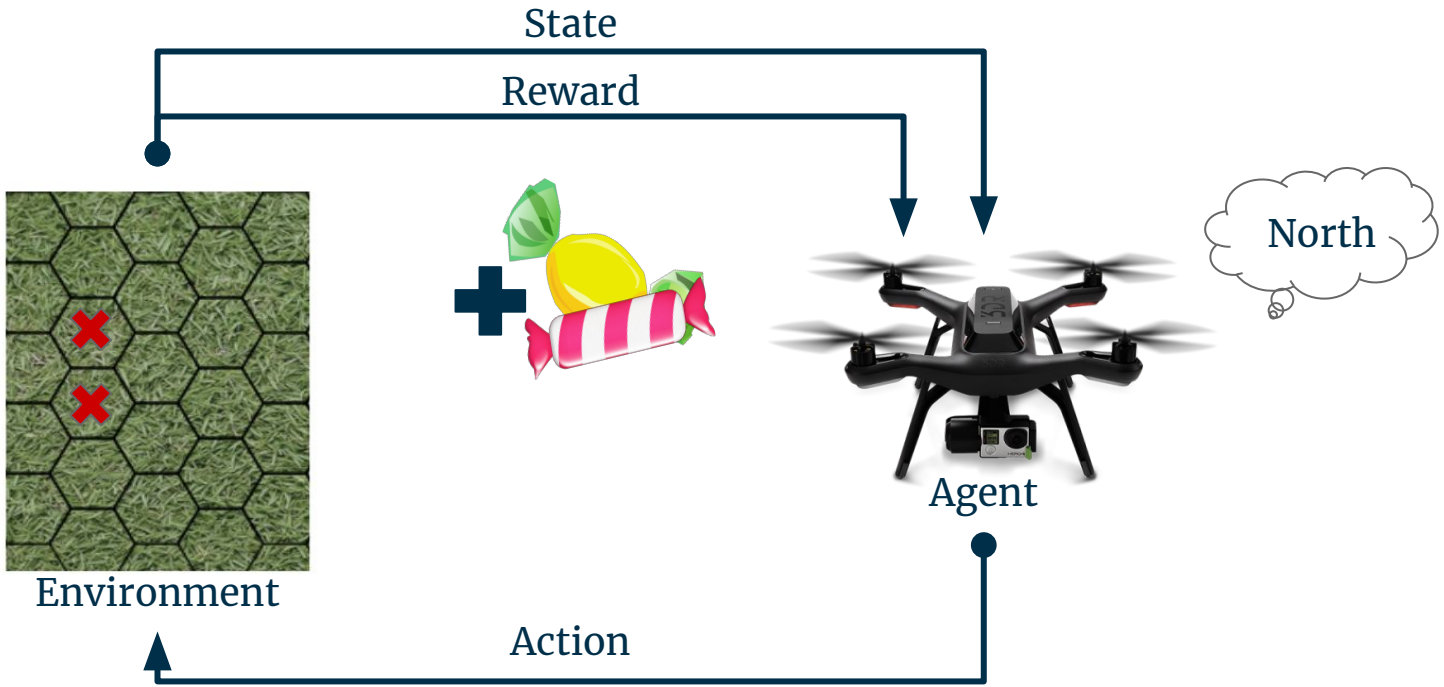


Source: UAV Coverage Using Hexagonal Tessellation

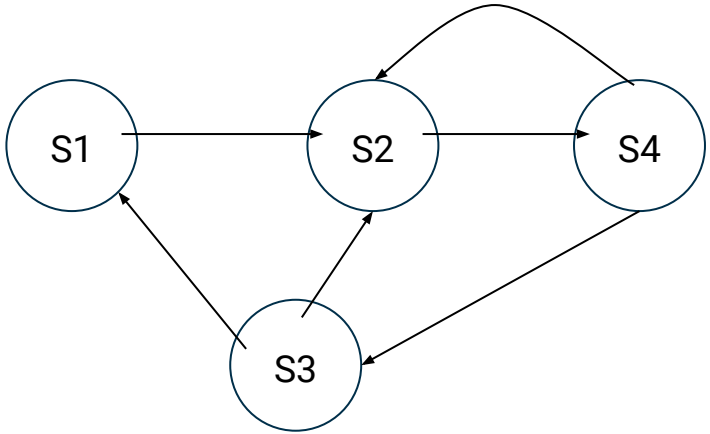
What is Area Coverage?



What is Reinforcement Learning?



What is Reinforcement Learning?



- Rewards impact the probability of actions
- Markov Decision Processes (MDP)
- Grid World

Naive Q-learning

For each episode:

For each step:

1. At time step t, we start from state Sk and pick action according to Q values, and apply ε-greedy policy

Exploration (Random move)

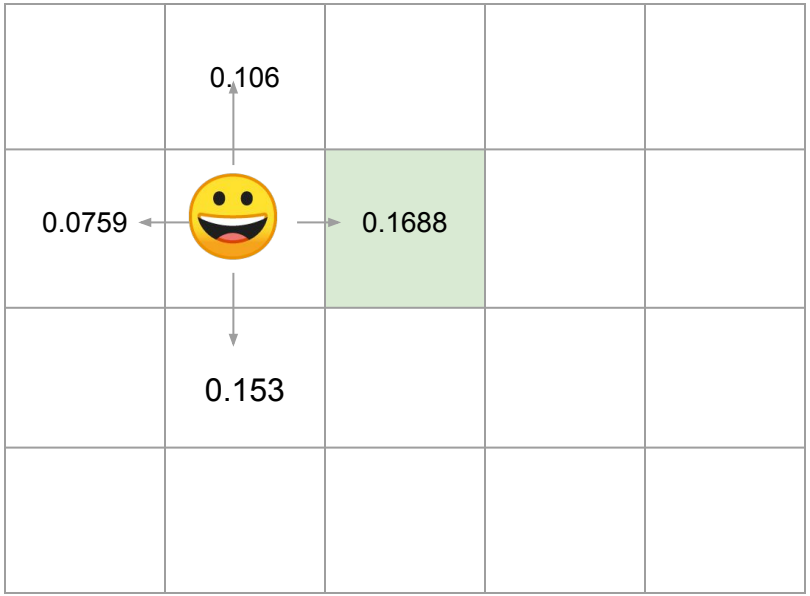
Exploitation (Greedy move)

$$\pi(s) = \begin{cases} a \in A, & \mathbb{P} = \epsilon \\ a \in \arg\max_{a\prime} Q_k(s_k, a\prime), & \mathbb{P} = 1 - \epsilon \end{cases}$$

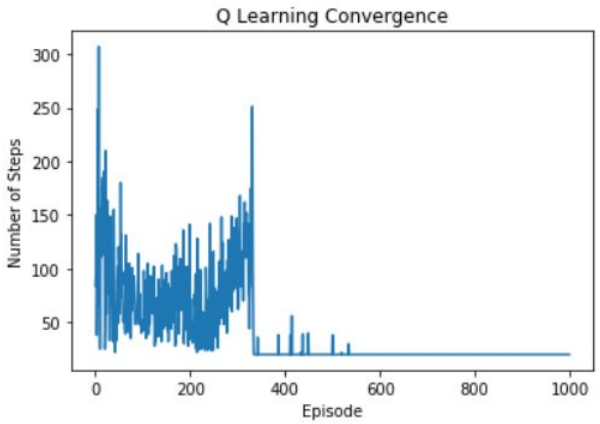
2. With action A_t, we observe reward R_{k+1} and get into the next state S_{k+1}.
3. $Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha(R_{t+1} + \gamma \max_{a \in A} Q(S_{t+1}, a) - Q(S_t, A_t))$
4. Repeat from step 1

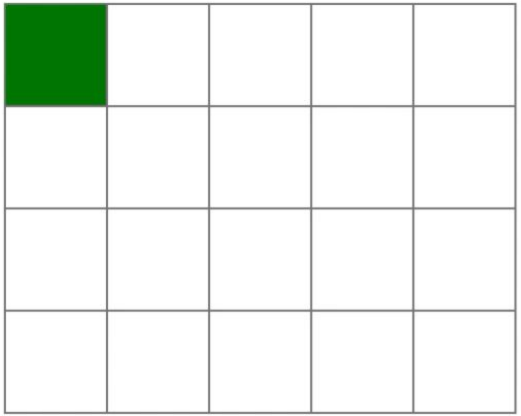
Stop if the agent covers the field and returns to the starting point

Find the Path based on the Q-table



Result





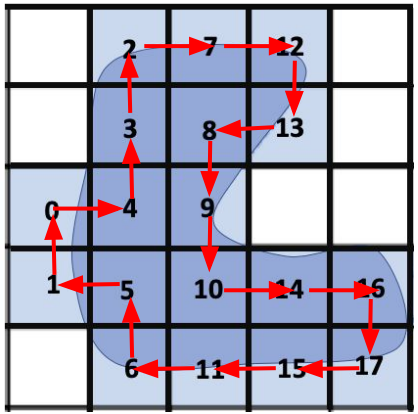
Path on 4x5 Grid

Our New Algorithm

Non-MDP to MDP

Training to find the shortest path is MDP

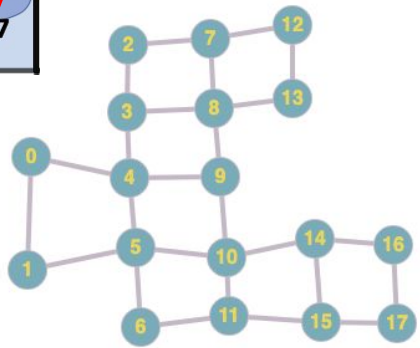
Avoid shortest path ⇔ visit more ⇔ Coverage



Graph Representation

Only consider the possible actions(vertices i, j are connected)

Working on any kinds of graph, not only Grid World



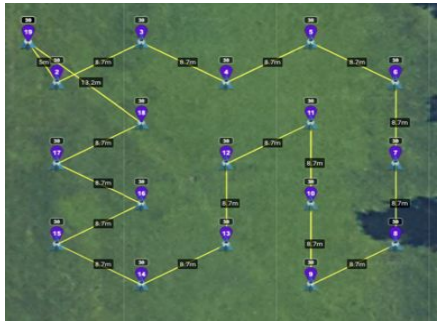
Our New Algorithm

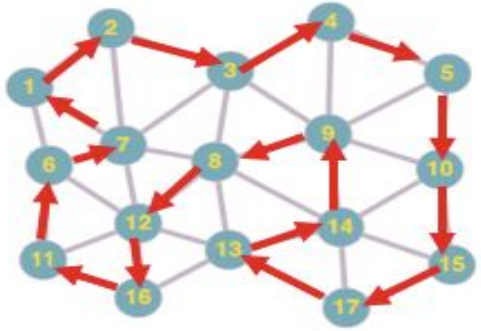
Improvement: 100% working in at least 20x20(400 Vertices) Grid World

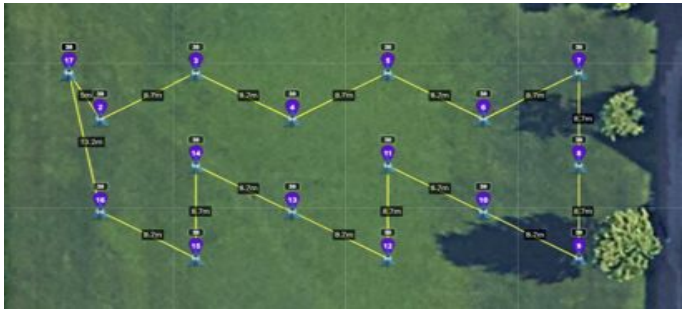
Algorithm 2 Q Learning with Graph-Based State Representations

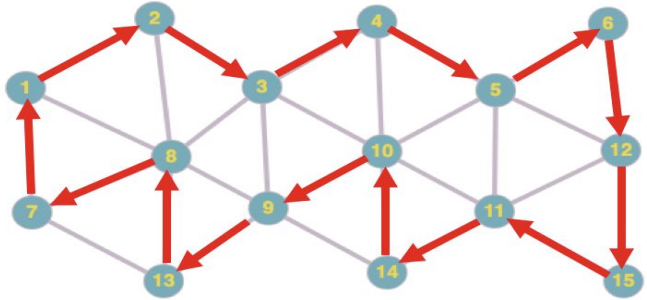
Initialize $\mathcal{S} = \{1, \ldots, V\}$
Initialize $\mathcal{R}$: $R[S, a]$, $\forall S, a$
Initialize $\mathcal{Q}$: $Q[S, a] = NA$, $\forall S, a$
while not reach the number of episodes **do**
    $S \leftarrow$ random $S \in \mathcal{S}$
    while $S \neq END$ **do**
        $a \leftarrow$ random a valid in the S
        $S' \leftarrow a$
        $Q[S, a] \leftarrow (1 - \alpha) \cdot Q[S, a] + \alpha \cdot (R[S, a] + \gamma \cdot \max Q[S', a'])$
        $S \leftarrow S'$
    end
end
$S = START$
while length(Path) $< V$ **do**
    Path.append(S)
    $S' \leftarrow S$
    $S \leftarrow$ argmin($Q[S, \ ]$)
    $Q[S', \ ] \leftarrow NA$
    $Q[\ , S'] \leftarrow NA$
end
Output: Path

Comparisons









What Are Our Next Steps?

One Drone Coverage

- Translate to GPS waypoints
- Graph without Hamiltonian circuit
- Search in the irregular field without waypoints, considering the energy, resolution, e.t.c. constraints(may be possible to translate these to edge weights)

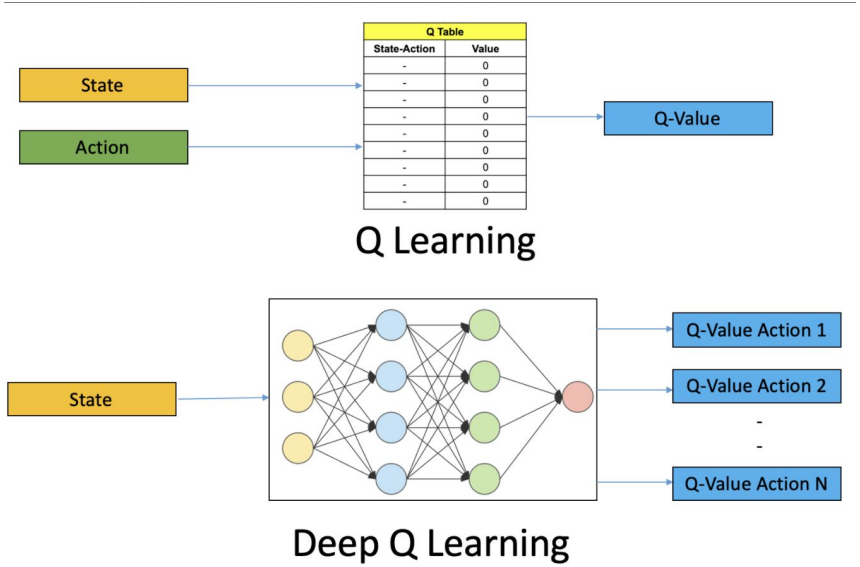
Multi-Drone Coverage

- Field decomposition
- Choosing the start points
- Multi-agent reinforcement learning

What Are Our Next Steps?

Optimizations in a Large Grid World:

- Deep Q Network(DQN)
- Non-Markov Decision Process(NMDP)



Source

Questions?