AA228 Final Project:

GitHub Repo: https://github.com/mmphan98/AA228_Aircraft_Landing.git

Landing a Plane on Final Approach with Q-Learning

Winter 2023

Poom Prasopsukh
Michael Phan

Acknowledgements and References

Course Staff:

Marc Schlichting
Hanna Yip
Sydney Katz

Relevant Related Work:

Open AI Gym Mountain Car
Vare and Sarkar’s Automated Aircraft Landing with RL

Overview

Problem Description

Landing a Plane on Final Approach

Goal: Maintain a stable trajectory throughout the final approach and successfully land the plane

Final Approach: Roughly defined to be <1 km away and 50 m above ground

The Plane: Cessna 172

Landing Success Criteria:

- Speed is below landing speed (65 knots, 33.4 m/s) and above stall speed (48 knots, 25 m/s)
- Vertical component of speed not to exceed a defined buffer (roughly 300 ft/min)
- Plane is located at the “runway”

Stochasticity: Impact of Wind Speed Variability

- Head/Tail wind only

Approach

Q-Learning with Epsilon Greedy Exploration

Algorithms taken from Kochenderfer's *Algorithms for Design Making*[3]

Parameters:

- Discount Factor (γ) = 1
- Learning Rate (α) = 0.3
- Exploration Parameter (ϵ) = *variable*

```julia
mutable struct QLearning
    𝒮 # state space (assumes 1:nstates)
    𝒜 # action space (assumes 1:nactions)
    γ # discount
    Q # action value function
    α # learning rate
end

lookahead(model::QLearning, s, a) = model.Q[s,a]

function update!(model::QLearning, s, a, r, s′)
    γ, Q, α = model.γ, model.Q, model.α
    Q[s,a] += α*(r + γ*maximum(Q[s′,:]) - Q[s,a])
    return model
end
```

Modeling of the Cessna 172

Implemented as a **mutable struct** for updates using a dynamics model, in the *continuous* space

Attitude Characteristics:

- **x** → x-coordinate (m)
- **y** → y-coordinate (m)
- **θ (th)** → pitch of the airplane (rad)
- **power** → throttle setting (kg)

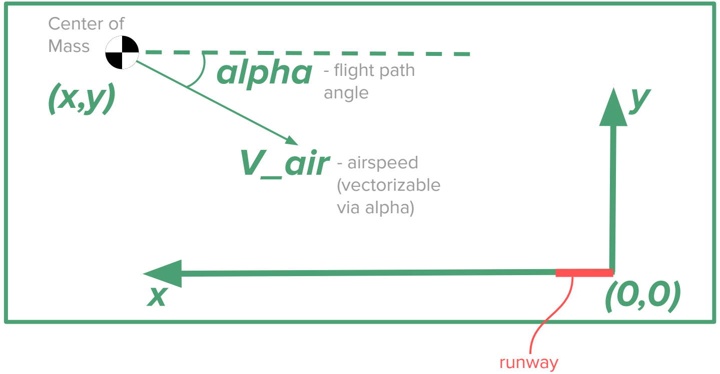
Flight Characteristics:

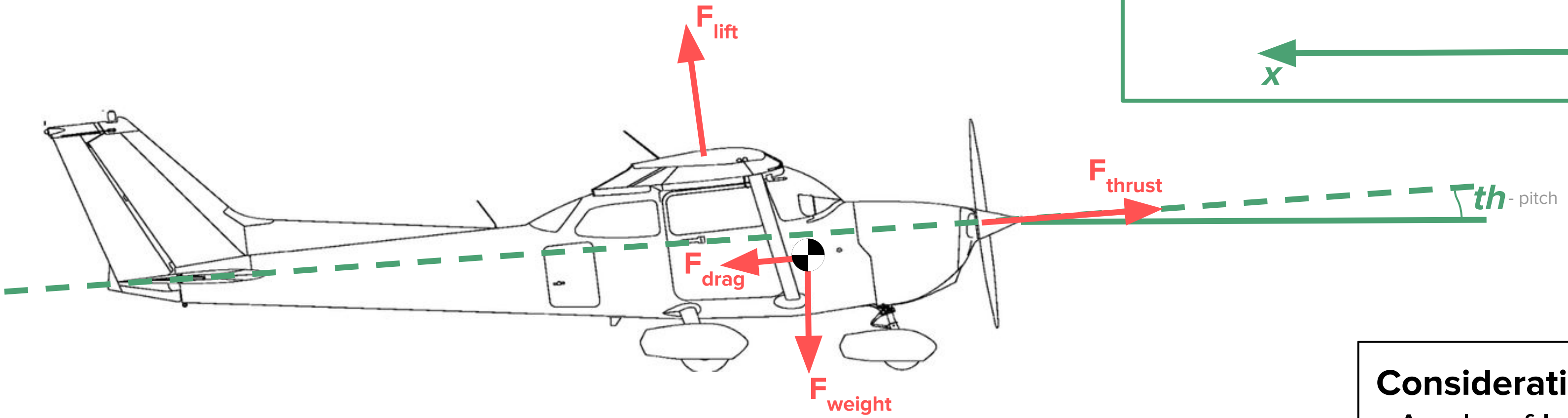
- **V_air** → airspeed of the airplane (m/s)
- ***α* (alpha)** → flight path angle relative to ground (rad)

State Tracking:

- **landed** → boolean to track weather the plane has landed or not (true/false)

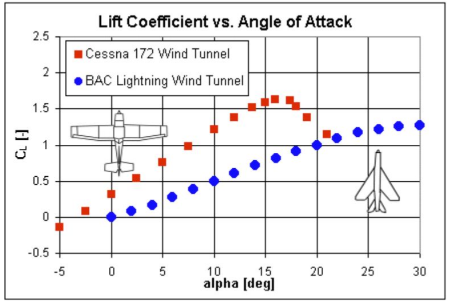
Modeling of the Cessna 172



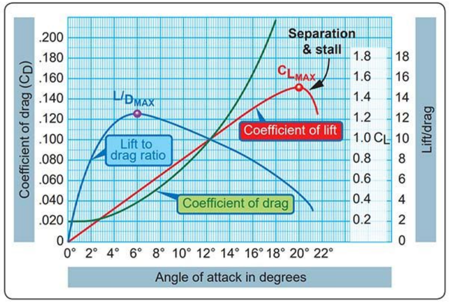


Considerations:
- Angle of Incidence (angle of attack when *th* = 0)

Neglections:
- No flaps
- Roll and Yaw
- Pitch Moments
- Transients between pitch/power settings



Lift Coefficient Curve vs Angle of Attack[4]*



Drag Coefficient Curve vs Angle of Attack[1]*

*Extracted via Web Plot Digitizer: https://apps.automeris.io/wpd/

State and Action Space Discretization

State Space: Total size of **114,114 states**

x → [-900: 0]m	Bucket Size = 50	Size = 19
y → [0: 50]m	Bucket Size = 5	Size = 11
V_air → [25: 60]m/s	Bucket Size = 1	Size = 36
α → [-0.13: 0.07]rad	Bucket Size = 0.01	Size = 21

Action Space: Total size of **9 actions**

th → up, same, down (rad)	Range: [-0.1745, 0.1745]rad	Increment = 0.01	Size = 3
power → up, same, down (kg)	Range: [20, 200]kg	Increment = 10	Size = 3

Dynamics Model as the Transition Model

Finite Time Step

- **dt** = 1 second

Sum of Forces in *x* and *y*

- **Fx** = $power_x - lift_x - drag_x = ma_x$
- **Fy** = $-weight + power_y + lift_y - drag_y = ma_y$

Update velocities based on acceleration vector

- $V_x = V_x^{(i-1)} + a_x$ ***dt**
- $V_y = V_y^{(i-1)} + a_y$ ***dt**
- **V_air** = $\sqrt{(V_x^2 + V_y^2)}$
- $\boldsymbol{\alpha}$ = $\arctan(V_y, V_x)$

Head/Tail Wind

- Only affects velocity relative to ground, *no affect* on **V_air**

Update State Space Variables

- **x, y, V_air, $\boldsymbol{\alpha}$**

Reward Model

Pitch changes	- 10
Power changes	- 10
Going out of bound	- 1000
Crashing / Stalling midair	- 1000
Deviating from optimal alpha	$- 60 * (\alpha_{target} - \alpha_{current}) / \alpha_{target}$
Distance from runway	$+ 120 * (dist_{max} - dist_{current}) / dist_{max}$
Crashing at runway	- 500
Landing	+ 5000

Results & Analysis

Q-Learning and Epsilon Greedy Parameters

Total Iterations = **400,000**

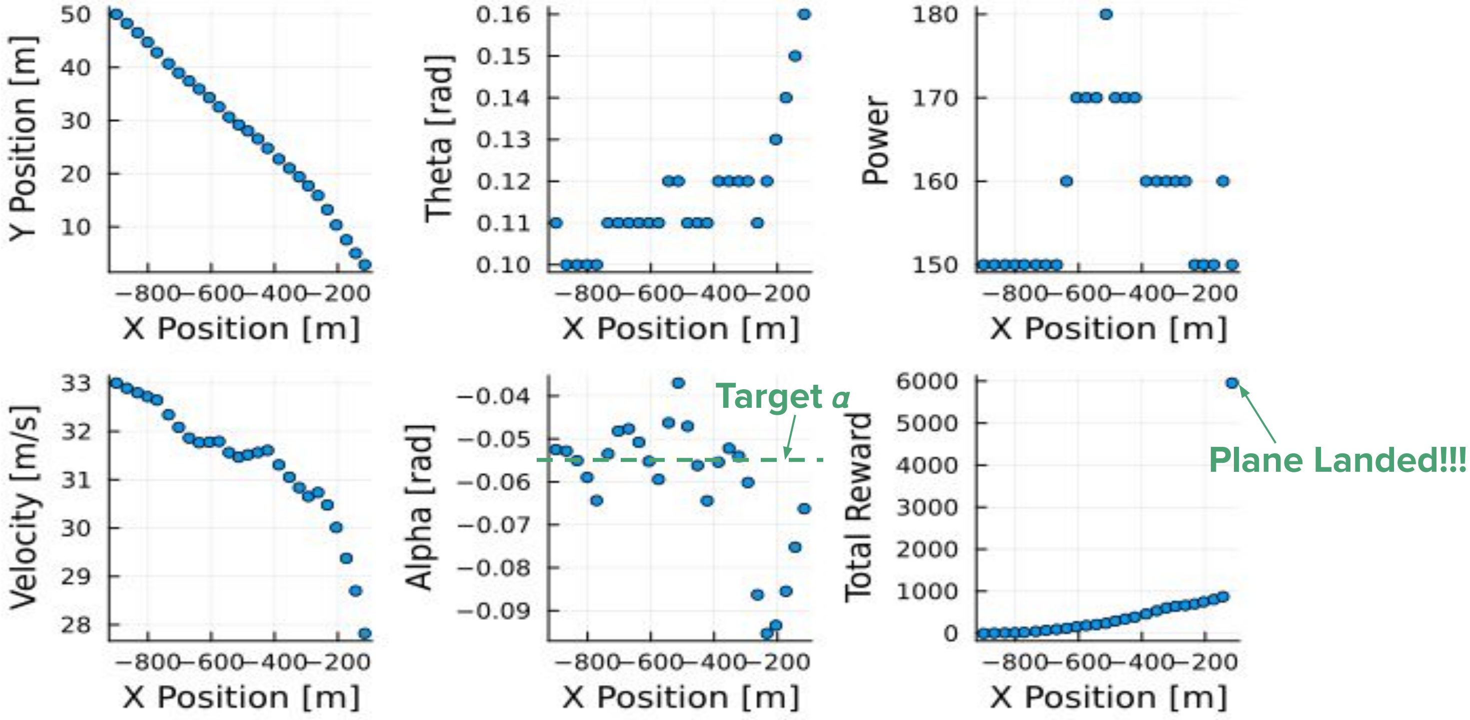
Q-Learning:

- Alpha (α) = **0.3**
- Gamma (γ) = **1**

ϵ-Greedy:

- 0 < Iterations < 280,000: Epsilon (ϵ) = **0.78 → 0.06** **(Decaying)**
- 280,000 ≤ Iterations ≤ 400,000: Epsilon (ϵ) = **0** **(Constant)**

Optimal Policy



Reward Progression vs Iterations

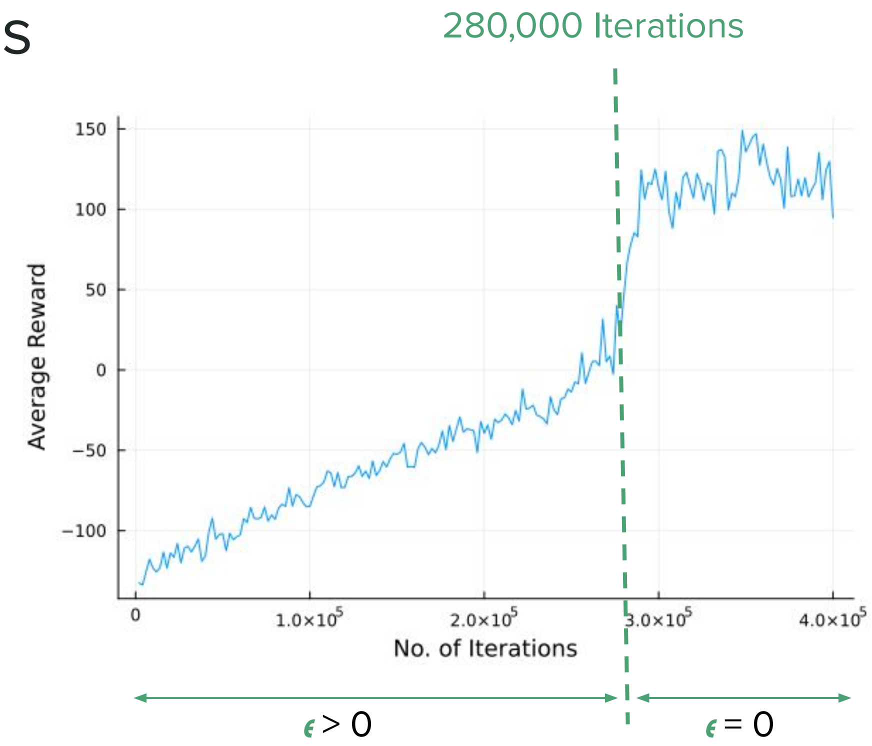
Average Reward

- For every 2000 iterations interval in the dataset

Converges Around +125

- Model still occasionally fails to land due to wind variability

Pitch changes	- 10
Power changes	- 10
Going out of bound	- 1000
Crashing / Stalling midair	- 1000
Deviating from optimal alpha	$-60 * (\alpha_{target} - \alpha_{current}) / \alpha_{target}$
Distance from runway	$+120 * (dist_{max} - dist_{current}) / dist_{max}$
Crashing at runway	- 500
Landing	+ 5000



Conclusion

Conclusion

- Plane landed!!!
- Q-Learning + ϵ-Greedy Limitations:
    - Simplified state/action space
    - Large time step limits possible chances for corrective actions
    - Lack of optimal policy for every state
- Potential Improvements
    - Consideration of transient effects and moments
    - Reward shaping to encourage more natural landing behavior
    - Improving model to 3-dimensions
    - Refining state/action space “buckets”

References

References

1. “Aircraft Power Curve: Private Pilot Online Ground School.” FLY8MA Online Flight Training, 13 Jan. 2021, https://fly8ma.com/topic/aircraft-power-curve/.

2. Chen, Mingtai. “Static Thrust Measurement for Propeller-Driven Light Aircraft.” Proceedings of the 2nd International Conference on Computer Application and System Modeling, 2012, https://doi.org/10.2991/iccasm.2012.165.

3. Kochenderfer, Mykel J., et al. Algorithms for Decision Making. MIT Press, 2022.

4. Scott, Jeff. “Finite and Infinite Wing.” Aerospaceweb.org | Ask Us - Finite and Infinite Wing, 14 Mar. 2004, https://aerospaceweb.org/question/aerodynamics/q0167.shtml.

5. Suharevs, Artūrs, et al. “Dynamic Model of Aircraft Landing.” Transport and Aerospace Engineering, vol. 3, no. 1, 2016, pp. 38–43., https://doi.org/10.1515/tae-2016-0005.

Thank You!

Poom Prasopsukh

Michael Phan