Analysis and simulation of The Lady in the Lake problem

Winston (Yuntao) Wu

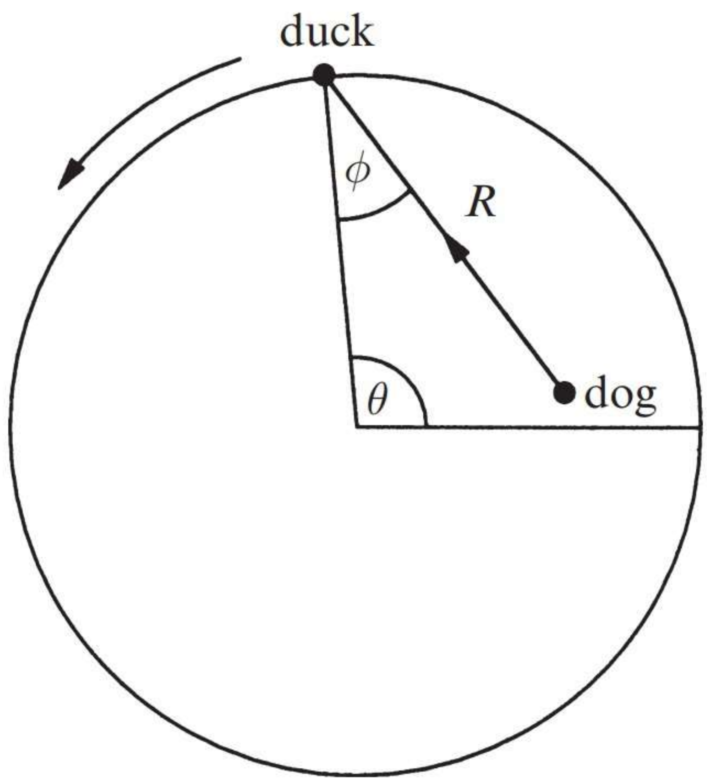
November 30, 2022

Motivation Problem

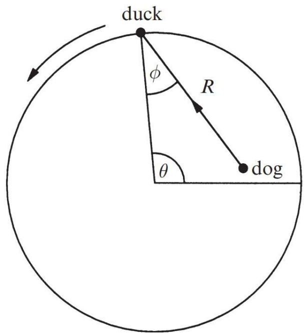
Problem 7.1.9 from (Strogatz), picture taken from the book directly.

Pond with radius $r = 1$
Duck: constant speed $v_{duck} = d\theta$ along the edge of the pond
Dog: constant speed $v_{dog} = kv_{duck}$ from center of the pond ($R = 1$ at the beginning)straight towards the duck

What does the dog do in the long run when $k = 1$ and $k = \frac{1}{2}$?

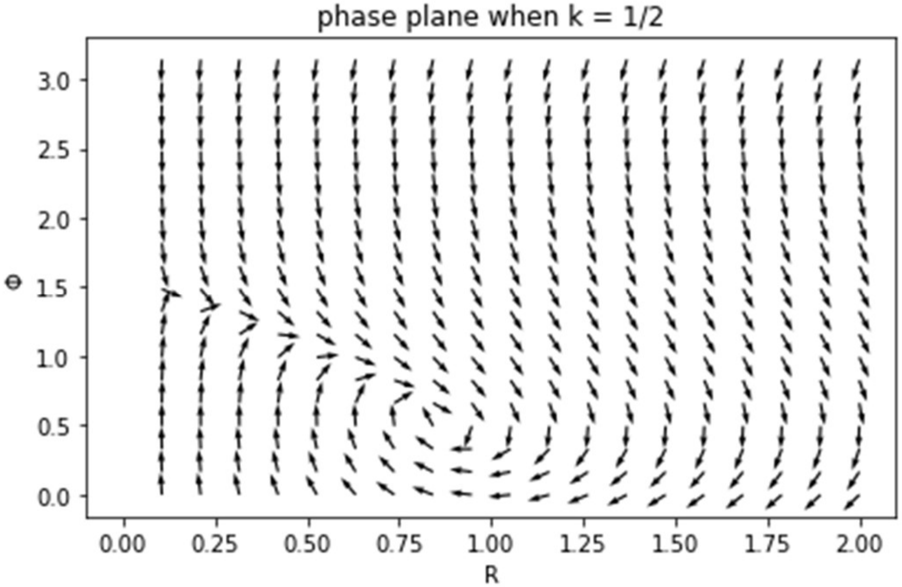


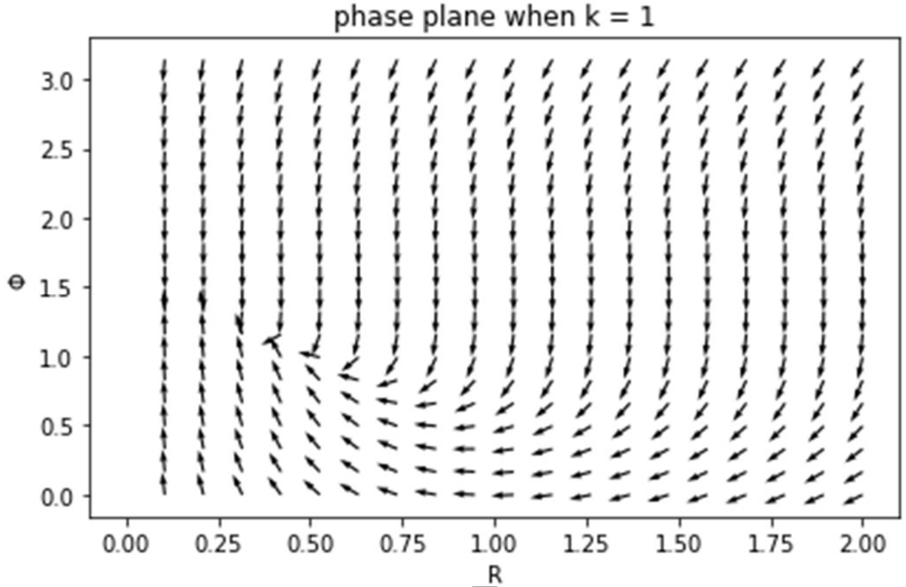
Motivation Problem: Solution



The system can be described by

$$\begin{cases} \dfrac{dR}{d\theta} = \sin\phi - k \\ \dfrac{d\phi}{d\theta} = \dfrac{\cos\phi}{R} - 1 \end{cases} \text{with } R \in [0,2],\ \phi \in \left[0, \frac{\pi}{2}\right]$$





When $k = \frac{1}{2}$, and starts from $R = 1$, attracted to $\phi = \frac{\pi}{6}$, $R = \frac{\sqrt{3}}{2}$

When $k = 1$, and starts from $R = 1$, no clear attraction, but still cannot reach $R = 0$

Motivation

The analysis of phase plane can give us some information about the long term behavior. However, it may not be able to solve the following questions easily.

- For what ks can the dog reach the duck as close as possible? (R cannot be zero in the analysis, so there is no definite number to define "closest".)
- What will happen if the speed of dog and duck are not constant?
- What will happen if the dog doesn't swim straight towards the duck?

Basic Definitions: Dynamical System

The motivation problem is a simple pursuit-evasion game, formulated as a nonlinear continuous-time autonomous dynamic system which can be described by:

$$\dot{\mathbf{x}} = f(\mathbf{x}, a),\ \mathbf{x}(0) = \mathbf{x}_0,$$

where $\mathbf{x} \in F^n$, F a field, $\mathbf{x}_0$ the initial condition and a a parameter.
It studies how trajectories and limiting behavior $\lim_{t\to\infty} \mathbf{x}(t)$ changes w.r.t. $\mathbf{x}_0$ and a.

Basic Definitions: Pursuit-Evasion Game

The Lady in the Lake problem is a pursuit-evasion game, formulated as a **two-person deterministic zero-sum differential game** with variable terminal time. The differential system is

$$\dot{\mathbf{x}} = f\big(t, \mathbf{x}, u^1(t), u^2(t)\big),\ \mathbf{x}(0) = \mathbf{x}_0.$$

This is nonautonomous. Parameters become player strategies dependent on time.

Terminal time: $T = \inf\{t \in \mathbb{R}^+, \mathbf{x}(t) \in \Lambda\}$, where Λ is the target set with boundary $\partial\Lambda$ defined by $l(t, x) = 0$

Objective function: $J(u^1, u^2) = \int_0^T g\big(t, \mathbf{x}, u^1(t), u^2(t)\big)dt + q\big(T, x(T)\big)$

The goal is to optimize the objective function.

Basic Definitions: Issacs Equation

Consider the minimax value of the objective function:

$$V(t,\mathbf{x}) = \min_{u^1} \max_{u^2} \left\{ \int_0^T g\big(t,\mathbf{x},u^1(t),u^2(t)\big)dt + q\big(T,\mathbf{x}(T)\big) \right\}$$

If V is C^1 in t and $\mathbf{x}$, then it satisfies the Issacs equation:

$$-\frac{\partial V}{\partial t} = \min_{u^1} \max_{u^2} \left\{ \frac{\partial V}{\partial \mathbf{x}} f\big(t,\mathbf{x},u^1(t),u^2(t)\big) + g\big(t,\mathbf{x},u^1(t),u^2(t)\big) \right\}$$

Let $H = \frac{\partial V}{\partial \mathbf{x}} f\big(t,\mathbf{x},u^1(t),u^2(t)\big) + g\big(t,\mathbf{x},u^1(t),u^2(t)\big)$, the following gives the solution to the costate function $p = \frac{\partial V}{\partial \mathbf{x}}$.

$$\frac{d}{dt}\left(\frac{\partial V}{\partial \mathbf{x}}\right) = -\frac{\partial H}{\partial \mathbf{x}}, \frac{\partial V}{\partial \mathbf{x}}(T) = \frac{\partial}{\partial \mathbf{x}}\Big(q\big(T,\mathbf{x}^*(T)\big)\Big) \text{(solve backwards in time)}$$

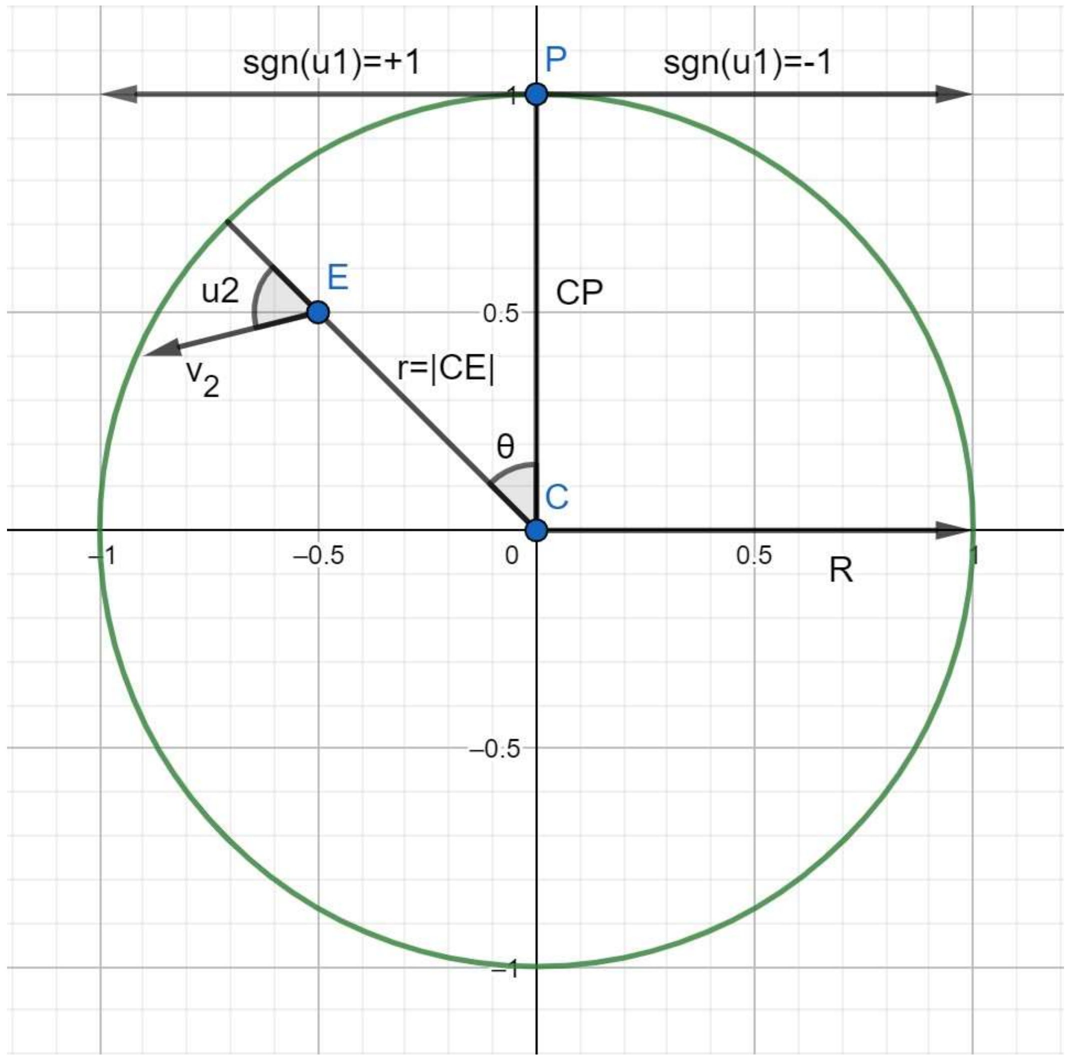
The Lady in the Lake problem

Circular pond with radius R

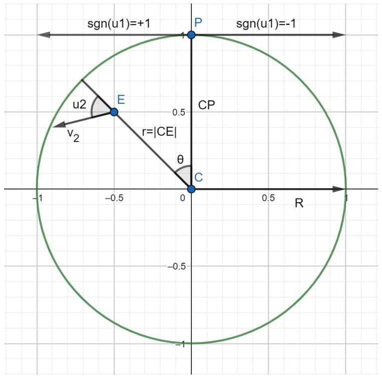
Pursuer $\mathbf{P}$ moves along the perimeter, with strategy $|u^1| \leq 1$, trying to catch the lady (evader) as she gets to the boundary.

Evader $\mathbf{E}$ starts from center C, with constant speed v_2, strategy u^2 representing direction she goes w.r.t. $\overrightarrow{CE}$, trying to reach the edge of the pond without being caught.

Let $\theta = \angle PCE$, $r = |CE|$



The Lady in the Lake problem: Analysis



The differential equations for the system is:

$$\dot{\mathbf{x}} = \begin{pmatrix} \dot{\theta} \\ \dot{r} \end{pmatrix} = \begin{cases} \dfrac{v_2 \sin u^2}{r} - \dfrac{u^1}{R} \\ v_2 \cos u^2 \end{cases}$$

Objective function $J = |\theta(T)|$ with $T = \inf\{t : r(t) = R\}$, $|\theta(t)| \le \pi$, $\forall t \in [0, T]$

We can see that $g \equiv 0$, $q\big(T, \mathbf{x}(T)\big) = |\theta(T)|$.

So the Issacs equation is:

$$0 = -\frac{\partial V}{\partial t} = \min_{u^1} \max_{u^2} \left\{ \frac{\partial V}{\partial \mathbf{x}} f\big(t, \mathbf{x}, u^1(t), u^2(t)\big) \right\} = \min_{u^1} \max_{u^2} \left\{ \frac{\partial V}{\partial \theta} \left(\frac{v_2 \sin u^2}{r} - \frac{u^1}{R} \right) + \frac{\partial V}{\partial r} \left(v_2 \cos u^2 \right) \right\}$$

It is separable in u^1 and u^2.

$$0 = \min_{u^1} \left\{ -\frac{u^1}{R} \frac{\partial V}{\partial \theta} \right\} + v_2 \max_{u^2} \left\{ \frac{1}{r} \frac{\partial V}{\partial \theta} \sin u^2 + \frac{\partial V}{\partial r} \cos u^2 \right\}$$

The Lady in the Lake problem: Analysis

Let $H = \frac{\partial V}{\partial \theta}\left(\frac{v_2 \sin u^2}{r} - \frac{u^1}{R}\right) + \frac{\partial V}{\partial r}\left(v_2 \cos u^2\right)$

The costate function is given by

$$\frac{d}{dt}\left(\frac{\partial V}{\partial \theta}\right) = -\frac{\partial H}{\partial \theta} = 0, \frac{\partial V}{\partial \theta}(T) = \frac{\partial}{\partial \theta}|\theta(T)| = \operatorname{sgn}(\theta(T))$$

$$\frac{d}{dt}\left(\frac{\partial V}{\partial r}\right) = -\frac{\partial H}{\partial r} = \frac{\partial V}{\partial \theta}\left(\frac{v_2 \sin u^2}{r^2}\right), \frac{\partial V}{\partial r}(T) = 0$$

The θ equation gives $\frac{\partial V}{\partial \theta}(t) = \operatorname{sgn}(\theta(T))$, $\forall t \in [0, T]$.

This will help us find the optimal strategies.

The Lady in the Lake problem: Analysis

$$0 = \min_{u^1}\left\{-\frac{u^1}{R}\frac{\partial V}{\partial \theta}\right\} + v_2 \max_{u^2}\left\{\frac{1}{r}\frac{\partial V}{\partial \theta}\sin u^2 + \frac{\partial V}{\partial r}\cos u^2\right\} (*)$$

$\min_{u^1}\left\{-\frac{u^1}{R}\frac{\partial V}{\partial \theta}\right\} = -\frac{1}{R}\left|\frac{\partial V}{\partial \theta}\right| = -\frac{1}{R}$ is achieved when $u^{1*} = \operatorname{sgn}\left(\frac{\partial V}{\partial \theta}\right) = \operatorname{sgn}\theta(T)$.

$\max_{u^2}\left\{\frac{1}{r}\frac{\partial V}{\partial \theta}\sin u^2 + \frac{\partial V}{\partial r}\cos u^2\right\} = \max_{u^2}\left\{\left(\frac{1}{r}\frac{\partial V}{\partial \theta}, \frac{\partial V}{\partial r}\right) \cdot \left(\sin u^2, \cos u^2\right)\right\}$ is achieved when the two vectors $\left(\frac{1}{r}\frac{\partial V}{\partial \theta}, \frac{\partial V}{\partial r}\right)$ and $\left(\sin u^2, \cos u^2\right)$ are parallel.

Let $\frac{1}{r}\frac{\partial V}{\partial \theta} = k \sin u^{2*}$, $\frac{\partial V}{\partial r} = k\cos u^2$, then $\max_{u^2}\left\{\left(\frac{1}{r}\frac{\partial V}{\partial \theta}, \frac{\partial V}{\partial r}\right) \cdot \left(\sin u^2, \cos u^2\right)\right\} = k$

Substitute u^{1*} and k back into $(*)$, $k = \frac{1}{Rv_2}$.

Then $\sin u^{2*} = \frac{Rv_2}{r}\frac{\partial V}{\partial \theta} = \frac{Rv_2}{r}\operatorname{sgn}\theta(T)$

The Lady in the Lake problem: Analysis

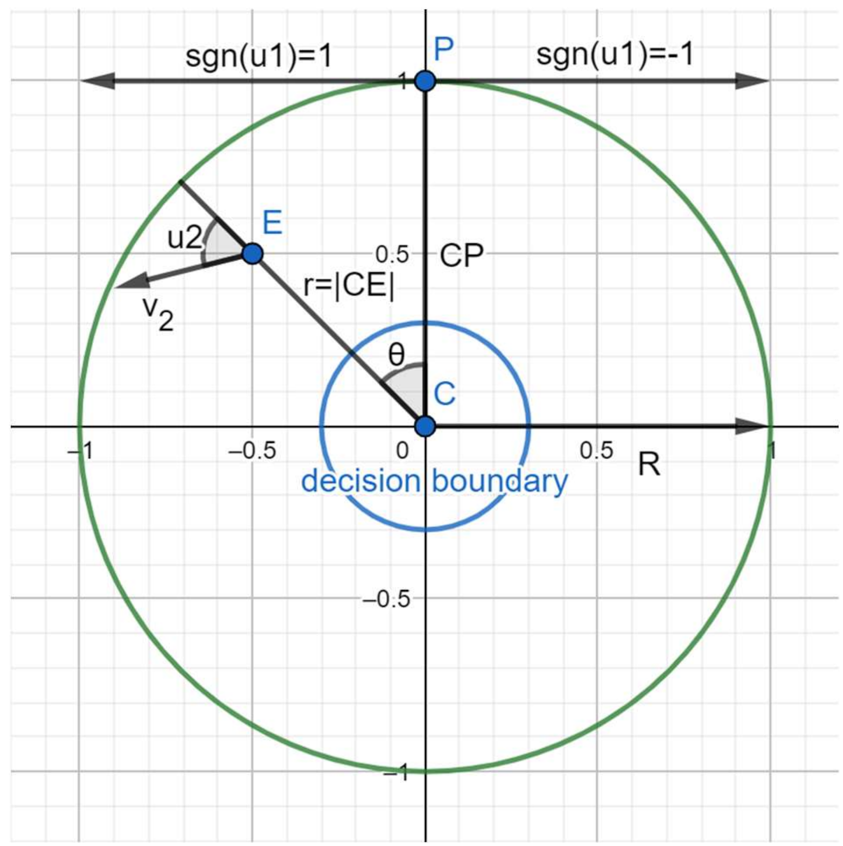
$$u^{1*} = \mathrm{sgn}\big(\theta(T)\big), \sin u^{2*} = \frac{Rv_2}{r}\,\mathrm{sgn}\,\theta(T)$$

What do these strategies mean?

$u^{1*} = \mathrm{sgn}\big(\theta(T)\big)$ means $\mathbf{P}$ should try to move in the direction to minimize θ

$\sin u^{2*} = \frac{Rv_2}{r}\,\mathrm{sgn}\,\theta(T)$ is only valid when $r \geq Rv_2$
$\mathbf{E}$ should try to maximize θ when $\mathbf{E}$ is outside of the decision boundary $\{\mathbf{x} : \|\mathbf{x}\| = Rv_2\}$.

When $r < Rv_2$, $\mathbf{E}$'s maximum angular velocity $\omega_{\mathbf{E}} = \frac{v_2}{r} \geq \frac{1}{R} \geq \omega_{\mathbf{P}}$, so $\mathbf{E}$ can always make θ as large as possible before reaching the decision boundary.



The Lady in the Lake problem: Analysis

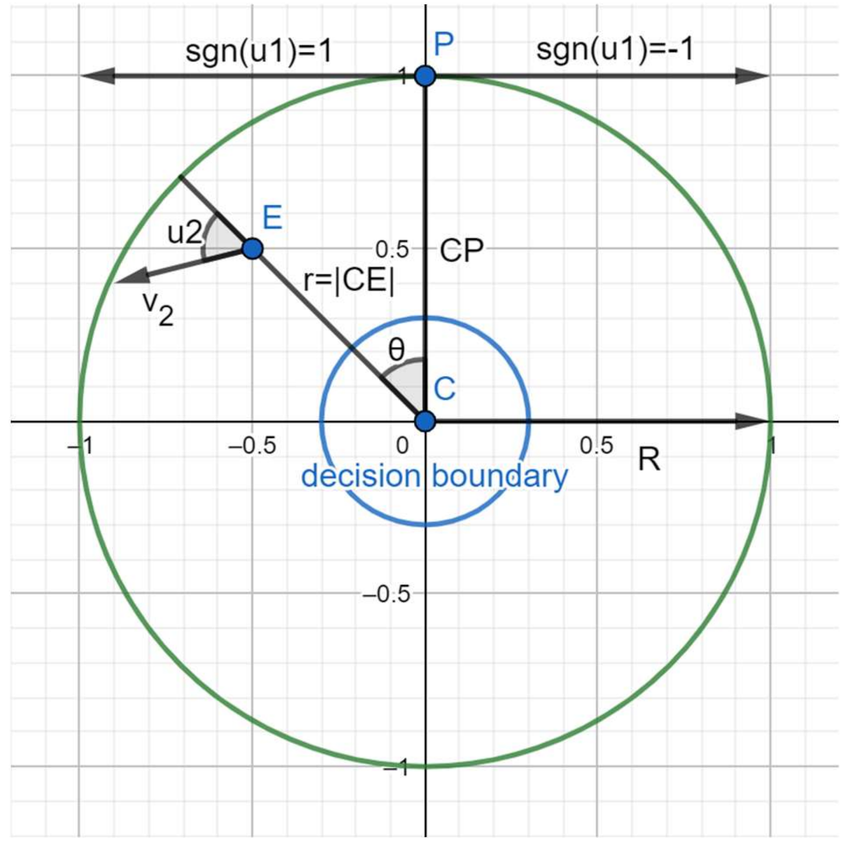
Under what conditions can $\mathbf{E}$ escape?

Ans: $\mathbf{E}$ can escape when $|\theta(T)| > 0$

T. Basar calculated $|\theta(T)| = \pi + \arccos v_2 - \frac{1}{v_2}\sqrt{1 - v_2^2}$.

$|\theta(T)| > 0$ if $v_2 > 0.21723$.

So, $\mathbf{E}$ should swim with speed $v_\mathbf{E} \geq 0.217 v_\mathbf{P}$ to escape.
And from the previous slide, $\mathbf{E}$ should try to maximize θ when reaching the decision boundary, and keep $\sin u^{2*} = \frac{R v_2}{r} \operatorname{sgn} \theta(T)$ at each r.



Inverse problem

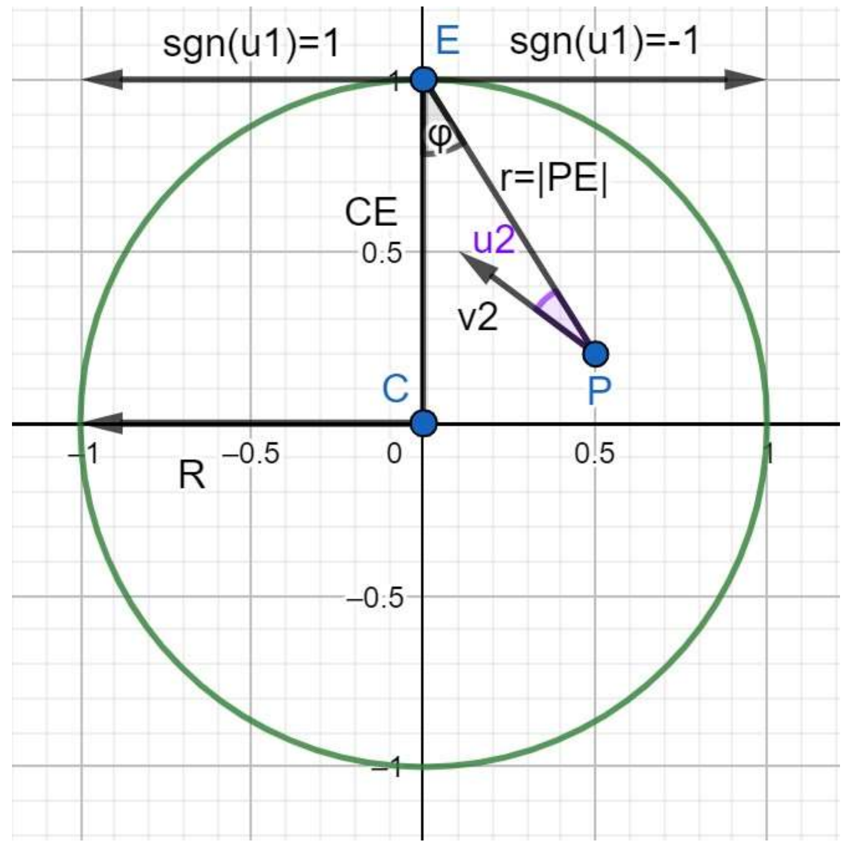
Motivation problem + strategies

Circular pond with radius R

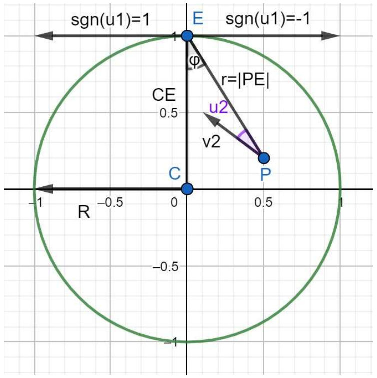
Evader $\mathbf{E}$ moves along the perimeter, with strategy $|u^1| \leq 1$ trying to survive as long as possible without being caught by $\mathbf{P}$.

Pursuer $\mathbf{P}$ starts from center C, with constant speed v_2, strategy u^2 representing direction she goes w.r.t. $\overrightarrow{PE}$, trying to catch $\mathbf{E}$ as fast as possible.

Let $\phi = \angle CEP$, $r = |PE|$



Inverse problem: Analysis



The differential equations for the system is:

$$\dot{\mathbf{x}} = \begin{pmatrix} \dot{\phi} \\ \dot{r} \end{pmatrix} = \begin{cases} \left(\dfrac{R\cos\phi}{r} - 1\right)u^1 - \dfrac{v_2 \sin u^2}{r} \\ u^1 \sin\phi - v_2 \cos u^2 \end{cases}$$

Objective function $J = r(T)$ with $T = \min\{\inf\{t : r(t) = 0\}, 15\ sec\}$

The Issacs equation is:

$$0 = \min_{u^2} \max_{u^1} \left\{ \frac{\partial V}{\partial \phi}\left(\left(\frac{R\cos\phi}{r} - 1\right)u^1 - \frac{v_2 \sin u^2}{r}\right) + \frac{\partial V}{\partial r}\left(u^1 \sin\phi - v_2 \cos u^2\right) \right\}$$

It is separable in u^1 and u^2.

$$0 = \max_{u^1}\left\{u^1\left(\frac{\partial V}{\partial \phi}\left(\frac{R\cos\phi}{r} - 1\right) + \frac{\partial V}{\partial r}\sin\phi\right)\right\} + v_2 \min_{u^2}\left\{-\frac{1}{r}\frac{\partial V}{\partial \phi}\sin u^2 - \frac{\partial V}{\partial r}\cos u^2\right\}$$

The optimal strategies are:

$u^{1*} = \operatorname{sgn}\left(\frac{\partial V}{\partial \phi}\left(\frac{R\cos\phi}{r} - 1\right) + \frac{\partial V}{\partial r}\sin\phi\right), (\sin u^{2*}, \cos u^{2*}) \parallel \left(-\frac{1}{r}\frac{\partial V}{\partial \phi}, -\frac{\partial V}{\partial r}\right)$, with reverse direction

Inverse problem: Analysis

Let $H = \frac{\partial V}{\partial \phi}\left(\left(\frac{R\cos}{r} - 1\right)u^1 - \frac{v_2 \sin u^2}{r}\right) + \frac{\partial V}{\partial r}\left(u^1 \sin\phi - v_2 \cos u^2\right)$

The costate function is given by

$$\frac{d}{dt}\left(\frac{\partial V}{\partial \phi}\right) = u^1 \frac{\partial V}{\partial \phi}\frac{R\sin\phi}{r} - u^1\frac{\partial V}{\partial r}\cos\phi\,, \frac{\partial V}{\partial \phi}(T) = \frac{\partial}{\partial \phi} r(T) = 0$$

$$\frac{d}{dt}\left(\frac{\partial V}{\partial r}\right) = \frac{\partial V}{\partial \phi}\left(\frac{R\cos\phi}{r^2}u^1 - \frac{v_2 \sin u^2}{r^2}\right), \frac{\partial V}{\partial r}(T) = 1$$

This time, it doesn't really give any helpful information for solving the system.

Inverse Problem: Analysis

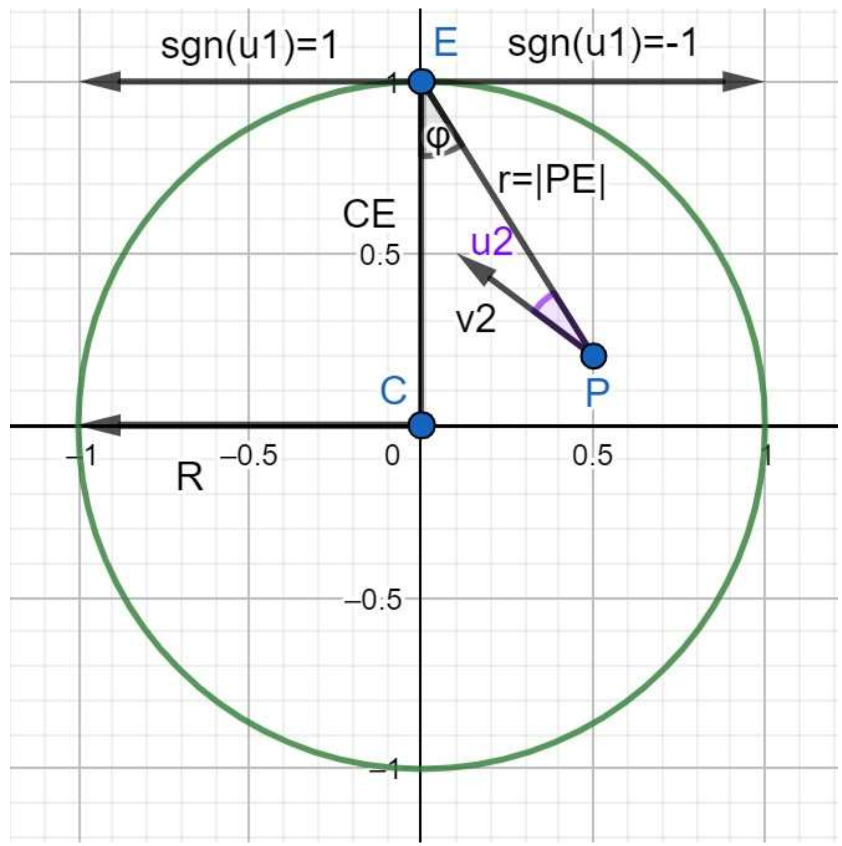
$$u^{1*} = \operatorname{sgn}\left(\frac{\partial V}{\partial \phi}\left(\frac{R\cos}{r} - 1\right) + \frac{\partial V}{\partial r}\sin\phi\right)$$

$$(\sin u^{2*}, \cos u^{2*}) \parallel \left(-\frac{1}{r}\frac{\partial V}{\partial \phi}, -\frac{\partial V}{\partial r}\right)$$

What do these strategies mean?

u^{1*} is the sign of a dot product between $\left(\frac{\partial V}{\partial \phi}, \frac{\partial V}{\partial r}\right)$ and $\mathbf{a} = \left(\left(\frac{R\cos}{r} - 1\right), \sin\phi\right)$.
$\mathbf{a}$ is the $(\dot{\phi}, \dot{r})$ for the simplified problem with speed $v_{\mathbf{P}} = 0$ along $\overrightarrow{PE}$. So $\mathbf{E}$ is trying to make ϕ larger so that $v_2 \cos u^2$ can be as small as possible and it will take longer for $\mathbf{P}$ to minimize r.

However, u^{2*} does not tell much.



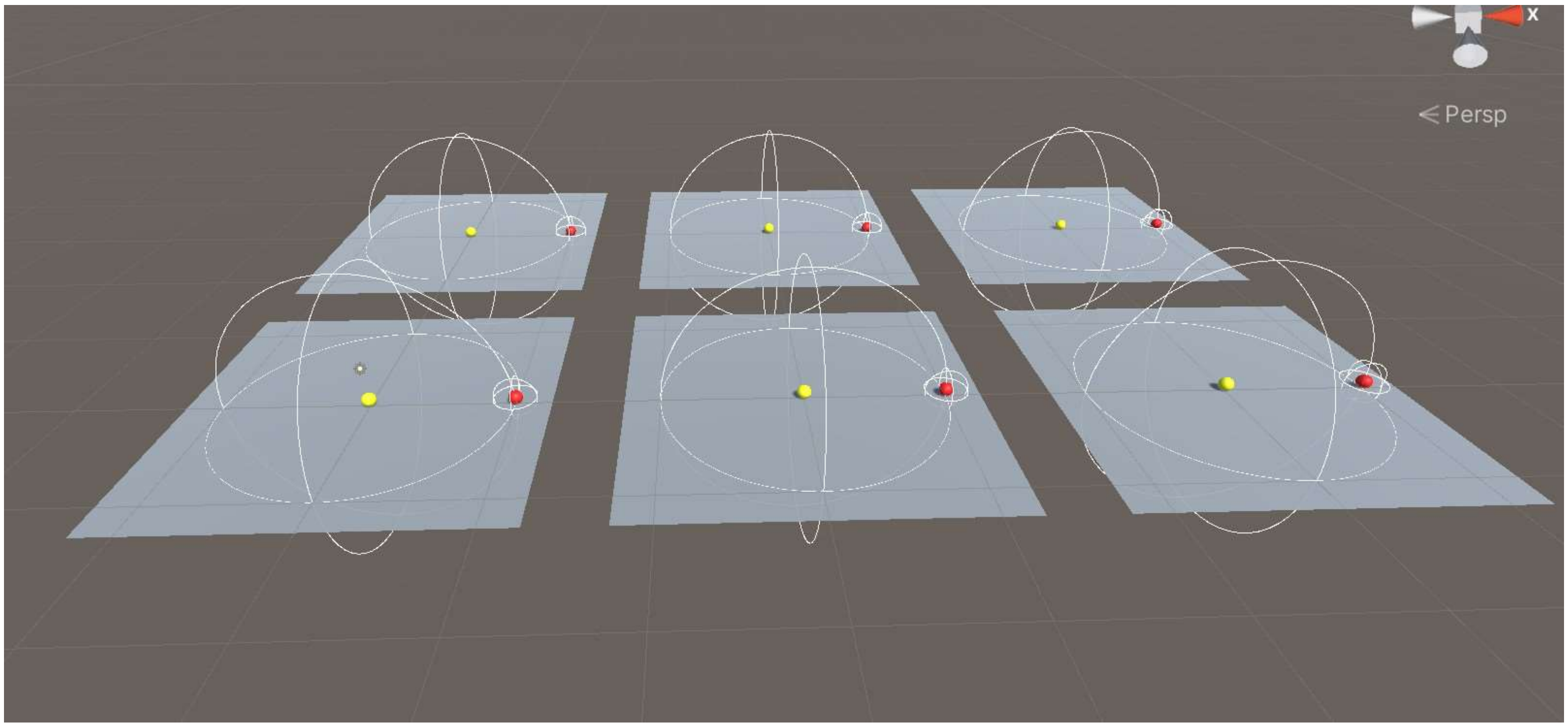
Tools for simulation

Unity with ml-agents is used to simulate the game.

- 6 training regions concurrently.
- Red sphere for $\mathbf{P}$, yellow sphere for $\mathbf{E}$.
- Large gizmos: boundary of the game $|\mathbf{x}| = R = 10$
- Small gizmos around $\mathbf{P}$: criteria for catching $||\mathbf{x_P} - \mathbf{x_E}|| \leq 1.5$ (account for non-negligible size).

Unity®

https://unity.com/



Tools for simulation

Proximal Policy Optimization is used to train the agents (players)

- Actor-critic (online policy) Reinforcement Learning: critic estimates the value function, actor updates the policy distribution.
- Better than A2C and TRPO in most continuous control environments.

Hyperparameters

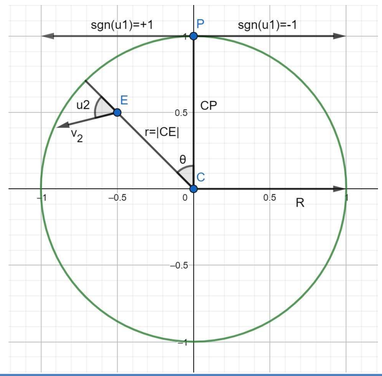
- Batch size: 512
- Buffer size: 5120
- Learning rate: 3.0×10^{-4}, decreasing linearly
- Maximum steps for each team: 10^6
- Neural network: 3-layer MLP with 512 neurons in each layer
- Discount factor: $\gamma = 0.99$

Self-play (for non-cooperative agents with inverse reward)

- Snapshots: taken every 20,000 steps
- Training team switches: every 100,000 steps
- Opponent's policy changes: every 2,000 steps, 50% chance using latest policy.

Original Problem: Setup

Let $l = \|PE\|$, $m = \frac{\overrightarrow{CE} \cdot \overrightarrow{CP}}{\|CE\|\|CP\|}$.



	Pursuer	Evader
Instantiation	Randomly on the circle	At origin
Observation	$\mathbf{s_P} = (x_\mathbf{P}, z_\mathbf{P}, x_\mathbf{E}, z_\mathbf{E}, l, m)$	$\mathbf{s_P} = (x_\mathbf{E}, z_\mathbf{E}, x_\mathbf{P}, z_\mathbf{P}, l, m)$
Action	$a_\mathbf{P} \in [-1,1]$, speed + direction	$a_\mathbf{E} \in [-1,1]$, direction
Reward	$+m$ at each step $+1000$ if successfully catches $\mathbf{E}$ -1000 if $\mathbf{E}$ escapes	$-m$ at each step $+1000$ if successfully escapes -1000 if caught

Given the Actions, position of $\mathbf{P}$ and velocity of $\mathbf{E}$:

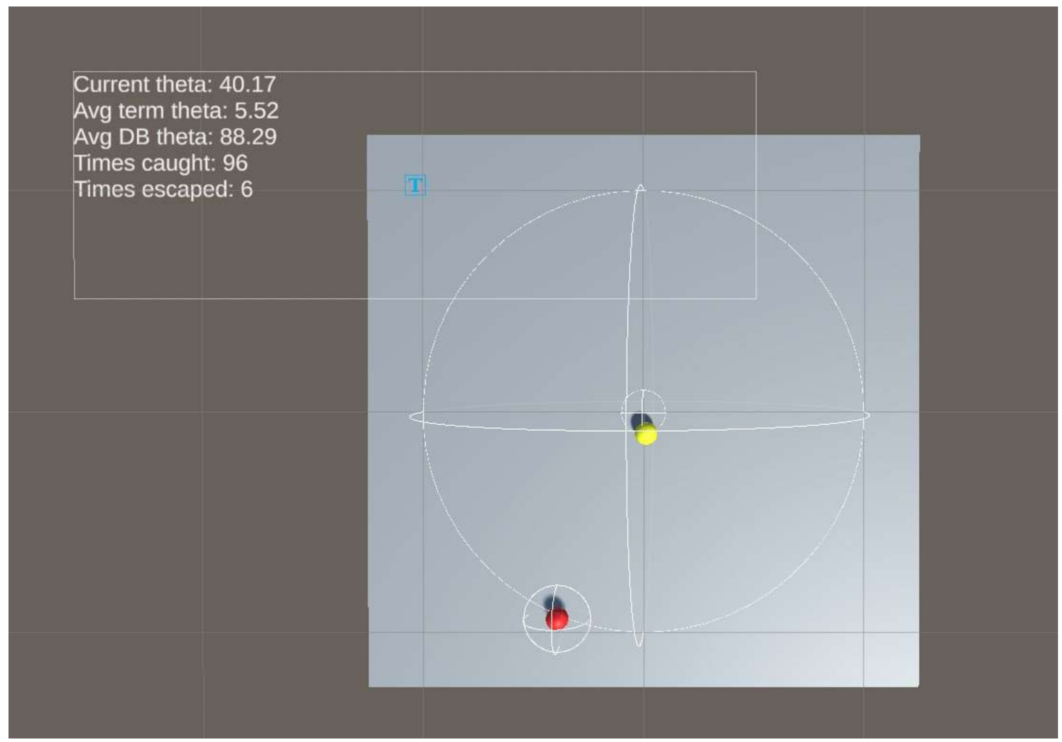
$$\mathbf{x_P} = (10\cos(\theta_\mathbf{P} + a_\mathbf{P} dt), 10\sin(\theta_\mathbf{P} + a_\mathbf{P} dt))$$
$$\mathbf{v_E} = (v_2\cos(\theta_\mathbf{E} + a_\mathbf{E}\pi), v_2\sin(\theta_\mathbf{E} + a_\mathbf{E}\pi))$$

Original Problem: Results

When $v_{\mathbf{E}} < 0.217 v_{\mathbf{P}}$:
Note that $a_{\mathbf{P}}$ is the angular velocity, the actual speed is $v_{\mathbf{P}} = a_{\mathbf{P}} R = 10 a_{\mathbf{P}}$.

So, we can take $v_{\mathbf{E}} = 1 < 0.217 \cdot 10 = 2.17$

Red is pursuer. Blue is evader.

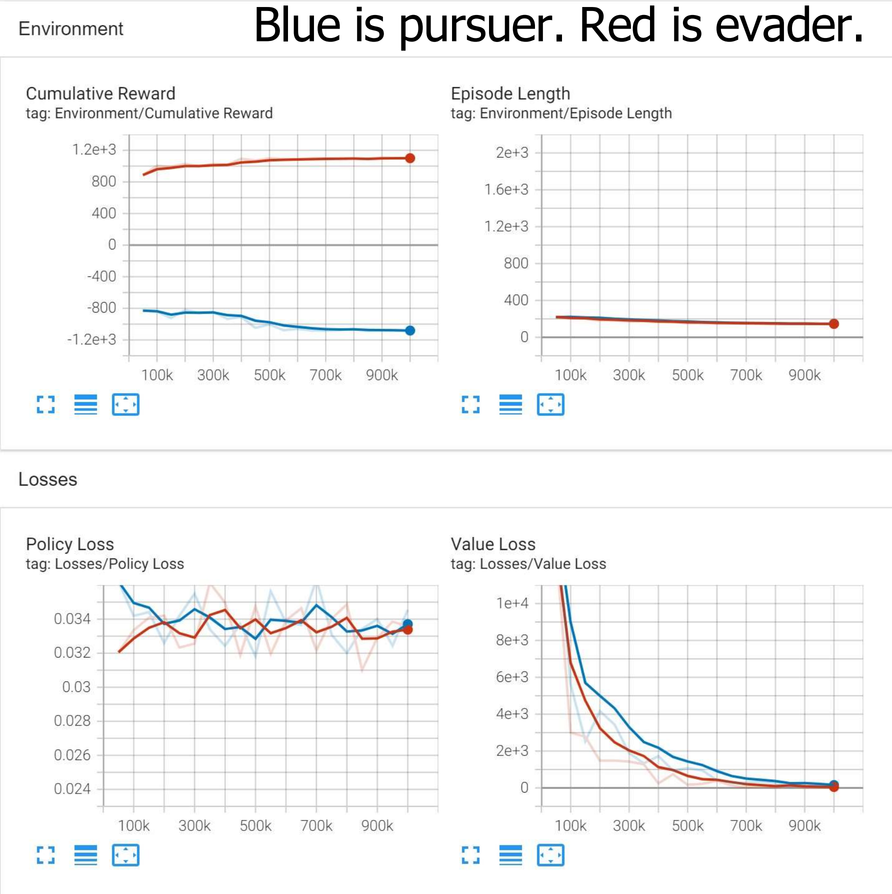


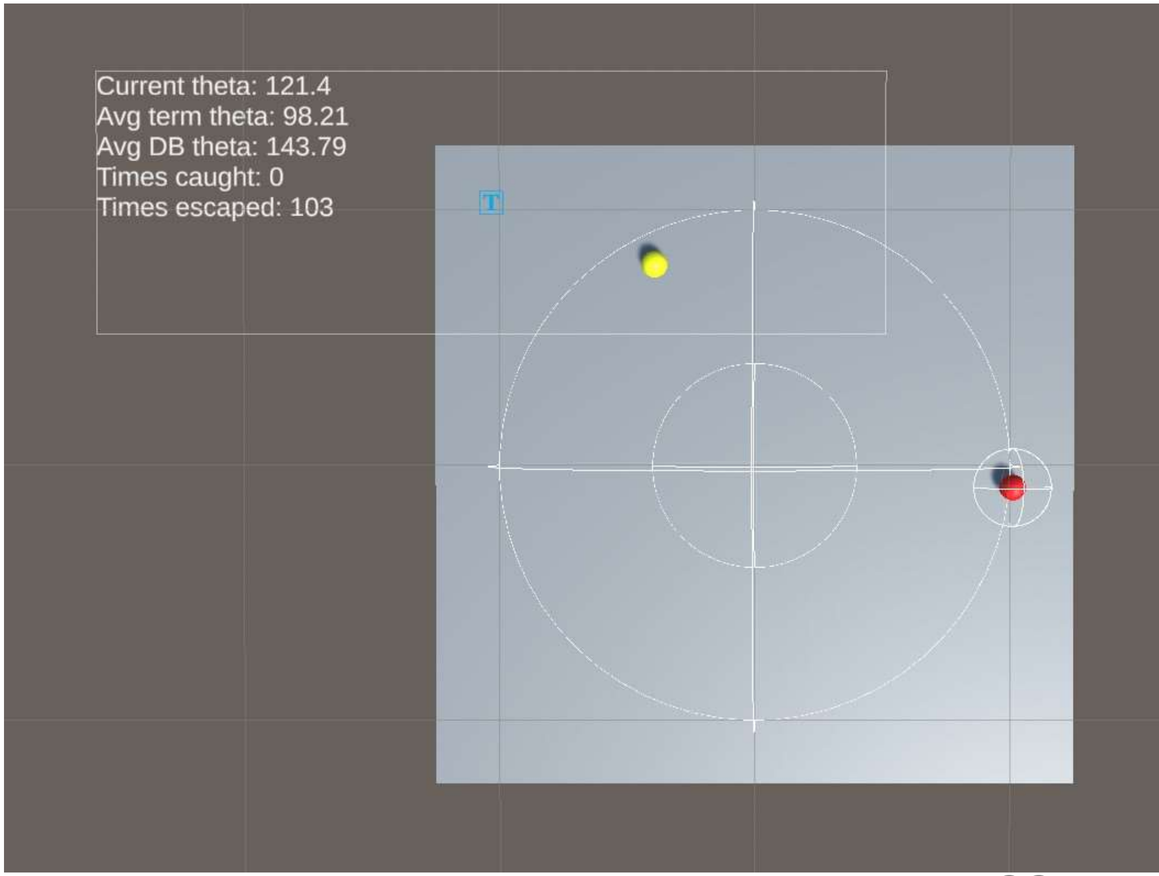
Original Problem: Results

When $v_{\mathrm{E}} > 0.217 v_{\mathrm{P}}$:

Take $v_{\mathrm{E}} = 4$

Blue is pursuer. Red is evader.

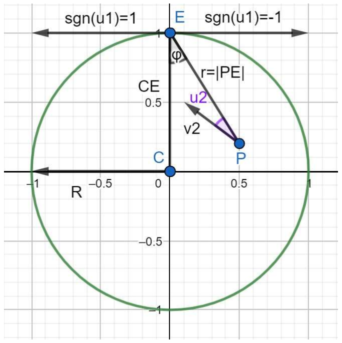




Inverse Problem: Setup

Let $r = \|PE\|$, $m = \frac{\overrightarrow{CE} \cdot \overrightarrow{CP}}{\|CE\|\|CP\|}$, $\theta = \text{atan2}\left(\overrightarrow{PE}\right)$.

$\mathbf{P}$ will be hard reset to origin if $\|\mathbf{x_P}\| > 10$.



	Pursuer	Evader
Instantiation	At origin	Randomly on the circle
Observation	$\mathbf{s_P} = (x_\mathbf{P}, z_\mathbf{P}, x_\mathbf{E}, z_\mathbf{E}, r, m)$	$\mathbf{s_P} = (x_\mathbf{E}, z_\mathbf{E}, x_\mathbf{P}, z_\mathbf{P}, r, m)$
Action	$a_\mathbf{P} \in [0.5,1.0] \times [-1,1]$, speed + direction	$a_\mathbf{E} \in [-1,1]$, speed + direction
Reward	$+0.5 - \frac{r}{20}$ at each step $+1000$ if successfully catches $\mathbf{E}$ -1000 if $\mathbf{E}$ survives for 15 sec	$+\frac{r}{20} - 0.5$ at each step $+1000$ if survives for 15 sec -1000 if caught

Given the Actions, velocity of $\mathbf{P}$ and position of $\mathbf{E}$:

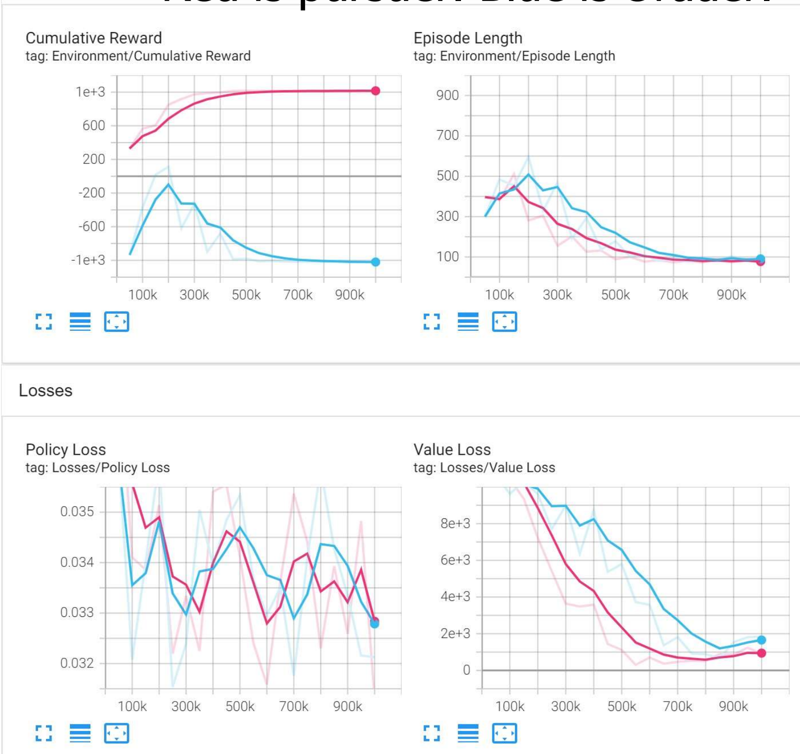
$$\mathbf{v_P} = (v_2 a_{\mathbf{P}1} \cos(\theta + a_{\mathbf{P}2}\pi), v_2 a_{\mathbf{P}1} \sin(\theta + a_{\mathbf{P}2}\pi))$$
$$\mathbf{x_E} = (10\cos(\theta_\mathbf{E} + a_\mathbf{E} dt), 10\sin(\theta_\mathbf{E} + a_\mathbf{E} dt))$$

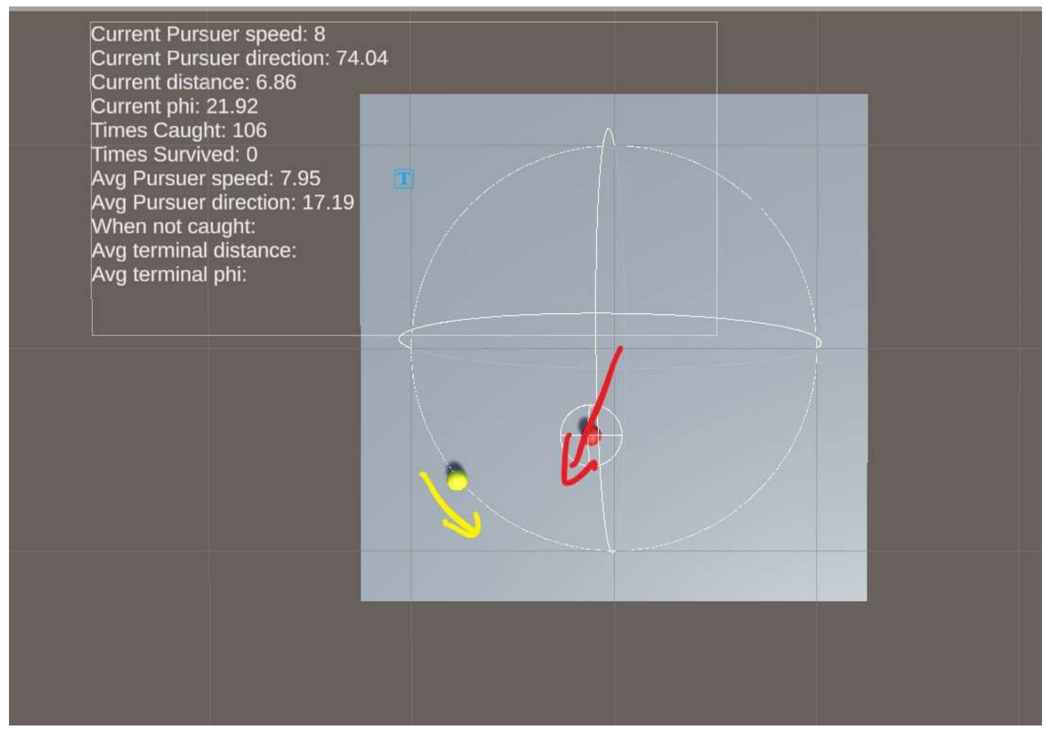
Inverse Problem: Results

Now, we don't know the optimal strategies for the agents.
So, for $v_{\mathbf{P}}$, instead of a constant speed, it is used as the maximum speed.

Take $v_{\mathbf{P}} = 8 = 0.8 v_{\mathbf{E}}$, this means that $\mathbf{P}$ can have speed $4 \sim 8$ based on my setup.

Red is pursuer. Blue is evader.

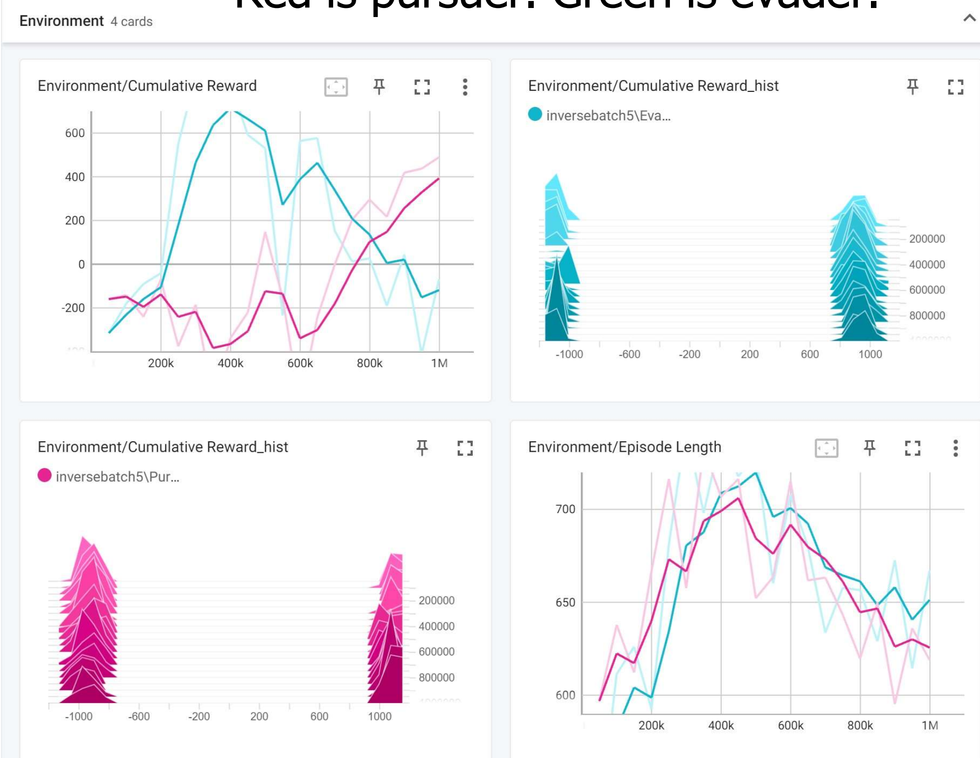


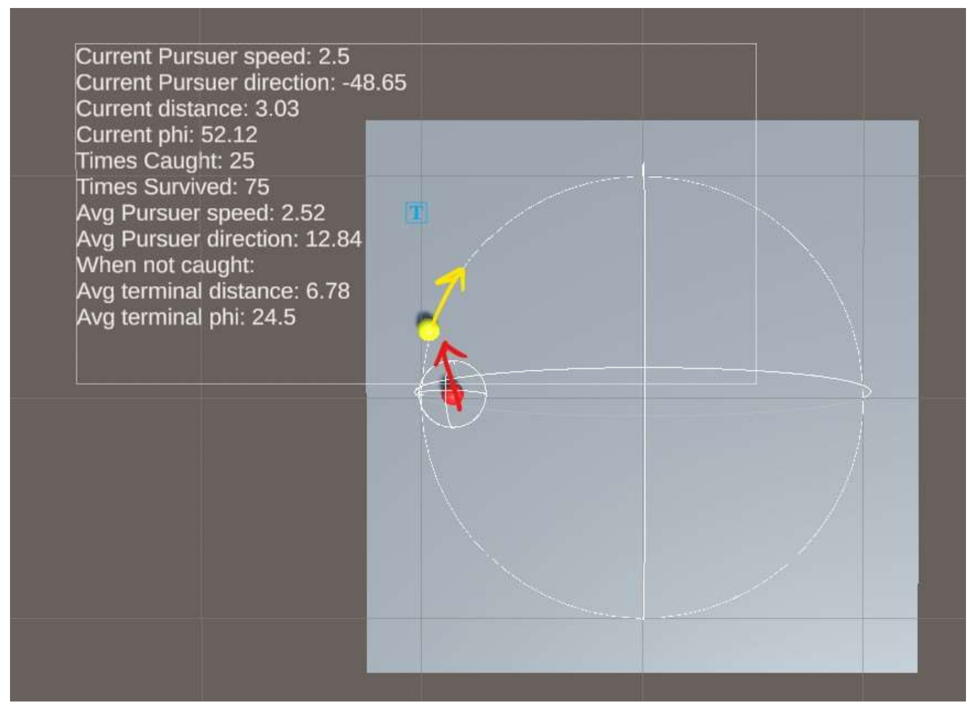


Inverse Problem: Results

Take $v_{\mathbf{P}} = 5 = 0.5v_{\mathbf{E}}$, this means that $\mathbf{P}$ can have speed $2.5{\sim}5$ based on my setup.

Red is pursuer. Green is evader.





Conclusion

What I have achieved:

- Analysis of “The Lady in the Lake” problem and its inverse problem.
- Reinforcement Learning simulation to verify the results of the original problem and analyze some behaviors of the inverse problem under different velocity conditions and different winners.

Limitations in simulation:

- Non-negligible size of the agents
- Relaxed terminal condition, can be improved by using collision detection.
- Potentially asynchronous calculation of angles, causing the game simulation to be not exactly zero-sum.
- Agents might learn better if the actions are speeds (v_x, v_z) directly instead of velocity + angle.

Limitations of the project:

- Unable to analytically solve the inverse problem. May need a different setup.
- Not tested with information advantages in either party.
- $v_{\mathbf{E}}$ can also be non-constant in the original problem for analysis.

References

- T. Basar and G. J. Olsder. Dynamic noncooperative game theory. SIAM Series Classics in Applied Mathematics, 2nd edition, 1999.
- Strogatz, Steven. Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering. CRC Press, 2018.
- Juliani, A., Berges, V., Teng, E., Cohen, A., Harper, J., Elion, C., Goy, C., Gao, Y., Henry, H., Mattar, M., Lange, D. (2020). Unity: A General Platform for Intelligent Agents. arXiv preprint arXiv:1809.02627. https://github.com/Unity-Technologies/ml-agents.
- Schulman, J., Wolski, F., Dhariwal, P., Radford, A. and Klimov, O., 2017. Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347.

Q&A

Questions?