Projet IBRID* : étude de la locomotion d'un poisson et contrôle visuel par apprentissage par renforcement

*de l'Identification Basée appRentIssage au contRôle basé moDèle.

Sardor Israilov, Li Fu, **Guillaume Allibert**, Christophe Raufaste, Médéric Argentina

Université Côte d'Azur
I3S UMR 7271, INPHYNI UMR 7010

7 mars 2022

Sommaire

Origine du projet

Appel à projet de recherche transdisciplinaire relevant des domaines de l'**EUR DS4H** et de l'**Académie 1** [1], impliquant plusieurs laboratoires **UCA**.
3 phases de sélections, avec décision finale en **juillet 2020**.
Partenaires : UMR INPHYNI et UMR I3S.

1. Réseaux, Information et Société Numérique

Origine du projet

Appel à projet de recherche transdisciplinaire relevant des domaines de l'**EUR DS4H** et de l'**Académie 1** [1], impliquant plusieurs laboratoires **UCA**.
3 phases de sélections, avec décision finale en **juillet 2020**.
Partenaires : UMR INPHYNI et UMR I3S.

Moyens accordés

- Une bourse de thèse → Sardor **ISRAILOV** (INSA Lyon)
- Un financement de 18 mois de post-doctorant → Li **FU** (LTDS, École Centrale de Lyon)
- Un environnement de 10k€

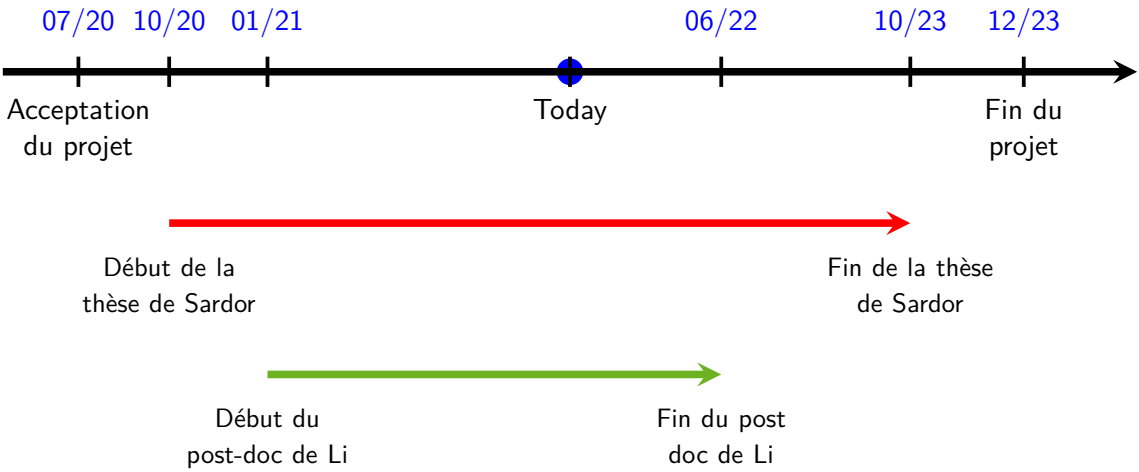
1. Réseaux, Information et Société Numérique

Objectifs du projet

Objectifs

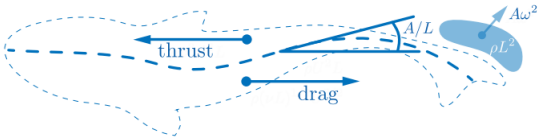
- Étude de la locomotion d'un poisson
- Contrôle par apprentissage par renforcement pour maximiser la force de poussée
- Contrôle par asservissement visuel par apprentissage

Calendrier du projet



Sommaire

Étude de la locomotion d'un poisson [Gazzola, 2014]



État de l'art (rapide !)

Force de poussée

$\rho\omega^2 A^2 L$

skin drag

$\rho U^2 \frac{1}{\sqrt{\mathrm{Re}}} L^2$

pressure drag

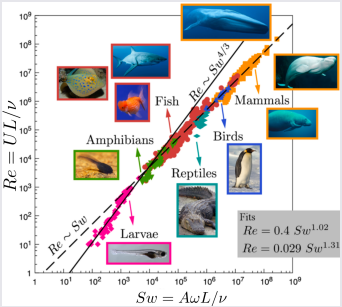
$\rho U^2 L^2$

Régime turbulent

$\left(\frac{\omega AL}{\nu}\right)^{4/3}$

Régime laminaire

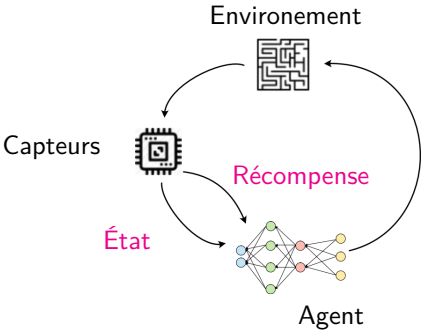
$\frac{\omega AL}{\nu}$



- ≈ 50 ans d'exp.
- ≈ 1000 données
- ≈ 8 ordres of mag.

Apprentissage par renforcement

L'Apprentissage par Renforcement [Sutton, 2018] -*Reinforcement Learning*-



https://www.youtube.com/watch?v=VMp6pq6_QjI

Apprentissage par renforcement

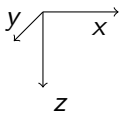
Challenges

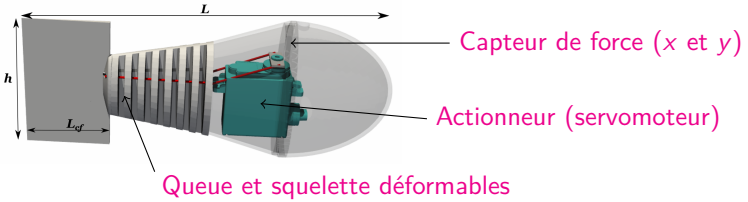
- Dilemme entre exploration (acquerir de nouvelles connaissances) et exploitation (utiliser ses connaissances)
- Choix de l'algorithme et des hyper-paramètres
- Généralisation face à des environnements différents

Domaines d'application

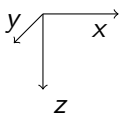
- Finance
- Santé
- Fouille de texte
- **Robotique**
- etc.

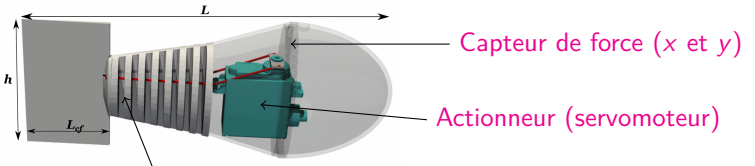
Apprentissage par renforcement : application à la nage

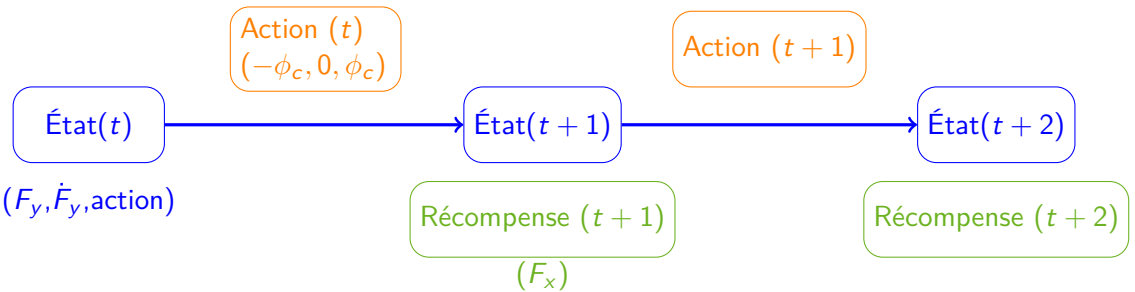




Apprentissage par renforcement : application à la nage

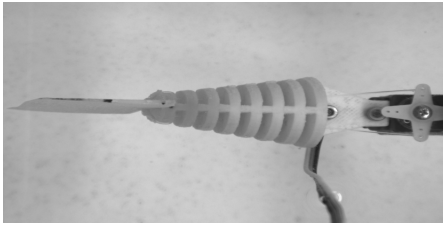


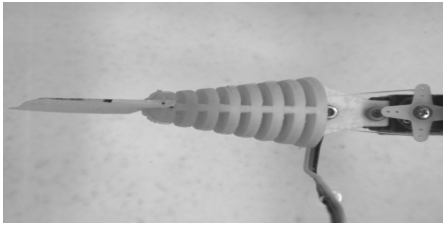




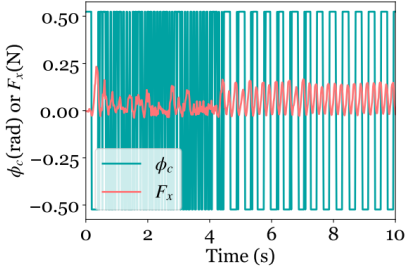
Apprentissage par renforcement : application à la nage

Quelques résultats expérimentaux

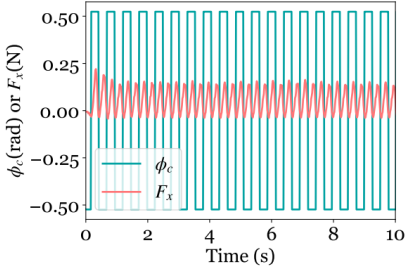




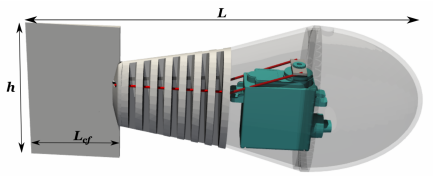
Après 5000 épisodes



Après 40000 épisodes



Apprentissage par renforcement : application à la nage



$$\dot{\phi} = \Omega \tanh\left(\frac{\phi_c - \phi}{\Delta\phi}\right)$$

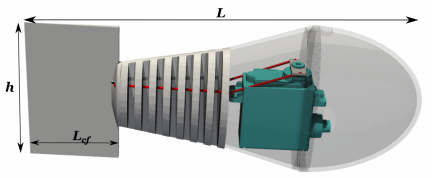
$$\ddot{\alpha} + 2\xi\omega_0\dot{\alpha} + \omega_0^2\delta_\alpha(1 - K_2\delta_\alpha^2) = 0$$

$$\delta_\alpha = \alpha - \alpha_c$$

$$\alpha_c = K_1\phi$$

$$F_y = -K_{\ddot{\alpha}}\ddot{\alpha} - K_{\dot{\alpha}}\dot{\alpha} \qquad F_x \approx F_y\alpha$$

Apprentissage par renforcement : application à la nage



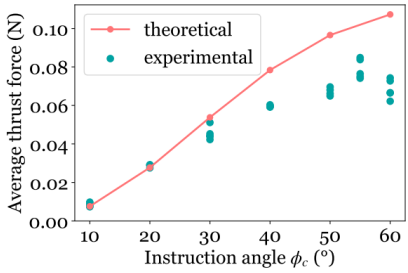
$$\dot{\phi} = \Omega \tanh\left(\frac{\phi_c - \phi}{\Delta\phi}\right)$$

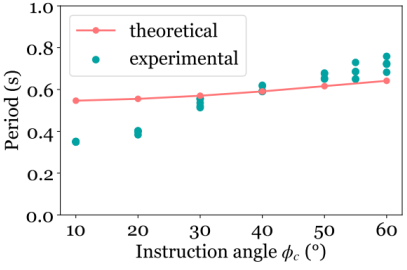
$$\ddot{\alpha} + 2\xi\omega_0\dot{\alpha} + \omega_0^2\delta_\alpha(1 - K_2\delta_\alpha^2) = 0$$

$$\delta_\alpha = \alpha - \alpha_c$$

$$\alpha_c = K_1\phi$$

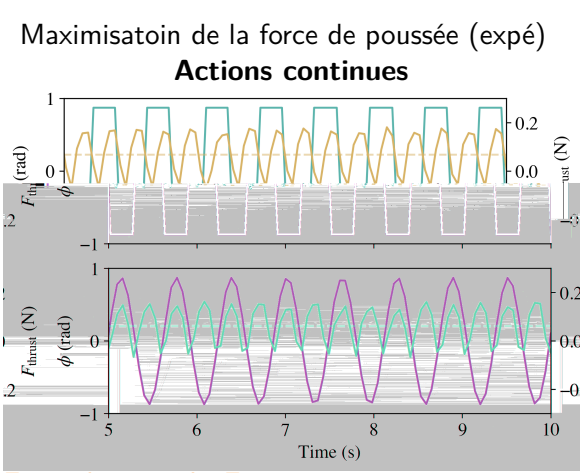
$$F_y = -K_{\ddot{\alpha}}\ddot{\alpha} - K_{\dot{\alpha}}\dot{\alpha} \qquad F_x \approx F_y\alpha$$





Apprentissage par renforcement : application à la nage

Maximisatoin de la force de poussée (expé)



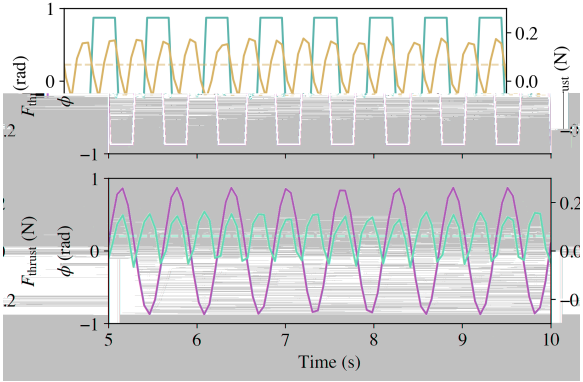
Force de poussée F_x

Angle de consigne ϕ_c

Apprentissage par renforcement : application à la nage

Maximisatoin de la force de poussée (expé)

Actions continues

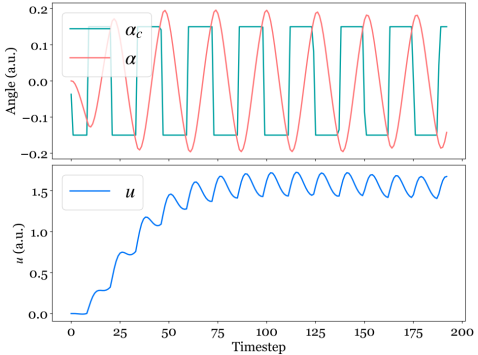


Force de poussée F_x

Angle de consigne ϕ_c

Maximisation de la vitesse (simu)

Actions continues

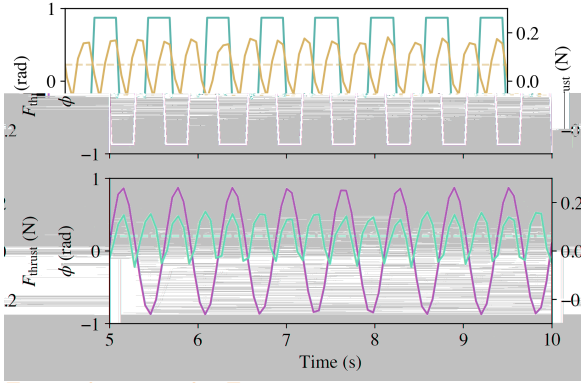


$$M\dot{u} = -|u|u - \ddot{\alpha}\alpha$$

Apprentissage par renforcement : application à la nage

Maximisatoin de la force de poussée (expé)

Actions continues

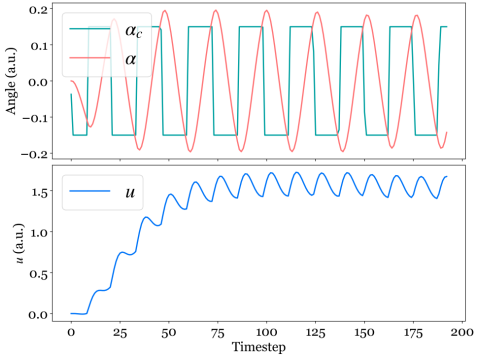


Force de poussée F_x

Angle de consigne ϕ_c

Maximisation de la vitesse (simu)

Actions continues

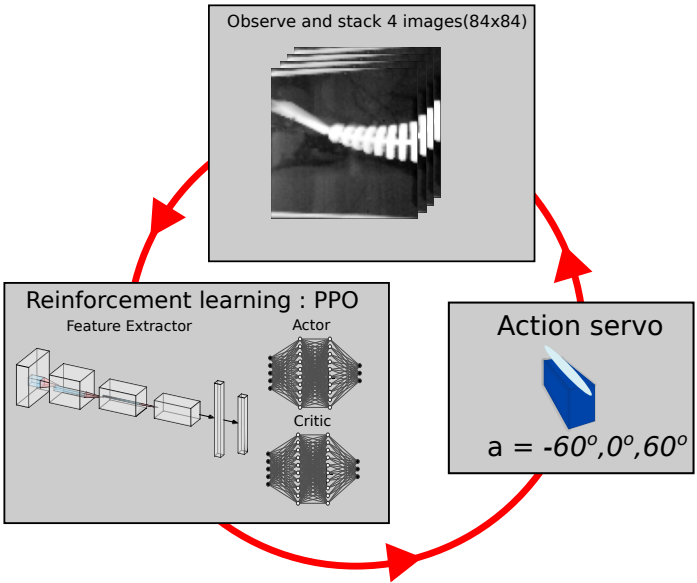


$$M\dot{u} = -|u|u - \ddot{\alpha}\alpha$$

Contrôle intermittent supérieur de 10% au contrôle harmonique

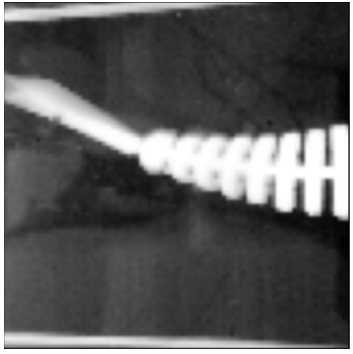
Sommaire

Vers un peu plus d'autonomie en utilisant la vision, CNN (1/2)

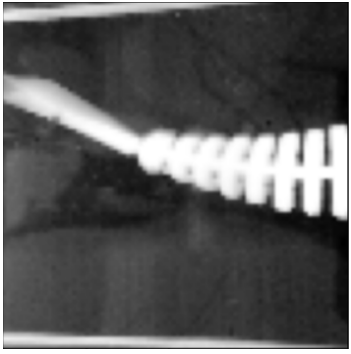


Vers un peu plus d'autonomie en utilisant la vision, CNN (2/2)

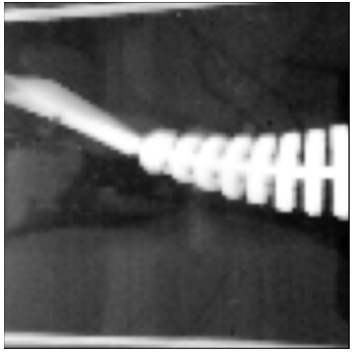
Après 3^{rd} épisodes



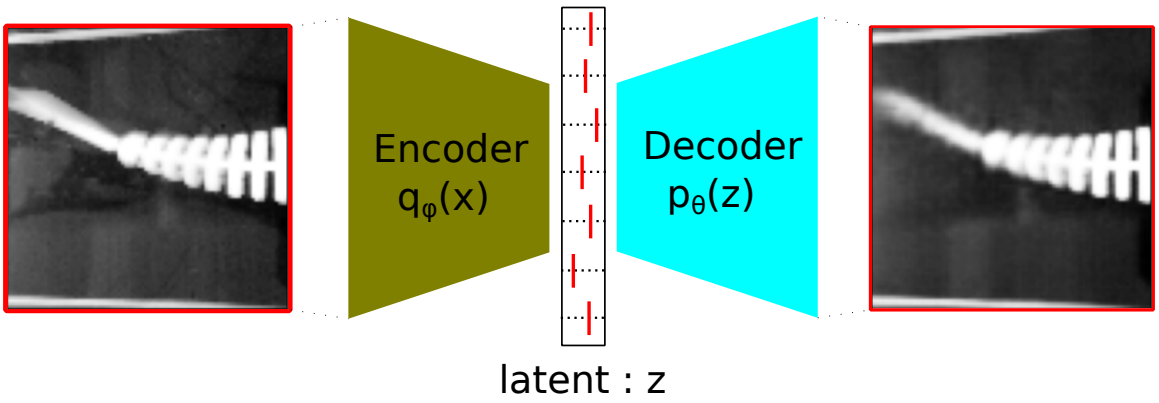
Après 21^{rd} épisodes



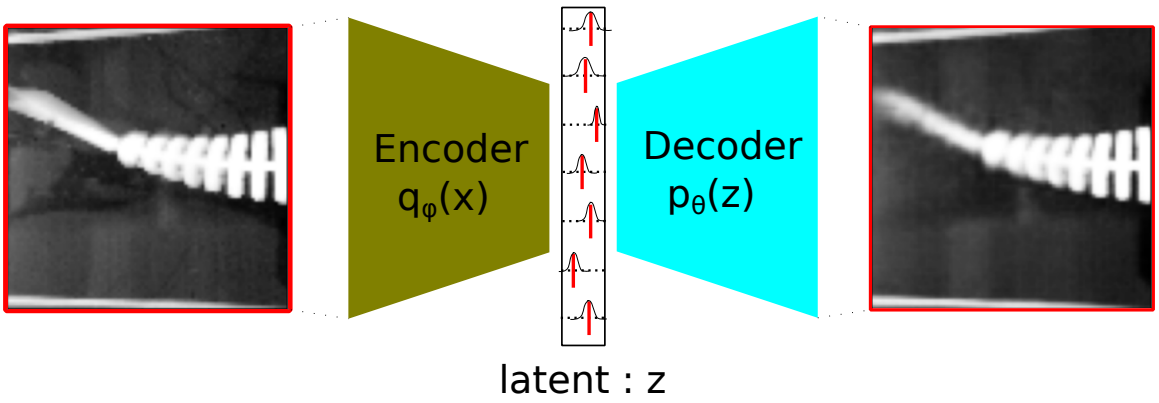
Après 156^{rd} épisodes



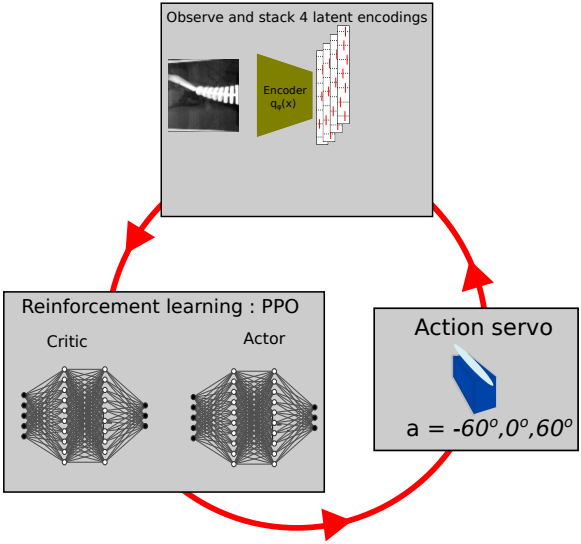
Vers un peu plus d'autonomie en utilisant la vision, VAE (1/3)



Vers un peu plus d'autonomie en utilisant la vision, VAE (2/3)



Vers un peu plus d'autonomie en utilisant la vision, VAE (3/3)



La suite...

Perspectives

- Asservissement visuel basé Apprentissage par renforcement
- Contrôle dans le long de l'axe x
- Contrôle dans le plan x, y
- Caméra embarquée

Merci pour votre attention

References

Mattia Gazzola, Médéric Argentina, Lakshminarayanan Mahadevan (2014)
Scaling macroscopic aquatic locomotion
Nature Physics 10(10) :758.

Richard Sutton, Andrew Barto (2018)
Reinforcement Learning, An introduction (Second Edition)
Bradford Books ISBN-13 978-0262039246.