

BirdBot : Prototype d'un agent réel apprenant à chanter avec un oiseau

Keywords

Recurrent Neural Network (RNN), Reservoir Computing, Sensorimotor Learning, Computational Neuroscience, Reinforcement Learning

Mots-Clés

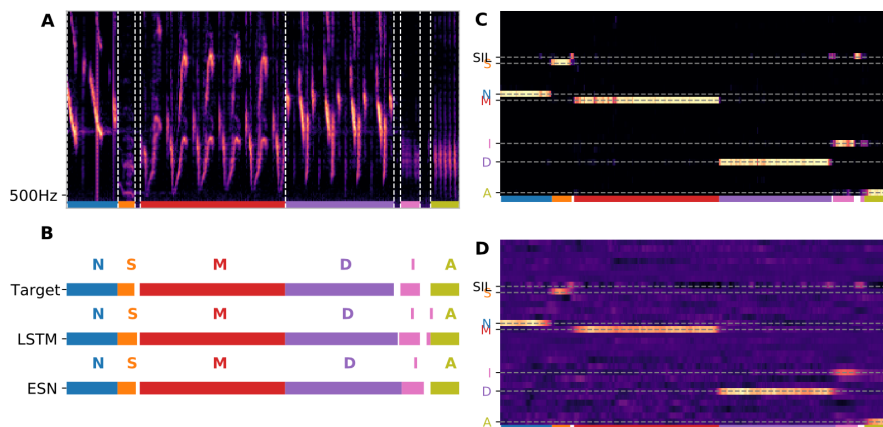
Réseau de neurones récurrent, réservoir calcul, apprentissage sensorimoteur, neurosciences computationnelles, apprentissage par renforcement

Encadrant

Xavier Hinaut (chercheur Inria)
xavier.hinaut@inria.fr

Équipe & Lieu du stage

Équipe Mnémosyne : Inria Bordeaux Sud-Ouest, LABRI & Institut des Maladies Neurodégénératives (Centre Broca Aquitaine, campus Carreire)
<https://team.inria.fr/mnemosyne>



Spectrogramme d'un chant de canari et reconnaissance par deux décodeurs (LSTM et ESN).

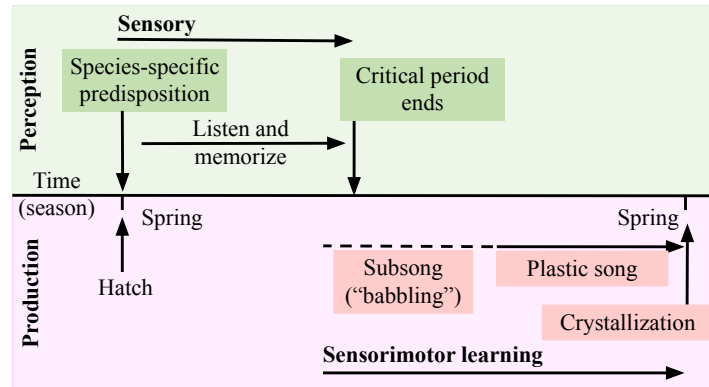
Introduction et contexte scientifique

L'étude des oiseaux chanteurs est pertinente pour comprendre comment les enfants (humains) apprennent une langue, ou plus généralement pour comprendre comment notre cerveau apprend et traite des séquences sonores (e.g. musique, ...). Les chants, en particulier chez les canaris, sont organisés sous forme de "chunks" : un chunk est donc une "méta-briques" composée de briques élémentaires. Quelques exemples pour le langage humain : une phrase est composée de plusieurs mots, un mot est composé de syllabes, et les syllabes sont elles mêmes composées de phonèmes.

Cette organisation en chunks est une caractéristique intéressante du chant des oiseaux car elle est omniprésente (et à plusieurs niveaux de hiérarchie) dans le langage humain.

Comprendre les processus cognitifs de segmentation et de traitement des chunks est une question fondamentale en neurosciences.

Comme les enfants apprenant à parler, les oiseaux apprennent à chanter en imitant leurs parents. En étudiant comment les oiseaux apprennent leur chant, nous comprendrons mieux comment les enfants apprennent le langage.

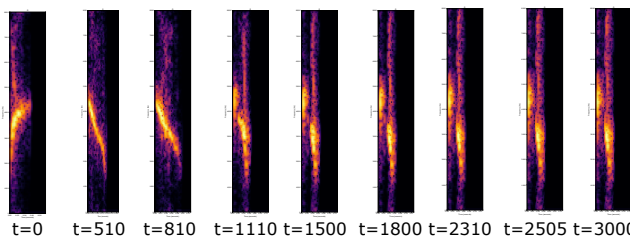
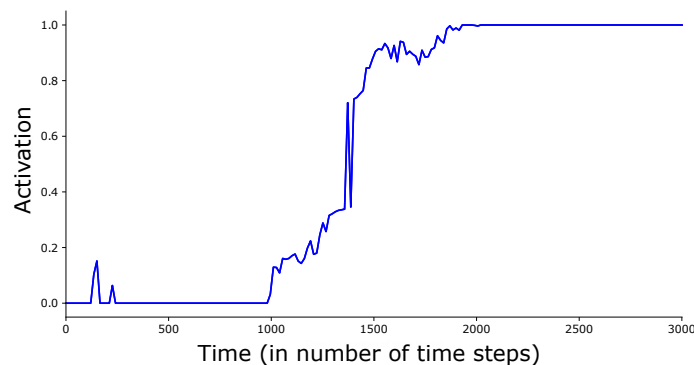


Phase de développement sur 1 an de l'apprentissage du chant chez un oiseau chanteur

Objectifs du stage

Nous avons déjà réalisé un agent virtuel qui est capable de catégoriser les différentes syllabes produites par un canari. Cet agent est ensuite capable de générer de vrais sons (à l'aide d'un GAN) afin d'essayer de reproduire ceux qu'il a entendus.

Le but du stage est de réaliser des expériences similaires mais en conditions réelles, c'est-à-dire d'émettre les sons grâce à une enceinte et d'écouter les sons produits avec un microphone. Cela sera fait dans différents environnements acoustiques, intérieurs, intérieurs bruyants, extérieurs, extérieurs avec des sons de villes, extérieurs avec d'autres chants d'oiseaux, etc. Le but ultime serait que le prototype fonctionne assez pour être utilisé pour apprendre avec d'autres oiseaux domestiques (en intérieur), et avec d'autres oiseaux (en extérieur).



Agent virtuel apprenant à imiter une syllabe en s'écoutant produire des sons

Méthodes

Nous utiliserons des réseaux de neurones récurrents (RNN) du type Reservoir Computing combinés avec d'autres méthodes d'apprentissage profond (Deep Learning) comme les GAN (Réseaux Antagonistes Génératifs), notamment les WaveGAN pour générer du son. Des méthodes d'apprentissage par renforcement pourront également être explorées et combinées.

Le projet sera réalisé en utilisant le langage Python à la vue du nombre important de bibliothèques existantes dans le domaine et pour profiter de l'expertise de l'équipe dans ce langage. Il sera utilisé la bibliothèque *ReservoirPy*¹ qui est développé dans l'équipe.

Compétences requises

- Bases solides en math-info et si possible en apprentissage artificiel ;
- Programmation en Python et bibliothèques scientifiques Numpy/Scipy ;
- Des connaissances en machine learning (dont réseaux de neurones) sont un plus, en particulier sur des techniques de Deep Learning (CNN, VAE, GAN) ;
- Intérêt pour la recherche en général et les neurosciences en particulier ;
- Capacité à travailler en équipe et à s'organiser dans un projet de recherche ;
- Bonne compréhension de l'Anglais à l'orale et à l'écrit.

Références bibliographiques

Markowitz JE, Ivie E, Kligler L, Gardner TJ (2013) Long-range Order in Canary Song. PLoS Comput Biol 9(5): e1003052.

<https://doi.org/10.1371/journal.pcbi.1003052>

Pagliarini, S., Leblois, A., & Hinaut, X. (2021, August). Canary Vocal Sensorimotor Model with RNN Decoder and Low-dimensional GAN Generator. In 2021 IEEE International Conference on Development and Learning (ICDL) (pp. 1-8). IEEE.

<https://hal.archives-ouvertes.fr/hal-03482372/>

Pagliarini, S., Trouvain, N., Leblois, A., & Hinaut, X. (2021). What does the Canary Say? Low-Dimensional GAN Applied to Birdsong.

<https://hal.inria.fr/hal-03244723/>

1 <https://github.com/reservoirpy/reservoirpy>