



Offre n°2025-08806

Doctorant F/H RÉSERVOIR DE CALCULS AVEC DES BACTÉRIES

Type de contrat : CDD

Niveau de diplôme exigé : Bac + 5 ou équivalent

Fonction : Doctorant

A propos du centre ou de la direction fonctionnelle

Le centre Inria de l'université de Bordeaux est un des neuf centres d'Inria en France et compte une vingtaine d'équipes de recherche. Le centre Inria est un acteur majeur et reconnu dans le domaine des sciences numériques. Il est au cœur d'un riche écosystème de R&D et d'innovation : PME fortement innovantes, grands groupes industriels, pôles de compétitivité, acteurs de la recherche et de l'enseignement supérieur, laboratoires d'excellence, institut de recherche technologique...

Contexte et atouts du poste

Description de la problématique de recherche

Contrairement aux approches traditionnelles « bottom-up » qui construisent des dispositifs biologiques pour le calcul au sein d'organismes vivants [1-3], ce projet de doctorat propose une stratégie novatrice « top-down ». L'objectif est d'exploiter des souches bactériennes dans un cadre de calcul par réservoir (reservoir computing, RC) pour résoudre des tâches computationnelles complexes.

Concevoir des dispositifs biologiques présente de nombreux défis : ces dispositifs imposent une charge métabolique importante aux cellules hôtes, sont difficiles à ajuster et sensibles au bruit [4, 5]. Le design de tels dispositifs s'inspire souvent des

systèmes biologiques de traitement de l'information, tels que les portes logiques, interrupteurs et perceptrons, déjà présents dans la nature [6, 7]. Cela soulève une question intrigante : au lieu de construire ces dispositifs de toutes pièces, pourrait-on exploiter directement les microorganismes naturels pour réaliser des tâches computationnelles complexes ?

En développant une approche de calcul par réservoir avec des souches bactériennes, cette recherche doctorale vise à démontrer le potentiel d'une approche « top-down » en biologie de synthèse. Elle pourrait ainsi avoir des implications pour résoudre des problèmes complexes habituellement pris en charge par des systèmes numériques. Ce projet a également pour but de poser les bases d'applications futures dans le diagnostic médical et d'explorer les synergies possibles entre le calcul par réservoir bactérien et des technologies émergentes comme le calcul neuromorphique et les matériaux vivants programmés.

Supervision

La thèse sera co-encadrée par Jean-Loup Faulon (HDR, PR) à l'INRAe et Xavier HINAUT (HDR, CR) à l'Inria.

Mission confiée

Domaine

Intelligence artificielle appliquée à des données biologiques.

Le calcul par réservoir (RC) est une branche de l'intelligence artificielle qui explore les capacités computationnelles de systèmes physiques, chimiques et biologiques [8]. Initialement développé comme une alternative aux réseaux de neurones artificiels classiques, en particulier les réseaux de neurones récurrents (RNN), le RC offre un processus d'entraînement plus efficace. Il existe deux types principaux de systèmes RC : le RC conventionnel et le RC physique.

Dans un cadre RC conventionnel typique, les données d'entrée sont fournies à un réservoir, les états des nœuds du réservoir sont enregistrés, et ces informations sont transmises à une couche de "post-lecture" - généralement un simple classificateur ou régresseur linéaire - qui interprète les états pour produire la sortie finale. Contrairement aux RNN avec poids entraînaibles, le RC conventionnel utilise un réservoir de RNN à poids fixes, ce qui le rend plus rapide à entraîner. Il est ainsi bien adapté aux applications telles que la prédiction de séries temporelles, le contrôle de systèmes dynamiques et les applications de l'Internet des objets (IoT) nécessitant un traitement rapide et efficace [9].

Dans le RC physique, le réservoir est remplacé par un objet physique. Le système fonctionne en introduisant d'abord des données d'entrée dans le réservoir physique,

qui les transforme en un état dynamique de haute dimension. Ces états transformés sont ensuite transmis à une couche post-lecture, similaire au RC conventionnel. Les applications du RC physique incluent des systèmes pionniers comme la machine à état liquide pour la reconnaissance de motifs dans l'eau [10], des développements en RC chimique pour des tâches de classification et la résolution d'équations différentielles [11], et des systèmes RC biologiques utilisant des neurones corticaux de rat cultivés sur des substrats micropatternés pour résoudre des tâches de classification [12].

Objectifs

Ce projet de doctorat vise à évaluer la faisabilité d'utiliser des souches bactériennes dans un cadre de calcul par réservoir (RC) et à explorer leur potentiel en tant que réservoirs. Comme d'autres méthodes d'apprentissage automatique, le RC repose sur des données d'entraînement comprenant des caractéristiques et des étiquettes, en cherchant à prédire les étiquettes à partir des caractéristiques. Dans une approche de réservoir bactérien, les caractéristiques du problème sont représentées par des nutriments fournis aux bactéries, et les réponses bactériennes sont mesurées par des observations phénotypiques. Ces mesures sont ensuite traitées par un régresseur ou un classificateur d'apprentissage automatique classique pour fournir des solutions aux tâches computationnelles.

Principales activités

Concrètement, le projet commencera (Année 1) en utilisant *E. coli* comme souche de test, entraînée sur des milieux enrichis en différents métabolites et en enregistrant les courbes de croissance pour diverses compositions de milieux. Les capacités computationnelles de ce réservoir d'*E. coli* ensuite comparées à des techniques classiques d'apprentissage automatique telles que la régression linéaire multiple (MLR), les machines à vecteurs de support (SVM) et le perceptron multicouche (MLP) pour diverses tâches de régression et de classification. Comme illustré dans la Figure 2 dans le document pdf et Faulon et al. [13], les résultats préliminaires obtenus sur un ensemble de taux de croissance d'*E. coli* avec différents nutriments (sucres, acides aminés, nucléotides) indiquent que cette approche peut être utilisée pour classer des motifs linéaires et non linéaires.

En guise d'application pratique, des réservoirs bactériens seront également testés (Année 1) pour la classification d'échantillons cliniques, en utilisant des souches bactériennes sauvages ou mutantes. Comme illustré dans la Figure 3 dans le document pdf, des données préliminaires suggèrent qu'un réservoir basé sur une souche mutante d'*E. coli* peut classer avec précision la sévérité d'échantillons de COVID-19, différenciant les cas modérés des cas graves selon les réponses des souches. Durant la seconde année, des réservoirs basés sur des modèles métaboliques à l'échelle du génome (GEMs) provenant de différentes espèces bactériennes ou mutantes (KO de gènes) seront explorés. Les données de croissance seront collectées expérimentalement pour les candidats les plus prometteurs. Ces

espèces et leurs données de croissance serviront ensuite à résoudre des tâches classiques de régression et de classification, avec des tests supplémentaires sur des échantillons cliniques (Année 3). Ici, au lieu d'utiliser une seule espèce, le système RC multi-espèces sera composé d'un ensemble d'espèces sauvages ou mutantes agissant comme réservoirs.

Pour les tests sur des échantillons cliniques, l'étudiant aura accès à des cohortes de cancer de la prostate et de COVID-19 fournies par les CHUs de Montpellier et Grenoble. D'autres échantillons liés à d'autres pathologies pourront être acquis durant le projet. Enfin, la possibilité de détecter des polluants environnementaux (par exemple dans l'eau) sera également étudiée dans le cadre du RC multi-espèces.

Pour les expériences simulées sur ordinateur, la bibliothèque ReservoirPy [14] - développée dans l'équipe Inria - sera favorisée :
<https://github.com/reservoirpy/reservoirpy>

Une version du sujet avec les figures est disponible à cette adresse :
<https://jfaulon.com/phd-subject-bacterial-reservoir-computer/>

Bibliography

1. Purnick, P. E. M. & Weiss, R. The second wave of synthetic biology: from modules to systems. *Nat. Rev. Mol. Cell Biol.* 10, 410–422 (2009).
2. Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403, 339–342 (2000).
3. Elowitz, M. B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* 403, 335–338 (2000).
4. Swain, P. S., Elowitz, M. B. & Siggia, E. D. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl. Acad. Sci.* 99, 12795–12800 (2002).
5. Borkowski, O., Ceroni, F., Stan, G. -B. & Ellis, T. Overloaded and stressed: whole-cell considerations for bacterial synthetic biology. *Curr. Opin. Microbiol.* 33, 123–130 (2016).
6. Hellingwerf, K. J., Postma, P. W., Tommassen, J. & Westerhoff, H. V. Signal transduction in bacteria: phospho-neural network(s) in *Escherichia coli*? *FEMS Microbiol. Rev.* 16, 309–321 (1995).
7. Scheres, B. & van Der Putten, W. H. The plant perceptron connects environment to development. *Nature* 543, 337–345 (2017).
8. Tanaka, G. et al. Recent advances in physical reservoir computing: A review. *Neural Netw.* 115, 100–123 (2019).
9. Chen, H. et al. Emerging memristors and applications in reservoir computing. *Front. Phys.* 19, 13401 (2024).
10. Fernando, C. & Sojakka, S. Pattern Recognition in a Bucket. in *Advances in Artificial Life* (eds. Banzhaf, W., Ziegler, J., Christaller, T., Dittrich, P. & Kim, J. T.) vol. 2801 588–597 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2003).

11. Baltussen, M. G., De Jong, T. J., Duez, Q., Robinson, W. E. & Huck, W. T. S. Chemical reservoir computation in a self-organizing reaction network. *Nature* 631, 549–555 (2024).
12. Sumi, T. et al. Biological neurons act as generalization filters in reservoir computing. *Proc. Natl. Acad. Sci.* 120, e2217008120 (2023).
13. Faulon, J.L. et al. Reservoir Computing with bacteria. *bioRxiv* DOI: 10.1101/2024.09.12.612674 (2024).
14. Trouvain, N., Pedrelli, L., Dinh, T. T., & Hinaut, X. (2020, September). Reservoirpy: an efficient and user-friendly library to design echo state networks. In *International Conference on Artificial Neural Networks* (pp. 494-505).

Avantages

- Restauration subventionnée
- Transports publics remboursés partiellement
- Congés: 7 semaines de congés annuels + 10 jours de RTT (base temps plein) + possibilité d'autorisations d'absence exceptionnelle (ex : enfants malades, déménagement)
- Possibilité de télétravail partiel et aménagement du temps de travail
- Équipements professionnels à disposition (visioconférence, prêts de matériels informatiques, etc.)
- Prestations sociales, culturelles et sportives (Association de gestion des œuvres sociales d'Inria)
-

Rémunération

2200€ brut mensuel

Informations générales

- **Thème/Domaine** : Neurosciences et médecine numériques
Biologie et santé, Sciences de la vie et de la terre (BAP A)
- **Ville** : Bordeaux
- **Centre Inria** : [Centre Inria de l'université de Bordeaux](#)
- **Date de prise de fonction souhaitée** : 2025-10-01
- **Durée de contrat** : 3 ans
- **Date limite pour postuler** : 2025-04-26

Contacts

- **Équipe Inria** : [MNEMOSYNE](#)

- **Directeur de thèse :**
Hinaut Xavier / xavier.hinaut@inria.fr

A propos d'Inria

Inria est l'institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 215 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3900 scientifiques pour relever les défis du numérique, souvent à l'interface d'autres disciplines. L'institut fait appel à de nombreux talents dans plus d'une quarantaine de métiers différents. 900 personnels d'appui à la recherche et à l'innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 200 start-up. L'institut s'efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

L'essentiel pour réussir

M2 recherche en biologie computationnelle, bioinformatique, mathématiques appliquées, ou autre discipline pertinente. Passion pour la programmation et l'apprentissage machine. Expérience avec Python.

Attention: Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.

Consignes pour postuler

Votre candidature doit comporter les documents suivants :

- CV
- Lettre de motivation
- Relevé de notes des années Master 1 et 2 (ou équivalent)
- Lettre(s) de recommandation si vous en avez if available

Sécurité défense :

Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

Politique de recrutement :

Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.