

MASTER INFORMATIQUE

Sujets de stages

26 novembre 2021

Liste des sujets

Équipe ADAGIo	4
Sujet ADAGIo 1 Extraction de pistes forestières de données LiDAR par apprentissage profond	5
Sujet ADAGIo 2 Approches locales sectorielles pour l'étude de surfaces digitales	7
Équipe BISCUIT	9
Sujet BISCUIT 1 Apprentissage par Renforcement sur Processeur Neuromorphique	10
Sujet BISCUIT 2 Apprentissage non-supervisé de flux optique avec des réseaux de neurones impul- sionnels	11
Sujet BISCUIT 3 Cartes auto-organisatrices et architectures reconfigurables	13
Sujet BISCUIT 4 Champs neuronaux sur circuits neuromorphiques	15
Sujet BISCUIT 5 Apprentissage par Renforcement Profond pour le Contrôle du Trafic Routier	17
Sujet BISCUIT 6 Agents collaboratifs intrinsèquement auto-organisés	18
Équipe Carbone	20
Sujet Carbone 1 Détection de motifs de virtualisation	21
Équipe COAST	22
Sujet COAST 1 Stragglers in Large-Scale Infrastructure: On the Impact of Hardware and Workloads	23
Équipe K	24
Sujet K 1 Un crawler de connaissances sur le web	25
Sujet K 2 Recommandation system for medical encoding	26
Équipe Larsen	27
Sujet Larsen 1 Quelles informations guident notre regard? Modélisation de l'attention visuelle.	28
Sujet Larsen 2 Planification d'actions avec durées incertaines	30
Sujet Larsen 3 Génération de chemin fluide pour robot aérien	32
Équipe MFX	33
Sujet MFX 1 Star-shaped distance interpolation	34
Sujet MFX 2 Propagation de son dans l'espace à l'aide de solutions fondamentales	36
Équipe Mocqua	37
Sujet Mocqua 1 A quantum stream language for polynomial time computation	38
Équipe Mosel / VeriDis	40
Sujet Mosel / VeriDis 1 Verifying cybersecurity properties using formal methods	41
Sujet Mosel / VeriDis 2 Efficient runtime verification of cyber-physical systems using formal methods	42
Sujet Mosel / VeriDis 3 Minimising the cost of diagnosing systems	43
Sujet Mosel / VeriDis 4 Distributed PlusCal	44
Sujet Mosel / VeriDis 5 Formal Models Validation – Validation of JeB	45
Sujet Mosel / VeriDis 6 Verifying the theory behind conflict-based quantifier instantiation for SMT solvers	47
Sujet Mosel / VeriDis 7 Evaluating SMT Encodings of Proof Obligations	48
Sujet Mosel / VeriDis 8 Efficient Proof Strategies for Cyclic Induction Reasoning	49
Équipe MULTISPEECH	51
Sujet MULTISPEECH 1 Semantic information in a speech recognition system : does the past help the present?	52
Sujet MULTISPEECH 2 Mitigating bias in texts using word removal and explanations	54
Sujet MULTISPEECH 3 Deep-learning-based voiced/unvoiced decision for pitch detection	56

Sujet MULTISPEECH 4	Target aware Hateful Language Detection	58
Sujet MULTISPEECH 5	Deep latent feature factorization for unsupervised speech separation	60
Sujet MULTISPEECH 6	Réparation de signaux audio par des modèles compacts prenant en compte la phase	62
Sujet MULTISPEECH 7	Evaluating the impact of speaker anonymization on emotion recognition . .	64
Sujet MULTISPEECH 8	Detection of audio segments	66
Sujet MULTISPEECH 9	Détermination tridimensionnelle de la forme du conduit vocal	67
Sujet MULTISPEECH 10	Inversion acoustique articulatoire de la parole à l'aide d'images IRM dy- namiques	69
Sujet MULTISPEECH 11	Investigating the speaker information carried on by the fundamental frequency	71
Sujet MULTISPEECH 12	Exploiting time-frequency correlations for variational autoencoder-based speech enhancement	72
Sujet MULTISPEECH 13	Dynamical variational autoencoders for unsupervised audiovisual speech enhancement	74
Sujet MULTISPEECH 14	Modélisation des accents régionaux pour une reconnaissance automatique de la parole inclusive	76
Sujet MULTISPEECH 15	Continuous 3D Sign language Generation	78
Équipe Neurorhythms		80
Sujet Neurorhythms 1	Détection adaptative de nouveautés : application à la caractérisation de l'activité cérébrale	81
Équipe Orpailleur		82
Sujet Orpailleur 1	A First Study in Causal Formal Concept Analysis	83
Sujet Orpailleur 2	Mining Molecular Structures for Discovering Antibiotics and Extensions	85
Sujet Orpailleur 3	New Research Directions in Numerical Pattern Mining	87
Sujet Orpailleur 4	Predicting biomolecule properties using machine learning approaches	89
Équipe SIMBIOT		91
Sujet SIMBIOT 1	Étude de l'influence du PRNG utilisé dans un processus d'apprentissage	92
Sujet SIMBIOT 2	Simulateur de conduite interactif pour la conception d'un véhicule électrique . . .	94
Sujet SIMBIOT 3	Simulation and implementation of trust management (TM) in VANET	96
Équipe Synalp		98
Sujet Synalp 1	Federated Learning pour l'Industrie 4.0	99
Sujet Synalp 2	Expérimentation des méthodes d'apprentissage auto-supervisé sur les modèles neu- ronaux multi-modaux : Application aux documents légaux.	100
Sujet Synalp 3	Biomedical Named Entity Recognition and Relation Extraction : implementing (then improving) a state-of-the-art architecture for the population of a knowledge graph	102
Équipe Tangram		104
Sujet Tangram 1	Tracking, validation et intégration temporelle de primitives de haut niveau issues de réseaux convolutionnels pour la localisation par l'image	105
Sujet Tangram 2	Gestion de contact d'un modèle 1D dans un réseau tubulaire	107
Équipe TYPES		109
Sujet TYPES 1	Bi-intuitionistic Bunched Implication Logic	110
Sujet TYPES 2	Formalizing the Undecidability of Relevance Logic and Non-Commutative Linear Logic	112

Équipe ADAGIo

<http://adagio.loria.fr/>

Sujet 1

Extraction de pistes forestières de données LiDAR par apprentissage profond

Proposé par : Philippe EVEN et Phuc NGO

Equipe : ADAGIo

Informations générales

Encadrants	Philippe EVEN	Phuc NGO
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy	
Téléphone	03 54 95 84 08	
Email	philippe.even@loria.fr	phuc.ngo@loria.fr
Bureau	A222	A221

Motivations

Le LiDAR (*Light Detection And Ranging*) est un outil de mesure 3D basé sur le balayage d'un signal laser sur une scène cible, et la réception des échos renvoyés par les obstacles touchés. En contexte forestier, le signal émis depuis un avion retourne plusieurs échos correspondant aux différents obstacles rencontrés avant de toucher le sol, depuis la canopée, jusqu'aux niveaux de végétation basse. Un pré-traitement de classification des points 3D déduits de chaque écho est effectué pour discriminer ceux correspondant au sol de ceux correspondant aux couches de végétation. Un modèle numérique de terrain (MNT), construit par interpolation des points de sol, fournit une visualisation du relief. Mais en contexte forestier, peu propices à la pénétration du laser, la forte disparité de densité des points de sol produit des approximations importantes dans ce MNT (voir Figure 1).

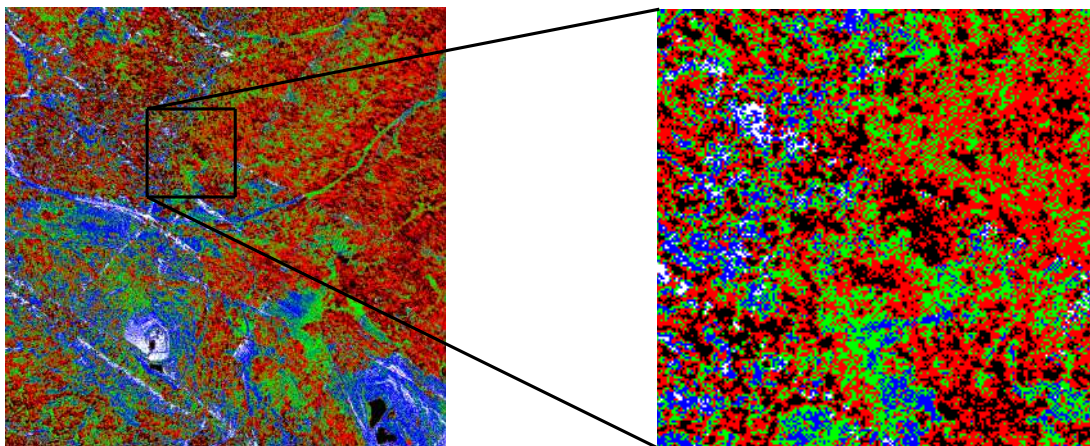


FIGURE 1 – Carte de densité dans un secteur de 1 km² aux abords du Saint-Mont, Vosges. Résolution : 1 m², noir : pas de point, rouge : de 1 à 7 points, vert : de 8 à 15 points, bleu : de 16 à 31 points, blanc : \geq 32 points.

Nous nous intéressons à la détection des structures linéaires, en particulier les pistes forestières, dans ces données LiDAR. Nous avons proposé récemment une nouvelle approche [1, 2] basée sur l'analyse de profils altimétriques à partir d'outils de géométrie discrète. Contrairement aux approches classiques axées sur l'analyse du MNT, elle exploite directement les points 3D classés "sol" pour s'affranchir des problèmes d'interpolation. Dans le cas des pistes forestières, nous avons proposé un modèle adapté au relief de montagne : un replat bordé de zones pentues (voir Figure 2).

Des tests menés à grande échelle sur l'ensemble du massif du Fossard (78 km²), puis sur des secteurs plus restreint pour lesquels des vérités de terrain sont disponibles, ont montré des performances de détection des pistes de l'ordre de 70% en couverture et en précision, pour un temps de traitement d'environ 30 minutes sur un ordinateur portable standard.

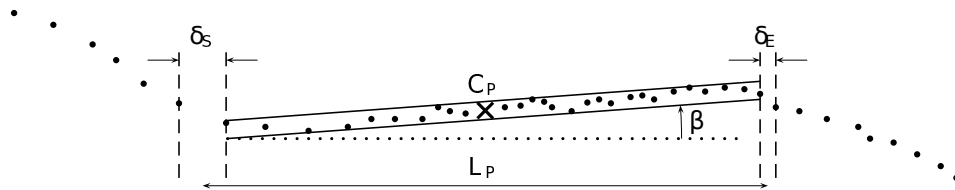


FIGURE 2 – Modèle de coupe d’une piste forestière calé sur un profil de points de sol; L_P : largeur de la piste, C_P : position du centre, β : inclinaison du plateau, δ_S et δ_E : précision de la détermination des bords.

Sujet

L’objectif de ce stage est d’étudier d’autres méthodes basées sur des réseaux de neurones profonds pour l’extraction des pistes forestières des données LiDAR.

Le stage commencera par une prise en main des outils de géométrie discrète et des algorithmes d’analyse de profils utilisés pour le développement de l’approche existante. En parallèle, différents modèles d’apprentissage profond seront étudiés à partir de la littérature, pour évaluer leur possible exploitation dans ce cadre, en comparaison d’approches existantes basées sur des réseaux de convolutions pour la localisation de routes [3, 4]. Suite à ce travail, une approche sera développée. Par défaut, l’exploitation d’un modèle de réseaux de neurones convolutifs de type Unet [5] pour segmenter les images du MNT (voir Figure 3) sera mise en œuvre. Les résultats obtenus seront alors comparés avec ceux de la méthode développée dans [2] pour évaluer les performances atteintes en temps d’exécution et en qualité de la détection.

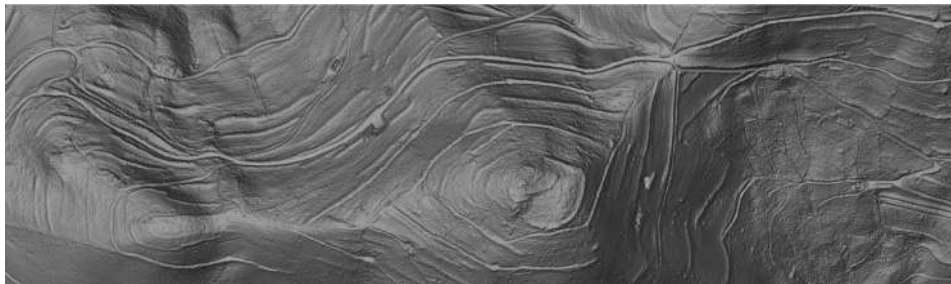


FIGURE 3 – Quelques pistes forestières visibles dans le MNT du Fossard, Vosges.

A partir de l’expérience acquise, des pistes pourront être proposées pour l’utilisation de réseau de neurones profonds afin d’optimiser le paramétrage des algorithmes de l’approche existante pour l’extraction d’autres types de structures linéaires des données LiDAR [6].

Cadre du travail

A la croisée des domaines de la géométrie discrète et de l’analyse d’images, ce stage comporte une phase de développement s’appuyant sur les langages Python et C++ avec une interface utilisateur basée sur Qt.

Références

- [1] Even, P., Ngo, P. Live extraction of curvilinear structures from LiDAR raw data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences (XXIV ISPRS Congress)*, V-2, 211–219. doi.org/10.5194/isprs-annals-V-2-2020-211-2020.(2020)
- [2] Even, P., Ngo, P. Automatic forest road extraction from LiDAR data of mountainous areas. J. Lindblad, F. Malmberg, N. Sladoje (eds), *Proc. of the First International Joint Conference on Discrete Geometry and Mathematical Morphology*, LNCS, 12708, Springer, Uppsala, Sweden, 93–106.(2021)
- [3] Liu, Q. and Kampffmeyer, M. and Jenssen, R. and Salberg, A.-B. Road Mapping In LiDAR Images Using A Joint-Task Dense Dilated Convolutions Merging Network. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. 10.1109/IGARSS.2019.8900082. (2019)
- [4] Salberg, A.-B. and Trier, Ø. D. and Kampffmeyer, M.. Large-scale mapping of small roads in LiDAR images using deep convolutional neural networks. *20th Scandinavian Conference on Image Analysis (SCIA)*. 10.1007/978-3-319-59129-2_17. (2017)
- [5] Ronneberger O., Fischer P., Brox T. U-Net : Convolutional Networks for Biomedical Image Segmentation. *LNCS. 9351*. 234-241. 10.1007/978-3-319-24574-4_28. (2018)
- [6] Even, P., Grzesznik, A., Gebhardt, A., Chenal, T., Even, P., Ngo, P. Interactive extraction of linear structures from LiDAR raw data for archaeomorphological structure prospection. *IAPRS, XLIII-B2-2021*, 153–161. doi.org/10.5194/isprs-archives-XLIII-B2-2021-153-2021. (2021)

Sujet 2

Approches locales sectorielles pour l'étude de surfaces digitales

Proposé par : Isabelle Debled-Rennesson, Fabien Feschet et Phuc NGO

Equipe : ADAGIo

Informations générales

Encadrants	Isabelle Debled-Rennesson ¹ , Fabien Feschet ² , Phuc Ngo ¹
Adresse ¹	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Adresse ²	LIMOS - UMR 6158 - Université de Clermont Auvergne
Téléphone (LORIA)	03 83 59 30 37 (I. Debled-Rennesson)
Téléphone (LORIA)	03 54 95 84 30 (P. Ngo)
Email ¹	isabelle.debled-rennesson@loria.fr, hoai-diem-phuc.ngo@loria.fr
Email ²	fabien.feschet@uca.fr
Bureau	A222, A221 (LORIA)

Motivations

Les droites digitales et les plans digitaux sont étudiés depuis plus de trente ans soit sous l'angle de la géométrie, de l'arithmétique ou de la théorie des mots. Si les définitions sont actuellement bien établies, la reconnaissance de morceaux de primitives n'est pas encore parfaitement comprise dès la dimension 3. On sait en dimension 2, grâce à la notion de couverture tangentielle [1], définir complètement le graphe de tous les segments digitaux d'une courbe digitale et obtenir grâce à elle des estimateurs de longueur, de normale ou de courbure qui sont convergents. La convergence permet d'assurer que les quantités calculées sur la discrétisation d'une forme continue convergent vers la quantité calculées sur la courbe réelle. De même, la longueur minimale en nombre de segments digitaux d'un cycle d'une courbe digitale peut être résolu via la couverture tangentielle en temps linéaire [2].

En revanche, à partir de la dimension 3, la situation devient plus délicate. On sait que le problème du recouvrement par un nombre minimum de plan digitaux d'une surface est NP-difficile [3] et on ne connaît pas de bonne définition d'une extension de la couverture tangentielle à la dimension 3, malgré quelques tentatives [4]. De même, on ne sait pas encore facilement comprendre la géométrie et l'arithmétique des morceaux de plans non rectangulaires. Dès lors, ce sont des méthodes de recherche locales, comme le plane probing [5], qui fournissent des constructions locales de plan (cf. figure 4). En revanche, ces constructions sont locales et hormis pour les polytopes digitaux, la répartition des morceaux de plans reste délicate sous l'angle de la combinatoire engendrée.

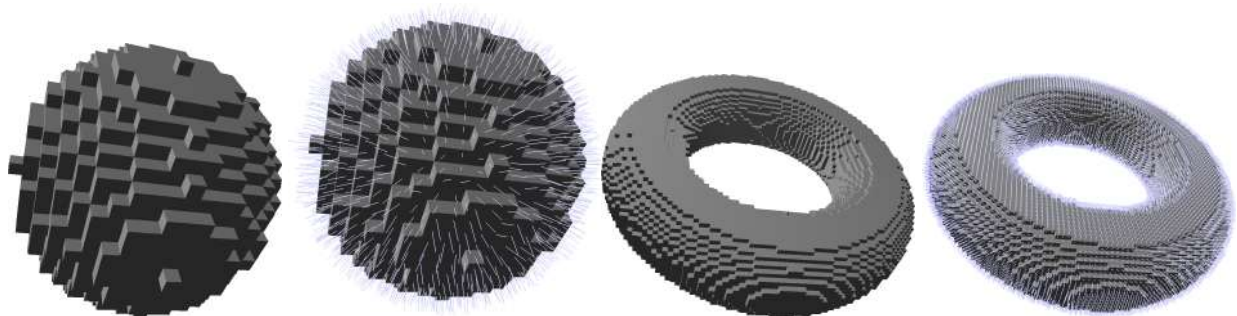


FIGURE 4 – Objets discrets et normales associées à chaque surfel, figures issues de [5].

Sujet

L'objectif de ce stage est d'étudier une nouvelle approche pour la structuration des voisinages locaux sur une surface digitale. Elle se base sur la décomposition d'un voisinage géodésique ou quasi-géodésique en secteurs

afin d'étudier par secteur les estimations locales de planéité, ou de normalité. Elle permet aussi d'étudier le champs normal d'une surface et d'en proposer des régularisations.

Une première étape consistera à décrire et mettre en œuvre un parcours efficace des secteurs généralisés autour d'un point notamment pour des problématiques de test de planéité locale. Un premier travail a été réalisé l'année passée, dans le cadre d'un stage de master, dans lequel l'intersection d'une boule d'un rayon fixé avec la surface a été considérée. D'autres stratégies sont à proposer et à tester. De plus, différentes notions de droites discrètes 3D [6, 7, 9, 8] pourront être utilisées ainsi que des algorithmes de construction et de reconnaissance de ces structures 3D [11]. Dans [10], les auteurs établissent des liens entre les morceaux de plans discrets et les segments de droites discrètes 3D (cf. figure 5), cette voie pourra aussi être explorée.

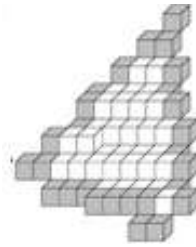


FIGURE 5 – Triangle 3D, figure issue de [10].

Une fois les voisinages établis, le second travail consistera à les utiliser pour partitionner les résultats d'un algorithme de type plane probing [5] afin d'extraire des plans moyens par secteur et ensuite pouvoir les fusionner pour comprendre la structure de la surface digitale de manière non locale, c'est-à-dire en dépassant le strict cadre du voisinage initial.

Enfin un objectif plus lointain, qui pourra être développé pendant une thèse, visera à établir des stratégies pour une construction efficace d'une extension de la couverture tangentielle à la dimension 3. Un effort particulier sera fait pour conserver une combinatoire raisonnable des morceaux de plans digitaux et des temps de calculs contenus.

Pré-requis

Bon niveau en programmation souhaité (idéalement C++, OpenGL)

Références

- [1] F. Feschet and L. Tougne. *Optimal Time Computation of the Tangent of a Discrete Curve : Application to the Curvature*. in DGCI 1999, 31–40.
- [2] F. Feschet and L. Tougne. *On the min DSS problem of closed discrete curves*. in Discret. Appl. Math., 151(1-3) : 138-153 (2005).
- [3] I. Sivignon and D. Coeurjolly. *Minimal Decomposition of a Digital Surface into Digital Plane Segments Is NP-Hard*. DGCI 2006 : 674-685.
- [4] E. Charrier, JO. Lachaud. *Maximal Planes and Multiscale Tangential Cover of 3D Digital Objects*. IWCI 2011 : 132-143.
- [5] JO. Lachaud, J. Meyron and T. Roussillon. *An Optimized Framework for Plane-Probing Algorithms*. Int. J. Math. Imaging Vis. 62(5) : 718-736 (2020).
- [6] D. Coeurjolly, I. Debled-Rennesson, O. Teytaud. *Segmentation and Length Estimation of 3D Discrete Curves*. Digital and Image Geometry 2000 : 299-317.
- [7] K. Voss. *Discrete Images, Objects, and Functions in Zn*. Springer, Berlin (1993).
- [8] L. Ibanez, C. Hamitouche, C. Roux. *A vectorial algorithm for tracing discrete straight lines in N-dimensional generalized grids*. IEEE Transactions on Visualization and Computer Graphics, vol. 7-2, 97-108 (2001).
- [9] F. Feschet, JP. Reveillès. *A Generic Approach for n-Dimensional Digital Lines*. DGCI 2006 : 29-40.
- [10] VE. Brimkov, RP. Barneva. *Graceful planes and lines*. Theor. Comput. Sci. 283(1) : 151-170 (2002).
- [11] DGtal : Digital Geometry tools and algorithms library <http://libdgtal.org>

Équipe BISCUIT

<http://biscuit.loria.fr/>

Sujet 1

Apprentissage par Renforcement sur Processeur Neuromorphique

Proposé par : Alain Dutech (Alain.Dutech@loria.fr),
Bernard Girau (Bernard.Girau@loria.fr)

Equipe : BISCUIT

Problématique

L'équipe BISCUIT, est une équipe du laboratoire Loria qui étudie de nouveaux paradigmes computationnels où les calculs sont adaptatifs, distribués et décentralisés, réalisés par des populations d'unités de calcul simples qui communiquent principalement avec leurs proches voisins. Ces propriétés sont compatibles avec la mise en oeuvre de principes d'auto-organisation non-supervisés, mais guidés, pour s'attaquer à des problèmes difficiles comme le calcul cognitif situé, la robotique autonome, l'allocation adaptative de ressources de calcul, etc.

Ces caractéristiques permettent également d'envisager une meilleure utilisation des processeurs dit "neuromorphiques" qui sont en train d'émerger (IBM Truenorth, Intel Loihi, etc.). Ces processeurs sont basés sur des principes neuro-inspirés qui respectent les contraintes des paradigmes que nous étudions, et peuvent bénéficier des mécanismes d'auto-organisation non supervisée et guidée que nous mettons au point, à la fois en termes d'applications et en termes de gestion des ressources neuromorphiques.

Sujet

L'objectif principal de ce stage de master est d'explorer l'utilisation de mécanismes situés à la croisée du calcul neuro-inspiré et de l'apprentissage par renforcement [1] dans le cadre de ces architectures dites *neuromorphiques*.

En s'inspirant d'un précédent travail de master, nous voulons poursuivre la définition, l'étude et l'implantation un algorithme d'apprentissage par renforcement distribué utilisant des populations de neurones impulsionsnels et compatible avec le processeur neuromorphique Loihi d'Intel [2]. Contrairement aux approches principales de la littérature utilisant un signal de récompense dans un cadre neuromorphique, notre algorithme doit être capable d'apprentissage séquentiel : la récompense à l'agent n'est donnée qu'à la fin de la trajectoire. En cela, l'approche de Frémaux est une inspiration intéressante [3]. Nos travaux précédents ont permis de ré-exprimer cette approche selon les contraintes neuromorphiques du processeur Loihi. Néanmoins, de nombreux aspects restent à étudier, notamment les problèmes de précision des poids synaptiques et de combinaison des traces d'activité synaptique. Enfin, un accent particulier sera mis sur la validation expérimentale du modèle sur des problèmes concrets de contrôle (robotique, saccades attentionnelles, etc).

Compétences

python, Linux, quelques notions de mathématiques (probabilités, équations différentielles), une connaissance préalable de l'apprentissage par renforcement et/ou du calcul neuronal impulsionsnel sera appréciée

Conditions de travail

Le stage se déroulera au sein du laboratoire Loria de Nancy. Le développement des modèles se fera principalement en python, y compris pour l'utilisation des bibliothèques spécifiques de programmation sur le processeur neuromorphique Intel Loihi.

Références

- [1] Sutton, R. and Barto, A. *Reinforcement Learning* Bradford Book, MIT Press, Cambridge, MA. (1998)
- [2] Davies, M., Srinivasa, N., Lin, T., Chinya, G., Cao, Y., Choday, S. H., Dimou, G., Joshi, P., Imam, N., Jain, S., Liao, Y., Lin, C., Lines, A., Liu, R., Mathaikutty, D., McCoy, S., Paul, A., Tse, J., Venkataramanan, G., Weng, Y., Wild, A., Yang, Y., and Wang, H.. *Loihi : A neuromorphic manycore processor with on-chip learning*. IEEE Micro, 38(1) :82-99, (2018).
- [3] Frémaux, Nicolas, Sprekeler, Henning and Gerstner, Wulfram. *Reinforcement learning using a continuous time actor-critic framework with spiking neurons* PLoS Comput Biol, vol 9(4), (2013).

Sujet 2

Apprentissage non-supervisé de flux optique avec des réseaux de neurones impulsionnels

Proposé par : Bernard Girau (contact : Bernard.Girau@loria.fr),
Adrien Fois

Equipe : BISCUIT

Motivations

Les neurones impulsionnels sont considérés comme la troisième génération de modèles de neurones artificiels. Ces modèles de neurones poussent le bio-mimétisme plus loin que leurs prédécesseurs en communiquant - à la manière des neurones biologiques - avec des impulsions produites dans le temps [5]. Une nouvelle dimension - la dimension temporelle - permet ainsi de transmettre l'information et de la traiter à la volée, de façon asynchrone.

Pour tirer pleinement profit de la puissance de calcul et de la très faible consommation énergétique induite, ces modèles de neurones impulsionnels peuvent être directement émulés *in silico*. C'est ce qu'ont notamment réalisé Intel ou IBM avec leurs processeurs neuromorphiques Loihi [1] et TrueNorth, respectivement. Loihi2 intègre un million de neurones impulsionnels et 120 millions de synapses programmables.

Dans la même veine bio-inspirée, les caméras événementielles gagnent en popularité [3]. Les caméras événementielles comme DVS (Dynamic Vision Sensor) fonctionnent de manière analogue à la rétine en transmettant l'information sous forme d'impulsion uniquement lorsqu'un changement local de luminosité - au niveau du pixel - est détecté. Ce traitement asynchrone de l'information visuelle apporte de grands avantages : 1) une vitesse d'échantillonnage près d'un million de fois supérieure à celle des caméras standard, 2) une latence d'une micro-seconde et 3) une plage dynamique de 130 décibels (les caméras standards n'ont que 60 dB). Le tout pour une consommation énergétique significativement inférieure à celle des caméras standards.

Lorsqu'un organisme équipé d'un système visuel est en mouvement dans son environnement, ou observe un objet en mouvement en restant statique, il perçoit un mouvement relatif entre lui-même et son environnement. Ce mouvement lui apparaît sous la forme de motifs spatio-temporels, appelé flux optique. Estimer le flux optique est une tâche essentielle pour l'organisme. Cette information lui permet de mieux estimer son propre déplacement et ainsi de mieux naviguer dans son environnement. Ces problématiques sont aussi transposables dans le domaine de la robotique autonome et des drones.

Ce stage d'une durée minimale de 5 mois, se situe à la croisée de ces différents domaines.

Sujet

Le but est d'utiliser une caméra événementielle et de traiter ses données avec un réseau de neurones impulsionnels équipé de règles d'apprentissage non supervisée [4, 2]. L'application visée est l'estimation de flux optique.

Il s'agira notamment de :

- mener une étude bibliographique sur les méthodes d'estimation de flux optique non-supervisée avec des réseaux de neurones impulsionnels
- adapter une règle d'apprentissage de type STDP, développée au sein de l'équipe[6], à l'application visée
- intégrer cette règle dans des réseaux de neurones impulsionnels
- proposer des adaptations compatibles avec une implémentation matérielle sur processeur neuromorphique
- implémenter et tester l'architecture avec Tensorflow

Cadre du travail

Une formation en informatique (avec des bases en intelligence artificielle) ou en neurosciences computationnelles est attendue, ainsi que de solides bases en programmation.

Le stage se déroulera au sein du laboratoire Loria, dans l'équipe Biscuit.

Références

- [1] M. Davies, N. Srinivasa, T.-H. Lin, Tsung-Han, G. Chinya, Y. Cao, S.H. Choday, G. Dimou, P. Joshi, N. Imam, S. Jain, Y. Liao, C.-K. Lin, A. Lines, R. Liu, D. Mathaikutty, S. McCoy, A. Paul, J. Tse, G. Venkataramanan, Y.H. Weng, A. Wild, Y. Yang and H. Wang. Loihi : A Neuromorphic Manycore Processor with On-Chip Learning. *IEEE Micro*, 38 (1), 2018.
- [2] T. Barbier, C. Teuliere, and J. Triesch. *ISpike Timing-Based Unsupervised Learning of Orientation, Disparity, and Motion Representations in a Spiking Neural Network*. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2021.
- [3] G. Gallego, T. Delbruck, G. Orchard, C. Bartolozzi, B. Taba and A. Censi, S. Leutenegger, A. Davison, J. Conradt, K. Daniilidis, D. Scaramuzza. Event-Based Vision : A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* , 2020.
- [4] F. Paredes-Valles, K. Schepera and G. De Croon. Unsupervised Learning of a Hierarchical Spiking Neural Network for Optical Flow Estimation : From Events to Global Motion Perception, *IEEE Transactions on Pattern Analysis and Machine Intelligence* , 2019.
- [5] A. Taherkhani, A. Belatreche, Y. Li, G. Cosma, L. Maguire, and T.M. McGinnity. A Review of Learning in Biologically Plausible Spiking Neural Networks. *Neural Networks*, 122, 2020.
- [6] A. Fois and B. Girau. *A Spiking Neural Architecture for Vector Quantization and Clustering*. 27th International Conference on Neural Information Processing (ICONIP), 2020.

Sujet 3

Cartes auto-organisatrices et architectures reconfigurables

Proposé par : Bernard Girau (contact : Bernard.Girau@loria.fr),
Horacio Rostro-Gonzalez

Equipe : BISCUIT

Contexte

Cette proposition de stage se situe dans le contexte d'une collaboration (projet ANR SOMA) allant des neurosciences computationnelles à la conception de circuits multi-coeurs reconfigurables, et dont le but principal est d'étudier l'apport possible des principes de plasticité structurelle neuronale dans le cadre de la conception d'architectures de calcul reconfigurables massivement distribuées.

Sujet

Les cartes de Kohonen ([4]) sont un modèle connu d'auto-organisation inspiré du cortex, et permettent une quantification vectorielle non supervisée de données dans laquelle les prototypes s'organisent selon des règles de voisinage pré-fixées (plasticité synaptique seule). D'autre part, le modèle Growing Neural Gas ([5]) permet une quantification vectorielle d'où émergent des relations de voisinage sans aucune contrainte (plasticité structurelle). Dans le cadre du projet SOMA où l'affectation de tâches aux différents noeuds de calcul d'un support manycore se fait par auto-organisation, nous avons développé des modèles intermédiaires entre les cartes de Kohonen et les Growing neural gas (CSOM, PSOM, SPCSOM, [2, 3]) dans lesquels les relations de voisinage entre prototypes peuvent s'adapter aux données tout en respectant des contraintes particulières, par exemple celles du placement et du routage matériel. Par ailleurs, des partenaires du projet SOMA ont défini une architecture multi-coeurs reconfigurable (carte SCALP, [1]) capable d'exploiter les principes d'auto-organisation matérielle. L'objectif du travail proposé ici est de faire le lien entre les modèles auto-organiseurs et l'architecture multi-coeurs issus du projet SOMA.



Après une étape de familiarisation avec les modèles auto-organiseurs CSOM, PSOM, SPCSOM d'une part, et la carte SCALP et son simulateur SCALPsim d'autre part, le travail qui sera réalisé se décomposera en deux phases.

- implantation des modèles de cartes auto-organisatrices à plasticité structurelle sur la carte SCALP (sur la base des implantations déjà proposées pour CSOM)
- analyse de leur comportement face à des données réelles et de leur efficacité au regard des contraintes de communication de la carte grâce au simulateur SCALPsim

Cadre du travail

Une partie du travail se faisant à l'aide d'un logiciel de simulation de la carte SCALP, une maîtrise de la conception logicielle est requise, mais une connaissance de la conception sur circuit numérique sera également nécessaire pour tester les modèles sur la carte SCALP (via l'utilisation du logiciel Vivado de Xilinx).

Références

- [1] F. Vannel, D. Barrientos, J. Schmidt, C. Abegg, D. Buhlmann, and A. Upegui. *SCALP : Self-configurable 3-d cellular adaptive platform*. Proceedings of the 2018 IEEE Symposium Series on Computational Intelligence (SSCI), November 2018, pp. 1307-1312.
- [2] B. Girau and A. Upegui. *Cellular self-organising maps*. Int. Workshop on Self-Organizing Maps and Learning Vector Quantization, Clustering and Data Visualization (WSOM), 2019.
- [3] A. Upegui, B. Girau, N. Rougier, F. Vannel, and B. Miramond. *Pruning self-organizing maps for cellular hardware architectures*. NASA/ESA Conf. on Adaptive Hardware and Systems (AHS), 2018.
- [4] T. Kohonen. The self-organizing map. *Neurocomputing*, 21(1-3), 1998.
- [5] B. Fritzke. A growing neural gas network learns topologies . In *Advances in Neural Information Processing Systems 7*, 1995.

Sujet 4

Champs neuronaux sur circuits neuromorphiques

Proposé par : Bernard Girau (contact : Bernard.Girau@loria.fr),
Horacio Rostro-Gonzalez

Equipe : BISCUIT

Motivations

Parmi les différentes alternatives aux architectures de Von Neumann, les approches neuromorphiques bénéficient actuellement des récents succès applicatifs des méthodes d'apprentissage profond (deep learning), et de l'implication des principaux fabricants de semi-conducteurs à travers la conception de puces neuromorphiques impressionnantes telles que l'IBM TrueNorth et l'Intel Loihi. De tels substrats matériels offrent une certaine flexibilité dans leur connectivité, voire une certaine adaptabilité grâce à une plasticité synaptique intégrée.

Dans l'équipe Biscuit, nous considérons qu'un moyen possible de promouvoir ces technologies neuromorphiques et d'élargir leur champ d'application potentiel est de pouvoir étendre cette flexibilité en plaquant les principes bio-inspirés de l'auto-organisation et de l'émergence de calcul au niveau des ressources matérielles neuromorphiques. Plus précisément, nous mettons en avant l'utilisation de deux familles de modèles neuronaux bio-inspirés, les champs neuronaux dynamiques et les cartes auto-organisatrices.

Sujet

Les champs neuronaux dynamiques (dynamic neural fields, DNF) sont des modèles bio-inspirés de populations neuronales (généralement 2D) modélisés à une échelle mésoscopique, où les unités neuronales interagissent en continu via des connexions excitatrices et inhibitrices. Les DNF ont été appliqués avec succès à l'attention visuelle [4], à la détection de nouveautés, à la reconnaissance faciale, au contrôle de robots, etc. Dans le contexte de la mise en œuvre numérique matérielle de DNF sur FPGA (field programmable gate arrays, circuits numériques reconfigurables), des versions impulsives (spiking) de DNF ont été définies [3], montrant des propriétés comportementales similaires avec une robustesse améliorée par rapport au bruit et aux distracteurs.

Ce sujet de Master vise à étudier comment les versions impulsives des champs neuronaux dynamiques peuvent être implantées sur des circuits neuromorphiques qui ont émergé récemment, tels que l'architecture Intel Loihi [1], en termes de calcul et de communication. La principale application illustrative consistera à suivre des cibles dans une scène visuelle capturée par une caméra impulsive DVS.

Après une étape de familiarisation avec les modèles neuronaux visés et le circuit Loihi d'Intel, le travail qui sera réalisé se décomposera en trois phases.

1. implantation de DNF impulsifs sur circuit neuromorphique
2. couplage avec une caméra impulsive pour le suivi de cible
3. étude de différents modes de communication des impulsions entre neurones (cf [2, 5])

Cadre du travail

Ce travail nécessite une connaissance de la conception logicielle et des notions de calcul neuronal bio-inspiré. Il s'appuiera sur des logiciels existants permettant la simulation de champs neuronaux dynamiques ainsi que sur les plateformes logicielles développées par Intel pour la simulation des implantations sur le circuit Loihi (le projet RV2.133.Girau a été approuvé par l'INRC, Intel Neuromorphic Research Community, et permet l'accès à ces logiciels confidentiels).

Références

- [1] A. Lines, P. Joshi, R. Liu, S. McCoy, J. Tse, Y. Weng, and M. Davies. *Loihi asynchronous neuromorphic research chip*. In 2018 24th IEEE International Symposium on Asynchronous Circuits and Systems (ASYNC), 2018.

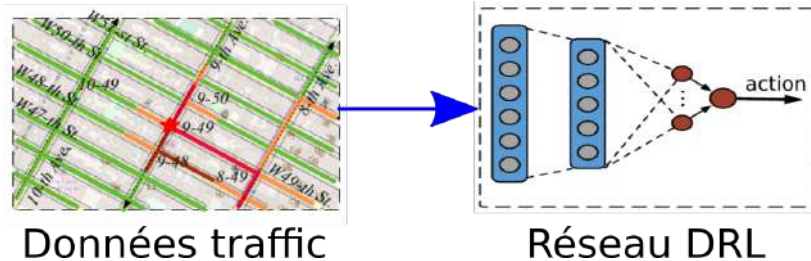
- [2] B. Chappet De Vangel, C. Torres-Huitzil, and B. Girau. *Randomly spiking dynamic neural fields*. Journal of Emerging Technologies in Computing Systems, 2014.
- [3] R. Vazquez, B. Girau and J.-C. Quinton. *Visual attention using spiking neural maps*. International Joint Conference on Neural Networks IJCNN, San José, United States, 2011.
- [4] J. Vitay, N. Rougier, F. Alexandre. *A Distributed Model of Spatial Visual Attention*. In : Wermter S., Palm G., Elshaw M. (eds) Biomimetic Neural Learning for Intelligent Robots. Lecture Notes in Computer Science, vol 3575. Springer, Berlin, Heidelberg, 2005.
- [5] B. Chappet De Vangel and B. Girau. *Stochastic and asynchronous spiking dynamic neural fields*. In *International Joint Conference on Neural Networks*, 2015.

Sujet 5

Apprentissage par Renforcement Profond pour le Contrôle du Trafic Routier

Proposé par : Alain Dutech (Alain.Dutech@loria.fr, members.loria.fr/ADutech), Joël Legrand (Joel.Legrand@loria.fr, http://joel-legrand.fr), KESK Heynrick??

Equipe : BISCUIT



Problématique

Les dernières années ont vu un usage accru de l'apprentissage automatique avec la mise au point de méthodes d'apprentissage profond. De nombreux domaines d'application ont bénéficié de ces progrès, notamment la vision par ordinateur, le traitement automatique des langues, la reconnaissance automatique de la parole, les jeux, etc. L'entreprise "Les Couleurs du Numérique", par l'entremise de sa solution KESK¹, s'intéresse à la régulation automatique du trafic routier. KESK explore l'utilisation de techniques d'apprentissage par renforcement profond (DRL) [1] dans ce domaine.

KESK dispose de données terrain issues de l'exploitation de caméras implantées sur différents carrefours d'une ville de la région parisienne. Les méthodes de DRL nécessitent énormément de données pour obtenir des résultats satisfaisant. Pour pallier cette difficulté, il est classique d'utiliser un simulateur pour fournir des données "synthétiques", se pose alors des problèmes inhérent à l'apprentissage par transfert. La problématique de ce travail de stage est donc d'utiliser un simulateur de trafic routier open-source (SUMO²) pour générer des données utilisée pour pré-apprendre un modèle. Ce modèle sera ensuite adapté et confronté aux données terrain.

Sujet

Ce travail s'appuiera principalement sur deux publications : [2] et [3]. Dans un premier temps, il s'agira de prendre en main le simulateur SUMO, de créer des environnements routiers et générer des données de trafic. Ensuite, nous testerons plusieurs algorithmes d'apprentissage par renforcement profond de l'état de l'art. Enfin, nous chercherons comment ces modèles appris avec des données synthétiques peuvent être adaptés aux données réelles.

Compétences

python, pytorch, Linux, Optimisation, quelques notions de mathématiques (probabilités, équations différentielles).

Conditions de travail

Le stage se déroulera au sein du laboratoire Loria de Nancy.

Références

- [1] V. François-Lavet, P. Henderson, R. Islam, M. Bellemare and J. Pineau. *An introduction to deep reinforcement learning* arXiv preprint arXiv :1811.12560, 2018.
- [2] N. Kheterpal, K. Parvate, C. Wu, A. Kreidieh, E. Vinitsky, and A. Bayen. *Flow : Deep reinforcement learning for control in sumo* EPiC Series in Engineering, Vol. 2, pages 134–151, 2018.
- [3] E. Walraven. *Traffic Flow Optimization using Reinforcement Learning*. Master's thesis, Delft University of Technology, 2014.

1. <https://www.kesk.eu>

2. <https://www.eclipse.org/sumo/>

Sujet 6

Agents collaboratifs intrinsèquement auto-organisés

Proposé par : Amine Boumaza (Amine.Boumaza@loria.fr,
<http://members.loria.fr/ABoumaza>),
Alain Dutech (Alain.Dutech@loria.fr,
<http://members.loria.fr/ADutech>)

Equipe : BISCUIT



Problématique

L'auto-organisation de systèmes complexes est une questions toujours largement ouverte et d'actualité. Dans de tels système, qui sont dit décentralisés, distribués et localisés, des agents avec des comportements "simples" peuvent parfois s'auto-organiser pour produire collectivement des comportements globaux aux propriétés intéressantes (robustesse aux défaillances individuelles, passage à l'échelle, etc). Si les premiers succès en la matière datent des années 1990 [1], il n'exite toujours aucun algorithme qui permet de définir les comportements locaux qui conduiront à l'émergence d'un comportement global défini à l'avance. On en est donc réduit à l'utilisation massive de simulations pour explorer l'effet de tel ou tel comportement local.

L'idée de ce sujet de stage est d'explorer l'apport des concepts de motivation intrinsèque pour aider ou guider l'auto(organisation de systèmes complexes. Depuis les travaux fondateurs de [2], de nombreuses variantes de motivation intrinsèques ont été proposées et appliquées. Nous voudrions ici étudier l'apport de l'utilisation de certaines de ces motivations intrinsèques pour la coordinations et l'auto-organisation d'un système complexe concret.

Sujet

Ce travail s'appuiera sur une plateforme robotique constituée de robots agents autonomes et indépendants, des robots Ergo Junior. Ils devront collaborer pour atteindre des objectifs nécessitant la participation des deux robots, comme par exemple déplacer un objet dans un conteneur manipulé par l'un des deux robots. Ces deux robots pourront, si le besoin s'en fait sentir, être simulés [3].

Ces deux robots sont vu comme deux agents apprenant autonomes dotés de motivations intrinsèques à définir. Il peut s'agir de curiosité, de biais à la performance, de biais à la nouveauté, etc (voir [4] comme inspiration). Le travail consistera essentiellement à explorer les avantages et les limitations de plusieurs types de motivations intrinsèques pour l'auto-organisation de ces deux robots.

Compétences

python, ROS, Linux, apprentissage automatique.

Conditions de travail

Le stage se déroulera au sein du laboratoire Loria de Nancy.

Références

- [1] M. Dorigo, V. Maniezzo, and A. Colorni. *Ant System : Optimization by a Colony of Cooperating Agents*. IEEE Transactions on Systems, Man, and Cybernetics-Part B, 26 (1) : 29–41, 1996.
- [2] P.Y. Oudeyer, F. Kaplan and V.Hafner. *Intrinsic motivation systems for autonomous mental development* IEEE Transactions on Evolutionary Computation, 2007.
- [3] Jessica Colombel, Jean-Luc Charles, Rémi Fabre, Jean-Baptiste Horel, Yoan Mollard, Alexandre Péré and Steve N’Guyen. *Robotique de manipulation avec Poppy Ergo Jr*. <https://learn.ros4.pro/fr/manipulation/ergo-jr>, visité le 7 octobre 2021.
- [4] Adrien Baranes and Pierre-Yves Oudeyer. *Active Learning of Inverse Models with Intrinsically Motivated Goal Exploration in Robots*. Robotics and Autonomous Systems, 61(1) :4973, January 2013. arXiv : 1301.4862.

Équipe Carbone

<http://carbone.loria.fr/>

Sujet 1

Détection de motifs de virtualisation

Proposé par : Guillaume Bonfante

Equipe : Carbone

Informations générales

Encadrants	Guillaume Bonfante
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Téléphone	03 54 95 84 61
Email	guillaume.bonfante@loria.fr
Bureau	B 248

Motivations

La virtualisation consiste à transformer un exécutable en un autre, tel que ce dernier ne soit pas fait d'instructions assembleur mais de meta-instructions, ou bytecode, interprété par un genre de machine virtuelle qu'il embarque. Ce sujet a pour but d'étudier la virtualisation de programmes et d'essayer de l'inverser dans des cas concrets, c'est-à-dire à partir d'un programme virtualisé, de retrouver le programme d'origine et de pouvoir ainsi plus facilement l'analyser statiquement, c'est-à-dire étudier son fonctionnement sans avoir à l'exécuter.

Sujet

Les malwares sont de plus en plus nombreux, et de plus en plus sophistiqués. Une des techniques pour empêcher les anti-malwares de détecter le fonctionnement d'un malware consiste à le virtualiser. Le fait d'inverser le processus de manière automatique serait donc un atout de taille dans la lutte contre les malwares.

L'objectif est de commencer à s'attaquer aux programmes qui s'appuieraient sur une machine virtuelle JAVA. C'est un excellent départ étudier la question. Un tel sujet peut bien sûr être poursuivi en thèse. L'approche pourra suivre de loin en loin celle qu'on peut trouver dans [1] ou encore [2].

Cadre du travail

L'entreprise Cyberdetect, issue de l'équipe CARBONE, commercialise un outil de détection de malwares se basant sur une approche différente des antivirus traditionnels, et appelée *analyse morphologique*. Comprendre mieux les processus de virtualisation et les déjouer serait utile dans le cadre de cette méthode de détection de malwares.

Références

- [1] Salwan, J., Bardin, S., Potet, M. : Symbolic deobfuscation : From virtualized code back to the original. In : Giuffrida, C., Bardin, S., Blanc, G. (eds.) Detection of Intrusions and Malware, and Vulnerability Assessment - 15th International Conference, DIMVA 2018, Saclay, France, June 28-29, 2018, Proceedings. Lecture Notes in Computer Science, vol. 10885, pp. 372–392. Springer (2018)
- [2] Saudel, F., Salwan, J. : Triton : A dynamic symbolic execution framework. In : Symposium sur la sécurité des technologies de l'information et des communications, SSTIC, France, Rennes, June 3-5 2015. pp. 31–54. SSTIC (2015)

Équipe COAST

<https://team.inria.fr/coast/>

Sujet 1

Stragglers in Large-Scale Infrastructure: On the Impact of Hardware and Workloads

Informations générales

Encadrants Thomas LAMBERT (COAST, LORIA), Shadi IBRAHIM (Myriads, IRISA)
Email thomas.lambert@inria.fr, shadi.ibrahim@inria.fr

Context

Many data-intensive applications nowadays run on large-scale shared clusters. The scale, hardware heterogeneity, and diversity of applications in those clusters result in performance variability. Performance variability causes stragglers (i.e., tasks performing relatively slower than other tasks) which result in a severe degradation in performance, especially as the completion time of a job strongly depends on the finishing times of these stragglers. Traces from production clusters at Facebook and Microsoft pointed out that stragglers can be 7-8 X slower than the median task and can increase the average job duration by up to 47 [1].

Objectives

Many research efforts have been dedicated to improving straggler mitigation in large-scale data-intensive clusters through using speculative execution [2], launching multiple copies (i.e., clones) of tasks belonging to small jobs [1], making speculative execution resource-aware [3], [4], and evaluating and improving stragglers detection [5]. In contrast, little work is focusing on measuring stragglers or understanding their impacts in large-scale production clusters.

The goal of this internship is to conduct a large-scale analysis to characterize stragglers in different production data-intensive clusters (through large-scale experiments and publicly available traces), and to precisely understand the effects of scale, workload characteristic and size, and hardware type on the number and durations of stragglers.

Cadre du travail

The internship takes place within the COAST team at LORIA/Inria Nancy - Grand Est.

Références

- [1] Ananthanarayanan, G., Ghodsi, A., Shenker, S., Stoica, I. Effective Straggler Mitigation : Attack of the Clones. In *NSDI'13*, pages 185–198. ACM, 2013.
- [2] Dean, J., Ghemawat, S. MapReduce : Simplified Data Processing on Large Clusters. In *Communications of the ACM*, 51(1) :107–113, 2008.
- [3] Ananthanarayanan, G., Kandula, S., Greenberg, A., Stoica, I., Lu, Y., Saha, B., Harris, E., Reining in the outliers in MapReduce clusters using Mantri In *OSDI'10*, pages, 265-278. 2010
- [4] Zhou, A. C., Phan, T. D., Ibrahim, S., He, B. Energy-efficient speculative execution using advanced reservation for heterogeneous clusters. In *ICPP'18*, pages 1-10. 2018
- [5] Phan, T. D., Pallez, G., Ibrahim, S., Raghavan, P. A new framework for evaluating straggler detection mechanisms in mapreduce. In *ACM Transactions on Modeling and Performance Evaluation of Computing Systems*, 4(3), pages 1-23. 2019

Équipe K

<https://www.loria.fr/fr/la-recherche/les-equipes/equipe-k/>

Sujet 1

Un crawler de connaissances sur le web

Proposé par : Mathieu d'Aquin et Emmanuel Nauer

Equipe : K

Informations générales

Encadrants	Mathieu d'Aquin, Emmanuel Nauer	
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy	
Téléphone	+33 (0)383593000	
Email	mathieu.daquin@loria.fr	emmanuel.nauer@loria.fr
Bureau	B 160	

Motivations

L'idée du web sémantique [1, 2] est de permettre l'utilisation de l'architecture du web pour transmettre des connaissances formalisées. Cela peut être vu comme un mariage entre les technologies du web, fournissant un accès universel aux documents, et celles de l'intelligence artificielle traditionnelle, qui s'intéresse à la représentation de connaissances et au raisonnement. Le résultat de ces travaux est non seulement le développement de milliers de modèles de connaissances sous la forme d'ontologies web [3, 4], mais aussi d'énormes graphes de connaissances³ tels que DBpedia.org, wikidata ou encore ceux utilisés par Google [5]. Néanmoins, du fait du grand nombre de ces ressources et qu'elles soient distribuées sur le web, leur utilisation reste difficile. Plusieurs moteurs de recherche pour ontologies ou graphes de connaissances ont été créés [6], mais ne sont plus opérationnels, ayant été développés sur la base de technologies inadaptées, avant l'ère du big data.

Sujet

L'objectif de ce stage est la création d'un crawler (robot d'exploration) permettant de collecter, sur le web, ces connaissances formalisées sous la forme de graphes de connaissances ou d'ontologies. C'est un sujet fondamentalement technique, qui requiert la mise en place d'une infrastructure pour la parallélisation et la priorisation de l'exploration de ressources. Il nécessite aussi des recherches à l'intersection de l'ingénierie des connaissances et de la recherche d'information concernant les choix de conception à faire pour ce crawler incluant :

- Comment reconnaître qu'une ressource web contient des connaissances ?
- Comment décider des prochains liens à explorer à partir d'une ressource contenant (ou pas) des connaissances ?
- Comment prendre en compte la charge des serveurs hébergeant les ressources dans l'exploration ?
- Comment gérer les échelles très variables des ressources de connaissances ?
- Comment stocker efficacement les connaissances trouvées et les informations nécessaires au fonctionnement du crawler ?

Cadre du travail

Ce stage sera réalisé dans le cadre de l'équipe K du LORIA, qui s'intéresse à l'intelligence artificielle symbolique au travers en particulier de la représentation et de l'ingénierie des connaissances et de différents types de raisonnements. Le développement du crawler de connaissances se place dans le cadre d'un programme plus vaste se focalisant sur l'exploitation de connaissances formalisés au travers d'entrepôts de connaissances, et sera donc appliqué dans des domaines variés au travers de plusieurs projets en cours ou à venir, en particulier en santé, en physique et en humanités numériques.

Références

- [1] Berners-Lee et al., The semantic web, in Scientific American, 254/5, 2001
- [2] d'Aquin and Motta, The epistemology of intelligent semantic web systems, Morgan & Claypool, 2016
- [3] Staab and Studer, Handbook on ontologies, Springer, 2010
- [4] OWL Working Group, OWL 2 Web Ontology Language Document Overview, W3C Recommendation, 2009
- [5] Zou, A survey on application of knowledge graph, Journal of Physics : Conference Series, 1487/1, 2020
- [6] d'Aquin et al., Semantic web search engines, in Handbook of Semantic Web Technologies, 2011

3. voir <https://lod-cloud.net/>

Sujet 2

Recommandation system for medical encoding

Proposé par : Nicolas Jay, Emmanuel NAUER

Equipe : K

General informations

Supervisors Nicolas JAY, Emmanuel NAUER
Address LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Email nicolas.jay@loria.fr emmanuel.nauer@loria.fr
Office B 152

Motivations

Unique in Europe, or even in the world, the National Health Data System has a considerable potential for analyzing and improving population state of health. For acute care hospitalizations, it is feeded each year by data from nearly 30 million stays coming from the “Programme de Medicalisation des systemes d’information” (PMSI). Mainly designed to finance healthcare settings, the PMSI is more and more used for research purposes, health monitoring, and care organization. The quality of PMSI data is therefore crucial for conducting reliable and relevant analyzes, whatever the purpose is. This quality is based on a synthesis and a coding work of the information contained in the medical file, work realized by humans coders according to established rules published in a national guide. The massive computerization of hospital medical records makes possible and necessary the development of coding support systems, which can rely on semi-structured heterogeneous data (biology results, drug prescriptions for example) and unstructured data (textual reports) contained in each record.

Research subject

The goal of this research is to build a knowledge base (from PSMI data) that will be used by a new coding recommendation system. The student will have to explore and adapt association rules extraction techniques [2, 1] to identify relevant coding rules, with a particular attention for reusing existing ontology knowledge to guide the rule extraction and the rule selection processes.

Two main limits of association rules extraction have to be outreached :

- the number of rules generated is often important and generates a heavy work of selection and validation for the domain experts ;
- the rule quality depends on the level of granularity with which the initial information is coded. A too fine granularity avoids the discovery of sufficiently general rules, whereas a too weak granularity causes the generation of trivial rules.

Two approaches will therefore be explored simultaneously : the contribution of external knowledge ahead and beside the datamining step, and the choice of interest measures [3] to select relevant rules. A global discovery strategy has to be elaborated by taking into account the interaction between these two approaches.

Work organization

The work will be carried out alternately at LORIA, in the team K, and at the University Hospital (CHRU) of Nancy, in the "Service d’Evaluation et d’Information médicales" under Nicolavs Jay’s responsibility. The student will work on data from the CHRU information system.

Références

- [1] SÃ©verac et al. Non-redundant association rules between diseases and medications : an automated method for knowledge base construction. BMC Medical Informatics and Decision Making, 2015.
- [2] Agrawal et al. Mining association rules between sets of items in large databases. in proceedings of the ACM SIGMOD Int’l Conference on Management of Data, 1993.
- [3] Guillet et al. Quality Measures in Data Mining. Studies in Computational Intelligence, 2007.

Équipe Larsen

<https://team.inria.fr/larsen/>

Sujet 1

Quelles informations guident notre regard ? Modélisation de l'attention visuelle.

Proposé par : Olivier Buffet et Vincent Thomas

Equipe : Larsen

Informations générales

Encadrants	Vincent Thomas	Olivier Buffet
Adresse	INRIA Nancy Grand Est, 615 rue du Jardin Botanique, 54600 Villers-lès-Nancy	
Téléphone	03 54 95 85 08	03 54 95 86 15
Email	vincent.thomas@loria.fr	olivier.buffet@loria.fr
Bureau	C125	C124

Motivation

Ce sujet de stage s'insère dans le projet interdisciplinaire PAAVUP(P) (Prédire l'Allocation de l'Attention Visuelle à l'aide des notions d'Utilité et de Probabilité (POMDP)) mené au sein de l'Université de Lorraine par des membres du LORIA et du laboratoire de psychologie et d'ergonomie PErSEUs. Ce projet a pour objectif de confronter les points de vue de l'informatique et de la psychologie cognitive sur la modélisation de l'attention visuelle. En psychologie cognitive, on considère classiquement que l'attention visuelle est composée de deux processus : (1) les processus ascendants (bottom-up), qui dépendent des caractéristiques de l'environnement, telles que la saillance des objets ; (2) les processus descendants (top-down), qui dépendent des objectifs de la personne, tels que la tâche en cours, et de ses connaissances préalables [4].

De nombreux travaux ont conduit à des implémentations de modèles quantitatifs et prédictifs de ce que les gens regardent. Cependant, ces modèles sont centrés sur les processus bottom-up, ce qui est problématique car les processus top-down guident majoritairement l'attention visuelle pour les situations du quotidien, qui sont dynamiques et multi-tâches. Les recherches en informatique offrent des pistes qui permettraient de penser un modèle prédictif principalement top-down. En effet, un processus top-down vise idéalement à faire des choix optimaux de prise d'information visuelle. Ce problème précis peut ainsi s'écrire sous la forme d'un problème de prise de décision en observabilité partielle pour lequel un agent autonome doit choisir les actions à faire dans le but de maximiser l'information qu'il a sur son environnement. Cela peut par exemple correspondre à la situation d'un robot devant choisir comment se déplacer dans son environnement pour recueillir le plus d'information à partir de ses capteurs. Dans le cadre du projet interdisciplinaire PAAVUP, on souhaiterait, à long terme, étudier s'il est possible de modéliser l'attention visuelle comme un problème de construction d'une stratégie optimale de saccades oculaires permettant de maximiser l'information utile d'une scène visuelle.

Sujet

Le sujet de ce stage consiste donc à modéliser et résoudre le problème de recherche d'information visuelle dans une scène sous la forme d'un processus de prise de décision en observabilité partielle. Le stage se fera en parallèle avec un étudiant de master 2 en psychologie cognitive amené à construire un dispositif expérimental et à recueillir des données auxquelles on pourra comparer les résultats obtenus par la stratégie que l'on cherchera à construire.

Dans un premier temps, le stagiaire devra se familiariser avec l'état de l'art concernant la prise de décision en observabilité partielle (cadre formel des POMDP) et des modélisations existantes des processus top-down de l'attention visuelle [1, 2].

Dans un second temps, le stagiaire devra

- modéliser le problème de recherche d'information dans une scène visuelle en se concentrant sur deux grands éléments : la tâche à résoudre (correspondant au dispositif expérimental que des sujets humains devront passer) et la manière dont l'œil peut acquérir de l'information visuelle en fonction de son point de fixation (largeur de la zone perceptive, modélisation du bruit de perception) ;
- proposer des techniques pour résoudre ce problème de recherche d'information en devant faire face à de larges espaces d'états, d'actions et d'observations (possiblement continus) [3] ;

- trouver des moyens de comparer les stratégies obtenues en calculant la solution au problème de prise de décision avec des trajectoires oculaires observées chez des sujets humains (différence de performance, distance entre les trajectoires optimales et les trajectoires observées).

Enfin, en fonction de l'avancée des travaux, on pourra réfléchir à la manière d'estimer des paramètres du modèle (taille de la zone perceptive, niveau de bruit, ...) à partir des trajectoires observées chez des sujets humains.

Cadre du travail

Ces méthodes seront mises en œuvre et évaluées d'abord dans un cadre de simulations dans le but de valider les stratégies d'acquisition d'informations visuelles qui auront pu être construites. Les stratégies obtenues seront ensuite comparées aux données recueillies au cours du stage de master en psychologie cognitive qui se déroulera en parallèle.

Le développement pourra se faire en Java mais pourra aussi dépendre des bibliothèques que nous pourrions utiliser pour le calcul de la stratégie optimale associée au problème. Le stage se déroulera au LORIA avec des réunions hebdomadaires et des réunions régulières seront aussi prévues avec l'ensemble des intervenants du projet (Sophie Lemonnier du laboratoire PErSEUs qui encadrera le stage de master de psychologie cognitive).

Références

- [1] Najemnik, J., Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*, 434(7031) : 387–391.
- [2] Butko, N. J., Movellan, J. R. (2008). I-POMDP : An infomax model of eye movement. In *7th IEEE International Conference on Development and Learning*.
- [3] Sunberg, Z. N., Kochenderfer, M. J. (2018). Online algorithms for POMDPs with continuous state, action, and observation spaces. In *Proceedings of the Twenty-Eighth International Conference on Automated Planning and Scheduling (ICAPS-18)*.
- [4] Wickens, C. D., Goh, J., Helleberg J., Horrey, W. J., Talleur, D. A. (2003). Attentional models of multitask pilot performance using advanced display technology. *Human factors*, 45(3) : 360–380.

Sujet 2

Planification d'actions avec durées incertaines

Proposé par : Vincent Thomas et Olivier Buffet

Equipe : Larsen

Informations générales

Encadrants	Vincent Thomas	Olivier Buffet
Adresse	INRIA Nancy Grand Est, 615 rue du jardin botanique, 54600 Villers-lès-Nancy	
Téléphone	03 54 95 85 08	03 54 95 86 15
Email	vincent.thomas@loria.fr	olivier.buffet@loria.fr
Bureau	C125	C124

Motivation

De nombreux travaux cherchent à contrôler l'état d'un système dynamique pour maximiser une fonction de performance, par exemple pour qu'un robot atteigne une situation en minimisant ses coûts de déplacement. Cependant, la durée des actions exécutées ou la durée des processus mis en œuvre dans l'environnement n'est souvent pas modélisée finement. Ceci peut devenir un obstacle lorsqu'on souhaite, par exemple, planifier les actions d'un robot qui doit agir avec un être humain. En effet, dans ce cas, le robot doit être capable d'estimer les durées (variables) des tâches effectuées par l'humain et de ses propres actions pour pouvoir agir correctement et au bon moment. Ceci est d'ailleurs d'autant plus vrai lorsque la durée d'une activité peut nous renseigner sur sa réussite ou son échec. Par exemple, la durée passée sur une tâche de réparation par un être humain peut constituer un indice sur l'action du robot qui aiderait le plus l'humain. Ce stage cherchera ainsi à voir comment représenter explicitement la notion de durée et en tirer parti pour proposer des algorithmes capables de résoudre des problèmes de prise de décision dans l'incertain (en particulier dans un cadre de collaboration homme-robot).

Sujet

Le cadre des MDP (*Markov Decision Process*) permet de représenter des problèmes de prise de décision séquentielle lorsque l'agent connaît l'état du monde mais que les actions qu'il effectue ont des résultats incertains. On souhaite dans un premier temps explorer ce cadre lorsque les actions effectuées ont des durées variables (par exemple des macro-actions constituées d'un certain nombre de sous-actions) ou lorsque certains processus ont des durées incertaines (mais que l'on peut quantifier). On abordera ensuite des problèmes plus complexes quand par exemple l'agent n'a plus accès à l'état du monde mais uniquement à des observations qui en dépendent (cadre des POMDP – *Partially Observable Markov Decision Process*).

L'objectif du stage consistera (i) à étudier les modèles de prise de décision markoviens intégrant explicitement la notion de durée et les algorithmes existants; (ii) à proposer et valider de nouveaux algorithmes permettant d'aborder ces problèmes tout en évitant une explosion du temps de calcul.

Cadre du travail

Dans un premier temps, il faudra se familiariser avec les formalismes MDP et POMDP et les formalismes intégrant explicitement la notion de durée (SMDP et POSMDP). Dans un second temps, on proposera et on cherchera à modéliser des problèmes de prise de décision en observabilité partielle pour lesquels raisonner sur la dimension temporelle est nécessaire pour agir optimalement. Dans un troisième temps, on proposera des algorithmes pour aborder ces problèmes. Ceux-ci seront ensuite implémentés et évalués expérimentalement et théoriquement.

On se concentrera tout particulièrement sur des problèmes de collaboration homme-robot et on étudiera dans quelle mesure l'approche proposée pourra s'appliquer sur les algorithmes de recherche d'équilibre de Nash sur lesquels nous travaillons actuellement [4].

Références

- [1] Sutton, R. S., Precup, D. and Singh S. (1999). Between MDPs and semi-MDPs : A framework for temporal abstraction in reinforcement learning. In *Artificial intelligence*, 112, 181–211, 1999.
- [2] Younes, H. L. S. and Simmons, R. G. (2004). Solving generalized semi-Markov decision processes using continuous phase-type distributions. In *Proceedings of the 19th National Conference on Artificial Intelligence (AAAI'04)*.
- [3] Lim, Z. L., Sun, L. and Hsu, D. (2011). Monte Carlo value iteration with macro-actions. In *Advances in Neural Information Processing Systems 24 (NIPS-2011)*.
- [4] You, Y., Thomas, V., Colas, F. and Buffet, O. (2021). Solving infinite-horizon Dec-POMDPs using Finite State Controllers within JESP. In *33rd International Conference on Tools with Artificial Intelligence (ICTAI-2021)*.

Sujet 3

Génération de chemin fluide pour robot aérien

Proposé par : Alexis Scheuer

Equipe : Larsen

Informations générales

Encadrant Alexis Scheuer
Adresse INRIA Nancy Grand Est, 615 rue du jardin botanique, 54600 Villers-lès-Nancy
Téléphone 03 83 59 20 84
Email Alexis.Scheuer@loria.fr
Bureau C121

Motivation

Un chemin est la partie géométrique d'un mouvement, sans aspect temporel. Générer des chemins fluides (C^2 ou plus), s'ils respectent les contraintes de mouvement d'un robot, permet un suivi par ce robot de l'ordre de dix fois plus précis que lorsque on utilise des chemins moins réguliers (C^1 ou moins).

Il est assez facile de déduire des contraintes de mouvement le mouvement associé à une commande en partant d'un état initial, et de trouver du coup l'état final de ce mouvement. À l'inverse, trouver la commande correspondant à un mouvement reliant un état initial à un état final donnés est beaucoup plus difficile (existence et unicité non garantie). Ce dernier problème est souvent traité par une recherche exhaustive, peu efficace.

Sujet

La bibliothèque Iseeml [1, 3] reste l'une des plus performante pour résoudre ce problème. Elle a été initialement prévue pour les chemins des robots de type voiture, mais elle peut être étendue à ceux de tous types de robots mobiles [2], en dimension deux (robot à roues ou à pattes) ou trois (avion ou hélicoptère).

La première extension a déjà été effectuée. De plus, à l'interface spécifique utilisant Qt (<http://www.qt.io>) s'ajoute maintenant une plateforme de comparaison de contrôleurs basée sur Qt et ROS (<http://www.ros.org>). Cette dernière permet de passer aisément de la simulation à un robot réel.

Cadre du travail

Le sujet de ce stage porte sur la seconde extension de la bibliothèque Iseeml. Les étapes pour atteindre cet objectif seront les suivantes :

1. il faudra tout d'abord se familiariser avec la méthode de génération de chemin, à l'aide des articles la décrivant (dont [1]), de la documentation générée par Doxygen pour la bibliothèque Iseeml [3] et du code associé, ainsi que des expérimentations avec les deux interfaces (Qt et Qt/ROS) ;
2. la méthode en dimension deux sera ensuite étendue à la troisième dimension, par une recherche (analytique plutôt qu'exhaustive) du plan le plus adapté pour chaque portion du chemin ;
3. la méthode obtenue sera validée à l'aide d'outils formels (comme Maple) et par des expérimentations.

Une connaissance de la programmation en C++ et des facilités en géométrie analytique sont préférables.

Références

- [1] Th. Fraichard and A. Scheuer. From Reeds and Shepp's to continuous-curvature paths. *IEEE Trans. on Robotics*, 20(6):1025-1035, décembre 2004.
- [2] A. Scheuer. La conduite des robots. In *Sciences & Info. prépa Hors-Série numéro 1 : Transformations et fonctions, Représentation et Analyse* (ISSN 1291-4932, ISBN 2-7298-0446-3), pp. 148-161. POLE / Ellipses, Paris, octobre 2000.
- [3] A. Scheuer. Documentation of ISeeML (Introducing a Smooth, Efficient and Easy-to-use Motion Library, <http://iseeml.loria.fr>). Technical Report RT-0396, 134 pages, Inria, Nancy (FR), octobre 2010.

Équipe MFX

<http://mfx.loria.fr/>

Sujet 1

Star-shaped distance interpolation

Proposé par : Jonàs Martínez

Equipe : MFX

Informations générales

Encadrants Jonàs Martínez
Adresse LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Téléphone 03 54 95 86 09
E-mail <mailto:jonas.martinez-bayona@inria.fr>
Bureau C 134

Motivations

A tessellation of a flat surface is the tiling of a plane using one or more geometric shapes, called tiles, with no overlaps and no gaps. *Periodic tilings* are made of periodically arranged tiles. Tessellations have found applications in multiple applications, such as manufacturing, art, and architecture. In some cases, it is interesting to be able to make a gradual transition between different types of tile geometries (see Figure 6).



FIGURE 6 – Progressive gradation between periodic tilings. Designed by Edmund Harriss
<https://maxwelldemon.com/>

In [1] we introduced a novel type of periodic tilings, allowing for robust interpolation and gradation of tile geometries. We obtained these tilings as Voronoi diagrams of regular lattices under star-shaped distance functions that generalize the Euclidean metric. A wide variety of geometries emerge when combining different distances and lattice types (see Figure 7). In addition, the distance functions can be smoothly varied across the underlying lattice. The paper can be downloaded from <https://hal.inria.fr/hal-02118846/file/paper.pdf>. An C++ implementation of the technique can be found at <https://github.com/mfx-inria/starshaped2d>.

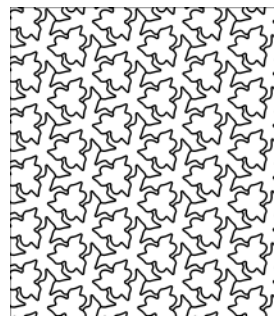
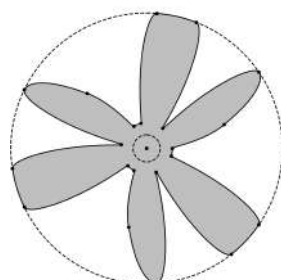


FIGURE 7 – Left : star-shaped distance with three-fold rotational symmetry. Right : resulting Voronoi diagram when using a honeycomb lattice.

Sujet

While more complex schemes are possible, in [1] we restricted considerations to simple linear interpolation between the star-shaped distances. Figure 8 shows an example of the result of applying linear interpolation between two different star-shaped distances.



FIGURE 8 – From left to right, linear interpolation between two star-shaped distances.

However, we observed that in some cases linear interpolation of the distance does not necessarily imply equally continuous changes in the tile geometry (see Figure 9), which is undesirable.

Travail demandé

The main objective of the internship is to study and develop better interpolation techniques to tackle the aforementioned issue. As a starting point, the simple case of interpolation between two tilings will be analyzed. We will explore ways to define interpolation paths that minimize, for instance, the changes in geometrical and/or mechanical properties.

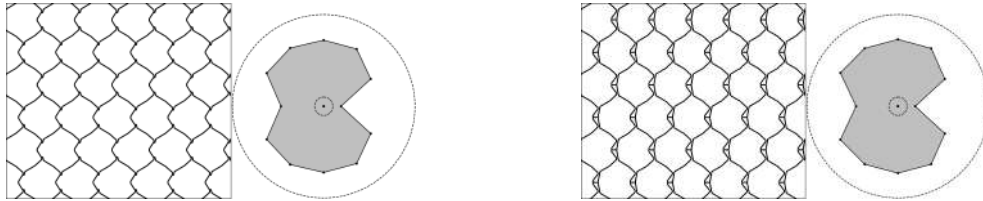


FIGURE 9 – Left : initial tiling. Right : small linear perturbation of the initial tiling distance. One can observe that this small perturbation translates to a discontinuous change of the geometry.

Cadre du travail

Le stage se déroulera à Inria Nancy (au LORIA), dans l'équipe MFX.

Références

- [1] *Star-shaped Metrics for Mechanical Metamaterial Design* ACM Transactions on Graphics, 2019.
<https://hal.inria.fr/hal-02118846>

Sujet 2

Propagation de son dans l'espace à l'aide de solutions fondamentales

Proposé par : Camille SCHRECK

Equipe : MFX

Informations générales

Adresse LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Email camille.schreck@inria.fr
Bureau C132

Motivations

Sujet

Grâce aux récentes avancées en Informatique Graphique, il est maintenant possible de créer des environnements 3D visuellement très réalistes. Mais le réalisme d'une scène virtuelle ne repose pas exclusivement sur la vue. Cependant la génération de données sensorielles relatives aux autres sens n'a été pour l'instant que peu étudiée. Notamment, le son est souvent indispensable pour obtenir un environnement réaliste. Malgré un intérêt grandissant pour le sujet (voir [1] par exemple), créer automatiquement les sons correspondants à une animation et les propager dans une scène virtuelle reste un défi.

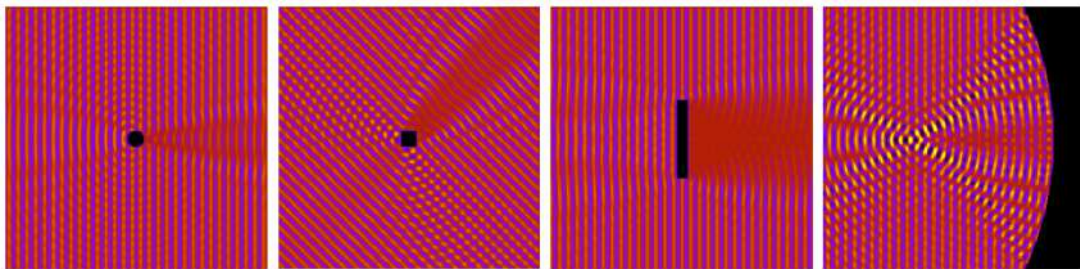


FIGURE 10 – Une vague plane dispersée par différents obstacles.

L'objectif de ce stage est d'utiliser des solutions fondamentales ([2]) afin de propager le son dans un environnement virtuel 3D, de façon similaire à ce qui a été fait pour les vagues surfaciques dans [3]. Le stagiaire devra adapter la méthode 2D proposée par [3] pour la 3D en tenant compte des différences entre les ondes surfaciques sur un liquide et les ondes acoustiques. Une interface permettant de traiter les entrées/sorties audios.

Le stagiaire doit avoir de bonnes compétences en C++ et mathématiques et être capable de communiquer en anglais.

Cadre du travail

Le stage se déroulera au Loria de Nancy dans l'équipe MFX.

Références

- [1] Liu, Shiguang and Manocha, Dinesh. *Sound synthesis, propagation, and rendering : a survey*. arXiv preprint arXiv :2011.05538, 2020.
- [2] Lee, Seongkyu. *The use of equivalent source method in computational acoustics*. Journal of Computational Acoustics, 2017.
- [3] Schreck, Camille and Hafner, Christian and Wojtan, Chris. *Fundamental Solutions for Water Wave Animation*. ACM Transactions on Graphics - Siggraph 2019.
- [4] Mehra, Ravish and Raghuvanshi, Nikunj and Antani, Lakulish and Chandak, Anish and Curtis, Sean and Manocha, Dinesh. *Wave-Based Sound Propagation in Large Open Scenes Using an Equivalent Source Formulation*. ACM Transactions on Graphics.

Équipe Mocqua

<http://mocqua.loria.fr/>

Sujet 1

A quantum stream language for polynomial time computation

Proposé par : Romain Péchoux et Vladimir Zamdzhiev

Equipe : Mocqua

Informations générales

Encadrants	Romain Péchoux, Vladimir Zamdzhiev
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Téléphone	03 54 95 84 06
Email	romain.pechoux@loria.fr, vladimir.zamdzhiev@inria.fr
Bureau	B 214

Motivations

In order to certify the correct execution of a program, it is important to be able to guarantee time and space bounds on its execution. Among the most popular techniques in the field are type systems using restrictions of linear logic but also methods based on semantics and amortized resource analysis. These methods have focused on the analysis of programs on discrete data and have thus captured the mainstream complexity classes (P, PSPACE, etc.). However, programs that manipulate non-discrete data (e.g. real numbers represented as infinite streams of digits, infinite superpositions of quantum states, etc.) have not been heavily studied by these methods and we propose to fill this gap.

Sujet

A very natural way to represent reals in a programming language is to encode them as streams (infinite sequences). The paper [1] provided a first characterisation of polynomial time on reals based on a linear type system and using the stream language introduced in [2]. The programmer can then write programs with coinductive data (which does not normalise by nature) while preserving polynomial time normalisation on inductive data. The characterization provided in [1] has some drawbacks, however. 1) the language is non-standard 2) stream duplication is forbidden; this makes the implementation of programs on the reals complex.

The objective of this internship is to develop a programming language that does not have these drawbacks. To this end, we will consider the DLAL (Dual Light Affine Logic) system [3] and try to extend it to streams while preserving the polynomial normalisation by combining it with systems allowing to ensure the productivity of programming languages [4, 5].

A direct application of these results can be made to quantum programming languages. For example, one can add coinductive data to the quantum language of [6]. Another application is to represent infinite superpositions of quantum states within a quantum programming language which is the final objective of this internship.

Objectives :

- Study the light affine logic based typing discipline.
- Combine this typing discipline with productivity methods.
- Obtain a characterization of polynomial time over stream data and characterize polynomial time over the real numbers (seen as a stream of digits).
- Study the case of a quantum programming language over an infinite dimensional state space.

Cadre du travail

The internship will take place in the loria laboratory within the Inria project mocqua.

Références

- [1] E. Hainry, D. Mazza, R. Péchoux. *Polynomial time over the reals with parsimony*. FLOPS 2020.
[2] D. Mazza, K. Terui. *Parsimonious Types and Non-uniform Computation*. ICALP(2) 2015.

- [3] P. Baillot, Kazushige Terui. *Light types for polynomial time computation in lambda calculus*. *Information and Computation* 207(1) : 41-62, 2009.
- [4] R. Clouston, A. Bizjak, H. B. Grathwohl, L. Birkedal. *Programming and Reasoning with Guarded Recursion for Coinductive Types*. FoSSaCS 2015.
- [5] Robert Atkey, Conor McBride. *Productive coprogramming with guarded recursion*. ICFP 2013.
- [6] Romain Péchoux, Simon Perdrix, Mathys Rennela, Vladimir Zamdzhiev. *Quantum Programming with Inductive Datatypes : Causality and Affine Type Theory*. FoSSaCS 2020.

Équipe Mosel / VeriDis

The Mosel/VeriDis team works on modeling notations and verification techniques for different kinds of software-intensive systems, including distributed and cyber-physical systems. Verification techniques of interest include model checking and synthesis, as well as automatic and interactive theorem proving.

- <https://team.inria.fr/veridis/>
- <https://mosel.loria.fr>

Sujet 1

Verifying cybersecurity properties using formal methods

Proposé par : Étienne André and Engel Lefauchaux

Equipe : Mosel / VeriDis

Informations générales

Encadrants	Étienne André, Engel Lefauchaux	
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy	
Téléphone	none	
Email	Etienne.Andre@loria.fr	Engel.Lefauchaux@loria.fr
Bureau	B 210	

Context

The pervasiveness of cyber-physical systems is highly increasing, raising many safety and security concerns. Formal methods aim at tackling those problems through the verification of formal properties on a model abstracting the real system.

A well-known formal model to reason about timed systems is *timed automata* [1], an extension of finite-state automata with continuous clocks measuring time. Timed automata have been extensively used to verify safety properties, but not so much security properties, with some exceptions [4, 3, 2].

Problem

The objective of the internship will be to study *opacity* properties for timed automata : given a timed automaton, is it possible to guarantee that the system is *opaque*, i.e., that some predefined secret is not leaked during the execution ?

A focus will be given on *control* (as in, e.g., [3]), i.e., to restrain some of the system behaviors (typically actions) so that it becomes opaque.

This research has theoretical aspects, as well as applications to cybersecurity.

Supervision

The internship is supervised by Étienne André and Engel Lefauchaux.

Depending on the applicant's wishes and on the opportunities of funding, a PhD thesis starting in September 2022 can be envisioned.

Références

- [1] Rajeev Alur, David L. Dill. A theory of timed automata. Theoretical Computer Science 126(2), pp. 183-235. Elsevier 1994.
- [2] Étienne André, Jun Sun. Parametric Timed Model Checking for Guaranteeing Timed Opacity. In : Proc. 17th Intl. Symp. Automated Technology for Verification and Analysis (ATVA 2019). LNCS 11781, pp. 115-130. Springer 2019.
- [3] Gilles Benattar, Franck Cassez, Didier Lime, Olivier H. Roux. Control and synthesis of non-interferent timed systems. Intl. J. Control 88(2), pp. 217-236, 2015.
- [4] Franck Cassez. The Dark Side of Timed Opacity. Proc. 3rd Intl. Conf. Advances in Information Security and Assurance (ISA 2009). LNCS 5576, pp. 21-30. Springer 2009.

Sujet 2

Efficient runtime verification of cyber-physical systems using formal methods

Proposé par : Étienne André

Equipe : Mosel / VeriDis

Informations générales

Encadrants	Étienne André, Mohammed Foughali	
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy	
Téléphone	none	
Email	Etienne.Andre@loria.fr	foughali@irif.fr
Bureau	B 210	

Context

The pervasiveness of cyber-physical systems is highly increasing, accompanied by associated safety concerns. Formal verification techniques usually require a (white-box) model, which is not often available, because some components are black-box, or because the entire system has no formal model. In addition, formal verification techniques for cyber-physical systems are often subject to state space explosion, often preventing a satisfactory scalability. Therefore, monitoring, as a lightweight yet feasible verification technique, can bring practical results of high importance for larger models.

Monitoring aims at analyzing the log of a concrete system, so as to deduce whether a specification (e.g., a safety property) is violated. Monitoring can be done offline (i.e., after the system execution, assuming the knowledge of the entire log), or online (at runtime, assuming a partial log).

More generally, runtime verification generalizes online monitoring by allowing an analysis in real-time.

Problem

The objective of the internship is to propose new methods for scalable runtime verification of complex cyber-physical systems.

The internship will consist in understanding the state of the art in monitoring using formal methods (e.g., [3, 1, 4, 2]), and then to propose new original techniques, addressing notably scalability.

Supervision

The internship is supervised by Étienne André and Mohammed Foughali.

Depending on the applicant's wishes and on the opportunities of funding, a PhD thesis starting in September 2022 can be envisioned.

Références

- [1] Mohammed Foughali, Saddek Bensalem, Jacques Combaz, Félix Ingrand. Runtime Verification of Timed Properties in Autonomous Robots. In : Proc. 18th ACM/IEEE Intl. Conf. Formal Methods and Models for System Design (MEMOCODE 2020). IEEE 2020.
- [2] Konstantinos Mamouras, Agnishom Chattopadhyay, Zhifu Wang. A Compositional Framework for Quantitative Online Monitoring over Continuous-Time Signals. In : Proc. 21st Intl. Conf. Runtime Verification (RV 2021). LNCS 12974, pp. 142-163, Springer 2021.
- [3] Masaki Waga, Étienne André, Ichiro Hasuo. Symbolic Monitoring against Specifications Parametric in Time and Data. In : I. Dillig, S. Tasiran (eds.), Proc. 31st Intl. Conf. Computer-Aided Verification (CAV 2019). LNCS 11561, pp. 520-539, Springer 2019.
- [4] Masaki Waga, Étienne André, Ichiro Hasuo. Model-bounded monitoring of hybrid systems. In : M. Maggio, J. Weimer, M. Al Farque, M. Oishi (eds.), Proc. 12th ACM/IEEE Intl. Conf. Cyber-Physical Systems (ICCPS). ACM 2021.

Sujet 3

Minimising the cost of diagnosing systems

Proposé par : Engel Lefauchaux

Equipe : Mosel / VeriDis

Informations générales

Encadrants Engel Lefauchaux
Adresse LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Téléphone 03 54 95 84 06
Email elefauch@inria.fr
Bureau B 235

Motivations

Cyber-physical systems are now omnipresent, and in particular can be found in many critical systems. In such systems, failures can lead to important human or economic losses. As such it is important to develop methods to accurately check that real systems correctly satisfy their specification.

Relying on formal methods is one such method, where a real system and a set of specifications are first translated into a mathematical model and formal properties which are then studied.

One traditional safety property that is studied through formal methods is *diagnosis* [1, 2] which aims at detecting specific, faulty behaviours within a model. Diagnosis of models including stochasticity (such as *Markovian models* [3, 4, 5]) is however still limited.

Subject

The goal of this internship will be to study the diagnosability of stochastic models.

We will be particularly interested in optimising the cost (in terms of sensors, actions, undetected faults, ...) of any action aimed at better detecting failures [6].

This research mainly has theoretical aspects, though it also has applications in safety verification.

Références

- [1] M. Sampath, R. Sengupta, S. Lafortune, K. Sinnamohideen and D. Teneketzis. Diagnosability of discrete-event systems. *IEEE Transactions on Automatic Control*, vol. 40, 1995.
- [2] Nathalie Bertrand, Serge Haddad and Engel Lefauchaux. A tale of two diagnoses in probabilistic systems. *Information & Computation*, 2019.
- [3] Wai-Ki Ching and Michael K. Ng. *Markov Chains : Models, Algorithms and Applications*. International Series in Operations Research & Management Science, Springer 2006.
- [4] Martin L. Puterman. *Markov Decision Processes : Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc., 1994.
- [5] <http://www.pomdp.org/index.html>
- [6] F. Cassez, S. Tripakis and K. Altisen. Synthesis Of Optimal-Cost Dynamic Observers for Fault Diagnosis of Discrete-Event Systems. *TASE'07*, 2007.

Sujet 4

Distributed PlusCal

Proposé par : Horatiu Cirstea et Stephan Merz

Equipe : Mosel / VeriDis

Administrative information

Supervisors Horatiu Cirstea, Stephan Merz
Address LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Telephone 03 54 95 84 78
Email horatiu.cirstea@loria.fr, stephan.merz@loria.fr
Office LORIA, B 208

Motivations

Formal verification techniques such as model checking or theorem proving help ensure correctness properties of algorithms and software systems. They can be applied at high levels of abstraction, allowing designers to identify errors that would be very costly to correct during later development stages.

However, languages used for formal modeling sometimes present a substantial gap from languages used in software development. For example, specifications written in TLA⁺ [2] are formulas that mix mathematical set theory and temporal logic, which may intimidate users. It is therefore desirable to introduce front-end languages that are more familiar to system designers, while still enabling the use of formal verification techniques.

Subject

PlusCal [1] is an algorithmic language intended for the description of distributed algorithms and systems at a high level of abstraction, providing the “look and feel” of imperative pseudo-code. The PlusCal translator generates a TLA⁺ specification from a PlusCal algorithm, and this specification can be verified using the TLC model checker or the TLAPS proof assistant. In particular, the design of PlusCal keeps the translation to TLA⁺ deliberately simple for maintaining a close correspondence with the original algorithm.

In order to achieve this simplicity, PlusCal imposes a number of restrictions. In particular, it only offers top-level parallel processes, making it difficult to model distributed systems where each node consists of several threads executing in parallel. In previous work, we proposed an extension called Distributed PlusCal [3] that removes this limitation. Distributed PlusCal also introduces primitives that support inter-process communication through message channels.

The objective of the proposed internship is to consolidate and extend Distributed PlusCal. As a first step, the student working on the subject will get familiar with the language by modeling additional algorithms and understanding how the existing translation handles features such as procedure calls or non-determinism in conjunction with sub-processes. The second objective is to add more flexibility to the implementation of inter-process communication : instead of using predefined FIFO and unordered channels as in the current version of the translator, it would be desirable to define communication semantics through standard TLA⁺ modules, where their properties can also be proved, and embed those in Distributed PlusCal. A third objective is to support processes having a parameterizable number of identical sub-processes, for example for modeling a variable number of processing cores. These extensions should be designed such that they maintain the close correspondence between the generated TLA⁺ specification and the PlusCal algorithm.

Prerequisites

Candidates should be familiar with formal methods and understand fundamental compilation techniques. Some experience in Java programming is expected. Previous knowledge of PlusCal/TLA⁺ and/or of distributed algorithms is a plus, but is not required for working on the subject.

Références

- [1] Leslie Lamport. The PlusCal Algorithm Language. In : 6th Intl. Coll. Theoretical Aspects of Computing (ICTAC 2009). LNCS 5684, pp. 36-60. Springer, 2009.
- [2] Leslie Lamport. Specifying Systems. Addison-Wesley (Boston, Mass.), 2002.
<http://lamport.azurewebsites.net/tla/tla.html>.
- [3] Heba Al Kayed, Horatiu Cirstea, Stephan Merz. An Extension of PlusCal for Modeling Distributed Algorithms. In : TLA⁺ Community Meeting 2020.

Sujet 5

Formal Models Validation – Validation of JeB

Proposé par : Jean-Pierre Jacquot

Equipe : Mosel / VeriDis

Informations générales

Encadrants	Jean-Pierre Jacquot
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Téléphone	03 72 74 50 92
Email	Jean-Pierre.Jacquot@loria.fr
Bureau	B 220 ou HP 705

Motivations

With the first computer programs came the first bugs. Building *correct* programs is still a challenge [2]. Formal methods answer this issue.

The *use* of formal methods is hampered by a serious obstacle : how to ensure that models developed and proved through costly efforts, are adequate models for the software to develop? This is the classic scientific problem of *model validation* [4, 6].

During his doctoral work [5], F. Yang has developed a tool, JeB, which allows us to execute Event-B specifications [3]. The basic idea is to translate the Event-B text (a mathematical model), into a JavaScript program (an executable model). The base translation strategy consists in a syntactic transformation of Event-B operators into function calls from a set-theoretical library.

Although the basic strategy was successful on many models, it failed on certain expressions. In his Master thesis, Y. Abegaz [1] has implemented extensions in JeB.

Subject

The general theme of the subject is *the validation of the validating tool*. Two kinds of work need to be done : **Semi-formal analysis.** The extension of JeB implements some transformations of the the Event-B structure. We need to get strong arguments to ensure that the transformations and their implementation is (formally) sound.

Practical validation. The extension of JeB was motivated by the validation of an Event-B model of the rules of checkers (the board game). While this particular model is too complicated to be used as a validation test-bed, other simpler board games such as *tic-tac-toe* or *connect-4* (*puissance 4*) can be used. An Event-B model of such a board game must be developed and used with JeB to verify that it is a good model of the game rules.

The work will have three main phases :

1. Getting acquainted with the tools (Rodin and Eclipse) and the issue of the correctness of tools used in formal methods (bibliographic work).
2. Analysing the transformations introduced by Abegaz to propose arguments for their correctness.
3. Modeling the rules of a simple board games in Event-B and validating the model with JeB.

Cadre du travail

The work will be executed in the Mosel-Veridis team. It requires a taste for using formal methods. Some knowledge about Java programming and Eclipse would ease the start of the work.

Références

- [1] Yeabsira Mintesnot Abegaz. *Validation of Formal Methods Event B*. Master thesis. Université de Lorraine. 2021.
- [2] Jean-Raymond Abrial. Formal methods in industry : achievements, problems, future. In *Proceedings of the 28th international conference on Software engineering*, ICSE '06, pages 761–768, New York, USA, 2006. ACM.
- [3] Jean-Raymond Abrial. *Modeling in Event-B : System and Software Engineering*. Cambridge University Press, 2010.
- [4] Atif Mashkoo and Jean-Pierre Jacquot. Validation of formal specifications through transformation and animation. *Requirements Engineering*, 22(4) :433–451, Nov 2017.
- [5] Faqing Yang. *A Simulation Framework for the Validation of Event-B Specifications*. Thèse, Université de Lorraine, November 2013.
- [6] Faqing Yang, Jean-Pierre Jacquot, and Jeanine Souquière. The Case for Using Simulation to Validate Event-B Specifications. In Pornsiri Muenchaisri Karl Leung, editor, *APSEC2012 - The 19th Asia-Pacific Software Engineering Conference*, pages 85–90, Hongkong, China, December 2012. The University of Hong Kong, IEEE.

Sujet 6

Verifying the theory behind conflict-based quantifier instantiation for SMT solvers

Proposé par : Sophie Turrett

Equipe : Mosel / VeriDis

Overall informations

Encadrants	Sophie Turrett
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Téléphone	03.54.95.84.78
Email	mailto:sophie.turrett@loria.fr
Bureau	B 208

Context

SMT solving is a well-known and very successful technique to check the satisfiability of formulas in first-order logic with theories. It has many applications, in particular in the domain of software verification. One thing that is notoriously difficult for SMT solvers is to find appropriate instances of quantified formulas to continue the search done by the ground SMT solver when it gets stuck.

Problem

There are several instantiation techniques that do not all offer the same theoretical guarantees. In that respect, conflict-based instantiation techniques are particularly interesting because they offer the guarantee to find an instance that contradicts the model currently proposed by the ground SMT solver, thus sparing a lot of exploration effort to the solver.

The theoretical problem behind conflict-based instantiation is called E-ground (dis)unification. It consists in finding how to instantiate a set of equations and disequations so that they are consequences of a given set of ground equations and disequations.

Subject

In the context of my work on quantifier instantiation, I would like the student to verify using Isabelle/HOL that the main theorems behind the existence of solutions to the E-ground (dis)unification problem is correct. The student carrying out the project will have the pen-and-paper proof as a guide to formalise this result.

Work environment

The student taking up this subject should be interested in automated reasoning and formal verification. Prior experience with an interactive theorem prover (Isabelle, Coq, Lean ...), ideally Isabelle/HOL would be an advantage. Knowledge of first-order logic is expected.

Références

- [1] Clark Barrett and Cesare Tinelli. Satisfiability Modulo Theories. In : Handbook of Model Checking, pp. 305-343. Springer, 2018.
- [2] Haniel Barbosa, Pascal Fontaine and Andrew Reynolds. Congruence Closure with Free Variables. TACAS, 2017.
- [3] Sophie Turrett, Pascal Fontaine, Daniel El Ouraoui, Haniel Barbosa. Lifting Congruence Closure with Free Variables to λ -free Higher-order Logic via SAT Encoding. SMT workshop, 2020.

Sujet 7

Evaluating SMT Encodings of Proof Obligations

Proposé par : Sophie Tourret and Stephan Merz

Equipe : Mosel / VeriDis

Overall informations

Encadrants	Sophie Tourret, Stephan Merz
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Téléphone	03.54.95.84.78
Email	mailto:sophie.tourret@loria.fr mailto:stephan.merz@loria.fr
Bureau	B 208

Context

The B method [1, 2] is a popular formalism for developing safety-critical systems, including control software for railbound transportation such as trains or subways. Its use relies on discharging proof obligations such as ensuring the preservation of invariants of a B model or verifying that a lower-level model is a correct refinement of a higher-level one. Tool support for B comes in the form of the interactive platforms Atelier B [5] and Rodin [6], which integrate different automatic back-end provers.

Problem

The degree of automation for discharging proof obligations depends to a significant extent on the strategy used for encoding them into the input language of back-end provers. SMT solvers [3] are among the best all-round proof tools for first-order logic over interpreted theories such as arithmetic. The standard encoding of the B language for SMT solvers relies on representing sets and relations as uninterpreted predicates and uses extensive rewriting for simplifying proof obligations.

Subject

In the context of a collaborative project with Clearys, the company maintaining the Atelier B platform, we wish to investigate the effect of encodings of B proof obligations for different back-end provers, including reasoning engines for higher-order logic such as Zipperposition [4] and SAT solvers. The student carrying out the proposed project will have access to a large base of proof obligations generated from Atelier B and will be able to choose a back-end solver. The objective is to identify a class of proof obligations for which the solver is successful, via a suitable encoding.

Work environment

The student taking up this subject should be interested in the application of automated reasoning tools to solving practical problems arising during system development. Programming experience, preferably in a functional programming language such as OCaml, and familiarity with parser generators is expected. Some experience with automatic solvers would be an advantage.

In principle, funding for a PhD thesis for the development of higher-order encoding and reasoning techniques for SMT solvers is available.

Références

- [1] Jean-Raymond Abrial. *The B-Book : Assigning Programs to Meanings*. Cambridge University Press, 1996.
- [2] Jean-Raymond Abrial. *Modeling in Event-B*. Cambridge University Press, 2010.
- [3] Clark Barrett and Cesare Tinelli. *Satisfiability Modulo Theories*. In : *Handbook of Model Checking*, pp. 305-343. Springer, 2018.
- [4] Petar Vukmirovic, Alexander Bentkamp, Jasmin Blanchette, Simon Cruanes, Visa Nummelin, Sophie Tourret. *Making Higher-Order Superposition Work*. In : *Proc. 28th Intl. Conf. Automated Deduction (CADE 2021)*. LNCS 12699, pp. 415-432. Springer, 2021.
- [5] <https://www.atelierb.eu>
- [6] <http://www.event-b.org/install.html>

Sujet 8

Efficient Proof Strategies for Cyclic Induction Reasoning

Proposé par : Sorin Stratulat

Equipe : Mosel / VeriDis

Overall information

Encadrants	Sorin Stratulat
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Téléphone	03.54.95.85.95
Email	sorin.stratulat@loria.fr
Bureau	B 215

Motivations

Induction reasoning is very helpful to prove properties about inductively defined data structures like the naturals and the lists. Implicit induction uses reductive reasoning techniques as rewriting for automatically proving properties about conditional specifications.

The SPIKE prover uses implicit induction to deal with many non-trivial applications [4]. An advantage of using SPIKE is that the generated proofs are sound by construction, i.e., no extra check of their soundness is required.

Cyclic induction is a more recent proof technique that can be applied for proof derivations involving repetitive steps. Since repetition may be an infinite process, cyclic pre-proofs help to give a finite representation by firstly stopping the development of repetitive nodes in the derivation graph, then linking them with the nodes that they repeat. The linking step may introduce cycles in the graph. The soundness of the technique consists in building some induction argument while analysing the formulas occurring along the cycles in the pre-proofs. In the current implementations of this approach, the soundness is checked anytime a new cycle is detected, which represents a costly and time-consuming operation.

E-CYCLIST [5] is a cyclic prover that can prove properties about specifications built in the frame of first-order logic with inductive definitions (FOL_{ID}) by implementing the sequent-based CLKID^ω inference system [1]. It also implements a more effective soundness checking approach [2, 3], for which (a subset of) CLKID^ω proofs are sound if some ordering and derivability constraints are satisfied. The ordering constraints are very similar to those used to validate implicit induction proofs. In spite of this, E-CYCLIST is still a prototype that can build proofs only for toy examples, of small sizes (tens of nodes).

Subject

Our main objective is to make E-CYCLIST deal with real-size applications. Therefore, it is compulsory to speed up its proof process. One solution is to build cyclic proofs that are sound by construction, by using proof strategies inspired from SPIKE. To reach this goal, the following tasks are given :

- become familiar with implicit induction reasoning and SPIKE;
- gain expertise on how SPIKE builds implicit induction proofs;
- become familiar with cyclic reasoning and E-CYCLIST;
- understand how E-CYCLIST builds cyclic proofs;
- propose and implement “sound by construction” proof strategies for E-CYCLIST;
- propose a set of non-trivial tests and examples that validate the previous step.

Work environment

SPKE and E-CYCLIST are developed in Ocaml (www.ocaml.org). Hence, a decent level of knowledge in Ocaml programming is required.

Références

- [1] J. Brotherston and A. Simpson. Sequent calculi for induction and infinite descent. *J. Logic and Computation*, 21(6), pp. 1177–1216, 2011.
- [2] S. Stratulat. Cyclic proofs with ordering constraints. In R. A. Schmidt and C. Nalon, editors, *Tableaux 2017 (26th Intl. Conf. Automated Reasoning with Analytic Tableaux and Related Methods)*, volume 10501 of *LNAI*, pp. 311–327. Springer, 2017.
- [3] S. Stratulat. Validating back-links of FOL_{ID} cyclic pre-proofs. In S. Berardi and S. van Bakel, editors, *CL&C'18 (7th Intl. Ws. Classical Logic and Computation)*, number 281 in EPTCS, pp. 39–53, 2018.
- [4] S. Stratulat. SPIKE, an automatic theorem prover – revisited. In *SYNASC 2020 : Proc. 22nd Intl. Symp. Symbolic and Numeric Algorithms for Scientific Computing*, IEEE Computer Society, pp. 93–96, 2020.
- [5] S. Stratulat. E-CYCLIST : Implementation of an Efficient Validation of FOL_{ID} Cyclic Induction Reasoning. In T. Kutsia, editors, *9th Intl. Symp. Symbolic Computation in Software Science*, number 342 in EPTCS, 2021.

Équipe MULTISPEECH

<https://team.inria.fr/multispeech/fr/>

Sujet 1

Semantic information in a speech recognition system : does the past help the present?

Proposé par : Irina ILLINA and Dominique FOHR

Equipe : MULTISPEECH

Informations générales

Encadrants	Irina ILLINA, Dominique FOHR	
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy	
Téléphone	03 54 95 84 90	
Email	illina@loria.fr	dominique.fohr@loria.fr
Bureau	C 147	

Motivations

Semantic and thematic spaces are vector spaces used for the representation of words, sentences or textual documents. The corresponding models and methods have a long history in the field of computational linguistics and natural language processing. Almost all models rely on the hypothesis of statistical semantics that states that : statistical schemes of appearance of words (context of a word) can be used to describe the underlying semantics. The most used method to learn these representations is to predict a word using the context in which this word appears [3], [4], and this can be realized with neural networks. These representations have proved their effectiveness for a range of natural language processing tasks. In particular, Mikolov's Skip-gram and CBOV models [3] and BERT model [1] have become very popular because of their ability to process large amounts of unstructured text data with reduced computing costs. The efficiency and the semantic properties of these representations motivate us to explore these semantic representations for our speech recognition system.

Robust automatic speech recognition (ASR) is always a very ambitious goal. Despite constant efforts and some dramatic advances, the ability of a machine to recognize the speech is still far from equaling that of the human being. Current ASR systems see their performance significantly decrease when the conditions under which they were trained and those in which they are used differ. The causes of variability may be related to the acoustic environment, sound capture equipment, microphone change, etc.

Sujet

Our speech recognition (ASR) system [5] is supplemented by a *semantic analysis* for detecting the words of the processed sentence that *could have been misrecognized* and for finding words having a *similar pronunciation and matching better the context* [2]. For example, the sentence « Silvio Berlusconi, prince de Milan » can be recognized by the speech recognition system as : « Silvio Berlusconi, prince de mille ans ». A good semantic context representation of the sentence could help to find and correct this error. This semantic analysis *re-evaluates (re-scores) the N-best transcription hypotheses* and can be seen as a form of dynamic adaptation in the case of noisy speech data. A semantic analysis is performed in combining predictive representations using continuous vectors. All our models are based on the powerful technologies of DNN. We use for this the BERT model. The semantic analysis contains the following two modules : the semantic analysis module and the module for re-evaluating sentence hypotheses (rescoring) taking into account the semantic information.

The semantic module improves significantly the performance of speech recognition system. But we would like to go beyond the semantic information of the current sentence. Indeed, sometimes the previous sentences could help to understand and to recognize the current sentence. The Master internship will be devoted to the innovative study of the *taking into account the past recognized sentences to improve the recognition of the current sentence*. Research will be conducted on the combination of semantic information from one or several past sentences with semantic information from current sentence to improve the speech recognition. As deep neural networks (DNNs) can model complex functions and get outstanding performance, they will be used in all our modeling. The performance of the different modules will be evaluated on artificially noisy speech data.

Required skills : background in statistics, natural language processing and computer program skills (Perl, Python).

Candidates should email a detailed CV with diploma.

Références

- [1] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. BERT : Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies, Volume 1 (Long and Short Papers), 2019.
- [2] S. Level, I. Illina, D. Fohr Introduction of semantic model to help speech recognition. In Proceedings of the International Conference on TEXT, SPEECH and DIALOGUE, 2020.
- [3] T. Mikolov, I. Sutskever, T. Chen, G. S. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. In Advances in Neural Information Processing Systems, pp. 3111-3119, 2013.
- [4] J. Pennington, R. Socher, and C. Manning. Glove : Global vectors for word representation. In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 1532-1543, 2014.
- [5] D. Povey, A. Ghoshal, G. Boulianne, L. Burget, O. Glembek, N. Goel, M. Hannemann, P. Motl?cek, Y. Qian, Y. Schwarz, J. Silovsky, G. Stemmer, K. Vesely. The Kaldi Speech Recognition Toolkit. In Proceedings of the ASRU, 2011.

Sujet 2

Mitigating bias in texts using word removal and explanations

Proposé par : Miguel Couceiro, Maxime Amblard and Irina Illina

Equipe : MULTISPEECH

Informations générales

Encadrants Miguel Couceiro, Maxime Amblard and Irina Illina
Adresse LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Téléphone 03 54 95 84 90
Email miguel.couceiro@loria.fr & maxime.amblard@loria.fr & illina@loria.fr

Motivations

NLP can be used to solve problems like information consumption, information graph completion, smart meta marking, query generation, query answering, idea summarization, conversational assistants for students, descriptive response assessment, and many other activities with social and economical impact on society [1], [3]. Applying modern NLP for real world applications demands interpretability, to make the system more transparent, explainable and robust. One of the major bias based problem in NLP is identification of contextual hate/fear speech. The identification of hate/fear speech is part of a larger campaign to reduce the damage caused by oppressive and abusive words [5]. With more data and more advanced algorithms, performance has increased [2], but the relative scarcity of hate/fear speech necessitates keyword sampling [7] or sampling from areas with exceptionally high concentrations of hate/fear speech. As a result, modern text classifiers fail to learn a model of hate/fear speech that applies to real-world scenarios [10]. Neural text classifiers achieve state-of-the-art performance in hate/fear speech detection, but are uninterruptible and can break when presented with unexpected inputs [6], [8]. As a result, contextualizing a model's treatment of identifier terms is challenging.

Sujet

In this work, we would be focusing on three prime objectives :

- **Objective 1** : Firstly, we would be focusing on doing literature review and looking at the existing work on hate or fear speech analysis. We would be looking at the existing the hate or fear speech dataset and the classes present in the dataset. In case the existing datasets don't fulfil our requirements, we can also collect new dataset from various social media platform if needed.
- **Objective 2** : Once we have the dataset finalised, we would work on our second objective of interpretability and explainability of text classification models for hate or fear speech detection. We would utilize various DNN language models such as LSTMs, CNNs and attention-based large pre-trained language models. We first determine a model's sensitivity towards group identifiers by examining the models themselves. Linear classifiers can be examined in terms of their most highly-weighted features. We apply explanation algorithm for this task of extracting these information from the fine-tuned methods. The Sampling and Occlusion (SOC) hierarchical description algorithm, which estimates model agnostic, context-independent post-hoc function value, is how we address this issue [4]. We would also use the heat map representation on the text inputs to learn at which part of the hate/fear speech the model is giving attention at the time of classification. We are willing to apply this approach on manually collected and annotated data on Hate/Fear speech data from Twitter/Reddit using APIs.
- **Objective 3** : Our final objective will be analysing the bias in the hate/fear speech prediction. We will use regularization techniques to reduce the attention given to group identifiers and heightens the importance of the more generalizable features of hate/fear speech, such as dehumanizing, fear and insulting language. Regularization lowers the false positive rate when maintaining or improving in-domain, out-of-sample classification accuracy. We will use simplest approach by removing group identifiers altogether words from both training and testing sentences. Other approach we can develop is using the SOC explanations extracted earlier.

Depending on the progression in these objectives and the availability of suitable datasets, we may also explore other settings, e.g., in fear and hate speech mitigation [9].

Required skills : background in statistics, natural language processing and computer program skills (Perl, Python).

Candidates should email a detailed CV with diploma.

Références

- [1] P.M. Addo, D. Guegan, and B. Hassani. Credit risk analysis using machine and deep learning models. *Risks*, 6(2) :38, 2018.
- [2] Óscar Garibo i Orts. Multilingual detection of hate speech against immigrants and women in twitter at semeval-2019 task 5 : Frequency analysis interpolation for hate in speech detection. In *Proceedings of the 13th International Workshop on Semantic Evaluation*, pages 460?463, 2019.
- [3] B. Iskandar. Terrorism detection based on sentiment analysis using machine learning. *Journal of Engineering and Applied Sciences*, 12(3) :691?698, 2017.
- [4] X. Jin, Z. Wei, J. Du, X. Xue, and X. Ren. Towards hierarchical importance attribution : Explaining compositional semantics for neural sequence models. *arXiv preprint arXiv :1911.06194*, 2019.
- [5] S. Mohan, A. Guha, M. Harris, F. Popowich, A. Schuster, and C. Priebe. The impact of toxic language on the health of reddit communities. In *Canadian Conference on Artificial Intelligence*, pages 51?56. Springer, 2017.
- [6] T. Niven and H. Kao. Probing neural network comprehension of natural language arguments. *arXiv preprint arXiv :1907.07355*, 2019.
- [7] A. Olteanu, C. Castillo, J. Boy, and K. Varshney. The effect of extremist violence on hateful speech online. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 12, 2018.
- [8] P. Saha, B. Mathew, K. Garimella, and A. Mukherjee. " short is the road that leads from fear to hate" : Fear speech in indian whatsapp groups. *arXiv preprint arXiv :2102.03870*, 2021.
- [9] P. Saha, B. Mathew, K. Garimella, and A. Mukherjee. "short is the road that leads from fear to hate" : Fear speech in indian whatsapp groups. *CoRR*, abs/2102.03870, 2021.
- [10] M. Wiegand, J. Ruppenhofer, and T. Kleinbauer. Detection of abusive language : the problem of biased datasets. In *Proceedings of the 2019 conference of the North American Chapter of the Association for Computational Linguistics : human language technologies*, volume 1 (long and short papers), pages 602?608, 2019.

Sujet 3

Deep-learning-based voiced/unvoiced decision for pitch detection

Proposé par : Stéphane ROSSIGNOL

Equipe : MULTISPEECH

Informations générales

Encadrants	Stéphane ROSSIGNOL, Vincent COLOTTE, Denis JOUVET	
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy	
Téléphone	03 87 76 47 73	
Email	stephane.rossignol@centralesupelec.fr	vincent.colotte@loria.fr
	denis.jouvet@loria.fr	
Bureau	B 306	

Motivations

Fundamental frequency (also called F_0 , or pitch) is one of the prosodic parameters, along with phone energy and phone durations. Prosody reflects various characteristics of the speaker and/or of the utterance, as for example the emotional state of the speaker, the modality of the utterance (question, statement, ...), some emphasis, ... Prosodic features are also involved in marking lexical stress. Hence the determination of the fundamental frequency is required in many applications including computer assisted language learning, development of speech synthesis systems, ...

Different algorithms have been developed in the past for computing the fundamental frequency. Some methods operate in the signal domain relying on autocorrelation methods [1], while some other methods operate in the spectral domain as for example the spectral comb approach [2] and the SWIPE (Sawtooth Waveform Inspired Pitch Estimator) method [3]. Some other algorithms combine processing in the spectral and in the signal domain [4].

On good quality speech signals, the various algorithms work quite well. However there exist some miss-detections (F_0 value associated to an unvoiced frame, or F_0 not detected for a voiced frame), as well as wrong estimations of the F_0 values. Moreover, errors get more frequent on lower quality speech data and on noisy speech data. Recent evaluations of a large set of pitch detection algorithms has confirm that the most frequent errors correspond to wrong voiced/unvoiced decisions [5].

Sujet

The goal of the proposed study is to develop and evaluate a robust and precise deep-learning-based approach that will handle the voiced/unvoiced decision, in place of the rather empirical signal processing based approaches that are currently used in the various pitch detection algorithms. The aim is to have a precise voiced/unvoiced decision leading to a significant performance improvement of the pitch detection algorithm, and to have an approach, which is robust to noise, and independent of the language.

Recurrent deep-learning models, such as LSTM-based models will be investigated. The input features will include some features previously used by pitch detection algorithms, plus any additional feature which may arise from a bibliography review. The approaches will be developed and evaluated using reference speech databases, for which the voiced/unvoiced decision (and the fundamental frequency values) are available. For the first set of experiments, the clean speech data will be artificially corrupted, through the addition of various noises at different signal to noise ratios. Assessment of the robustness of voiced/unvoiced decision approach will be measured on another reference database, containing clean and noisy data in another language. Final evaluations will include taking into account the F_0 values measured on the voiced segments, and provided by available algorithms.

Cadre du travail

The internship will be located at CentraleSupélec Metz.

Références

- [1] A. De Cheveigné, H. Kawahara. YIN, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America*, 111(4), 1917-1930.
- [2] P. Martin. Comparison of pitch detection by cepstrum and spectral comb analysis. In *ICASSP-1982, IEEE International Conference on Acoustics, Speech, and Signal Processing*.
- [3] A. Camacho. SWIPE : A sawtooth waveform inspired pitch estimator for speech and music. Doctoral dissertation, University of Florida.
- [4] A. Sorin, T. Ramabadran, D. Chazan, R. Hoory, M. McLaughlin, D. Pearce, ..., Y. Zhang. The ETSI extended distributed speech recognition (DSR) standards : client side processing and tonal language recognition evaluation. In *ICASSP-2004, IEEE International Conference on Acoustics, Speech, and Signal Processing*.
- [5] D. Jouviet, Y. Laprie. Performance analysis of several pitch detection algorithms on simulated and real noisy speech data. In *EUSIPCO-2017, 25th European Signal Processing Conference*.

Sujet 4

Target aware Hateful Language Detection

Proposé par : Tulika Bose

Equipe : MULTISPEECH

Informations générales

Encadrants	Tulika Bose, Irina Illina, Dominique Fohr
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Email	tulika.bose@loria.fr, irina.illina@loria.fr, dominique.fohr@loria.fr
Bureau	C 140

Motivations

With the widespread use of social media, the problem of hateful language against different communities have increased tremendously in the recent years. Such language is manifested as derogatory comments, insults, threats or the propagation of negative stereotypes against target groups belonging to different race, religion, gender, and the like. The onus of curbing the spread of such language mostly lies with the owners of social media platforms. Since manual moderation of hateful content alone in the ever increasing pool of social media comments is infeasible, automated hateful language classifiers are employed to assist the manual moderators. In recent years, deep learning based classifiers have emerged as the state-of-the-art mechanisms for hateful language detection [1]. However, hateful language is a complex phenomenon and building classifiers that can learn the complexities of hateful language is a challenging task. Moreover, due to the sparsity of hateful content in real world, most of the hateful language corpora are crawled through keyword based sampling. This results in biased datasets [2] and the classifiers tend to learn spurious correlations present in the dataset.

As mentioned in [3], the presence of targets is an important criteria for a comment to be considered hateful and this differentiates it from merely offensive language. For example, the comment (taken from [3]) ‘XYZ is too real, major f*cking respect for XYZ...’ is labeled as offensive for the use of an abusive word, but not hateful. However, the comment ‘f*ck you n*gger...’ is labeled as hateful as it is targeted against a particular community. As such, a hateful language classifier can be provided with additional features informing the model about the targets present in the comment. To this end, a sequence prediction task can be performed prior to the hateful language classification, that is trained to predict targets. The representations from the sequence tagger can subsequently be used as additional features for hateful language classification. This would enrich the classifier to learn more structured information about hate rather than learning from the biases present in the dataset, resulting in accurate detection of hate.

Sujet

The objective of this work is to build a hateful language classifier that is well informed about the target-oriented nature of hateful language and that does not rely on spurious features of biased datasets. We hypothesize that such a classifier would be able to learn the nuances of hateful language much more efficiently and make accurate classification. Since the available hateful language datasets may not be always annotated at the level of targets, the list of common hate-targets from the Hatebase⁴ lexicon can be used for this purpose to automatically annotate the tokens of comments. Subsequently, we would like to build a sequence predictor that learns to predict the targets in the available dataset. The learned representations from the sequence predictor can then be provided as external features to the hateful language classifier. Due to the presence of these additional features, the classifier should learn to make decisions based on the presence of targets in comments.

Cadre du travail

The work plan for the internship is as follows : At the beginning, the intern will perform a short literature review on the recent developments in hateful language classification with deep neural networks. The student will then write a python based script to use the Hatebase target lexicons for automatically annotating the available hateful language datasets at the token level and prepare the datasets for the sequence prediction task. After this, the student will train a sequence tagger to predict the targets in the datasets. Relevant modifications

4. <https://hatebase.org/>

in the sequence tagger would need to be performed as required for the task. Subsequently, the student will build a hateful language classifier that can fuse the representations from the sequence tagger while training the model. Both the sequence predictor and the classifier can be initialized with the pre-trained BERT model [4]. The performance of such a target aware classifier will be evaluated on multiple hateful language datasets and compared with classifiers that do not make use of target level information from comments.

Required Skills : The candidate should have a theoretical and a moderate practical experience with Deep Learning, including a good practice in Python and an understanding of deep learning libraries like Pytorch. The intern shall be required to modify an existing Pytorch repository and further implement the work in Pytorch.

Keywords : hate-targets, sequence-prediction, target awareness, deep-learning

Références

- [1] P. Badjatiya, S. Gupta, M. Gupta, and V. Varma. “Deep Learning for Hate Speech Detection in Tweets”. *Proceedings of the 26th International Conference on World Wide Web Companion*, 2017, pp. 759–760.
- [2] M. Wiegand, J. Ruppenhofer, and T. Kleinbauer. “Detection of Abusive Language and the Problem of Biased Datasets”. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies, Volume 1 (Long and Short Papers)*, 2019, pp. 602–608. Association for Computational Linguistics.
- [3] T. Davidson, D. Warmusley, M. Macy and I. Weber. “Automated Hate Speech Detection and the Problem of Offensive Language”. *ICWSM*, 2017.
- [4] J. Devlin, M. Chang, K. Lee, and K. Toutanova. “BERT : Pre-training of Deep Bidirectional Transformers for Language Understanding”. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies, Volume 1 (Long and Short Papers)*, 2019, 4171–4186. Association for Computational Linguistics.

Sujet 5

Deep latent feature factorization for unsupervised speech separation

Proposé par : Paul MAGRON, Mostafa SADEGHI

Equipe : MULTISPEECH

Informations générales

Encadrants	Paul Magron	Mostafa Sadeghi
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy	
Email	paul.magron@inria.fr	mostafa.sadeghi@inria.fr
Bureau	C 141	C 136

Motivations

Speech separation consists in isolating the signals that correspond to each speaker from an acoustic mixture where several persons might be speaking. This task is an important preprocessing step in many applications such as hearing aids or vocal assistants based on automatic speech recognition.

State-of-the-art separation systems rely on supervised deep learning, where a network is trained to predict the isolated speakers' signals from their mixture [1, 2]. However, these approaches are costly in terms of training data and have a limited capacity to generalize to unseen speakers.

Sujet

The goal of this internship is to design a fully unsupervised system for speech separation, which is more data-efficient than supervised approaches, and applicable to any mixture of speakers. To that end, we propose to combine deep autoencoders (AEs) with nonnegative matrix factorization (NMF) models. NMF consists in decomposing a given input matrix (usually : an audio spectrogram) as the product of two interpretable factors (usually : a dictionary of spectra and a temporal activation matrix). This family of methods has been extensively researched before the era of deep learning [3], but it is limited since real-world audio spectrograms usually don't comply with the factorization assumption.

Therefore, we propose to leverage AEs as a tool to learn a latent representation of the data which is regularized using NMF. Such a system can be cast as an instance of *transform learning* [4] : the key idea is to apply a (learned) transform to the data so that it better complies with a desirable property - here, factorization. Once trained, the resulting system operates in three stages : (i) the (mixture) audio spectrogram is projected through the encoder into some latent space ; (ii) this latent representation is factorized efficiently using an NMF algorithm, which provides a latent feature for each speaker ; (iii) these latent features are passed through the decoder to retrieve a spectrogram for each speaker.

Such a system is promising since it is fully unsupervised (it can be trained without knowledge of specific mixtures), it yields an interpretable decomposition of the latent representation, and it can serve as a basis for other applications (including speaker diarization, speech enhancement or voice conversion).

A good practice in Python and basic knowledge about deep learning, both theoretical and practical (e.g., using PyTorch) are required. Some notions of audio/speech signal processing and machine learning is a plus.

Cadre du travail

The trainee will be supervised by Paul Magron (Chargé de Recherche Inria) and Mostafa Sadeghi (Researcher, Inria Starting Faculty Position), and will benefit from the research environment and the expertise in audio signal processing of the MULTISPEECH team. This team includes many PhD students, post-docs, trainees, and permanent staff working in this field, and offers all the necessary computational resources (GPU and CPU, speech datasets) to conduct the proposed research.

Références

- [1] D. Wang and J. Chen, Supervised Speech Separation Based on Deep Learning : An Overview *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 10, pp. 1702-1726, 2018.
- [2] Y. Luo and N. Mesgarani, Conv-TasNet : Surpassing Ideal Time-Frequency Magnitude Masking for Speech Separation *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 8, pp. 1256-1266, 2019.
- [3] T. Virtanen, Monaural Sound Source Separation by Nonnegative Matrix Factorization With Temporal Continuity and Sparseness Criteria *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 3, pp. 1066-1074, 2007.
- [4] D. Fagot, H. Wendt and C. Févotte, Nonnegative Matrix Factorization with Transform Learning *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018.

Sujet 6

Réparation de signaux audio par des modèles compacts prenant en compte la phase

Proposé par : Antoine DELEFORGE, Paul MAGRON

Equipe : MULTISPEECH

English version available at : https://team.inria.fr/multispeech/m2_internship_phase/

Informations générales

Encadrant	Antoine Deleforge	Paul Magron
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy	
Téléphone	03 54 95 86 44	
Email	antoine.deleforge@inria.fr	paul.magron@inria.fr
Bureau	C 144	C 141

Motivations

Le domaine du rehaussement de signaux audio a connu d'important progrès ces dernières années. Les méthodes conventionnelles se basent généralement sur l'estimation d'un masque appliqué à une représentation temps-fréquence à valeurs complexes du signal (ex : la transformée de Fourier à court terme). Ces méthodes utilisent des modèles de l'*amplitude* du signal à rehausser, mais ne tiennent généralement pas compte de sa *phase*, se contentant d'utiliser celle du signal bruité en entrée. Bien que cette approche donne des résultats satisfaisants lorsqu'il s'agit de réduire un bruit de niveau modéré, elle génère d'importantes distorsions dans le cas de bruit impulsifs importants (par exemple, des bruits de claviers lors d'une vidéo-conférence ou quelqu'un qui tousse), et n'est plus du tout valide lorsque des échantillons ou des morceaux entiers de signaux sont complètement manquants (ex : saturation, cliques, ou craquements dans des enregistrements audio endommagés).

En outre, la plupart des méthodes récentes du domaine s'appuient sur l'apprentissage supervisé (deep learning) et requièrent donc des bases de données d'entraînement très volumineuses. Ces approches se généralisent mal à des signaux de types inconnus ou éloignés de ceux utilisés à l'entraînement.

Sujet

Ce stage se focalisera sur la reconstruction d'échantillons manquants dans des **signaux audio fortement dégradés**. Il combinera des modèles de faible complexité ne requérant que peu ou pas d'apprentissage, et des approches novatrices venant du domaine de la *restitution de phase (phase retrieval)*. Ce domaine a connu de récentes percées méthodologiques venant de l'échantillonnage comprimé (*compressed sampling* [1]), de la complétion de matrices de rang faible [2], de la programmation semi-définie [3] et de l'optimisation non-convexe [4]. Bien que puissantes, ces techniques n'ont reçu que très peu d'attention en traitement du signal audio à ce jour. Leur potentiel pour le rehaussement audio a été démontré dans des travaux préliminaires effectués dans l'équipe [5, 6].

Pour mener à bien ce stage, le candidat devra avant tout posséder de **fortes compétences et une forte affinité pour les mathématiques**, notamment en traitement du signal, algèbre linéaire et optimisation. Des compétences en programmation Python ou Matlab sont aussi requises, et une expérience ou affinité pour l'audio est un plus.

Cadre du travail

Ce stage de master s'inscrit dans le cadre du projet DENISE : "Attaque de problèmes difficiles en audio par des approches inverses non-linéaires économes en données" (2021-2024) financé par l'Agence Nationale de la Recherche (ANR). Le projet DENISE vise à explorer l'applicabilité à l'audio de percées récentes dans les domaines des problèmes inverses non-linéaires.

Le stage sera encadré par Antoine Deleforge et Paul Magron (Chargés de Recherche Inria), et s'appuiera sur des travaux de recherche préliminaire dans cette direction [5, 6]. Il bénéficiera de l'environnement de recherche et de l'expertise en traitement du signal audio de l'équipe MULTISPEECH, comptant de nombreux doctorants, post-doctorant, stagiaires, et permanents travaillant dans ce domaine.

Références

- [1] Emmanuel J Candès and Michael B Wakin. An introduction to compressive sampling. *IEEE signal processing magazine*, 25(2) :21–30, 2008.
- [2] Emmanuel J Candes, Yonina C Eldar, Thomas Strohmer, and Vladislav Voroninski. Phase retrieval via matrix completion. *SIAM review*, 57(2) :225–251, 2015.
- [3] Irène Waldspurger, Alexandre d'aspremont, and Stéphane Mallat. Phase recovery, maxcut and complex semidefinite programming. *Mathematical Programming*, 149(1-2) :47–81, 2015.
- [4] Emmanuel J Candes, Xiaodong Li, and Mahdi Soltanolkotabi. Phase retrieval via wirtinger flow : Theory and algorithms. *IEEE Transactions on Information Theory*, 61(4) :1985–2007, 2015.
- [5] Antoine Deleforge and Walter Kellermann. Phase-optimized k-svd for signal extraction from underdetermined multichannel sparse mixtures. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 355–359. IEEE, 2015.
- [6] Antoine Deleforge and Yann Traonmilin. Phase unmixing : Multichannel source separation with magnitude constraints. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 161–165. IEEE, 2017.

Sujet 7

Evaluating the impact of speaker anonymization on emotion recognition

Proposé par : Hubert Nourtel & Denis Jovet

Equipe : MULTISPEECH

Informations générales

Encadrants	Hubert Nourtel	Denis Jovet
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy	
Téléphone	03 54 95 86 26	
Email	hubert.nourtel@loria.fr	denis.jovet@loria.fr
Bureau	C046	C147

Motivations

Over the last decade, great progress has been made in automatic speech recognition. This is due to the maturity of deep learning techniques, to the availability of very large datasets, and to the increase in computational power. Consequently, the use of speech recognition has spread in many applications, such as virtual assistants (as for instance Apple's Siri, Google Now, Microsoft's Cortana, or Amazon's Alexa) which collect, process and store personal speech data in centralized servers, raising serious concerns regarding the privacy of the data of their users. However, speech recognition technology still has limited performance in adverse conditions (e.g., noisy environments, reverberated speech, strong accents, etc.) and thus, there is a need for performance improvement. This can only be achieved by using large speech corpora that are representative of the actual users and of the various usage conditions. There is therefore a strong need to share speech data for improved training that is beneficial to all users, while preserving the privacy of the users, which means at least keeping the speaker identity and voice characteristics private. Note that when sharing data, users may want not to share data conveying private information at the linguistic level (e.g., phone number, person name, ?). Such privacy aspects also need to be taken into account, but they are out-of-the scope of this project.

In the ANR project DEEP-PRIVACY⁵, we investigate speaker anonymization using a voice conversion system. A voice conversion system modifies a speech signal to sound as if pronounced by another speaker [1]. In this context, the aim of the internship proposal will be to investigate the impact of the application of speaker anonymization system with respect to the emotions present in expressive speech data. Note that emotions may be considered either as ?utility? information that should remain present in the data, or as ?privacy? information one would like to hide.

Sujet

The goal during the internship will be to measure how much emotion information is present in the anonymized speech data, and to compare that with emotion information present in the original data (i.e., before anonymization). One possibility for measuring the emotion information present in some speech data is to use an emotion detection system. Some experiments have already been conducting using emotion detection systems based on SVM (Support Vector Machine) [2] and on DNN (Deep Neural Network) [3]. Preliminary results have been reported in [4] using the IEMOCAP expressive speech corpus [5].

The goal of the internship will be to investigate further the DNN-based approaches to improve their performance for detecting emotions in speech signals. Investigations will be related to both the architecture of the DNN, and the training protocol. As the size of the expressive speech corpus used in the anonymization experiments is limited, we will investigate pre-training it on some other speech corpora, expressive or not. One difficulty stems from the fact that various expressive speech corpora do not exhibit the same sets of emotions. A possible solution will be to handle that in a way similar to multilingual training [6].

If time and data permit it, it will be interesting to evaluate the impact of the voice anonymization process with respect to emotion information, on several expressive speech corpora, and in several languages; and / or to

5. <https://project.inria.fr/deepprivacy/>

conduct some listening test with respect to human perception of emotions, to complement the above experiments based on automatic detection of emotions.

Cadre du travail

The internship will be done in the multispeech team.

Required skills : Good knowledge of machine learning approaches, in particular with deep-learning based approaches; and experience with deep-learning frameworks.

Références

- [1] F. Fang, X. Wang, J. Yamagishi, I. Echizen, M. Todisco, N. Evans, and J.-F. Bonastre. Speaker Anonymization Using X-vector and Neural Waveform Models. 10th ISCA Speech Synthesis Workshop, 2019.
- [2] M. Tahon, and L. Devillers. Towards a small set of robust acoustic features for emotion recognition : challenges. IEEE/ACM transactions on audio, speech, and language processing, vol. 24, n° 1, 2015, pp. 16-28.
- [3] R. Pappagari, T. Wang, J. Villalba, N. Chen, and N. Dehak. X-vectors meet emotions : A study on dependencies between emotion and speaker recognition. ICASSP 2020, IEEE International Conference on Acoustics, Speech and Signal Processing, 2020, pp. 7169-7173.
- [4] H. Nourtel, P. Champion, D. Jouvét, A. Larcher, and M. Tahon. Evaluation of Speaker Anonymization on Emotional Speech. 1st ISCA Symposium on Security and Privacy in Speech Communication, Nov 2021, Virtual, Germany.
- [5] C. Busso, M. Bulut, C.-C. Lee, A. Kazemzadeh, E. Mower, S. Kim, J. N. Chang, S. Lee, and S. S. Narayanan. IEMOCAP : Interactive emotional dyadic motion capture database. Journal of Language Resources and Evaluation, 2008.
- [6] R. Fér, P. Matějka, F. Grézl, O. Plchot, K. Veselý, and J.H. Šernocký. Multilingually trained bottleneck features in spoken language recognition. Computer Speech & Language, vol. 46, 2017, pp. 252-267.

Sujet 8

Detection of audio segments

Proposé par : Dominique FOHR

Equipe : MULTISPEECH

Informations générales

Encadrants	Dominique FOHR
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Téléphone	03 83 59 20 27
Email	dominique.fohr@loria.fr
Bureau	C 143

Motivations

More and more videos are produced each day. For broadcasting these videos, it is necessary to generate subtitles (sous-titres) or to perform dubbing (doublage). The subtitles allow people who are hard of hearing or foreigners to better understand the video. For generating the subtitles, two steps are needed : first the detection of the audio segments where people are speaking. Then the transcription is performed. For dubbing, another step is required : the diarization. It consist in detecting and clustering the segments uttered by the same speaker.

Robust automatic speech recognition (ASR) is always a very ambitious goal. Despite constant efforts and some dramatic advances, the ability of a machine to recognize the speech is still far from equaling that of the human being. Current ASR systems see their performance significantly decrease when the conditions under which they were trained and those in which they are used differ. The causes of variability may be related to the acoustic environment, sound capture equipment, microphone change, etc.

Sujet

The aim of this internship is to use and modify the tools developed for speech recognition to speed up the tasks of subtitling or dubbing. The first task to perform will be the detection of the audio segments where speech is present. The main problems are the background noises, the background music or the presence of songs. The algorithm should be robust to these adverse conditions. The method will be based on deep neural networks. Recurrent neural networks (RNN [1]), convolutional neural networks (CNN [3]) or attention based models will be studied. The internship will be carried out with ECOSM, a company specialized in subtitling.

Required skills : background in statistics, computer program skills (Perl, Python, Tensorflow, Pytorch).

Candidates should email a detailed CV with diploma.

Références

- [1] Thad Hughes and Keir Mierle. Recurrent neural networks for voice activity detection ICASSP, 2013
- [2] Joohyung Lee, Youngmoon Jung, Hoirin Kim Dual Attention in Time and Frequency Domain for Voice Activity Detection Interspeech, 2020
- [3] Cheng Yu , Kuo-Hsuan Hung, I-Fan Lin, Szu-Wei Fu, Yu Tsao, Jehi-weih Hung Waveform-based Voice Activity Detection Exploiting Fully Convolutional networks with Multi-Branched Encoders arxiv 2020
- [4] Ruixi Lin, Charles Costello, Charles Jankowski, Vishwas Mruthyunjaya Optimizing Voice Activity Detection for Noisy Conditions Interspeech 2019

Sujet 9

Détermination tridimensionnelle de la forme du conduit vocal

Proposé par : Yves Laprie

Equipe : MULTISPEECH

Informations générales

Encadrants	Yves Laprie
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Téléphone	03 83 59 20 36
Email	yves.laprie@loria.fr
Bureau	C 146

Motivations

La synthèse articulatoire [2] repose sur l'utilisation de simulations acoustiques numériques appliquées à la géométrie instantanée du conduit vocal. La géométrie est prédite à partir de la suite de phonèmes à articuler. Le conduit vocal a une géométrie tridimensionnelle complexe qui évolue en fonction des sons et cela signifie que la forme en un point du conduit vocal, par exemple au niveau du palais dur, dépend du son à produire. L'an dernier nous avons développé un algorithme de prédiction de la forme du conduit vocal dans le plan médiaosagittal (le plan vertical avant-arrière au milieu de la tête) qui utilise l'apprentissage profond appliqué à une grande base de données d'IRM dynamiques bidimensionnelles [1]. Cet algorithme prédit une forme dans le plan médio-sagittal à laquelle il manque la troisième dimension dans la direction transverse (d'abord coronale dans la cavité buccale puis évoluant vers la direction axiale quand on se rapproche du larynx). L'imagerie IRM ne permet pas actuellement d'acquérir des images tridimensionnelles dynamiques et nous utilisons donc des heuristiques pour approcher la forme dans la direction transverse, avec le défaut de commettre des erreurs qui se répercutent sur l'acoustique du signal généré.

Sujet

L'objectif du projet est donc d'optimiser la détermination de la forme transverse. Malgré son importance ce problème a été peu étudié [3] et nous explorerons deux approches qui peuvent être complémentaires. La première est d'utiliser des images statiques 3D qui correspondent à une centaine de positions articulatoires. Pour chacune des images 3D, et dans chaque plan perpendiculaire à l'axe du conduit vocal, la partie de l'image qui correspond à la colonne d'air du conduit vocal sera détournée manuellement puis automatiquement après apprentissage automatique. Les formes ainsi détournées serviront à leur tour de base d'apprentissage pour déterminer l'aire du conduit vocal dans la direction transverse.

Une seconde approche consistera à utiliser des images dynamiques dans 2 plans orthogonaux, l'un des deux étant systématiquement le plan médio-sagittal, et le second dans la direction coronale ou axiale en fonction de point cible dans le conduit vocal. L'avantage d'utiliser des images dynamiques est de traiter de la parole naturelle au lieu de positions articulatoires statiques un peu artificielles. Le travail consistera à combiner un nombre d'acquisitions suffisant pour couvrir tout le conduit vocal ce qui nécessitera un travail de détournement puis de fusion des différentes acquisitions. La première approche permettra d'amorcer le détournement automatique de l'aire transverse.

Cadre du travail

Ce sujet se fera en étroite collaboration avec le laboratoire IADI, en particulier Pierre-André Vuissoz qui est spécialiste de l'acquisition et le traitement de données IRM.

Nous disposons déjà de données IRM statiques tridimensionnelles et nous acquerrons des données dynamiques pour travailler sur la seconde approche.

Références

- [1] Vinicius Ribeiro, Karyna Isaieva, Justine Leclere, Pierre-André Vuissoz, and Yves Laprie. Towards the prediction of the vocal tract shape from the sequence of phonemes to be articulated. In *iNTERSPEECH 2021*, Brno, Czech Republic, August 2021. URL <https://hal.inria.fr/hal-03360113>.
- [2] Rachel Alexander, Tanner Sorensen, Asterios Toutios, and Shrikanth Narayanan. A modular architecture for articulatory synthesis from gestural specification. *The Journal of the Acoustical Society of America*, 146(6) :4458–4471, 2019. URL <https://doi.org/10.1121/1.5139413>.
- [3] R.S. McGowan, M.T. Jackson, and M.A. Berger. Analyses of vocal tract cross-distance to area mapping : an investigation of a set of vowel images. *JASA*, 131(1) :424–434, 2012. URL <https://doi.org/10.1121/1.3665988>

Sujet 10

Inversion acoustique articulatoire de la parole à l'aide d'images IRM dynamiques

Proposé par : Yves Laprie

Equipe : MULTISPEECH

Informations générales

Encadrants	Yves Laprie
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Téléphone	03 83 59 20 36
Email	yves.laprie@loria.fr
Bureau	C 146

Motivations

La synthèse articulatoire imite le processus de production de la parole en générant d'abord la forme du conduit vocal à partir d'une séquence de phonèmes à prononcer, puis le signal acoustique en résolvant les équations de l'aéro-acoustique. Par rapport à d'autres approches de la synthèse de la parole qui offrent un niveau de qualité très élevé, l'intérêt est avant tout de contrôler l'ensemble du processus de production, au-delà du seul signal acoustique.

L'inversion acoustique articulatoire est le processus inverse qui consiste à retrouver la forme du conduit vocal à partir du signal acoustique. En associant l'inversion acoustique articulatoire aux outils de synthèse articulatoire qui progressent rapidement il deviendra possible d'étudier la production de la parole en profondeur pour aborder notamment la question des troubles de la production de la parole, d'expliquer les difficultés rencontrées par les malentendants lors de l'acquisition de la langue orale et de fournir des retours audiovisuels articulatoires pour la remédiation.

La quasi-totalité des travaux actuels en inversion reposent sur l'utilisation de données issues de l'articulographie électro-magnétique (EMA ElectroMagnetic Articulography) qui donne la position de capteurs collés sur la langue et les autres articulateurs facilement accessibles. Du point de vue des techniques d'inversion proprement dites l'apprentissage profond est largement utilisé parce qu'il permet d'exploiter efficacement les corpus de données EMA. À l'heure actuelle, l'approche LSTM (LongShort-Term Memory) et sa variante bidirectionnelle donne les meilleurs résultats [1].

Malgré leur très bonne précision géométrique, et parce que les données EMA ne peuvent couvrir que la partie du conduit vocal la plus proche de la bouche, les approches actuelles ne permettent pas de retrouver la géométrie complète du conduit vocal alors que l'on sait par exemple que le larynx joue un rôle déterminant sur l'acoustique du conduit vocal. En pratique, cela limite considérablement l'intérêt des techniques d'inversion puisque les résultats ne peuvent pas être utilisés pour reconstruire le signal de parole. L'objectif de ce master est de lever ce verrou et de retrouver toute la géométrie du conduit vocal. Pour cela nous nous proposons d'utiliser les données de l'IRM dynamique que nous pouvons acquérir à Nancy au laboratoire IADI.

Sujet

L'IRM dynamique fournit des images bidimensionnelles dans le plan médiosagittal à 50Hz de très bonne qualité et le signal de parole acquis à l'aide d'un microphone optique et débruité très efficacement (exemples disponibles sur <https://artspeech.loria.fr/resources/>). Nous prévoyons d'utiliser des corpus déjà acquis ou en cours d'acquisition. Ces corpus représentent un volume de données très grand (plusieurs centaines de milliers d'images) et il est donc nécessaire de les prétraiter afin d'identifier le contour des articulateurs impliqués dans la production de la parole (mandibule, langue, lèvres, vélum, larynx, épiglotte). L'an dernier nous avons développé une approche du suivi du contour des articulateurs dans les images IRM qui donne de très bons résultats [2]. Les contours suivis automatiquement peuvent donc être utilisés pour réaliser l'inversion. Dans l'optique d'utiliser l'inversion afin d'analyser le comportement individuel d'un articulateur qui peut être défaillant, chaque articulateur est suivi indépendamment des autres. L'objectif est de réaliser l'inversion en utilisant vraisemblablement l'approche LSTM sur les données d'un petit nombre de locuteurs pour lesquels il

existe des données en quantité suffisante. Cette approche devra être adaptée à la nature des données et afin de pouvoir identifier la contribution de chacun des articulateurs.

Cadre du travail

Ce projet sera mené conjointement au laboratoire IADI (INSERM U1254) et au Loria (équipe Inria MultiSpeech) qui collaborent ensemble depuis plusieurs années sur l'imagerie du conduit vocal et l'étude de la production de la parole.

Références

- [1] Maud Parrot, Juliette Millet, and Ewan Dunbar. Independent and Automatic Evaluation of Speaker-Independent Acoustic-to-Articulatory Reconstruction. In *Interspeech 2020 - 21st Annual Conference of the International Speech Communication Association*, Proceedings of INTERSPEECH 2020, 21st Annual Conference of the International Speech Communication Association, Shanghai / Virtual, China, October 2020. URL <https://hal.archives-ouvertes.fr/hal-03087264>.
- [2] Karyna Isaieva, Yves Laprie, Nicolas Turpault, Alexis Houssard, Jacques Felblinger, and Pierre-André Vuissoz. Automatic Tongue Delineation from MRI Images with a Convolutional Neural Network Approach. *Applied Artificial Intelligence*, 34(14) :1115–1123, October 2020. URL <https://hal.archives-ouvertes.fr/hal-02962336>.

Sujet 11

Investigating the speaker information carried on by the fundamental frequency

Proposé par : Pierre Champion & Denis Juvet

Equipe : MULTISPEECH

Informations générales

Encadrants	Pierre Champion	Denis Juvet
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy	
Téléphone	03 54 95 86 26	
Email	pierre.champion@loria.fr	denis.juvet@loria.fr
Bureau	C040	C147

Motivations

Besides the linguistic content itself, a speech signal conveys a lot of information about the speaker, including its identity, gender, age, as well as stress, emotion, ? thus making speech a very personal data. Because of that, there are some studies that investigate how speech data can be anonymized. The ANR project DEEP-PRIVACY⁶ is one such project. One goal of the project is to anonymize speech data to allow for an easier sharing of speech data, which is necessary for large data collection for improving speech technologies.

Currently, the most relevant approach is based on a voice conversion system. A voice conversion system modifies a speech signal to sound as if pronounced by another speaker [1]. In the referred approach, the speech signal is decomposed into three streams of data : one corresponding to the fundamental frequency values, one corresponding to linguistic content (for example, probabilities of phone units), and a last one corresponding to the speaker information. The speaker information is represented by a speaker embedding vector, called x-vector [2]. For anonymization, this x-vector is replaced by a target x-vector, and a speech signal is reconstructed using this target x-vector and the two other streams (fundamental frequency and linguistic content). This makes the reconstructed speech signal sound as if pronounced by the speaker corresponding to the target x-vector.

Sujet

In this context, the goal of the internship is to investigate how much speaker information is carried by the fundamental frequency values.

This will be investigated through deep-learning based speaker recognition or verification experiments, using fundamental frequency values as main input data, possibly with some additional information. The research work will concern the coding of the fundamental frequency information, the choice of additional information, as well as optimizing the deep-learning architecture used to process the data.

As currently the best speaker recognition / verification⁷ performance is achieved using x-vector representation; the fundamental frequency information will ultimately be merged with the x-vector representation, to measure its impact on the speaker verification performance.

Cadre du travail

The internship will be done in the multispeech team.

Required skills : Good knowledge of machine learning approaches, in particular with deep-learning based approaches; and experience with deep-learning frameworks.

Références

- [1] F. Fang, X. Wang, J. Yamagishi, I. Echizen, M. Todisco, N. Evans, and J.-F. Bonastre. Speaker Anonymization Using X-vector and Neural Waveform Models. 10th ISCA Speech Synthesis Workshop, 2019.
- [2] D. Snyder, D. Garcia-Romero, G. Sell, D. Povey, and S. Khudanpur. X-vectors : Robust DNN embeddings for speaker recognition. ICASSP 2018, IEEE International Conference on Acoustics, Signal and Speech Processing, 2018, pp.5329-5333.

6. <https://project.inria.fr/deepprivacy/>

7. Speaker recognition means recognizing a speaker among a set of known speakers; whereas speaker verification refers to the task of checking if the speech signal matches with the claimed speaker identity.

Sujet 12

Exploiting time-frequency correlations for variational autoencoder-based speech enhancement

Proposé par : Mostafa SADEGHI, Xavier ALAMEDA-PINEDA, Antoine DELEFORGE

Equipe : MULTISPEECH

Informations générales

Encadrants Mostafa SADEGHI, Xavier ALAMEDA-PINEDA, Antoine DELEFORGE
Adresse LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Email {mostafa.sadeghi, xavier.alameda-pineda, antoine.deleforge}@inria.fr

Motivations

Speech enhancement aims to improve the quality and intelligibility of noisy speech signals, which is a fundamental problem in speech signal processing, with several important applications, including automatic speech recognition systems and hearing aids. A recent class of speech enhancement methods combines the expressive power of deep neural networks with traditional model-based signal processing approaches to arrive at an unsupervised approach [1-3]. They work in the time-frequency domain, using short-time Fourier transform (STFT), where in the training phase, the statistical distribution of clean speech is learned based on some generative model, e.g. variational autoencoder (VAE). The learned model is then combined with a noise model, e.g. based on non-negative matrix factorization (NMF), to estimate clean speech signal at test time. Compared to supervised approaches, where a deep neural network is learned to map a noisy speech signal to a clean estimate, unsupervised approaches have potential to achieve better generalization performance.

Sujet

Early unsupervised speech enhancement approaches processed speech spectra time frames independently, ignoring their statistical correlations [1,2]. Recently, some dynamical VAEs have been developed for speech enhancement that properly model time correlations [3]. However, by assuming an isotropic (diagonal) covariance matrix for the Gaussian generative distribution of each time frame, they ignore frequency-wise correlations, which are important for efficient speech modeling. In this internship project, we are going to develop a speech generative model that simultaneously exploits time and frequency correlations. This will be achieved by considering an anisotropic (non-diagonal) covariance matrix for the distribution of each spectrum time frame within a dynamical VAE framework. We will also derive a corresponding inference algorithm to estimate clean speech at test time.

Background in statistical signal processing, machine learning, and programming languages and deep learning frameworks (Python, PyTorch) are favored. Interested candidates should send an email to the supervisors with a detailed CV and transcripts.

Cadre du travail

The intern will join the Multispeech team, where the necessary computational resources (GPU and CPU, speech datasets) to carry on the proposed research are available.

Références

- [1] S. Leglaive, L. Girin, and R. Horaud, A variance modeling framework based on variational autoencoders for speech enhancement *IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, Aalborg, Denmark, September 2018.

- [2] M. Sadeghi, S. Leglaive, X. Alameda-Pineda, L. Girin, and R. Horaud, Audio-visual speech enhancement using conditional variational auto-encoders *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 1788-1800, 2020.
- [3] L. Girin, S. Leglaive, X. Bie, J. Diard, T. Hueber, and X. Alameda-Pineda, Dynamical variational autoencoders : A comprehensive review *arXiv preprint arXiv :2008.12595*, 2020.

Sujet 13

Dynamical variational autoencoders for unsupervised audiovisual speech enhancement

Proposé par : Mostafa SADEGHI, Romain SERIZEL

Equipe : MULTISPEECH

Informations générales

Encadrant	Mostafa SADEGHI	Romain SERIZEL
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy	
Email	mostafa.sadeghi@inria.fr	romain.serizel@loria.fr

Motivations

Audiovisual speech enhancement is defined as the task of improving the quality and intelligibility of a noisy speech signal by utilizing the complementary information provided by the visual modality, i.e., lip movements of speaker [1]. Visual modality is especially important in high noise situations, and it's less affected by acoustic noise. As such, it's exploited in several practical applications, including hearing assistive devices. Numerous works have already studied the integration of visual modality with audio modality to improve the performance of speech enhancement. While the majority of audiovisual speech enhancement algorithms rely on deep neural networks and supervised learning, they require very large audiovisual datasets with diverse noise instances to have good generalization performance.

A recently introduced audiovisual speech enhancement approach is based on unsupervised learning [2,3], where in a training phase, the statistical distribution of clean speech is learned using a clean audiovisual dataset. This is done using a deep generative model, e.g. variational autoencoder (VAE). Then, at test (inference) time, the learned distribution is combined with a noise model, based on, e.g. non-negative matrix factorization (NMF), to estimate the clean speech signal from the available noisy speech observations. As the training phase involves no noise signals, this approach is unsupervised, and as such, it has more potential for generalization than its supervised counterpart.

Sujet

Up to now, all the developed unsupervised audiovisual speech enhancement methods are based on independently processing of speech spectrogram time frames, which violates the correlations between consecutive time frames. Recently, some variants of VAEs have been introduced and tested on speech enhancement that model time dynamics of speech spectra [4]. It has been shown that with dynamical VAEs, the speech enhancement performance can be greatly improved. However, the computational complexity of inference algorithms is also increased with the introduction of dynamical models.

The main objective of this internship project will be to develop some dynamical VAE-based generative models for audiovisual speech enhancement. This will require careful investigation and development of audiovisual fusion strategies tailored to dynamical models, along with efficient inference algorithms for speech signal estimation.

Background in statistical signal processing, probabilistic machine learning, optimization, and programming languages and deep learning frameworks (Python, PyTorch) are favored. Interested candidates should send an email to the supervisors with a detailed CV and transcripts.

Cadre du travail

The intern will join the Multispeech team, where the necessary computational resources (GPU and CPU, speech datasets) to carry on the proposed research are available.

Références

- [1] D. Michelsanti, Z.-H. Tan, S.-X. Zhang, Y. Xu, M. Yu, D. Yu, and J. Jensen, An overview of deep-learning-based audio-visual speech enhancement and separation *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 1368?1396, 2021.
- [2] S. Leglaive, L. Girin, and R. Horaud, A variance modeling framework based on variational autoencoders for speech enhancement *IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, Aalborg, Denmark, September 2018.
- [3] M. Sadeghi, S. Leglaive, X. Alameda-Pineda, L. Girin, and R. Horaud, Audio-visual speech enhancement using conditional variational auto-encoders *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 1788?1800, 2020.
- [4] L. Girin, S. Leglaive, X. Bie, J. Diard, T. Hueber, and X. Alameda-Pineda, Dynamical variational autoencoders : A comprehensive review *arXiv preprint arXiv :2008.12595*, 2020.

Sujet 14

Modélisation des accents régionaux pour une reconnaissance automatique de la parole inclusive

Proposé par : Emmanuel VINCENT

Equipe : MULTISPEECH

Informations générales

Encadrants	Emmanuel VINCENT, Anne BONNEAU	
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy	
Téléphone	03 83 59 30 83	
Email	emmanuel.vincent@inria.fr	anne.bonneau@loria.fr
Bureau	C 147	

Motivations

La reconnaissance automatique de la parole (RAP) conversationnelle a vu un progrès considérable au cours de la dernière décennie, avec un taux d'erreur sur les mots aujourd'hui proche de celui des humains [1]. Cela s'explique par la maturité des réseaux de neurones profonds mais surtout par l'accroissement de la taille des corpus d'apprentissage propriétaires collectés par les grands opérateurs, de l'ordre de 10 000 h ou plus pour le français. Le coût de collecte et d'annotation de tels corpus est inaccessible à des PME. De plus, quelle que soit la taille du corpus, les divers accents régionaux sont peu représentés par rapport à la forme dominante de la langue parlée, ce qui se traduit par une perte de qualité de la RAP pour les personnes parlant ces accents.

Diverses méthodes ont été développées récemment pour traiter ce problème [2, 3]. En particulier, l'approche d'adaptation de modèle développée par Multispeech réduit le taux d'erreur jusqu'à 30% pour la parole anglaise accentuée [4].

Sujet

L'objectif du stage est d'augmenter la robustesse de la RAP aux accents régionaux de la langue française. Dans un premier temps, le/la stagiaire évaluera l'approche [4] sur la parole française accentuée. Dans un deuxième temps, il/elle cherchera à l'améliorer en augmentant le lexique (qui décrit les prononciations possibles de chaque mot) par l'ajout de règles issues de l'étude phonétique de la langue [5] ou extraites de façon supervisée [6] de corpus non-propriétaires dédiés à l'étude des accents tels que PFC [7]. Il comparera ces règles à celles pouvant être extraites de façon non-supervisée du corpus Mozilla Common Voice [8], qui comporte 800 h de parole transcrite sous forme textuelle mais dont les divers accents ne sont pas annotés. Ce lexique pourra éventuellement être spécifique à chaque locuteur. Dans un troisième temps, il/elle cherchera à généraliser cette approche à des systèmes de RAP basés sur les réseaux de neurones de bout-en-bout.

Cadre du travail

Ce stage s'inscrit dans le cadre de la Chaire Industrielle IVOIRE entre l'équipe Inria/LORIA Multispeech et la société Vivoka. Fondée en 2015 et récompensée par deux CES Innovation Awards, Vivoka a créé et commercialise le Voice Development Kit (VDK), la toute première solution permettant à une entreprise de concevoir une interface vocale de manière simple, autonome et rapide. Qui plus est, cette interface est embarquée : elle peut être déployée sur des terminaux sans connexion Internet et préserve totalement la vie privée. La Chaire IVOIRE vise à optimiser cette technologie afin de rivaliser avec les technologies actuelles les plus performantes.

Le/la stagiaire sera co-encadré(e) par Emmanuel Vincent (Directeur de Recherche Inria), Anne Bonneau (Chargée de Recherche CNRS) et un membre de l'équipe R&D de Vivoka. Il/elle aura l'opportunité de passer un ou plusieurs jours par mois au sein de l'équipe R&D de Vivoka. Il/elle bénéficiera ainsi à la fois des compétences de l'équipe Multispeech sur la RAP et la phonétique et de celles de Vivoka sur les besoins des utilisateurs.

Références

- [1] W. Xiong, J. Droppo, X. Huang, F. Seide, M. L. Seltzer, A. Stolcke, D. Yu and G. Zweig. Toward human parity in conversational speech recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 25(12) :2410–2423, 2017.
- [2] X. Yang, K. Audhkhasi, A. Rosenberg, S. Thomas, B. Ramabhadran, and M. Hasegawa-Johnson. Joint modeling of accents and acoustics for multi-accent speech recognition. *2018 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, pp. 5989–5993, 2018.
- [3] S. Sun, C.-F. Yeh, M.-Y. Hwang, M. Ostendorf, and L. Xie. Domain adversarial training for accented speech recognition. *2018 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, pp. 4854–4858, 2018.
- [4] M. A. T. Turan, E. Vincent, and D. Jouvét. Achieving multi-accent ASR via unsupervised acoustic model adaptation. *Interspeech*, pp. 1286–1290, 2020.
- [5] C. van Bael and S. King. An accent-independent lexicon for automatic speech recognition. *International Congress of Phonetic Sciences (ICPhS)*, pp. 1165–1168, 2003.
- [6] U. Nallasamy, M. Garbus, F. Metze, Q. Jin, T. Schaaf, and T. Schultz. Analysis of dialectal influence in pan-arabic ASR. *Interspeech*, pp. 1721–1724, 2011.
- [7] <https://www.projet-pfc.net/>
- [8] <https://commonvoice.mozilla.org/>

Sujet 15

Continuous 3D Sign language Generation

Proposé par : Seyed Ahmad Hosseini, Slim Ouni

Equipe : MULTISPEECH

Informations générales

Encadrants	Seyed Ahmad Hosseini, Slim Ouni
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Téléphone	03 54 95 84 06
Email	seyed-ahmad.hosseini@loria.fr slim.ouni@loria.fr
Bureau	C 140

Motivations

Sign language is the language of communication for the Deaf community, a rich visual language with complex grammatical structures. Sign language translation aims to improve communication between signers and non-signers while allowing each party to use their preferred language. As with any other natural language, sign languages have their own unique linguistic and grammatical structures, which often do not have a one-to-one mapping to their spoken language counterparts. As such, this problem truly represents a machine translation task. Prior work on Sign Language Translation (SLT) improves the translation performance by having a mid-level sign gloss tokenization [1]. More recent work proposed Transformers architectures to translate from spoken language sentences to continuous sign pose sequences in an end-to-end manner [2]. Saunders et al. [3] proposed separating the sign language production task into two distinct jointly-trained sub-tasks. The first translation subtask translates from spoken language to sign language representation, with explicit gloss supervision. The animation sub-task produces expressive sign language sequences that closely resemble the gloss representation.

Sujet

According to the motivation of the work, the goal of the internship has a multidisciplinary nature defining the different facades of the task such as Computer Vision (CV) mainly for corpus acquisition, Neural Machine Translation (NMT) for the translation task and 3D real-time digital human interpreters. In this internship, the focus will be on two major objectives :

Objective 1 : Neural networks benefit from large quantities of labeled training data. However, in many settings labeled data is much harder to come by than unlabeled data. Therefore, the Master internship will be investigating the State-of-the-art pose estimation approaches of sign interpretations from broadcast footage since it can be considered as alternate sources of data [4] [5]. Pose Estimation is the task of detecting the human figures in images or video. Sign languages are visual languages produced by the movement of hands, face and body. Theoretically, skeletal poses contain all the relevant information required to understand signs produced in videos. Given the variety and quantity of the signed content, they are an ideal resource for training sign language translation models that generalize to large domains of discourse. During the internship, we will perform human pose estimation methods on already collected anonymized news footage with picture-in-picture sign language interpretations to obtain human skeletal features, namely 3D Skeleton.

Objective 2 : In this work, we will use these skeletal features as spatial embedding to train Seq2Seq [1] or Transformer networks [2] from the text to sign translation task and investigating aspects of unsupervised methods for speech to sign are further assignments of the internship.

Prérequis

A good practice in Python or C++, good knowledge of theoretical and practical machine learning, and a first experience with deep-learning frameworks are recommended. Some knowledge of computer games engines such as unreal engine is a plus.

Références

- [1] Necati Cihan Camgoez, Simon Hadfield, Oscar Koller, Hermann Ney, Richard Bowden. *Neural Sign Language Translation*. In IEEE Conference on computer vision and pattern recognition (CVPR), 2018.
- [2] Ben Saunders, Necati Cihan Camgoz, and Richard Bowden. *Progressive Transformers for End-to-End Sign Language Production*. Computer Vision (ECCV), 2020.
- [3] Ben Saunders, Necati Cihan Camgoez, Richard Bowden. *Mixed SIGNALs : Sign Language Production via a Mixture of Motion Primitives*. 2021.
- [4] Helen Cooper and Richard Bowden. *Learning Signs from subtitles : A Weakly Supervised Approach to Sign Language Recognition*. 2019.
- [5] Necati Cihan Camgoz, Ben Saunders, Guillaume Rochette, Marco Giovanelli, Giacomo Inches, Robin Nachtrab-Ribback, Richard Bowden *Content4All Open Research Sign Language Translation Datasets*. 2021.

Équipe Neurorhythms

<http://neurorhythms.loria.fr/>

Sujet 1

Détection adaptative de nouveautés : application à la caractérisation de l'activité cérébrale

Proposé par : Laurent BOUGRAIN

Equipe : Neurorhythms

Informations générales

Encadrants	Laurent Bougrain
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Téléphone	03 83 59 20 54
Email	laurent.bougrain@loria.fr
Bureau	C045

Motivations

Les interfaces cerveau-ordinateur actuelles [1] (Brain-Computer Interfaces ou BCI en anglais) requièrent une calibration i.e., un apprentissage du classifieur spécifique pour chaque utilisateur car les caractéristiques de l'activité électroencéphalographique (EEG) sont très variables d'une personne à l'autre. Pour des BCI basées sur l'activité cérébrale motrice, cette phase de calibration nécessite habituellement des signaux EEG lors d'intentions de mouvements (classe 1) et de repos (classe 2) pour l'utilisateur cible. Or, il n'est pas toujours possible pour des questions de temps, de coût, de fatigue ou de contexte d'obtenir des facilement des exemples d'une classe. Nous ne pouvons donc pas dans ce cas utiliser les méthodes classiques d'apprentissage puisque nous n'aurons pas à l'avance les données des deux classes du sujet.

Sujet

Dans ce projet, nous allons explorer les approches d'apprentissage dites one-class [2] où l'idée va être de construire un modèle de reconnaissance décrivant les caractéristiques d'une classe (ici les propriétés d'une classe de signaux EEG). L'autre classe sera ensuite identifiée par détection de rupture (ici lors d'un changement de propriété des signaux EEG de la classe 1). De plus, nous étudierons les approches d'apprentissages adaptatives [3], où l'idée va être de construire de manière adaptative un modèle, en partant de zéro et en l'améliorant au fur et à mesure de l'acquisition de nouvelles données.

Cadre du travail

Le travail se déroulera au LORIA dans l'équipe Neurorhythms qui travaille sur la modélisation et l'analyse expérimentale des rythmes neuronaux normaux et pathologiques pour les applications médicales et la neurorobotique.

Références

- [1] M. Clerc, L. Bougrain, F. Lotte. *Brain-computer interfaces. Vol 1, foundations and methods*. Wiley, 2016.
- [2] L. Ruff et al. *A Unifying Review of Deep and Shallow Anomaly Detection*. Proc. IEEE, vol. PP, p. 1-40, févr. 2021.
- [3] S. Kumar, F. Yger, et F. Lotte. *Towards Adaptive Classification using Riemannian Geometry approaches in Brain-Computer Interfaces*. IEEE International Winter Conference on Brain-Computer Interfaces, Jeon-seong, South Korea, 2019.

Équipe Orpailleur

<http://orpailleur.loria.fr/>

Sujet 1

A First Study in Causal Formal Concept Analysis

Proposé par : Alexandre Bazin, Miguel Couceiro and Amedeo Napoli

Equipe : Orpailleur

Informations générales

Encadrants	Alexandre Bazin, Miguel Couceiro and Amedeo Napoli
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Email	{alexandre.bazin,miguel.couceiro,amedeo.napoli}@loria.fr
Équipe	Orpailleur (https://orpailleur.loria.fr/)

Motivation

The study of causal relations is central in applied science [5]. Experimental protocols are often implemented in order to manipulate a potential cause –an object, a state or a process– in such a way that the effects of such a cause can be inferred. However, when objects cannot be directly manipulated, one has to rely on purely observational data. For example, in medicine, a physician examining a patient has only access to observational data. Then the physician tries to connect a set of symptoms and a possible disease, e.g., if the patient has fever –the effect– the physician will search for a potential infection –the cause.

Studying causality from data involves two related tasks : (i) discovering a “potential causal structure” relying on a (set of) causal relations between two sets of variables, (ii) inferring causal effects for understanding how does the cause lead to the effect. Inferring causal relations from observational data is a challenging task. Most of the approaches in the literature are probabilistic [8, 9], while some recent algorithmic approaches based on approximations of Kolmogorov complexity are gaining attention [2, 3, 6, 7].

A set of causal relations forms the causal structure. In the following, we are interested in the discovery of a potential causal structure in observational data. A causal structure can be presented to humans to help them understand a situation described by data and support decision making or it can be used to automatically select important variables in data [10]. Moreover, most of the work about causality focuses on univariate *causal inference* where the task consists in determining whether a causal relation exists between two variables x and y , and deciding its direction, i.e., what is the *cause* and what is the *effect*. In such a case the causal structure is represented by a directed acyclic graph called the *causal diagram*.

However, in practice effects involve an interaction between multiple causes, leading to *multivariate* causal relations. In such a case, causal relations are existing between sets X and Y of variables. The corresponding causal diagrams should to be properly formalized and studied.

Master Thesis Subject

In this Master Thesis, we are interested in studying how a causal structure can be formalized with “Formal Concept Analysis” (FCA), a mathematical framework used in classification and knowledge discovery [4]. FCA is applied to numerical data table with objects in rows and attributes in columns, and outputs a set of so-called formal concepts that can be ordered within a concept lattice. Moreover, FCA is particularly well-suited to the discovery and representation of rules and implications.

In [1], a first formalization of causality in FCA is proposed, where it is shown how to deal with the concept lattice and implications as potential candidates for representing the causal diagrams and causal relations. The objective of this Master Thesis is to rely on this first formalization and to study the related algorithms and examples. Then the student will propose possible new algorithms for dealing with causality, complete the existing implementation, and carry out experiments.

Environment

The research work will be carried out at the LORIA Lab in the Orpailleur Team. The supervisors of the student have a good experience in knowledge discovery, Formal Concept Analysis, and in the study of causality.

Moreover, the team is involved in a research project, namely HYAIAI, where causality is one of the main research subject at the moment.

Références

- [1] A. Bazin, M. Couceiro, M.-D. Devignes, and A. Napoli. Steps Towards Causal Formal Concept Analysis. *International Journal on Approximate Reasoning*, 2021. Submitted.
- [2] K. Budhathoki and J. Vreeken. MDL for causal inference on discrete data. In *2017 IEEE International Conference on Data Mining (ICDM)*, pages 751–756. IEEE, 2017.
- [3] K. Budhathoki and J. Vreeken. Origo : causal inference by compression. *Knowledge and Information Systems*, 56(2) :285–307, 2018.
- [4] B. Ganter and R. Wille. *Formal Concept Analysis : Mathematical Foundations*. Springer, 1999.
- [5] S. Kleinberg. *Why : A guide to finding and using causes*. O’Reilly Media, 2015.
- [6] A. Marx and J. Vreeken. Identifiability of cause and effect using regularized regression. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 852–861, 2019.
- [7] O. Mian, A. Marx, and J. Vreeken. Discovering fully oriented causal networks. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 2021.
- [8] J. M. Mooij, J. Peters, D. Janzing, J. Zscheischler, and B. Schölkopf. Distinguishing cause from effect using observational data : methods and benchmarks. *The Journal of Machine Learning Research*, 17(1) :1103–1204, 2016.
- [9] E. Sgouritsa, D. Janzing, P. Hennig, and B. Schölkopf. Inference of cause and effect with unsupervised inverse regression. In *Artificial intelligence and statistics*, pages 847–855, 2015.
- [10] K. Yu, X. Guo, L. Liu, J. Li, H. Wang, Z. Ling, and X. Wu. Causality-based feature selection : Methods and evaluations. *ACM Computing Surveys*, 53(5) :1–36, 2020.

Sujet 2

Mining Molecular Structures for Discovering Antibiotics and Extensions

Proposé par : Miguel Couceiro, Bernard Maigret, Amedeo Napoli et Claire Theobald

Equipe : Orpailleur

Informations générales

Encadrants Miguel Couceiro, Bernard Maigret, Amedeo Napoli, Claire Theobald
Adresse LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Email {miguel.couceiro, bernard.maigret, amedeo.napoli, claire.theobald}@loria.fr
Équipe Orpailleur (<https://orpailleur.loria.fr/>)

Motivation

Nowadays, there are crucial needs in the research and design of new antibiotics for facing the growing resistance of bacteria against standard antibiotics⁸. For fighting such resistance to antibiotics, one should be able to deal with so-called “active molecules” while, at the same time, only a few active molecules are currently designed. Actually, most of the antibiotics are deriving from well-known antibiotics structures, and a drug based on such structures will be faced to resistance problems more or less quickly. Then it is necessary to work in a different way in taking advantage of recent advances in machine learning, and in particular in deep learning.

A common hard point in all drug methodologies is to be able to deal with very large amounts of molecular data in an efficient way w.r.t. time and space complexity. It should be noticed here that molecular structures are usually represented as labeled graphs. In addition, domain knowledge should be used to deal with such complex data structures. Then an objective is to be able to detect among large sets of original molecules which are the molecular structures –as whole molecules or parts of molecules– that can be identified to be reused as valuable antibiotics with reasonable computational resources.

Numerical classifiers especially based on deep learning methods are good candidates to be tested and used for dealing with the search for new antibiotics. Indeed such classifiers are rather efficient and can deal with large and complex numerical data but also with graph data. These classifiers are based on supervised learning and can be trained with existing bases of known antibiotics. They will be aimed at identifying “target substructures” in a collection of new molecular candidates showing a potential antibiotic activity. Such substructures can be then reused in drug design for building a new generation of antibiotics. One first challenge is to adapt numerical classifiers to deal with these complex structures to identify potential active substructures.

However, such a classification process cannot be carried out without taking into account domain knowledge, and in this case, knowledge units about chemical structures, drug design, and antibiotics structures. This is a second challenge that should be faced, to be able to work with numerical classifiers guided by chemical domain knowledge.

Finally, a third challenge is related to the “explanations” of the results which is a very important aspect in such a classification process. We will then follow recent work on the production of explanations [6, 1] for providing the specialists in drug design potential “molecular explanations” in terms of atoms, bonds, substructures. . . explaining why a given molecular substructure can be used as a support for building an antibiotics.

Master Thesis Subject

Two preceding and recent research works will also be used as a basis for this Master Thesis. The first one [2] presents the basic problem of molecule classification using two modern classifiers, namely DeepChem⁹ [4, 5] and ChemProp [8]. This research report will be used a guide for studying and adapting these classifiers for dealing with the molecular databases. In particular, we will be interested in mining real-world data and two main repositories, MDDR¹⁰ and PubChem¹¹.

8. See <http://www.who.int/mediacentre/news/releases/2017/bacteriaantibiotics-needed/fr/>

9. <https://deepchem.io/>

10. MDL Drug Data Report : <http://www.akosgmbh.de/accelrys/databases/mddr.htm>

11. The PubChem Project : <https://pubchem.ncbi.nlm.nih.gov/>

The second Master Thesis [3] is more about the production of explanations and fairness of algorithms. In his work, the author has developed an interface prototype where the potential antibiotics substructures can be visualized and “compound explanations” are provided in graphical terms, where for example, atoms and bonds responsible of some classification aspects are highlighted positively or negatively.

There are several objectives in this Master Thesis. The first objective is to study the existing work and then to extend these work, in a theoretical and practical way, for example in dealing with new molecular databases and using alternative explainers.

Moreover, a new experiment will also be carried out on the recognition of galaxies in astronomical data [7]. The main idea here is to understand how the explanation process about classification can be reused in another context, here the recognition of galaxies in astronomical images, and to take advantage of the explanations for improving the work of the classifiers.

Environment

The research work will be carried out at the LORIA Lab in the Orpailleur Team. The supervisors of the student have a good experience in classification and classifiers, in knowledge discovery, in chemistry and drug design.

Références

- [1] G. Alves, M. Amblard, F. Bernier, M. Couceiro, and A. Napoli. Reducing Unintended Bias of ML Models on Tabular and Textual Data. In *8th IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, pages 1–10. IEEE, 2021.
- [2] C. Bellanger. Fouille de données moléculaires. Master’s thesis, Université de Lorraine, 2021.
- [3] F. Bernier. A Study about Explainability and Fairness in Machine Learning and Knowledge Discovery. Master’s thesis, Nancy-Telecom, Université de Lorraine, 2021.
- [4] S. Chan, H. Shan, T. Dahoun, H. Vogel, and S. Yuan. Advancing Drug Discovery via Artificial Intelligence. *Trends in Pharmacological Sciences*, 40(8) :592–604, 2019.
- [5] B. Ramsundar, P. Eastman, P. Walters, V. Pande, K. Leswing, and Z. Wu. *Deep Learning for the Life Sciences*. O’Reilly Media, 2019.
- [6] C. Rudin, C. Chen, Z. Chen, H. Huang, L. Semenova, and C. Zhong. Interpretable Machine Learning : Fundamental Principles and 10 Grand Challenges. *CoRR*, abs/2103.11251, 2021.
- [7] C. Theobald, B. Arcelin, F. Pennerath, B. Conan-Guez, M. Couceiro, and A. Napoli. A Bayesian Convolutional Neural Network for Robust Galaxy Ellipticity Regression. In *Proceedings of the Machine Learning and Knowledge Discovery in Databases European Conference (ECML-PKDD)*, Lecture Notes in Computer Science 12979, pages 135–150. Springer, 2021.
- [8] K. Yang, K. Swanson, W. Jin, C. W. Coley, P. Eiden, H. Gao, A. Guzman-Perez, T. Hopper, B. Kelley, M. Mathea, A. Palmer, V. Settels, T. S. Jaakkola, K. F. Jensen, and R. Barzilay. Analyzing Learned Molecular Representations for Property Prediction. *Journal of Chemical Information and Modeling*, 59(8) :3370–3388, 2019.

Sujet 3

New Research Directions in Numerical Pattern Mining

Proposé par : Miguel Couceiro, Tatiana Makhalova and Amedeo Napoli

Equipe : Orpailleur

Informations générales

Encadrants	Miguel Couceiro, Tatiana Makhalova and Amedeo Napoli
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Email	{miguel.couceiro, tatiana.makhalova, amedeo.napoli}@loria.fr
Équipe	Orpailleur (https://orpailleur.loria.fr/)

Motivation

The objective of pattern mining is to discover a small set of interesting patterns that describe together a large portion of a dataset and can be easily interpreted and reused. Actually pattern mining encompasses a large variety of algorithms in knowledge discovery and data mining aimed at analyzing datasets [5]. Present approaches in pattern mining are aimed at discovering an interesting *pattern set* rather than a set of individually interesting patterns, where the quality of patterns is evaluated w.r.t. both the dated and other patterns. One common theoretical basis for pattern set mining relies on the Minimum Description Length principle (MDL), which is applied to many types of patterns and in particular to itemsets [6].

In this Master Thesis subject, we are interested in numerical pattern mining, which is aimed at analyzing numerical datasets. To date, one of the most common ways to mine numerical pattern sets relies on the application of itemset mining to binarized datasets. In parallel, clustering algorithms have been extensively developed, and many different and efficient approaches have been proposed. However, numerical pattern mining is more interested in the description of a group of objects in terms of a set of attributes related to these objects, while clustering focuses more on the detection of these groups of objects based on their commonalities as measured by a similarity or a distance.

Contrasting the recent advances in pattern mining and clustering, algorithms for mining numerical data appear to be insufficiently explored. This leads us to study and to revisit this problem and to propose a new algorithm, namely MINT for “Mining INTeresting Numerical Pattern Sets” [1].

Master Thesis Subject

In this Master Thesis subject, we propose to study an MDL-based approach to numerical pattern set mining based on MINT. MINT computes numerical patterns as m -dimensional hyper-rectangles which are products of m intervals, where the intervals are related to the attributes and their values. MINT does not need to explore the pattern space in advance and the complexity of the algorithm is low. Moreover, MINT is based on MDL and outputs a small set of non-redundant informative patterns, while a series of experiments shows that MINT is efficient and outputs sets of patterns of very high quality (diversity, non-redundancy). In addition, MINT is most of the time more efficient and outputs more concise and better fitting patterns, when compared to competitors IPD [2], REALKRIMP [7], and SLIM [3].

Thus, in this Master Thesis subject, we are interested in first studying the basics of the MINT and second to improve its implementation, to complete and extend the experiments already carried out. In particular, the present code could be improved if combined with other very efficient codes developed for itemset mining such as those introduced in [4].

Environment

The research work will be carried out at the LORIA Lab in the Orpailleur Team. The supervisors of the student have a good experience in knowledge discovery, pattern mining, and clustering, which are directly related to this Master Thesis subject.

Références

- [1] T. Makhalova, S. O. Kuznetsov, and A. Napoli. Mint : MDL-based approach for Mining INTeresting Numerical Pattern Sets. *Data Mining and Knowledge Discovery*, 2021. To appear.
- [2] H.-V. Nguyen, E. Müller, J. Vreeken, and K. Böhm. Unsupervised interaction-preserving discretization of multivariate data. *Data Mining and Knowledge Discovery*, 28(5-6) :1366–1397, 2014.
- [3] K. Smets and J. Vreeken. Slim : Directly Mining Descriptive Patterns. In *Proceedings of SDM Conference*, pages 236–247. SIAM, 2012.
- [4] L. Szathmary, P. Valtchev, A. Napoli, R. Godin, A. Boc, and V. Makarenkov. A fast compound algorithm for mining generators, closed itemsets, and computing links between equivalence classes. *Annals of Mathematics and Artificial Intelligence*, 70(1–2) :81–105, 2014.
- [5] J. Vreeken and N. Tatti. Interesting patterns. In C. C. Aggarwal and J. Han, editors, *Frequent Pattern Mining*, pages 105–134. Springer, 2014.
- [6] J. Vreeken, M. Van Leeuwen, and A. Siebes. Krimp : mining itemsets that compress. *Data Mining and Knowledge Discovery*, 23(1) :169–214, 2011.
- [7] J. Witteveen, W. Duivesteijn, A. Knobbe, and P. Grünwald. RealKrimp – Finding Hyperintervals that Compress with MDL for Real-Valued Data. In *Proceedings of the International Symposium on Intelligent Data Analysis (IDA)*, LNCS 8819, pages 368–379. Springer, 2014.

Sujet 4

Predicting biomolecule properties using machine learning approaches

Proposé par : Yannick Toussaint and Kira Weissman

Equipe : Orpailleur

Informations générales

Encadrants	Yannick Toussaint	Kira Weissman
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy	Molecular and Structural Enzymology Group UMR 7365 CNRS-UL : IMoPA
Téléphone	03 83 59 20 91	+33 3 72 74 66 63
Email	yannick.toussaint@loria.fr	kira.weissman@univ-lorraine.fr

Motivation and context

Over the last several years, European nations have jointly agreed to prioritize the transition towards a more innovative, resource-efficient economy. The overall aim is to establish a low-emissions economic system which reconciles food security, demands for sustainable agriculture and fisheries, and the use of renewable biological resources for industrial purposes, while simultaneously preserving biodiversity and increasing the health of the environment.

In this context, the Biomolecules project from the "Lorraine Université d'Excellence" initiative aims to develop an extensive and ambitious public/private collaborative research program anchored in biomolecules, which directly responds to local, national and European Green Deal priorities.

In this ambitious context, one major aspect of this program is related to the identification of biomolecule properties which, of course, condition their uses in certain domain. Many initiatives already built datasets of molecules and that may be used for training learning algorithms. So data science may help in efficiently predicting molecule properties so that biologists more easily target their experiments.

This master's thesis aims at using machine learning approaches for predicting biomolecule properties : anti-biotic, antioxydant...but also much more precise prediction on which population of bacteria resist or not...

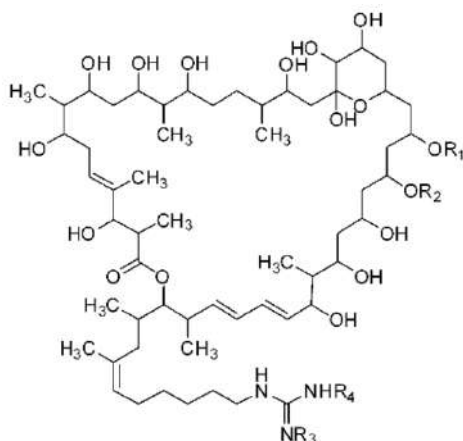


FIGURE 11 – Chemical structures of 40-membered Guanidine containing polyhydroxyl macrolides

Objectives

The goal of the thesis is to predict biomolecule properties using machine learning. The work will be organised into the following steps :

- Reading few papers to acquire some background in biology, molecule, properties, structure. The challenging part of this work comes from the complexity of molecules : complexity in term of structural organisation of the molecule as different representations can be considered (1-D, 2-D or 3-D), in term of composition (some components are well-known to be linked with some properties) . .
- State of the art on machine learning approches for molecule properties prediction : rules, random forest and deep-learning. As the representation of the molecule as a graph is a pretty natural way, graph embeddings suit well the structural organisation of a mollecule. Many attempts in using graph neural network architectures such as convolution graph neural networks or recurrent graph neural networks give promising results.
- Building a list of free/non-free resources to train learning systems (from papers). A subset of molecules will be selected for training and testing the results, in accordance with our local expertise expertise and goal.
- Comparing and then selecting an encoding approach for data in order to feed Machine learning algorithms.
- Experiments and analysis of results.

LORIA will bring its expertise on Machine learning. The IMopa lab will bring its expertise on biology on the question of representing or encoding a molecule and for analysing of the results.

The student will be located at LORIA (Nancy) and a meeting will take place every week between the two laboratories.

Important : Before applying, students interested by this subject should absolutly send a mail with a curriculum vitae to Yannick Toussaint. Having the experience of at least one course in deep learning is required and a real motivation for this field is expected. They should also be fluent in reading and writing english.

Références

- [1] R Agrawal, T Imielski, and A Swami. Mining association rules between sets of items in large databases. In *proceedings of the ACM SIGMOD Int'l Conference on Management of Data*, pages 207–216, 1993.
- [2] Fabrice Guillet and Howard J. Hamilton. *Quality Measures in Data Mining (Studies in Computational Intelligence)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2007.
- [3] François SÁ©verac, Erik A. Sauleau, Nicolas Meyer, Hassina LefÁ©vre, Gabriel Nisand, and Nicolas Jay. Non-redundant association rules between diseases and medications : an automated method for knowledge base construction. *BMC Medical Informatics and Decision Making*, 15 :29, 2015.

Équipe SIMBIOT

[http ://simbiot.loria.fr/](http://simbiot.loria.fr/)

Sujet 1

Étude de l'influence du PRNG utilisé dans un processus d'apprentissage

Proposé par : S. Contassot-Vivier, J-F. Couchot

Equipe : SIMBIOT

Informations générales

Encadrants	S. Contassot-Vivier, J-F. Couchot
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Email	contasss@loria.fr jean-francois.couchot@univ-fcomte.fr
Mots clés	Générateurs de nombres pseudo-aléatoires, apprentissage automatique

Motivations

Un générateur de nombres pseudo-aléatoires, (PRNG en anglais) est un algorithme qui peut produire une séquence de nombres dont les propriétés sont proches de celles des nombres générés aléatoirement (à l'aide de dés non pipés par exemple). Les nombres doivent être par exemple indépendants les uns des autres, leur observation ne doit pas permettre de deviner quelle sera la composition d'une séquence ultérieure. Les PRNGs sont utilisés dans de nombreuses applications et notamment en apprentissage automatique.

Dans ce type d'applications, l'influence du choix du PRNG sur la qualité des résultats finaux est rarement évaluée. Intuitivement, on suppose que ce choix peut influencer la qualité finale des réseaux obtenus (remplacement récent de MTGP32 [1] par Philox [2] dans PyTorch). Mais qu'en est-il précisément ?

Quelques études ont abordé ce point mais dans des contextes un peu différents [3],[4].

Certains PRNGs [5],[6],[7] sont extrêmement rapides à fournir une suite de bits. D'autres sont cryptographiquement sûrs [8],[9], c'est à dire qu'avec une puissance de calcul raisonnable il n'est pas possible de trouver le $n + 1$ ème bit si l'on ne connaît que les n premiers bits. Enfin certains passent les tests statistiques les plus durs comme le test TestU01 [10]. Rares sont ceux qui arrivent à posséder les trois propriétés. Le laboratoire FEMTO-ST et le LORIA ont une forte expérience de recherche sur les PRNGs [11],[12],[13],[14]. Notamment, plusieurs travaux ont été menés sur la conception et la mise en œuvre de générateurs pseudo-aléatoires rapides et robustes sur CPU et FPGA.

Dans ce contexte, nous souhaitons étudier l'influence de la qualité statistique d'un PRNG dans un processus d'apprentissage automatique.

Sujet

Le premier objectif de ce stage est d'établir une bibliographie sur les PRNG couramment utilisés dans les processus d'apprentissage.

Le second objectif est de choisir un exemple d'apprentissage relativement simple, dans lequel le PRNG interne pourra facilement être remplacé par un autre.

Le troisième objectif sera de sélectionner un ensemble de PRNG à comparer.

Enfin, le quatrième objectif sera de définir et mettre en œuvre un protocole expérimental permettant d'évaluer l'influence du PRNG interne sur la qualité de l'apprentissage obtenu.

Cadre du travail et profil attendu

Le travail demandé se déroulera dans l'équipe Simbiot du Loria. Les développements seront réalisés sous environnement Linux. Les langages utilisés seront python ou C/C++ selon l'outil d'apprentissage choisi.

Références

- [1] Mutsuo Saito and Makoto Matsumoto. Variants of mersenne twister suitable for graphic processors. *ACM Trans. Math. Softw.*, 39(2) :12 :1 12 :20, 2013.
- [2] John K. Salmon, Mark A. Moraes, Ron O. Dror, and David E. Shaw. Parallel random numbers : as easy as 1, 2, 3. In Scott Lathrop, Jim Costa, and William Kramer, editors, *Conference on High Performance Computing Networking, Storage and Analysis, SC 2011, Seattle, WA, USA, November 12-18, 2011*, pages 16 :1 16 :12. ACM, 2011.
- [3] M. Hussain and N. Fujimoto. Effect of the Pseudorandom Number Generators on the Standard Particle Swarm Optimization on a GPU. In *International Conference on Computational Science and Computational Intelligence 2018*. IEEE Xplore.

- [4] A. Augusto Alves, Anton Poctarev, Ralf Ulrich. Counter-based pseudorandom number generators for COR-SIKA 8 - A multi-thread friendly approach. *EPJ Web Conf.* 251 03039 (2021). DOI:10.1051/epjconf/202125103039
- [5] Pierre L'Ecuyer. Maximally equidistributed combined tausworthe generators. *Mathematics of Computation of the American Mathematical Society*, 65(213) :203–213, 1996.
- [6] Makoto Matsumoto and Takuji Nishimura. Mersenne twister : a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 8(1) :3–30, 1998.
- [7] Pierre L'Ecuyer. Tables of maximally equidistributed combined lfsr generators. *Mathematics of Computation of the American Mathematical Society*, 68(225) :261–269, 1999.
- [8] Lenore Blum, Manuel Blum, and Mike Shub. A simple unpredictable pseudo-random number generator. *SIAM Journal on computing*, 15(2) :364–383, 1986.
- [9] Robert McEvoy, James Curran, Paul Cotter, and Colin Murphy. *Fortuna : cryptographically secure pseudo-random number generation in software and hardware.* 2006.
- [10] Pierre L'Ecuyer and Richard Simard. Testu01 : A c library for empirical testing of random number generators. *ACM Trans. Math. Softw.*, 33(4) :22 :1–22 :40, August 2007.
- [11] Mohammed Bakiri, Christophe Guyeux, Jean-François Couchot, and Abdelkrim Kamel Oudjida. Survey on hardware implementation of random number generators on FPGA : theory and experimental analyses. *Computer Science Review*, 27 :135–153, 2018.
- [12] Mohammed Bakiri, Jean-François Couchot, and Christophe Guyeux. CIPRNG : A VLSI family of chaotic iterations post-processings for \mathbb{F}_2 -linear pseudorandom number generation based on zynq mpsoc. *IEEE Trans. on Circuits and Systems*, 65-I(5) :1628–1641, 2018.
- [13] Sylvain Contassot-Vivier, Jean-François Couchot, Mohammed Bakiri, Pierre-Cyrille Héam. Fast and robust PRNGs based on jumps in N-cubes for simulation, but not exclusively for that. The 2019 International Conference on High Performance Computing & Simulation, Jul 2019, Dublin, Ireland.
- [14] Sylvain Contassot-Vivier, Jean-François Couchot, Christophe Guyeux, and Pierre-Cyrille Héam. Random walk in a n-cube without hamiltonian cycle to chaotic pseudorandom number generation : Theoretical and practical considerations. *I. J. Bifurcation and Chaos*, 27(1) :1–18, 2017.
- [15] Melissa E O'Neill. Pcg : A family of simple fast space-efficient statistically good algorithms for random number generation. *ACM Transactions on Mathematical Software*, 2014.

Sujet 2

Simulateur de conduite interactif pour la conception d'un véhicule électrique

Proposé par : Vincent CHEVRIER

Equipe : SIMBIOT

Informations générales

Encadrants	Vincent Chevrier,
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Téléphone	03 54 95 84 06
Email	<code>prenom.nom@loria.fr</code>
Bureau	C 025

Contexte : le projet HY2CAR - Hybridized Hydrogen Car

Le transport routier est un des principaux secteurs de consommation d'énergie et d'émissions gazeuses et thermiques, en particulier, le transport personnel urbain et péri-urbain. Le projet Hybridized Hydrogen Car (HY2CAR) a pour ambition de mettre au point un véhicule personnel adapté à ce mode de transport moins gourmand en énergie que la plupart des véhicules électriques actuels. Les émissions de ce véhicule, dont le coût doit rester abordable, seront ainsi réduites.

Le concept du projet repose sur l'hybridation directe d'une pile à combustible de 10 kW par des supercapacités (21 kW), correspondant à une puissance maximale de la source hybride formée de 31 kW, par gestion partagée de la puissance selon la demande énergétique du véhicule. Le concept, adapté à des trajets courts en péri-urbain, a été validé à échelle réduite par modélisation et essais par les demandeurs. L'ambition à terme du projet est de mettre au point le véhicule électrique depuis la conception de la source hybride de 31 kW, son implantation dans le véhicule et la gestion optimale des demandes de puissance en fonction du trajet, avant essais sur route. Le consortium regroupant des équipes des sciences de l'ingénieur, humaines et sociales, et économiques, évaluera également l'accessibilité de HY2CAR par le citoyen, la viabilité économique de la filière ainsi que les risques liés au fonctionnement du véhicule (sécurité, émissions etc.).

La puissance étant faible, il faut donc la gérer au mieux et notamment anticiper les phases demandant de la puissance. Cette gestion dépend de l'état de charge de la source électrique (et donc du comportement passé du véhicule), du trajet et du profil routier (phase d'accélération, arrêt, pente de la route, etc), et du comportement du conducteur (un conducteur « nerveux » n'utilisera pas la puissance de la voiture de la même manière qu'un conducteur « calme »). Nous avons déjà mis en place l'acquisition de données relatives au trajet effectués sous forme de trace GPS, nous disposons d'un modèle de comportement de la source électrique et d'un modèle physique simple du véhicule.

Cette année, nous souhaitons explorer la piste de simulation interactive pour intégrer le comportement du conducteur dans l'étude et la gestion de l'énergie du véhicule.

Sujet

Il s'agit de mettre en place un simulateur de conduite interactif adapté aux besoins du projet HY2Car. En effet, la simulation est une alternative attractive à l'expérimentation réelle ne serait-ce que pour des raisons de coût, de temps, de sécurité, etc. Nous souhaitons donc un simulateur de conduite qui reproduise les caractéristiques essentielles du véhicule de test réel dans des situations variées mais aussi en reprenant des données capturées sur le véhicule réel (refaire en simulation un parcours réel par exemple).

L'idée retenue est de coupler un simulateur de conduite (si possible existant) à des modèles de comportement énergétique et mécanique du véhicule et ainsi mettre en place une chaîne d'acquisition d'informations, de stockage et d'analyse. Les informations recueillies serviront d'une part à mieux comprendre le fonctionnement de la source électrique et décider d'un fonctionnement du véhicule de manière optimale et de faire des recommandations au conducteur.

Une première étape du stage vise à définir les besoins du projet, lesquels permettront de sélectionner un simulateur candidat, et à terme d'évaluer la solution construite.

Le stage s'appuyera les étapes suivantes :

- Définition (avec les membres du projet) des informations qui doivent être collectées (vis-à-vis des étapes ultérieures d'analyse)
- Définitions de critères techniques caractérisant la solution logicielle en termes d'utilisation, installation, maintenance, évolutivité, etc.
- Définition des critères souhaités quant à l'expérience conducteur (c'est-à-dire tous les aspects qui font qu'une personne se sentira à l'aise avec la simulation et qui permettront d'obtenir un comportement proche de ce qu'il serait avec une « vraie » voiture)
- Etude des simulateurs de conduites proposé de manière libre/gratuite, évaluation de leur adéquation aux besoins
- sélection d'un simulateur candidat et développement d'une solution.

Cadre du travail

Le projet Hy2Car rassemble différentes laboratoires et équipes universitaires. Le stage aura lieu au Loria dans l'équipe Simbiot (l'une des équipes partenaires du projet au LORIA. Il est à dominante informatique.

Cependant, la personne recrutée sera amenée à interagir avec certains chercheurs d'autres laboratoire que le LORIA (pour comprendre les attentes du projet, voir le véhicule réel) ou d'autres équipes du LORIA (pour intégrer le simulateur à l'existant et aux besoins).

Profil attendu

La personne retenue aura un niveau Master 2 en informatique ou équivalent, des compétences en intégrations logicielle et simulateur de conduite seront un atout.

Au vu du caractère interdisciplinaire du projet, il est souhaitable d'être curieux, de ne pas hésiter à interroger, à savoir travailler en équipe à l'interface de plusieurs domaines et de faire preuve de synthèse.

Sujet 3

Simulation and implementation of trust management (TM) in VANET

General information

Supervisors Ye-Qiong SONG, Runbo SU
Adresse LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Telephone 03 54 95 84 64
Email ye-qiong.song@loria.fr runbo.su@loria.fr
Office C 024/C027

Background

VANET (Vehicular Ad-hoc NETwork) is a wireless vehicular network that enables data to exchange between vehicles through the On-Board Unit (OBU) and Road Side Unit (RSU) infrastructures. Despite numerous security mechanisms applied in VANET, such as ITS-G5 system of PKI-based security system, malicious vehicles can propagate false information or modify/delete messages due to the dynamic nature of the VANET network [1]. In this regard, a considerable amount of trust management (TM) models have been proposed to maintain the reliability of VANET network and at the same time identify malicious intrusions. However, these models still suffer from various limitations [4] : standardized message utilization is missing, countermeasures against malicious attacks are insufficiently exploited, etc.

Description

This internship aims to develop a dynamic TM model suitable for VANET to secure the VANET system by using two types of standardized messages, namely, CAM (Cooperative Awareness Message) and DENM (Decentralised Event Notification Message) [2]. Both CAM and DENM are standardized by ETSI (European Telecommunications Standards Institute) to exchange traffic information in a cooperative manner : Each vehicle broadcasts CAM messages periodically containing its speed, position, type, acceleration, etc., to inform about the traffic situation, and DENM messages will be generated when a specific event is detected. The feasibility of the current TM models is demanding since the majority of them do not take CAM and DENM into consideration. Furthermore, as stated before, malicious attacks are often not sufficiently addressed, and thus a TM model that shows effectiveness and robustness facing attacks is needed.

The plan of action :

- Bibliographic study of TM [3] and TM in VANET [4].
- Design of a TM model that enables dynamical trust evaluation of vehicles and their behaviors.
- Numerical simulation by using NS-3 [6, 5], as well as SUMO [7].
- Implementation of proposed TM in VANET using OBU and RSU cards of neoGLS, which are (<https://www.neogls.com/>) available at LORIA.

Références

- [1] Hasrouny, Hamssa and Samhat, Abed Ellatif and Bassil, Carole and Laouiti, Anis, *VANet security challenges and solutions : A survey*, Vehicular Communications, 7,7-20,2017.
- [2] Santa, José and Pereñíguez, Fernando and Moragón, Antonio and Skarmeta, Antonio F, *Experimental evaluation of CAM and DENM messaging services in vehicular communications*, Transportation Research Part C : Emerging Technologies, 46, 18-120,2014.
- [3] Sharma, Avani and Pilli, Emmanuel S and Mazumdar, Arka P and Gera, Poonam, *Towards trustworthy Internet of Things : A survey on Trust Management applications and schemes*, Computer Communications, 2020
- [4] Hussain, Rasheed and Lee, Jooyoung and Zeadally, Sherali, *Trust in VANET : A survey of current solutions and future research opportunities*, IEEE transactions on intelligent transportation systems, 22(5), 2553-2271,2020.
- [5] Campanile, Lelio and Gribaudo, Marco and Iacono, Mauro and Marulli, Fiammetta and Mastroianni, Michele, *Computer network simulation with ns-3 : A systematic literature review* Electronics, 9(2), 272, 2020
- [6] Wikipedia Network Simulator — Wikipedia, 2020, http://fr.wikipedia.org/w/index.php?title=Network_Simulator&oldid=174772816, En ligne ; Page disponible le 16-septembre-2020

- [7] Liu, Weiwei and Wang, Xichen and Zhang, Wenli and Yang, Lin and Peng, Chao, 2016 22nd Asia-Pacific Conference on Communications (APCC), Coordinative simulation with SUMO and NS3 for Vehicular Ad Hoc Networks, 2016, 337-341.

Équipe Synalp

<http://synalp.loria.fr/>

Sujet 1

Federated Learning pour l'Industrie 4.0

Proposé par : Christophe Cerisara et Alexandre Voisin

Equipe : Synalp

Informations générales

Encadrants	Christophe Cerisara et Alexandre Voisin	
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy	
Téléphone	03 54 95 84 06	
Email	cerisara@loria.fr	alexandre.voisin@univ-lorraine.fr
Bureau	B 208	

Motivations

Aujourd'hui, les modèles de deep learning sont généralement entraînés sur des fermes de GPU, comme ceux accessibles au LORIA, à savoir les supercalculateurs Jean Zay, EXPLOR et Grid5000. Toutefois, la construction, l'utilisation et la maintenance de tels centres de calcul sont très coûteux et polluants. De plus, les données sur lesquelles le modèle est entraîné doivent être stockées au même endroit, ce qui pose des problèmes de vie privée (peut-on rassembler toutes les données médicales des français sur un serveur?) et de propriétés intellectuelles (peut-on déposer toutes les données des usines de Continental et de Michelin sur un serveur si l'on veut entraîner un modèle des procédés de fabrication des pneus?).

Une approche moderne pour résoudre partiellement ces deux problèmes consiste à laisser les données sur les serveurs de leurs propriétaires, et à entraîner des modèles locaux avec les moyens de calcul déjà disponibles sur site, tout en transférant et fusionnant les modèles entre ces serveurs pour progressivement construire un modèle global partagé : c'est le Federated Learning [1].

Sujet

Ce sujet consiste à étudier l'état de l'art du federated learning, proposer une méthode adaptée à l'industrie du futur, l'implémenter et la valider expérimentalement. Plus précisément, la question de recherche principale du stage concerne les spécificités du federated learning pour les systèmes industriels, dans lesquels les données issues de systèmes répartis géographiquement (au sein d'une même usine ou de plusieurs) sont utilisées localement pour entraîner un réseau profond local. Tout comme pour l'IoT, l'une des difficultés principales est liée à la faible puissance de calcul disponible localement, qui ne permet pas d'y déployer des modèles très complexes. De plus, les systèmes vieillissent au fil du temps, et les modèles locaux entraînés sur le long terme s'adapteront à cette dérive mais oublieront (cf. *catastrophic forgetting* [2]) leurs comportements passés à cause de leur capacité limitée. C'est donc au modèle central de prendre en charge ce rôle de mémorisation de la dérive des données due au vieillissement des systèmes sur le long terme, et d'exploiter cette connaissance de manière pertinente lorsqu'il met à jour les réseaux locaux.

L'objectif du stage est de proposer une solution à ce problème et de la valider expérimentalement sur des corpus de données que nous fournissons.

Cadre du travail

Les compétences requises pour le stage sont une bonne connaissance des modèles profonds et de pytorch. Le stage sera co-encadré par Christophe Cerisara du LORIA et Alexandre Voisin du CRAN, les deux laboratoires étant situés sur le campus de Vandoeuvre à 50m l'un de l'autre.

Références

- [1] Keith Bonawitz, Hubert Eichner, Wolfgang Grieskamp, Dzmitry Huba, Alex Ingerman, Vladimir Ivanov, Chloé Kiddon, Jakub Konečný, Stefano Mazzocchi, Brendan McMahan, Timon Van Overveldt, David Petrou, Daniel Ramage, and Jason Roselander. Towards federated learning at scale : System design. In A. Talwalkar, V. Smith, and M. Zaharia, editors, *Proceedings of Machine Learning and Systems*, volume 1, pages 374–388, 2019.
- [2] Prakhar Kaushik, Alex Gain, Adam Kortylewski, and Alan L. Yuille. Understanding catastrophic forgetting and remembering in continual learning with optimal relevance mapping. *CoRR*, abs/2102.11343, 2021.

Sujet 2

Expérimentation des méthodes d'apprentissage auto-supervisé sur les modèles neuronaux multi-modaux : Application aux documents légaux.

Proposé par : Jean-Charles LAMIREL - Bart LAMIROY

Equipe : Synalp

Informations générales

Encadrants	Jean-Charles LAMIREL, BART LAMIROY, Aurélie VALLEREAU, El-Azhar JEBBARI
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-les-Nancy
Téléphone	03 54 95 84 06
Email	jean-charles.lamirel@loria.fr bart.lamiroy@loria.fr aurelie.vallereau@batt.eu el-azhar.jebbari@loria.fr
Bureau	B 208

Motivations

Sujet

L'apprentissage profond est devenu un composant essentiel de la plupart des systèmes intelligents. La capacité d'apprendre des modèles riches à partir de l'abondance de données aujourd'hui a fait de l'utilisation des Réseaux Neuronaux Profonds (RNP) une approche convaincante dans la majorité des tâches. Cependant, les approches traditionnelles d'apprentissage supervisé dépendent fortement de la quantité de données d'apprentissage annotées disponibles. Bien qu'une large quantité de données soient aujourd'hui disponibles, très peu d'entre elles sont annotées. Par conséquent, l'apprentissage supervisé à partir de données labélisées a presque atteint sa saturation.

C'est là que les méthodes auto-supervisées jouent un rôle essentiel en exploitant les progrès de l'apprentissage profond sans avoir besoin d'annotations coûteuses et en apprenant des représentations de caractéristiques à partir de données non labélisées.

L'apprentissage auto-supervisé a été appliqué avec succès sur plusieurs types de données : image, vidéo, son, signaux, texte etc ... On distingue principalement deux types d'apprentissage auto-supervisé : l'apprentissage contrastif [3] [4] et l'apprentissage génératif avec les modèles Generative Adversarial Networks (GANs) [1] [2].

Concernant l'apprentissage contrastif, il s'agit de construire d'abord une représentation pertinente des données dans un espace latent et d'ensuite réaliser un apprentissage sur une tâche spécifique. La représentation des données se fait à partir d'une approche discriminative dont le but est de rapprocher les données similaires (données positives) et d'éloigner celles qui sont dissimilaires (données négatives). L'apprentissage du modèle à différencier les échantillons positifs et négatifs s'opère alors grâce à des prétextes d'apprentissage. Il s'agit des stratégies utilisées pour apprendre des représentations de données en utilisant des pseudo-étiquettes. Le choix du prétexte d'apprentissage dépend fortement du problème à résoudre. Ce choix est crucial car il introduit des biais qui impactent directement sur les performances de l'apprentissage pour la tâche spécifique à considérer.

Le Document Intelligence (DI) est un domaine d'application très porteur qui fait référence à des techniques de lecture, de compréhension et d'analyse automatique de documents commerciaux. Certains documents, comme les documents contractuels représentent un véritable challenge dans ce domaine en raison de la diversité des mises en page et des formats, de la mauvaise qualité des images de documents numérisés ainsi que de la complexité de structures des modèles attendus. Un défi majeur est donc celui d'élaborer des modèles neuronaux multimodaux qui prennent en compte cette complexité.

Dans la littérature, certaines propositions ont été récemment faites en ce sens, tel que le modèle LayoutLM [5]. LayoutLM est un modèle auto-supervisé multimodal qui représente conjointement le texte et la mise en page des documents pour des tâches de compréhension d'images de documents. Ce modèle a cependant été conçu pour l'analyse de documents commerciaux assez élémentaires (reçus, bons de commande, etc...) et a par conséquent deux limitations importantes. (1) Il n'est pas adapté traitement des documents longs à cause du nombre fixe de tokens en entrée du modèle. (2) Il ne prend pas en compte la typologie des textes.

Cadre du travail

Nous avons donc élaboré un nouveau modèle multimodal adapté à la compréhension de documents longs. Ce modèle apprend de manière auto-supervisée une représentation conjointe du texte et de la mise en page, cela grâce à plusieurs prétextes d'apprentissage utilisés de manière simultanée. Le but du stage est donc étudier les prétextes d'apprentissage qui permettraient d'obtenir la représentation multimodale la plus pertinente des documents pour ensuite l'adapter à une tâche de segmentation fine de documents longs. Les expérimentations seront menées en exploitant les nombreuses bases de données de documents légaux qui sont disponibles sur le Web.

Références

- [1] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. & Bengio, Y. Generative adversarial nets. *Advances In Neural Information Processing Systems*. **27** (2014)
- [2] Kim, T., Cha, M., Kim, H., Lee, J. & Kim, J. Learning to discover cross-domain relations with generative adversarial networks. *International Conference On Machine Learning*. pp. 1857-1865 (2017)
- [3] Caron, M., Misra, I., Mairal, J., Goyal, P., Bojanowski, P. & Joulin, A. Unsupervised learning of visual features by contrasting cluster assignments. *ArXiv Preprint ArXiv :2006.09882*. (2020)
- [4] Chen, T., Kornblith, S., Norouzi, M. & Hinton, G. A simple framework for contrastive learning of visual representations. *International Conference On Machine Learning*. pp. 1597-1607 (2020)
- [5] Xu, Y., Xu, Y., Lv, T., Cui, L., Wei, F., Wang, G., Lu, Y., Florencio, D., Zhang, C., Che, W. & Others LayoutLMv2 : Multi-modal pre-training for visually-rich document understanding. *ArXiv Preprint ArXiv :2012.14740*. (2020)

Sujet 3

Biomedical Named Entity Recognition and Relation Extraction : implementing (then improving) a state-of-the-art architecture for the population of a knowledge graph

Proposé par : Joël Legrand et Adrien Coulet

Equipe : Synalp

Informations générales

Encadrants	Joël Legrand, Adrien Coulet
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Téléphone	03 54 95 84 06
Email	joel.legrand@loria.fr adrien.coulet@inria.fr
Bureau	B 104

Motivations

Ce projet de Master s'inscrit dans la continuité du Projet ANR PractiKPharma (practikpharma.mystrikingly.com/) qui a conduit à la constitution de deux ressources importantes, un corpus de textes manuellement annotés (PGxCorpus) et un graphe de connaissances (PGxLOD), pour le domaine de la pharmacogénomique (que nous abrégons par PGx). Ce domaine de la pharmacologie étudie l'impact des facteurs génétiques sur la différence de réponse individuelle aux médicaments [1]. L'intérêt du corpus PGxCorpus est d'offrir la possibilité d'entraîner (ou d'affiner l'entraînement) des modèles de traitement du langage naturel (TAL) pour l'identification de relations entre 2 ou 3 des éléments suivants : {facteur génétique, médicament, réponse au médicament} [2]. Le graphe de connaissances PGxLOD est la ressource qui nous permet de regrouper, structurer et comparer de telles relations [3].

Une des particularités du corpus PGxCorpus réside dans la présence d'entités disjointes et/ou imbriquées. Relativement peu présentes dans les corpus généralistes, de telles entités sont indispensables pour capturer la richesse des concepts utilisés dans le domaine biomédical. Pour autant, peu de travaux ont été conduits pour les extraire automatiquement. La plupart des modèles existants se contentent de les exclure des corpus d'entraînement et de test. Pourtant leur considération est particulièrement intéressante pour notre domaine d'application car l'imbrication permet de constituer des relations ternaires, à partir d'une relation binaire + une relation d'imbrication. Quelques approches ont été proposées pour résoudre le problème d'imbrication. Pour les entités disjointes, une première approche consiste à étendre le classique schéma d'annotation BIO (*Begin*, *Intermediate*, *Other*) de sorte que le traditionnel modèle CRF linéaire (Linear Conditional Random Field) puisse être utilisé pour des structures disjointes [8]. Cependant, ces modèles souffrent grandement d'ambiguïtés lors du décodage dues à l'utilisation de ces nouveaux schémas d'annotation. Une autre approche consiste à utiliser une représentation des entités sous la forme d'un hypergraphe afin de réduire le niveau d'ambiguïté [5]. Ces systèmes encodent de manière inexacte l'ensemble de l'espace des entités disjointes, en conservant les ambiguïtés durant la phase d'apprentissage, puis s'appuient sur des heuristiques pour gérer ces ambiguïtés durant le décodage. Une extension de ce modèle, en deux étapes, a été proposée afin de tirer parti des progrès des méthodes d'apprentissage profond. Des réseaux récurrents de type LSTM ont été proposés pour la phase de construction de l'hypergraphe [6] et pour la phase de reconstruction des entités disjointes [7]. Plus récemment, des modèles à base de l'architecture transformer ont été proposés, par exemple [4]. Notons que certaines de ces approches comme celle de [4] prennent en considération la représentation de la syntaxe des phrases pour cette tâche.

Sujet

Le projet du stage de master est de valoriser le corpus de textes pour entraîner les modèles qui permettront la reconnaissance d'entités nommées et l'extraction de relations, et cela notamment en alimentant PGLOD (le graphe de connaissances) avec des relations de très bonne qualité. Nos travaux précédents [11], nous ont permis

de définir des architectures pour ces deux tâches de TAL, notamment en nous appuyant sur l'architecture transformeur [9] et son implémentation BERT [10].

Le premier objectif du Master est de finaliser l'implémentation et l'évaluation de l'ensemble du pipeline que nous avons défini. Une fois cette implémentation finalisée, elle servira à extraire un grand nombre de relations binaires de bonne qualité qui pourront venir enrichir notre graphe de connaissances.

Le second objectif est d'améliorer la qualité de la reconnaissances d'entité en mieux considérant les entités discontinues. Nous proposons d'améliorer notre *pipeline* en lui permettant de considérer les entités discontinues. Pour cela, nous proposons de partir de [4] et des auteurs qui tirent parti des transformeur et de la syntaxe, pour reconstruire les entités discontinues. L'approche choisie pourra être testée, réglée finement à l'aide de PGxCorpus et éventuellement adaptées pour proposer une solution nouvelle, éventuellement motivée par la particularité des textes bio-médicaux. La nouvelle version du pipeline pourra alors permettre de venir enrichir le graphe de connaissances avec des relations de meilleure qualité.

Cadre du travail

Le stage se déroulera au sein de l'équipe Synalp du laboratoire Loria de Nancy.

Références

- [1] Xie, Hong-Guang and Frueh, Felix W *Pharmacogenomics steps toward personalized medicine*, Personalized medicine, 2005
- [2] J. Legrand, R. Gogdemir, C. Bousquet, K. Dalleau, M.-D. Devignes, W. Digan, C.-J. Lee, N.C. Ndiaye, N. Petitpain, P. Ringot, M. Smaïl-Tabbone. *PGxCorpus, a manually annotated corpus for pharmacogenomics*. Scientific data 7, no. 1 (2020) : 1-13.
- [3] P. Monnin, J. Legrand, G. Husson, P. Ringot, A. Tchechmedjiev, C. Jonquet, A. Napoli, A. Coulet. *PGxO and PGxLOD : a reconciliation of pharmacogenomic knowledge of various provenances, enabling further comparison*. BMC bioinformatics 20.4 (2019) : 1-16.
- [4] F. Li, L. Zhichao, Z. Meishan, J. Donghong. *A Span-Based Model for Joint Overlapped and Discontinuous Named Entity Recognition*. Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, 2021
- [5] A.O. Muis, W. Lu *Learning to recognize discontinuous entities* arXiv preprint arXiv :1810.08579, 2018
- [6] B. Wang, W. Lu *Neural segmental hypergraphs for overlapping mention recognition* arXiv preprint arXiv :1810.01817, 2018
- [7] B. Wang, W. Lu *Combining spans into entities : A neural two-stage approach for recognizing discontinuous entities* arXiv preprint arXiv :1909.00930 ,2019
- [8] J. Xu, Y. Zhang, J. Wang, Y. Wu, M. Jiang, E. Soysal, H. Xu. *UTH-CCB : the participation of the SemEval 2015 challenge-Task 14* proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)
- [9] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin *Attention is all you need*, Advances in Neural Information Processing Systems, 2017
- [10] J. Devlin, M-W Chang, K. Lee, K. Toutanova *BERT : Pre-training of Deep Bidirectional Transformers for Language Understanding*, Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies, Volume 1 (Long and Short Papers, 2019
- [11] W. Hafiane, J. Legrand, Joel , Y. Toussaint, A. Coulet *Experiments on transfer learning architectures for biomedical relation extraction* arXiv preprint arXiv :2011.12380, 2020

Équipe Tangram

<http://magrit.loria.fr/> et <http://team.inria.fr/tangram>

Sujet 1

Tracking, validation et intégration temporelle de primitives de haut niveau issues de réseaux convolutionnels pour la localisation par l'image

Proposé par : M.-O. Berger & G. Simon

Equipe : Tangram

Informations générales

Encadrants Marie-Odile BERGER, Gilles SIMON
Adresse LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Téléphone 03 54 95 85 01
Email marie-odile.berger@loria.fr & gilles.simon@loria.fr
Bureau C036 & C034
Homepages <https://members.loria.fr/gsimon> & <https://members.loria.fr/moberger/>

Motivations

Les méthodes basées sur les réseaux convolutionnels (CNN) permettent désormais la détection de primitives de haut niveau de manière largement indépendante des conditions d'illumination ou du point de vue adopté pour la prise de vue. La détection d'objets [1] a largement bénéficié de ces méthodes tandis que des méthodes prometteuses émergent pour la détection de lignes d'horizon ou de points de fuite [2] et commencent à concurrencer les méthodes traditionnelles [3]. Alors que le calcul de la pose de la caméra reposait il y a encore peu de temps uniquement sur les points d'intérêt détectés dans l'image, il est maintenant possible d'utiliser les objets détectés dans l'image pour la localisation [4], ce qui est particulièrement intéressant dans les environnements peu texturés. Par ailleurs, la détection de la ligne d'horizon et des points de fuite permet d'apporter des informations sur la composante rotationnelle de la pose.

Dans beaucoup d'applications, notamment de réalité augmentée, l'utilisateur est libre de se déplacer dans la scène et on peut ainsi avoir à traiter des vues globales de la scène ou au contraire des vues rapprochées de certains objets. Le nombre d'indices disponibles peut alors être petit et il importe que la localisation puisse tirer partie au mieux de ces indices, même s'ils sont relativement imprécis. Par ailleurs la localisation doit être faite de manière continue temporellement, ce qui implique de pouvoir mettre en correspondance ces indices au cours du temps.

Sujet

Le sujet consistera à poser les bases d'un système permettant l'intégration temporelle de ces indices de haut niveau avec les classiques points d'intérêt au sein d'un système de localisation par l'image.

Les indices considérés dans ce travail seront les objets détectés, les lignes d'horizon et des points de fuite (PF). Les points suivants seront abordés :

- mise en correspondance temporelle des objets : à un instant donné il peut exister plusieurs correspondants possibles. Dans le cas de la mise en correspondance de points, la validation est la plupart du temps réalisée en considérant la cohérence de la photométrie. Le problème est plus complexe dans le cas de primitives de haut niveau comme les objets car il faut comparer des structures plus complexes dont la forme n'est pas précisément définie puisque la détection fournit seulement une boîte englobante. La similarité des descripteurs CNN [5] pourra être utilisée entre deux images ainsi que la stabilité temporelle de la trajectoire des objets.
- mise en correspondance des points de fuite : les détecteurs de PF génèrent souvent de nombreux points qu'il est difficile d'apparier sur la seule base de leur coordonnées. Des stratégies visant à rectifier les images à partir des PF afin de valider leur ressemblance, et donc la pertinence de la mise en correspondance des PF, seront étudiées

- en lien avec des travaux en cours dans l'équipe, le candidat contribuera à la quantification de l'incertain sur les indices détectés par CNN. Des stratégies d'intégration de ces indices dans la procédure de localisation, avec les traditionnels points d'intérêt, devront ensuite être proposées, l'objectif étant d'intégrer de manière préférentielle les indices de haut niveau apportant une information complémentaire de celle des points présents. Des techniques d' *explanability* [6] pourront être utilisées pour identifier les zones de l'image ayant contribué à la détection des indices. Diverses manières d'intégrer ces indices dans la localisation, soit comme contrainte ou directement dans une fonction de coût, seront étudiées.

Cadre du travail

Le ou la stagiaire sera intégré dans l'équipe TANGRAM (anciennement MAGRIT) et bénéficiera des compétences de l'équipe sur la localisation en général, et sur la localisation à partir d'objets. Il sera aussi intégré au projet MOVEON, en collaboration avec le DFKI (Kaiserslautern), qui vise à développer de nouvelles méthodes de localisation visuelle.

Le sujet proposé est à la confluence de l'apprentissage et de la géométrie. Nous recherchons un ou une candidate avec des connaissances (ou un fort intérêt) dans les domaines suivant : traitement d'image, vision par ordinateur et apprentissage.

Références

- [1] J. Redmon, S. Divvala, R. Girshick, A. Farhadi *You Look Only Once : Unified, real-time object detection.* CVPR, 2016.
- [2] S. Workman, M. Zhai, N. Jacobs *Horizon Lines in the Wild* BMVC, 2016.
- [3] G. Simon, A. Fond, M.-O. Berger *A Contrario Horizon-first Vanishing Point Detection Using Second Order Grouping Laws.* ECCV, 2018.
- [4] V. Gaudillière, G. Simon, M.-O. Berger *Perspectibve-2-Ellipsoid : Bridging the GAP Between Object Detection and 6-DoF Camera Pose.* IEEE Robotics and Automation Letter, 2020.
- [5] N. Sunderhauf et al. *Place Recognition with ConvNet Landmarks : Viewpoint-Robust, Condition-Robust, Training Free.* Robotics Sciences and Systems Conference, 2015.
- [6] Ramprasaath et al. *Grad-CAM : Visual Explanations from Deep Networks via Gradient-based Localization.* CVPR, 2017.

Sujet 2

Gestion de contact d'un modèle 1D dans un réseau tubulaire

Proposé par : Erwan Kerrien et Pierre-Frédéric Villard

Equipe : Tangram

Informations générales

Encadrants	Erwan Kerrien, Pierre-Frédéric Villard	
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy	
Téléphone	03 83 59 20 67	
Email	erwan.kerrien@loria.fr	pierrefrederic.villard@loria.fr
Bureau	C034	

Motivations

Un cathéter est un tube long, fin et flexible que le médecin fait naviguer dans le réseau vasculaire afin d'établir une route depuis l'aîne jusqu'au cerveau. Cette route servira par la suite à acheminer les micro-outils jusqu'au site pathologique et traiter l'AVC. Un cathéter peut être modélisé par un modèle 1D géométriquement constitué par un ensemble de points reliés par des segments et dont le comportement mécanique obéit à une formulation du type Cosserat [1]. La navigation fait intervenir de nombreux contacts entre le cathéter et la paroi des vaisseaux sanguins qui aujourd'hui sont difficiles à simuler de manière rapide et fidèle à la réalité. L'objectif de ce stage est d'étudier et implémenter la gestion de contacts entre le cathéter et les vaisseaux sanguins.

Sujet

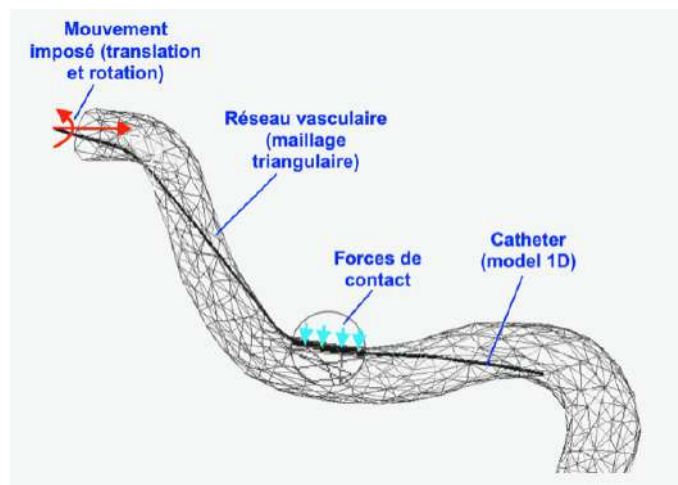


FIGURE 12 – Illustration de la problématique

La modélisation physique est un domaine bien établi qui fait appel à des notions et techniques avancées à la fois en physique, mathématiques et informatique (voir par exemple le cours de D. James [2], un peu ancien mais qui constitue toujours une excellente introduction au domaine). Parmi les forces à simuler, les forces de réponse à une collision posent un problème particulièrement difficile car elles dépendent de la géométrie qui varie pendant la simulation, notamment à cause de ces mêmes forces de collision. Par conséquent, leur gestion – qui recouvre la détection et le calcul de la force de rappel – va être impactée à la fois par la méthode de résolution du système mécanique à simuler, et par la représentation géométrique choisie pour les objets en contact. Dans notre cas particulier, l'objectif est de pouvoir simuler le contact entre un cathéter, objet 1D dont le comportement obéit au modèle mécanique de Cosserat, et une surface représentée de manière implicite. On suivra dès lors 3 étapes :

1. Réaliser un état de l'art sur les méthodes permettant de gérer les contacts entre un objet dynamique et une structure statique avec des contraintes de temps interactif (voir [3] pour une base de travail récente).
2. Implémenter une méthode de gestion contact entre le modèle 1D utilisé dans le projet et une structure tubulaire modélisée de manière discrète par un maillage triangulaire (voir Figure 1)
3. Concevoir et implémenter une méthode de gestion contact entre ce même modèle 1D et la structure cette fois modélisée de manière continue par une surface implicite [4].

Cadre du travail

Ce stage se fera dans le contexte du projet PreSPIN [5] financé par l'ANR. Ce projet consiste à concevoir des outils de simulation pour la planification et le traitement des AVC ischémiques [6]. Plus particulièrement, le stage porte sur la partie simulation de la navigation d'un cathéter [7] dans les vaisseaux sanguins. Ce dernier point est le sujet de la thèse de doctorat de Radhouane Jilani, avec lequel le ou la stagiaire devra travailler en étroite collaboration.

Références

- [1] J. Spillmann. CORDE : Cosserat rod elements for the animation of interacting elastic rods. Thèse de doctorat de l'université Albert Ludwig de Freiburg im Breisgau, 2008. <https://freidok.uni-freiburg.de/fedora/objects/freidok:5775/datastreams/FILE1/content>
- [2] D. James. Physically Based Modeling and Interactive Simulation. *Cours du département d'informatique de Carnegie Mellon University*, 2003. <http://www.cs.cmu.edu/~djames/pbmis/spring03.html>
- [3] B. Wang et al. A Large-scale Benchmark and an Inclusion-based Algorithm for Continuous Collision Detection. *ACM Transactions on Graphics*, 40(5) :1-16 (No 188), 2021. <https://doi.org/10.1145/3460775>
- [4] E. Kerrien et al. Blood vessel modeling for interactive simulation of interventional neuroradiology procedures. *Medical Image Analysis*, 35 :685–698, 2017. https://hal.inria.fr/hal-01390923/file/medima_20161008.pdf
- [5] <https://project.inria.fr/prespin/>
- [6] https://fr.wikipedia.org/wiki/Accident_vasculaire_cérébral#Ischémique
- [7] <https://fr.wikipedia.org/wiki/Cathéter>

Équipe TYPES

<https://types.loria.fr/>

Sujet 1

Bi-intuitionistic Bunched Implication Logic

Proposé par : Didier GALMICHE & Daniel MERY

Equipe : TYPES

Informations générales / General Information

Encadrants	Didier GALMICHE, Daniel MERY	
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy	
Téléphone		
Email	Didier.Galmiche@loria.fr	Daniel.Mery@loria.fr
Bureau	B 224	

Motivations

BI Logic [5] is a resource logic that combines additive and multiplicative connectives. In the standard version the additive connectives are interpreted as connectives of intuitionistic logic (IL) and the multiplicative connectives are interpreted as connectives of Multiplicative Intuitionistic Linear Logic (MILL). In other words, BI logic is a combination of IL and MILL, and a conservative extension of each of these logics. BI and its variants are the logical kernel of others resource logics : separation logics, tree logics, context logics, ambients. . . There exist several proof systems for BI, mainly a purely syntactic sequent calculus, called LBI [5], and a labelled tableaux systems, called TBI [2].

The bi-intuitionistic logic (Bi-int) is also another conservative extension of intuitionistic logic in which one adds a new connective, called co-implication or exclusion, that behaves as the dual of the implication [1, 8]. If it is easy to extend the standard sequent calculus for IL by adding rules for the co-implication, that are symmetric rules w.r.t. implication rules, the obtained sequent calculus does not satisfy the cut-elimination property. That is why new proof systems have been proposed these last years and they satisfy the cut-elimination property [3, 4, 6].

Sujet / Subject

The goal of this subject is, by starting from the BI logic, to extend this logic with the connective of exclusion (co-implication) of Bi-int, in order to define a new logic, called Bi-intuitionistic BI logic (Bi-BI).

A first step consists in formalizing, in a precise and detailed way, this logic Bi-BI (syntaxe, sémantique, expressivité) and then to propose proof/refutation calculi, in the spirit of the internal calculi (without labels) and external calculi (with labels) that have been studied for BI and its variants.

A second step consists in studying the possible use of Bi-BI as an intermediate logic for the translation of labelled proofs in the TBI system into equivalent proofs without labels in the LBI system. In fact, it is often less easy to translate a proof in an external system into an equivalent proof in an internal system than to do the reverse translation. For instance, even in the restricted logical fragment of IL, the current techniques need an intermediate step involving the Bi-intuitionistic logic (Bi-int) and its exclusion connective [7].

In the context of BI, there is no proof translation that allows one to transform a proof in TBI into an equivalent proof in LBI.

Cadre du travail / Working Context

The internship will take place in the TYPES team at LORIA, UMR 7503.

Références

- [1] T. Crolard. Subtractive logic. *Theoretical Computer Science*, 254(1-2) :151–185, 2001.
- [2] D. Galmiche and D. Méry. Semantic labelled tableaux for propositional BI without bottom. *Journal of Logic and Computation*, 13(5) :707–753, 2003.
- [3] D. Galmiche and D. Méry. A connection-based characterization of bi-intuitionistic validity. In *23rd Int. Conference on Automated Deduction, CADE'2011*, LNAI 6803, pages 268–282, Wrocław, Poland, 2011.

- [4] R. Goré, L. Postniece and A. Tiu. Cut-elimination and proof-search for Bi-intuitionistic logic using nested sequents. In *7th Conference on Advances in Modal Logic, AiML'2008*, AiML 7, pages 43–66, Nancy, France, 2008.
- [5] P.W. O'Hearn and D. Pym. The Logic of Bunched Implications. *Bulletin of Symbolic Logic*, 5(2) :215–244, 1999.
- [6] L. Pinto and T. Uustalu. Proof search and counter-model construction for Bi-intuitionistic propositional logic with labelled sequents. In *Int. Conference on Analytic Tableaux and Related Methods, TABLEAUX'2009*, LNAI 5607, pages 295–309, Oslo, Norway, 2009.
- [7] L. Pinto and T. Uustalu. Relating sequent calculi for bi-intuitionistic propositional logic. In *Proceedings 3rd Int. Workshop on Classical Logic and Computation, CL&C'2010*, EPTCS 47, pages 57–72, Brno, Czech Republic, 2010.
- [8] C. Rauszer. An algebraic and Krike-style approach to a certain extension of intuitionistic logic. *Dissertationes Mathematicae*, 168, 1980.

Sujet 2

Formalizing the Undecidability of Relevance Logic and Non-Commutative Linear Logic

Proposé par : Dominique LARCHEY-WENDLING

Equipe : TYPES

Informations générales / General Information

Encadrants	Dominique LARCHEY-WENDLING
Adresse	LORIA, Campus Scientifique - BP 239, 54506 Vandœuvre-lès-Nancy
Téléphone	
Email	dominique.larchey-wendling@loria.fr
Bureau	B 222

Motivations

The (<https://github.com/uds-psl/coq-library-undecidability>) Coq Library of Undecidability Proofs that we have been actively developing for nearly three years collects a significant amount of formalized and mechanized undecidability results concerning basic models of computations like Turing machines, counter (or Minsky) machines, stack machines, μ -recursive functions, the FRACTRAN programming language, λ -calculus, etc., but also decision problems about strings like the word problem for string rewrite (or semi-Thue) systems, the Post correspondence problem, unification, or decision problems about natural numbers like the unsolvability of Diophantine equations, i.e. <http://dx.doi.org/10.4230/LIPIcs.FSCD.2019.27> (Hilbert tenth problem) [1]. See [2] (<https://members.loria.fr/DLarchey/files/papers/coqpl20-final6.pdf>) for a quick summary of the contents and structure of the library of undecidability proofs.

This library is centered around the notion of *synthetic reduction* which amounts to implement and verify the correctness of many-one reductions in a computer language that is both very expressive and allows only computable total functions to be written. In the computability literature, many-one reductions are the main tool to transport undecidability results from one seed problem to another target problem.

Among the already implemented results are also undecidability results for various logics like first order logic provability and (<https://arxiv.org/abs/2004.07390>) finite satisfiability [3], but also about linear logic, i.e. the (<https://dl.acm.org/doi/10.1145/3293880.3294096>) undecidability of entailment for linear logic, both in intuitionistic [4] and classical flavor.

Sujet / Subject

Concerning linear logic, most of the undecidability results derive from many-one reductions from counter (or Minsky) machines, starting with [5]. In ([https://doi.org/10.1016/0168-0072\(92\)90075-B](https://doi.org/10.1016/0168-0072(92)90075-B)), Lincoln *et al.* also consider the case of *non-commutative* linear logic of which, in contrast with the commutative version, the undecidability is derived from that of the word problem for string rewrite systems, also called semi-Thue systems. It should also be noted that the (<http://www.jstor.org/stable/2274261>) landmark paper of Urquhart [6] also starts from the word problem to establish the undecidability of a whole family of relevance logics.

The word problem is already proved undecidable in the Coq undecidability library, directly derived from that of the Halting problem for Turing machines. In this subject, we naturally propose to use this seed to implement the undecidability proofs of Urquhart for relevance logic, and of Lincoln *et al.* for non-commutative linear logic, hopefully following the given pen and paper proofs. The implementation part will be done in the (<https://coq.inria.fr>) Coq proof assistant. Training and help with the assistant will be provided by the supervisor.

Cadre du travail / Working Context

The internship will take place in the TYPES team at LORIA, UMR 7503.

Références

- [1] D. Larchey-Wendling and Y. Forster. Hilbert’s Tenth problem in Coq. *Formal Structures for Computation and Deduction, FSCD 2019*.
- [2] Y. Forster and D. Larchey-Wendling and A. Dudenhefner and D. Kirst *et al.* A Coq Library of Undecidable Problems. *The Sixth International Workshop on Coq for Programming Languages, CoqPL 2020*.
- [3] D. Kirst and D. Larchey-Wendling. Trakhtenbrot’s Theorem in Coq. *International Joint Conference on Automated Reasoning, IJCAR 2020*.
- [4] Y. Forster and D. Larchey-Wendling. Certified Undecidability of Intuitionistic Linear Logic via Binary Stack Machines and Minsky Machines. *Certified Programs and Proofs, CPP 2019*.
- [5] P. Lincoln and J. Mitchell and A. Scedrov and N. Shankar. Decision problems for propositional linear logic. *Annals of Pure and Applied Logic*, 56(1) :239–311, 1992.
- [6] A. Urquhart. The Undecidability of Entailment and Relevant Implication. *The Journal of Symbolic Logic*, 49(4) :1059–1073, 1984.