A propos du centre ou de la direction fonctionnelle
The Inria Lille - Nord Europe Research Centre was founded in 2008 and employs a staff of 360, including 300 scientists working in sixteen research teams. Recognised for its outstanding contribution to the socio-economic development of the Hauts-de-France region, the Inria Lille - Nord Europe Research Centre undertakes research in the field of computer science in collaboration with a range of academic, institutional and industrial partners.

The strategy of the Centre is to develop an internationally renowned centre of excellence with a significant impact on the City of Lille and its surrounding area. It works to achieve this by pursuing a range of ambitious research projects in such fields of computer science as the intelligence of data and adaptive software systems. Building on the synergies between research and industry, Inria is a major contributor to skills and technology transfer in the field of computer science.

Contexte et atouts du poste
Applications are welcome for a post-doctoral position at Inria as part of the HYAIAI Inria challenge. More specifically, this position is part of a collaboration between the Lacodam and Sequel, Inria teams. The post-doc will share her time between the two groups (Lacodam in Rennes, Sequel in Lille). She will actively participate in the activities of HYAIAI, in particular reporting on her work in the HYAIAI meetings.

Mission confiée
The main purpose of this post-doctoral fellowship is to port the principles of interpretable AI and ML to the domain of the visual question answering. Attaining such a goal requires us to overcome other challenges such as understanding what makes a good explanation in a multi-modal setting.

Principales activités
The post-doctoral researcher in charge of this project will work on methods to explain the output of a VQA system in order to understand why it erred (or not).

In this regard, we aim at deploying post-hoc interpretability modules that can provide hints on the logic behind a VQA module for a given case. Techniques such as LIME [Ribeiro et. al., 2016], SHAP [Lundberg and Lee, 2017], or Anchors [Ribeiro et. al., 2018] provide the foundations to generate such explanations for any type of black-box model. That said, those techniques deliver explanations in terms of an interpretable space that depends on the data type (e.g., tabular data, texts, images) and is usually very different from the representation space of the black-box. For instance, pixel channels and feature maps in images are replaced by super-pixels (image segments) in explanations, whereas word embeddings are substituted by word occurrences. No research work until now has tried to reconcile those representations in a multimodal setting, thus the main challenge is to find a common interpretable representation for explanations that encompasses both images and texts in a VQA setting. We are particularly interested in representations that resort to semantically meaningful units such as known objects or patterns (e.g., pointy borders, a nose, limbs, vehicles) as studied in network dissection. This would allow us to yield explanations in terms of the presence or absence of those objects. Such explanations could be enhanced with semantic knowledge, such as a taxonomy, in order to signal interesting associations automatically, e.g., pointy borders in a bird may refer to its beak. Other recent

Informations générales
- Thème/Domaine : Optimisation, apprentissage et méthodes statistiques
- Ville : Lille
- Centre Inria : CRI Lille - Nord Europe
- Date de prise de fonction souhaitée : 2020-06-01
- Durée de contrat : 2 ans
- Date limite pour postuler : 2020-06-30

Contacts
- Equipe Inria : SEQUEL
- Recruteur : Galarraga Del Prado Luis / luis.galarraga-del-prado@inria.fr

A propos d’Inria
Inria est l’institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 200 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3500 scientifiques pour relever les défis du numérique, souvent à l’interface d’autres disciplines. L’institut fait appel à de nombreux talents dans plus d’une quarantaine de métiers différents. 900 personnels d’appui à la recherche et à l’innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 180 start-up. L’institut s’efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l’économie.

Consignes pour postuler
To apply for the position, the candidate must send an email to the list of contacts below. The email should include:
- A CV
- A statement letter explaining the candidate’s motivations to apply for the position
- At least two recommendation letters

Contacts
- Philippe Preux (philippe.preux@inria.fr)
- Luis Galarraga (luis.galarraga@inria.fr)

Sécurité défense :
Ce poste est susceptible d’être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L’autorisation d’accès à une zone est délivrée par le chef d’établissement, après avis ministériel favorable, tel que défini dans l’arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l’annulation du recrutement.
approaches [Shi et al., 2019] [Yi et al., 2018] have focused on the extraction of knowledge from the input before defining a reasoning program to execute. This knowledge may be a start in the definition of an interpretable space for explanations of VQA systems.

Traditional post-hoc interpretability modules do not make any assumptions about the architecture of the model they try to explain, i.e., they are model-agnostic. In the context of VQA systems, a possible solution is to drop this assumption and mine neuron activation patterns that identify the instances for which the system fails. We could use contrast pattern mining techniques for this purpose.

Our research will make use of publicly available VQA datasets such as GQA, VQA-E, CVQA, and CLEVR.

Compétences
We are searching for motivated candidates with a PhD degree in Computer Science and with competences in machine learning (preferably with focus on deep learning). Knowledge in data mining, e.g., sequence and itemset mining, will be also appreciated.

The candidate should be proficient in written and spoken English (at least B2 level according to the CEFR system).

Avantages
- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking (after 6 months of employment) and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

Rémunération
Gross monthly salary (before taxes) : 2653 €