Offer #2024-07776

Post-Doctoral Research Visit F/M Machine Learning for translation between formal mathematics libraries.

Contract type : Fixed-term contract

Level of qualifications required : PhD or equivalent

Fonction : Post-Doctoral Research Visit

About the research centre or Inria department

The Inria research centre in Lyon is the 9th Inria research centre, formally created in January 2022. It brings together approximately 300 people in 17 research teams and research support services.

Its staff are distributed in Villeurbanne, Lyon Gerland, and Saint-Etienne.

The Lyon centre is active in the fields of software, distributed and high-performance computing, embedded systems, quantum computing and privacy in the digital world, but also in digital health and computational biology.

Context

The postdoc will take place within the framework of the défi Inria LLM4Code, which **objectives are to apply machine learning techniques to the automatic generation of code,** and in particular for proofs assistants libraries, such as Coq libraries.

The research will be carried out in the LIP laboratory of ENS Lyon under the local supervision of Cyril Cohen, within the CASH team. However, there will be frequent meetings, and regular in person meetings, with participants from the other work packages of the LLM4Code project.

Travel expenses are covered within the limits of the scale in force.

Assignment

Assignments:

With the help of people from the projects and their collaborators, the recruited person will work on the automated translations from, to and between libraries of formalized mathematics using modern AI techniques, and in particular Large Language Models.

Such translations may require nontrivial alteration of the proof assistants at use, such as the creation of plugins, the use of meta-programming, or rephrasing libraries in ways that are better suited for machine learning.

It will also require to set up training, fine-tuning and/or RAG on existing models, and evaluate the relevance of each method.

The expected outcome of this work are research publications and software production.

For a better knowledge of the proposed research subject:

The state of the art, bibliography and scientific references are fast evolving. We suggest to read the recent survey A Survey on Deep Learning for Theorem Proving by Zhaoyu Li, Jialiang Sun, Logan Murphy, Qidong Su, Zenan Li, Xian Zhang, Kaiyu Yang, Xujie Si, with particular focus on publications about Coq or Lean.

Collaboration:

The recruited person will be mainly in connection with Cyril Cohen, but will also work remotely with other members of the LLM4Code project.

Responsibilities:

The person recruited is responsible for setting up the software infrastructure to conduct the experiments and take initiatives for the choices of software, and available hardware, within the limits allowed by Inria.

Main activities

Main activities:

- Determine how to train, finetune, or use RAG techniques on, existing models, and combine these methods, to achieve the desired translation.
- Setup the infrastructure to implement and reproduce these methods.
- Implement extensions of Coq or Lean in order to extract relevant data.
- Document the resulting software.
- Collect outputs, run benchmarks and publish scientific papers.

Additional activities:

- Design user interfaces to make the resulting tool widely usable.
- Determine a long term maintainance scheme for these.

Skills

Technical skills and level required :

- Know the basics of at least one proof assistant, preferably one based on dependent type theory.
- Understand the difference between various machine learning techniques.
- Have hands on practice in training models.

Languages :

- Advanced English.

Relational skills :

- Communicate about your results.
- Work autonomously, while talking about strategic decisions with other members of the project.
- Listen to feedback.

Other valued appreciated :

- Expertise in LLM.
- Expertise in Coq or Lean.
- Git, GitHub, Hugging Face, WandB, and other platforms for code sharing, model sharing, etc.

Benefits package

- 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 (90 days / year) and flexible organization of working hours (except for intership)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage under conditions

Remuneration

2788 € gross salary / month

General Information

- **Theme/Domain :** Architecture, Languages and Compilation
- **Town/city :** Lyon
- **Inria Center :** Centre Inria de Lyon
- **Starting date :** 2024-09-01
- **Duration of contract :** 2 years
- **Deadline to apply :** 2024-06-30

Contacts

- **Inria Team :** CASH
- **Recruiter :**
  Cohen Cyril / cyril.cohen@inria.fr

About Inria

Inria is the French national research institute dedicated to digital science and technology. It employs 2,600 people. Its 200 agile project teams, generally run jointly with academic partners, include more

than 3,500 scientists and engineers working to meet the challenges of digital technology, often at the interface with other disciplines. The Institute also employs numerous talents in over forty different professions. 900 research support staff contribute to the preparation and development of scientific and entrepreneurial projects that have a worldwide impact.

The keys to success

The candidate should demonstrate an interest in the construction of formalized mathematical libraries and machine learning. Although an in-depth training in proof assistants is not expected, it is important to understand the general principles and have a willingness to learn more. Similarly, it is not necessary to be an expert in AI, but it is expected to have sufficient knowledge to quickly update oneself (either independently or through interaction with project members), and to have the willingness to implement artificial learning techniques in practice.

Warning : you must enter your e-mail address in order to save your application to Inria. Applications must be submitted online on the Inria website. Processing of applications sent from other channels is not guaranteed.

Instruction to apply

Applications must be submitted online via the Inria website. Processing of applications submitted via other channels is not guaranteed.

Defence Security :
This position is likely to be situated in a restricted area (ZRR), as defined in Decree No. 2011-1425 relating to the protection of national scientific and technical potential (PPST).Authorisation to enter an area is granted by the director of the unit, following a favourable Ministerial decision, as defined in the decree of 3 July 2012 relating to the PPST. An unfavourable Ministerial decision in respect of a position situated in a ZRR would result in the cancellation of the appointment.

Recruitment Policy :
As part of its diversity policy, all Inria positions are accessible to people with disabilities.