



Offre n°2025-09101

PhD Position F/M Query-specific incremental maintenance algorithms on words and trees

Le descriptif de l'offre ci-dessous est en Anglais

Type de contrat : CDD

Niveau de diplôme exigé : Bac + 5 ou équivalent

Fonction : Doctorant

A propos du centre ou de la direction fonctionnelle

The Inria University of Lille centre, created in 2008, employs 360 people including 305 scientists in 15 research teams. Recognised for its strong involvement in the socio-economic development of the Hauts-De-France region, the Inria University of Lille centre pursues a close relationship with large companies and SMEs. By promoting synergies between researchers and industrialists, Inria participates in the transfer of skills and expertise in digital technologies and provides access to the best European and international research for the benefit of innovation and companies, particularly in the region.

For more than 10 years, the Inria University of Lille centre has been located at the heart of Lille's university and scientific ecosystem, as well as at the heart of Frenchtech, with a technology showroom based on Avenue de Bretagne in Lille, on the EuraTechnologies site of economic excellence dedicated to information and communication technologies (ICT).

Contexte et atouts du poste

The task of incremental maintenance on dynamic data studies how we can efficiently evaluate queries and then quickly update their results whenever the data is modified. This PhD proposal focuses on settings where the queries are evaluated over data represented as a tree or as a word: this covers many application settings, such as XML documents, JSON records, parse trees for source code, text, event sequences, etc.

On data of this form, the simplest kind of queries to consider are Boolean queries that ask whether the data belongs to a specified formal language L : for instance L can be a regular languages on words. Then, if w is the input word, we can test in time $O(|w|)$ whether $w \in L$. When w is modified, e.g., by changing characters, we want to know whether $w \in L$ after each modification. The naive algorithm is to test membership to L by reading the whole word after each update, with update complexity $O(|w|)$; but in many cases we can do better. For instance, consider the language $b^*(ab^*ab^*)^*$ of words on $\{a, b\}$ with an even number of a 's. We can maintain membership to this language in $O(1)$ — simply by keeping a count of the number of a 's.

This concrete example illustrates a fundamental insight: for a given problem on trees and words (e.g., dynamic membership to a regular language), the complexity of incremental maintenance problems can differ depending on the specific query that we are interested in maintaining. Further, they can also depend on the update language that we use. For instance, in the streaming setting where the word w can only be modified by appending new characters at the right, then it is trivial to maintain dynamic membership to any language represented as a word automaton. The goal of this PhD proposal is to achieve a deeper understanding of the complexity of incremental maintenance problems on trees and words, with algorithms and lower bounds that are tailored to the specific query that we want to evaluate.

The research of this PhD thesis will be mostly theoretical in nature. However, the candidate may also explore some practical application settings in which the proposed research is relevant and which we can also supervise. One application is the problem of earliest query answering on words and trees, where we discover a tree in streaming and must produce query answers as early as possible when they are certain, i.e., are true no matter what comes next in the document. Another broad application domain to explore is that of incremental parsing, as we expect that some of the algorithms proposed on 2trees can relate to the task of maintaining the parsing of context-free grammars on an

input word. One last potential application could be to implement some algorithms within the Rsonpath project. This project aims at providing a very efficient query engine for JsonPath. In particular some dedicated tasks make it necessary to efficiently enumerate the solution of queries in streaming; this is well into the scope of this thesis, and motivates the search for query classes admitting efficient algorithms.

This PhD will take place in the LINKS / D-DAL team at Inria Lille, it will be co-supervised by Antoine Amarilli (Advanced Research Position at Inria Lille) and Charles Paperman (Associate professor at Université de Lille).

Regular travel is expected when necessary to present articles at conferences and do research visits, and will be approved by the supervisors and funded by the team funds.

Mission confiée

The person hired on this position will be in charge of accomplishing a PhD on the topic of "Query-specific incremental maintenance algorithms on words and trees"

Principales activités

The candidate will get acquainted with the state of the art on the PhD topic, perform original research in interaction with the thesis supervisors and other collaborators, write research articles describing the research results, and publish the research and present it at conferences and meetings where necessary.

Compétences

- fluency in scientific English
- scientific reasoning and writing skills
- written and oral communication skills
- skills in mathematics and theoretical computer science
- ability to reflect on and document complex subjects and to choose one's own research questions
- ability to work in a team

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 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

2200€ gross monthly salary

Informations générales

- **Thème/Domaine** : Représentation et traitement des données et des connaissances
Systèmes d'information (BAP E)
- **Ville** : Villeneuve d'Ascq
- **Centre Inria** : [Centre Inria de l'Université de Lille](#)
- **Date de prise de fonction souhaitée** : 2025-09-01
- **Durée de contrat** : 3 ans
- **Date limite pour postuler** : 2025-08-09

Contacts

- **Équipe Inria** : [LINKS](#)
- **Directeur de thèse** :
Amarilli Antoine / antoine.a.amarilli@inria.fr

A propos d'Inria

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

L'essentiel pour réussir

- know how to communicate with supervisors about difficulties encountered
- stay motivated over the long term when studying complex problems
- effective time management

- interest in abstract and fundamental questions, and in advancing understanding of the basics of computer science
- be motivated by discovering and writing up new scientific results

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

Consignes pour postuler

Please send your CV and cover letter

Sécurité défense :

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

Politique de recrutement :

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