

STAGE LIFO (action DOING)

Titre : Mises à jour de bases de données graphe avec informations incomplètes

Disciplines : informatique (bases de données).

Stage dans le cadre de l'action **DOING (Données Intelligentes)** du **GDR MADICS** et du réseau régional **DIAMS** (Réseau Thématique de Recherche Données, Intelligence Artificielle, Modélisation et Simulation).

Encadrantes :

- Mirian Halfeld Ferrari (LIFO, Orléans), mirian@univ-orleans.fr
- Jacques Chabin (LIFO, Orléans), jacques.chabin@univ-orleans.fr

Sujet du stage : Les connexions sont partout dans le monde d'aujourd'hui: non seulement entre les personnes, les objets, les endroits ou les entreprises, mais aussi dans la science du vivant, la bioinformatique, la détection de fraudes, etc.

Les observations, structurées comme des réseaux ont leurs propres règles d'interconnexion. Quand ces connexions deviennent aussi importantes que les données, c'est-à-dire, les interrogations sur les données sont aussi essentielles que celles sur la topologie du réseaux, les bases de données graphe s'imposent comme un outil crucial.

Dans ce contexte, la question de la cohérence de la base doit évoluer pour prendre en compte non seulement les contraintes d'intégrité, mais aussi des contraintes liées à la topologie du graphe. De plus, cette cohérence doit être maintenue après chaque mise à jour (une insertion ou une suppression).

Dans [1], nous avons proposé une politique pour la mise à jour cohérente des bases de données ayant des informations incomplètes. Ces informations incomplètes sont des données inconnues, représentées par des valeurs nulles qui peuvent être liées entre elles (par exemple, nous savons que Paul est à un forum. Nous ne savons pas quel est ce forum, mais nous savons qu'il a lieu à Paris). La manipulation de ces informations incomplètes est fondamentale aujourd'hui avec la croissance des applications qui impliquent l'intégration ou l'échange de données.

Le but de cet stage serait de proposer une adaptation de la politique de mise à jour présentée dans [1] aux graphes d'attributs. Un graphe d'attributs est un graphe où les noeuds et les arcs peuvent avoir des attributs/propriétés associés (par exemple, un noeud ETUDIANT a comme attributs (Nom, Age, Adresse)) et sont utilisés comme modèle pour des SGBD comme Neo4J.

Dans ce cadre nous attendons:

- Définir un formalisme pour exprimer les différents types de contraintes sur un graphe d'attributs (avec une étude préalable des propositions existantes)
- Proposer de nouveaux algorithmes de mises à jour (en rapport avec la politique présentée dans [1])

Lieu du stage: Nous souhaitons que le stagiaire soit physiquement au LIFO, même si des périodes de télétravail peuvent être envisagées/ programmées.

Profil souhaité: l'étudiant doit être intéressé par l'approfondissement de ses connaissances dans le domaine de bases de données et intelligence artificielle. Niveau **master**.

Durée du stage: de 4 à 6 mois, financé par les fonds d'un contrat CIFRE Orléans.

CANDIDATURE:

Etape 1: Les candidats doivent envoyer (1) Curriculum Vitae et (2) Relevé de notes aux encadrants (emails ci-dessus) avant le **5 DECEMBRE 2021**

Etape 2: Candidats sélectionnés seront contactés (vers le 16 décembre) pour un entretien par visio, début janvier (l'entretien inclura une petite discussion sur un article, donné en amont)

REFERENCES

[1] Jacques Chabin, Mirian Halfeld Ferrari, Dominique Laurent: **Consistent updating of databases with marked nulls**. Knowl. Inf. Syst. 62(4): 1571-1609 (2020)