

Résumé de l'article “The Dichotomy of Evaluating Homomorphism-Closed Queries on Probabilistic Graphs”

Cet article a été publié en 2022 dans le journal “Logical Methods in Computer Science” (LMCS), revue de référence en logique mathématique et informatique théorique, en accès ouvert “diamant”. Il fait suite à un article en 2020 au congrès “International Conference on Database Theory” (ICDT), la conférence européenne de référence en théorie des bases de données, dont les actes sont également en libre accès. L'article ICDT'20 en question a été primé par le prix du meilleur article à cette conférence.

Ce travail s'intéresse à l'évaluation de requêtes sur des bases de données relationnelles probabilistes, c'est-à-dire dont chaque fait est annoté avec une probabilité d'existence, indépendamment des autres faits. Pour une requête booléenne (dont la réponse est oui ou non), la tâche à résoudre est de déterminer la probabilité que la requête soit satisfaite, sur la distribution de probabilités que la base de données représente. Ce travail s'intéresse à des requêtes très générales, et suppose simplement qu'elles sont closes sous homomorphismes : cette hypothèse s'applique en particulier aux requêtes conjonctives et unions de requêtes conjonctives (UCQ) pour lesquelles cette tâche avait été étudiée par des travaux antérieurs. Notre étude se limite toutefois, pour des raisons techniques, à des signatures relationnelles d'arité 2, c'est-à-dire à des graphes plutôt qu'à des hypergraphes.

Nous démontrons que, pour toute requête close sous homomorphismes sur les graphes, si elle n'est pas logiquement équivalente à une UCQ, alors le problème d'évaluation probabiliste est intractable ($\#P$ -difficile). Ce résultat très général garantit qu'il n'y a aucune requête de cette classe pour laquelle ce problème peut être résolu avec une meilleure complexité : il concerne en particulier les requêtes récursives exprimées dans des langages tels que Datalog.

Ce travail est une collaboration avec İsmail İlkan Ceylan de l'université d'Oxford, démarrée lors d'une visite de recherche à Oxford en 2019. La question de recherche initiale a été proposée par İsmail au cours de cette visite où nous avons eu des discussions préliminaires, mais je suis l'auteur principal de l'article, ayant conçu la majeure partie de la preuve par la suite et assuré la plus grande part de la rédaction (sauf les figures). J'ai poursuivi le développement de cette preuve dans un travail subséquent à ICDT'23 (article en seul auteur).