# Stage de recherche – LIP6

## Vérification de propriétés sur les graphes temporels - vers un métathéorème

**Mots-clefs:** graphe, algorithme, logique temporelle, sommets jumeaux, graphe $k$-régulier.

**Lieu:** LIP6, CNRS – SU UPMC, Paris.

**Encadrants:** Nathalie Sznajder <nathalie.sznajder@lip6.fr> et Binh-Minh Bui-Xuan <buixuan@lip6.fr>.

## Contexte:

Les bases de données à la fois orientées graphe et à historique sont au coeur de l'hyperconnectivité contemporaine. De manière informelle, un graphe temporel est une suite de graphes partageant le même ensemble de sommets. Ici, l'ordre d'apparition des arêtes dans cette suite représente les instants temporels où ces arêtes sont enregistrées. Nous nous intéressons à ces instants temporels parce qu'ils constituent une information aussi importante que le reste de la donnée. En effet, ce type de donnée datée est important pour des applications aussi diverses que dans le transport routier [13], la traversée de graphe [10], la centralité d'un groupe [15], le routage adhoc périodique [4], et plus généralement la modélisation de la dynamicité des graphes [6, 14].

Afin d'introduire le contexte de ce que l'on appelle méta-théorème en vérification de graphe, nous allons d'abord nous restreindre aux graphes classiques, statiques. Les formules logiques exprimant une propriété sur ces graphes peuvent se caractériser selon l'interprétation des variables:

- FO (First Order logic): les variables désignent uniquement les sommets et les arêtes, pas les sous-ensembles de ceux-ci.
- MO$_V$ (Monadic second Order logic, Vertex set): en plus de FO, un sous-ensemble de variables désignent les sous-ensembles de sommets.
- MO$_{VE}$ (idem, Vertex and Edge sets): en plus de MO$_V$, certaines variables désignent les sous-ensemble d'arêtes.
- enfin, notons que nous ne nous intéresserons pas aux formules mal définies, ni aux formules hors MO$_{VE}$.

Nous pouvons citer les exemples suivants de formules logiques exprimant des problèmes classiques sur les graphes: test de $k-$régularité (FO); existence d'une paire de sommets jumeaux (FO); ensemble dominant (MO$_V$); et couplage parfait (MO$_{VE}$). Notons également que les formules exprimables en FO le sont en MO$_V$, et celles exprimables en MO$_V$ le sont en MO$_{VE}$.

A titre d'exemple dans l'énoncé du théorème qui suit, nous allons rappeler la définition de la largeur de clique, pas celle de la largeur arborescente ni de la largeur de jumeaux. En effet, bien que similaires, les deux dernières largeurs demandent une définition plus complexe. Un graphe de largeur de clique bornée par $k$ est défini comme celui qui peut être obtenu par les opérations suivantes: création d'un graphe à sommet unique; colorer un graphe à sommet unique avec une couleur parmi $k$ couleurs; union disjointe de deux graphes colorés; dans un graphe coloré, remplacez la couleur $i$ par $j$; dans un graphe coloré, pour $i \neq j$, ajouter une arête entre tout couple d'un sommet de couleur $i$ et d'un sommet de couleur $j$. Dans ce contexte, l'énoncé typique d'un méta-théorème est le suivant:

**Modèle de méta-théorème:** *Toute formule exprimable en logique MO$_{VE}$/MO$_V$/FO peut être vérifiable en temps polynomial sur les graphes de largeur bornée, pour une définition de largeur telle que la largeur arborescente, la largeur de clique, ou la largeur de jumeaux.*

L'étude de tels méta-théorèmes algorithmiques s'inscrit dans une ligne de recherche vivace [2, 12], couvrant des résultats fondamentaux de Courcelle et al. [8, 9] aux récentes études de Bonnet et al. [3], en passant par des livres de type *"textbook"* de Courcelle [7] et Flum-Grohe [11].

**Vérification de propriétés de graphes exprimables dans une logique temporelle:** Pour simplifier, nous décrivons uniquement le cas fini. On appelle un graphe temporel (ou graphe de temps varié, *time varying graph*) une suite $\mathcal{G} = (G_i)_{i \in [\![0,\tau[\![}$ de graphes $G_i = (V, E_i)$, $0 \leq i < \tau$, sur le même ensemble de sommets $V$, avec $E_i \subseteq \binom{V}{2}$ et $\tau$ un entier positif ou nul. Pour mieux faire la différence avec le graphe temporel $\mathcal{G}$, on appelle aussi un graphe $G_i, 0 \leq i < \tau$ un graphe "snapshot". Soit $T = [\![0, \tau[\![$. Dans ce qui suit, tout élément de l'intervalle $T$ sera appelé un instant, tout élément de $V$ sera appelé un sommet, et tout élément de $E_i, 0 \leq i < \tau$ sera appelé arête. Notons pour qu'une arête soit bien définie dans une formule logique, l'instant $i$ doit la précéder dans la formule. On peut donc décrire des propriétés sur les graphes temporels dans $FO$ dans laquelle les variables peuvent également désigner les instants. Comme précédemment, nous pouvons donc définir plusieurs logiques sur les graphes temporels, allant de $FO$ à MO$_{VET}$ en passant par MO$_T$ ainsi que d'autres indices mixant $V$, $E$ et $T$, dans lesquelles l'indice $T$ indique que l'on s'autorise à utiliser des symboles de variables désignant des ensembles d'instants. Cependant, il est notoire que l'algorithmique des graphes temporels est particulièrement difficile; par exemple, résoudre le couplage parfait peut se faire en temps polynomial pour les graphes statiques, mais devient NP-difficile lorsqu'on le généralise aux graphes temporels [1]. Avant même de nous poser la question de méta-théorème pour les graphes temporels, beaucoup de questions restent encore à explorer. En remarquant que la notion d'intervalles de temps, si elle semble définissable dans $MO_T$ car elle nécessite de désigner des ensembles d'instants, peut en fait déjà être définie dans $FO$ en désignant uniquement les bornes de l'intervalle, on peut déjà se poser les questions suivantes:

**Question 1:** *Si la propriété porte sur un intervalle de temps, on peut vérifier cette propriété sur les différents "snapshots" en utilisant l'algorithme de vérification des graphes statiques. Existe-il cependant une façon de factoriser son expression globale afin de permettre une vérification plus rapide de la propriété?*

**Question 2:** *D'un point de vue algorithmique, y a-t-il une différence entre la vérification d'une formule FO (sur les intervalles) et celle d'une formule MO$_T$ (instants séparés)?*

On connaît déjà des cas où la réponse à la Question 1 est positive. Si nous revenons à la formule exprimant l'existant d'une paire de sommets jumeaux (FO pour les graphes statiques), des travaux existent autour de deux généralisations: jumeaux éternels et $\Delta$-jumeaux. Pour les deux généralisations, la réponse à la Question 1 est positive, via une structure de donnée inspirée des arbres bicolores [5]. Quant à la Question 2, on connaît des cas où la réponse est négative (c'est-à-dire que MO$_T$ n'est pas plus difficile à vérifier que FO). Pour le cas des jumeaux éternels, bien que les expressions soient différentes dans ces deux logiques, nous pouvons décider les deux cas par le même algorithme polynomial décrit dans [5].

Ceci dit, la réponse à ces questions, plus particulièrement à la Question 2, n'est pas toujours évidente. Si nous revenons au test de $k$-régularité (FO pour les graphes statiques), sa généralisation au cas temporel peut s'exprimer en FO ainsi: existe-t-il un instant $i_0$ à partir du quel les graphes $G_i, i_0 \leq i < i_0 + \Delta$ sont tous $k$-réguliers pendant $\Delta$ instants? Le même problème en MO$_T$ demande de tester la $k$-régularité dans un sous ensemble d'instants, au lieu d'un sous-intervalle. Il est peu clair comment déterminer le niveau de difficulté dans la vérification de l'une ou de l'autre de ces deux formules.

Notre stage s'adresse exactement à cette généralisation des problèmes connus pour les graphes statiques aux cas moins évidents dans les graphes dynamiques. Nous proposons de commencer le stage par un travail de recherche en algorithmique et programmation sur la vérification des propriétés de graphe temporels exprimables en $FO$, puis en $MO_T$. Comme présenté ci-dessus, le cas de $k$-régularité dans les graphes temporels semble peu évident, bien que la complexité du cas statique est linéaire de façon triviale. Ce problème pourra nous servir d'exemple de base avant d'étudier les logiques $FO$ et $MO_T$. Dans le cadre du stage, nous pouvons à la fois étudier le problème d'un point de vue théorique, ainsi que proposer des prototypes d'implémentation pratique. Plus spécifiquement, nous nous intéressons à la question suivante:

**Question du stage:** *Existe-il un paramètre de largeur de graphes temporels tel que la vérification d'une formule exprimable en FO/MO$_T$ sur les graphes de largeur bornée s'effectue en temps polynomial?*

## References


[1] J. Baste, B.-M. Bui-Xuan, and A. Roux. Temporal matching. *Theoretical Computer Science*, 806:184–196, 2020.

[2] B. Bergougnoux, J. Dreier, and L. Jaffke. A logic-based algorithmic meta-theorem for mim-width. In *34th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 3282–3304. SIAM, 2023.

[3] E. Bonnet, E.J. Kim, S. Thomassé, and R. Watrigant. Twin-width I: tractable FO model checking. *Journal of the ACM*, 69(1):3:1–3:46, 2022.

[4] B.-M. Bui-Xuan, A. Ferreira, and A. Jarry. Computing shortest, fastest, and foremost journeys in dynamic networks. *International Journal of Foundations of Computer Science*, 14(2):267–285, 2003.

[5] B.-M. Bui-Xuan, H. Hourcade, and C. Miachon. Computing temporal twins in time logarithmic in history length. In *9th International Conference on Complex Networks and their Applications*, volume 944 of *SCI*, pages 651–663, 2020.

[6] A. Casteigts, P. Flocchini, E. Godard, N. Santoro, and M. Yamashita. Expressivity of time-varying graphs. In *19th International Symposium on Fundamentals of Computation Theory*, pages 95–106, 2013.

[7] B. Courcelle and J. Engelfriet. *Graph Structure and Monadic Second-Order Logic - A Language-Theoretic Approach*. Cambridge University Press, 2012.

[8] B. Courcelle, J. Engelfriet, and G. Rozenberg. Handle-rewriting hypergraph grammars. *Journal of Computer and System Sciences*, 46(2):218–270, 1993.

[9] B. Courcelle, J. Makowsky, and U. Rotics. Linear Time Solvable Optimization Problems on Graphs of Bounded Clique-Width. *Theory of Computing Systems*, 33(2):125–150, 2000.

[10] T. Erlebach, M. Hoffmann, and F. Kammer. On Temporal Graph Exploration. In *42nd International Colloquium on Automata, Languages, and Programming*, volume 9134 of *LNCS*, pages 444–455, 2015.

[11] J. Flum and M. Grohe. *Parameterized Complexity Theory*. Springer, 2006.

[12] P.A. Golovach, G. Stamoulis, and D.M. Thilikos. Model-checking for first-order logic with disjoint paths predicates in proper minor-closed graph classes. In *34th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 3684–3699. SIAM, 2023.

[13] S. Gupta, A. Kosowski, and L. Viennot. Exploiting hopsets: Improved distance oracles for graphs of constant highway dimension and beyond. In *46th International Colloquium on Automata, Languages, and Programming*, volume 132 of *LIPIcs*, pages 143:1–143:15, 2019.

[14] M. Latapy, T. Viard, and C. Magnien. Stream graphs and link streams for the modeling of interactions over time. *Social Network Analysis and Mining*, 8(61), 2018.

[15] I. Tsalouchidou, R. Baeza-Yates, F. Bonchi, K. Liao, and T. Sellis. Temporal betweenness centrality in dynamic graphs. *International Journal of Data Science and Analytics*, pages 1–16, 2019.