

Inférence approchée dans les réseaux bayésiens à l'aide de l'optimisation

Thématique : optimisation combinatoire, inférence probabiliste, réseau bayésien, réseau de fonctions de coûts

Équipe d'accueil : DECISION / Statistique et Algorithmique pour la Biologie

Laboratoire d'accueil : LIP6 / Mathématiques et Informatique Appliquées de Toulouse, INRA

Lieu : Paris / Castanet-Tolosan (près de Toulouse), France

Encadrants : Pierre-Henri Wuillemin pierre-henri.wuillemin@lip6.fr Tel : 01 44 27 71 48

Simon de Givry simon.de-givry@inra.fr Tel : 05 61 28 50 74

Gratification : \approx 550 euros / mois

Contexte

Le problème étudié est celui de l'inférence approchée dans un modèle probabiliste (Peyrard, 2018). Il s'agit d'un problème d'énumération de solutions de complexité (P-SPACE) supérieure au problème d'optimisation consistant à identifier une solution la plus probable (problème NP-difficile). Plusieurs approches ont été développées pour effectuer un calcul d'inférence exact ou approché. L'approche exacte repose habituellement sur la programmation dynamique qui nécessite un espace mémoire exponentiel en un paramètre du modèle graphique (la largeur d'arbre de la décomposition arborescente associée au graphe). Les méthodes approchées exploitent l'échantillonnage ou des algorithmes par propagation de messages (*méthodes variationnelles*). L'équipe Décision conduit des recherches sur les modèles relationnels probabilistes. Elle met en oeuvre ses méthodes dans une librairie C++ conçue pour faciliter le développement d'applications utilisant des modèles graphiques tels que les réseaux Bayésiens (Naim, 2007) (**aGrum** <http://agrum.gitlab.io>). L'équipe SaAB mène des travaux en optimisation combinatoire dans les sciences du vivant et développe un outil C++ d'optimisation dans les réseaux de fonctions de coûts (Cooper, 2010) qui a remporté plusieurs compétitions sur les modèles graphiques probabilistes (**toulbar2** <http://www.inra.fr/mia/T/toulbar2>).

Sujet

L'objectif du stage est d'étudier l'apport des méthodes exactes en optimisation combinatoire pour effectuer une inférence approchée. Pour un réseau bayésien partiellement observé, l'objectif est de fournir les probabilités marginales approchées de toutes les variables non observées. Des méthodes exploitant la largeur d'arbre ont été développées en optimisation qui autorisent des largeurs bien plus grandes que pour l'inférence exacte classique (Sanchez, 2009; Allouche, 2015; Jégou, 2017). Il s'agira d'utiliser ces méthodes de manière itérative en fixant successivement chaque variable aléatoire non observée. Le travail de recherche portera sur la réutilisation partielle des calculs d'une requête d'optimisation à l'autre. Une piste d'amélioration est d'énumérer un nombre restreint de solutions optimales à l'instar de (Flerova, 2016). Une comparaison entre l'approche obtenue combinant **aGrum** et **toulbar2** et les approches existantes sera effectuée sur des benchmarks issus de la communauté des modèles graphiques.

Bibliographie

N. Peyrard, M.J. Cros, S. de Givry, A. Franc, S. Robin, R. Sabbadin, T. Schiex, M. Vignes
Exact or approximate inference in graphical models: why the choice is dictated by the treewidth, and how variable elimination can be exploited. <https://arxiv.org/abs/1506.08544>, mars 2018.

P. Jégou, H. Kanso, C. Terrioux.
Adaptive and Opportunistic Exploitation of Tree-Decompositions for Weighted CSPs.
In Proc. of ICTAI 2017, pages 366-373, Boston, MA, USA, 2017.

N. Flerova, R. Marinescu, R. Dechter.
Searching for the M Best Solutions in Graphical Models.
Journal of Artificial Intelligence Research (JAIR), volume 55, pages 889-952, 2016.

D. Allouche, S. de Givry, G. Katsirelos, T. Schiex, M. Zytnicki.
Anytime Hybrid Best-First Search with Tree Decomposition for Weighted CSP.
In Proc. of CP-15, pages 12-28, Cork, Ireland, 2015.

M.C. Cooper, S. de Givry, M. Sanchez, T. Schiex, M. Zytnicky, T. Werner.
Soft Arc-consistency revisited. In journal of Artificial Intelligence, 2010.

M. Sanchez, D. Allouche, S. de Givry, T. Schiex.
Russian doll search with tree decomposition.
In Proc. of IJCAI'09, Pasadena (CA), USA, 2009.

P. Naim, P.H. Wuillemin, P. Leray, O. Pourret, A. Becker.
Réseaux bayésiens. Eyrolles, pp.424, 2007.