Proceedings of the Seventh Annual Symposium on Combinatorial Search (SoCS 2014)

STLS: Cycle-Cutset-Driven Local Search For MPE

Alon Milchgrub
The Hebrew University of Jerusalem, Israel
alon.milchgrub@mail.huji.ac.il

Rina Dechter*
University of California, Irvine
dechter@ics.uci.edu

Abstract

In this paper we present *Stochastic Tree-based Local Search* or $STLS$, a local search algorithm combining the notion of cycle-cutsets with the well-known Belief Propagation to approximate the optimum of sums of unary and binary potentials. This is done by the previously unexplored concept of traversal from one cutset to another and updating the induced forest, thus creating a local search algorithm, whose update phase spans over all the forest variables.We study empirically two pure variants of $STLS$ against the state-of-the art GLS^+ scheme and against a hybrid.

Introduction

The problem of optimizing discrete multivariate functions or "energy functions" described as sums of potentials on (small) subsets of variables is one of fundamental importance and interest in a wide variety of fields, such as computer vision and graphical models. In the context of the latter, conditional probability tables (CPT) are used to describe the relations between the variables of a model. Instances of this problem arise in the form of Most Probable Explanation (MPE) problems, where finding a maximum of such energy functions composed of the CPTs translates to finding an assignment of maximum probability given some partial assignment as evidence.

Background

Definition 1 (Energy Minimization Problem). let $\bar{x} = x_1, \ldots, x_N$ be a set of variables over a finite domain $\mathcal{D}$, let $\varphi_i : \mathcal{D} \to \mathbb{R}$ for $i \in \{1, \ldots, N\}$ be unary potentials, and let $\psi_{i,j} : \mathcal{D}^2 \to \mathbb{R}$ for a subset of pairs $E \subseteq \{\{i,j\} : 1 \leq i < j \leq N\}$ be binary potentials, then *the problem of energy minimization* is finding

$$\bar{x}^* = argmin_{\bar{x}} \sum_i \varphi_i(x_i) + \sum_{\{i,j\} \in E} \psi_{i,j}(x_i, x_j)$$

Definition 2 (Cycle-Cutset). Let $G = (V, E)$ be an undirected graph. A *cycle-cutset* in G is a subset C of V, such that the graph induced on $V' = V \setminus C$ is acyclic.

*This work was sponsored by the NSF grant IIS-1065618 and in part by the United States Air Force under Contract No. FA8750-14-C-0011 under the DARPA PPAML program.

Algorithm 1 pseudo code of STLS/STLS*.
STLS and STLS* differ in the restart and the cutset generation procerdures (see text). $is\_stagnated$ is set to TRUE once no change has been made by both the BP and the local search stages during a given number of iterations.

Input: Graph $G = (V, E)$ annotated with potentials φ_i and $\psi_{i,j}$.
Output: An assignment $\bar{x}$ which achieves the minimum energy found in time t.

$\bar{x} \leftarrow$ `InitializeValues`
while $runtime < t$ **do**
  $C \leftarrow$ `GenerateCutset(`G`)`
  $F \leftarrow V \setminus C$;
  `// Alternate BP on forest variables and local search on cutset variables`
  repeat
    $\bar{x}|_F \leftarrow$ `BP_min_sum(`$G, \bar{x}|_C, F$`)`
    $\bar{x}|_C \leftarrow$ `SubsidiaryLocalSearch(`$G, \bar{x}, C$`)`
  until *no change in* $\bar{x}$;
  if $is\_stagnated$ **then**
    $\bar{x} \leftarrow$ `InitializeValues`
    $is\_stagnated \leftarrow$ FALSE
  end
end

Given an instance of the energy minimization problem, a primal graph can be built, where every variable x_i is assigned a vertex, and two vertices x_i and x_j are connected if there exists a potential $\psi_{i,j}$. If the resulting graph is acyclic, the problem can be solved efficiently using Belief Propagation (BP) (Pearl 1988). If the graph is not acyclic, a cycle cutset can be generated and an optimal assignment to the forest variables given the assignment to the cutset variables can be found in a method known as "*cutset-conditioning*"(Pearl 1988; Dechter 2013).

STLS: Stochastic Tree-based Local Search

(Pinkas and Dechter 1995) suggested iteratively conditioning on a different cutset and finding exact optimal solution on the rest variables as a possible scheme for dealing with cycles in the graph. The algorithm can additionally perform regular local search on the cutset variables. However, they did not go further to establish the capabilities of this method. The operation of of $STLS$ is given in Algorithm 1. In ev-

Set (# inst.)		STLS Best	STLS Ratio	STLS* Best	STLS* Ratio	Hybrid Best	Hybrid Ratio	GLS+ Random Best	GLS+ Random Ratio	GLS+ Best	GLS+ % of best
Grids (21)	mean	2 (1)	1	2 (4)	1	**17** (1)	**1.02**	0 (0)	0.75	0 (15)	95%
	max	3 (2)	**1.04**	3 (9)	**1.04**	**15** (2)	1.03	0 (0)	0.86	0 (8)	
CSP (29)	mean	15 (3)	**1.19**	14 (7)	1.18	13 (7)	1.17	20 (5)	1	**21** (4)	77%
	max	19 (1)	**1.19**	18 (5)	**1.19**	16 (11)	1.18	20 (5)	1	**21** (4)	
Protein. (9)	mean	2 (4)	**1**	0 (0)	0.99	3 (1)	**1**	**5** (2)	**1**	**5** (2)	**100%**
	max	**6** (1)	**1**	1 (2)	0.99	5 (2)	**1**	5 (1)	**1**	5 (2)	
SGM. (90)	mean	32 (29)	0.94	0 (6)	0.85	37 (40)	0.99	53 (33)	**1**	**57** (32)	**100%**
	max	37 (32)	0.96	30 (23)	0.96	42 (43)	0.99	**55** (35)	**1**	52 (35)	

Table 1: Statistics for average and maximal results over 10 runs on sets from PIC2011. The values presented refer to the results obtained after 1 minute for the Segmentation set and 3 minutes for all other sets. Best is the number of the instances for whom the algorithm achieved the best result (and second best in parenthesis). Ratio is the average ratio of the result obtained by the algorithm to that of classic GLS$^+$. For GLS$^+$ the average ratio of the result to the best overall result is presented.

ery iteration an optimal assignment to the forest variables is generated given the values of the cutset variables using *BP min-sum*. Therefore, the energy of the system can not increase from iteration to iteration and the resulting algorithm is a local search algorithm finding the optimal solution on all the forest variables in every iteration.

Two variants of $STLS$ were tested, one which used the cutset selection procedure of (Becker, Bar-Yehuda, and Geiger 2000) and a simple restart scheme, and a more complex variant named $STLS^*$, which uses previous assignments to estimate a "certainty index". The assignment history and the certainty index are used to heuristicly generate new cutsets and in the initialization of restarts.

Experiments

The two variants of $STLS$ were compared to GLS^+ (Hutter, Hoos, and Stützle 2005), another local search algorithm, considered for the last decade to be the state-of-the-art, as well as to a simple hybrid of $STLS$ and GLS^+. All algorithms were run 10 times on problems from the Grids, CSP, ProteinFolding and Segmentation problem sets of the PASCAL2 Probabilistic Inference Challenge (PIC2011)[1] (see (Lee, Lam, and Dechter 2013) for a summary of the statistics of these benchmark sets). The resulting energies were sampled after 0.1, 1, 10, 60, 120, and 180 seconds (Segmentation problems only up to 1 minute), and all the results of a given instance were linearly normalized to the interval $[0, 1]$, mapping the worst result to 0 and the best to 1. In $STLS$ and $STLS^*$ the variables were initialized to an *undefined* value, thus ignored until obtaining a valid value. GLS^+ was initialized either randomly or using the customary Mini-Bucket heuristic. The cutset variables were updated using the Hopfield Model activation function as the local search algorithm mentioned in line 7 of Algorithm 1. See results in Table 1.

Results and Future Work

As can be seen in Table 1, although the GLS^+ variants do manage to produce the best results more often, especially in the Segmentation benchmark, the average ratios of the average and maximal results obtained by the various algorithms to those of the classic GLS^+ range from slight superiority for GLS^+ (ratio < 1) on the Segmentation benchmark to significant dominance of the $STLS$ algorithms on the CSP benchmark. This implies that while GLS^+ manages to produce the best results in many cases, it does not significantly outperform $STLS$ and it struggles considerably on some instances. The overall improved performance of the hybrid implies that either algorithm acts as a mechanism for avoiding local optima by the other, and suggests the union of both algorithms as a promising trail for further improvement.

In future work, the algorithm should be extended to handle potentials of higher arity than 2 as well. Importantly, $STLS$ yields strong local optima (conditional optimal on every forest), and therefore, in the limit it is as good as min-sum belief propagation in quality, while it can be more effective computationally. Comparing with specific loopy belief propagation scheme is left for future work as well.

[1] http://www.cs.huji.ac.il/project/PASCAL/index.php

References

Becker, A.; Bar-Yehuda, R.; and Geiger, D. 2000. Randomized algorithms for the loop cutset problem. *J. Artif. Int. Res.* 12(1):219–234.

Dechter, R. 2013. *Reasoning with Probabilistic and Deterministic Graphical Models: Exact Algorithms*. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers.

Hutter, F.; Hoos, H. H.; and Stützle, T. 2005. Efficient stochastic local search for mpe solving. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence*, IJCAI'05, 169–174. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Lee, J.; Lam, W.; and Dechter, R. 2013. Benchmark on daoopt and gurobi with the pascal2 inference challenge problems. In *DISCML*.

Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Pinkas, G., and Dechter, R. 1995. Improving connectionist energy minimization. *J. Artif. Int. Res.* 3(1):223–248.