Stochastic Constraint Optimisation with Applications in Network Analysis

Anna L. D. Latour[1,2] *, Behrouz Babaki[3], Daniël Fokkinga[1], Marie Anastacio[1],
Holger Hoos[1,4], and Siegfried Nijssen[5]

[1] Leiden University, Leiden, the Netherlands
[2] National University of Singapore (NUS), Singapore
[3] Polytechnique Montréal, Montreal, Canada
[4] University of British Columbia (UBC), Vancouver, Canada,
[5] UCLouvain, Louvain-la-Neuve, Belgium

Abstract

We present an extensive study of methods for exactly solving *stochastic constraint (optimisation) problems* (SCPs). Each of these methods combines *weighted model counting* (WMC) with constraint satisfaction and optimisation. They use knowledge compilation to address the model counting problem; subsequently, either a *constraint programming* (CP) solver or *mixed integer programming* (MIP) solver is used to solve the overall SCP. We first explore a method that is generic and decomposes constraints on probability distributions into a multitude of smaller, local constraints. For the second method, we show that many probability distributions in real-world SCPs have a monotonic property, and how to exploit that property to develop a global constraint. We discuss theoretical (dis)advantages of each method, with a focus on methods that employ *ordered binary decision diagrams* (OBDDs) to encode the probability distributions, and then evaluate their performances experimentally. To configure the space of parameters of these approaches, we propose to use the framework of *programming by optimisation* (PbO). The results show that a CP-based pipeline obtains the best performance on well-known data mining and frequent itemset mining problems.

This is an extended abstract of a publication in the Artificial Intelligence Journal (AIJ), 2022 [7]. We presented a preliminary version of this work at the International Workshop on Model Counting (MCW), 2020 [5].

---

*Corresponding author, latour@nus.edu.sg.

Summary

Making decisions under uncertainty is an important problem in business, governance and science. We find examples in the fields of planning and scheduling, as well as in fields like data mining and bioinformatics.

Many of these problems can be formulated on *probabilistic networks*. Examples are social networks [1], where we are uncertain about how likely people are to adopt ideas from others, and power grid networks [2], where we are uncertain about the reliability of power lines. We can represent the uncertainty by associating probabilities with edges or nodes.

Consider the *spread of influence* problem [4] in more detail. Given a social network of people (vertices) that have stochastic relationships (edges), we want to use word-of-mouth advertisement to turn friends of our customers into new customers. We start this spread-of-influence campaign by distributing at most k free product samples to members of the network. Each node may convince a neighbour with a probability indicated on the edge towards its neighbour. What is the k-sized set S of most influential nodes in this network?

This is an instance of a general class of problems, called *stochastic constraint (optimisation) problems* (SCPs). SCPs have the following characteristics: (1) they involve (Boolean) random variables and (Boolean) decision variables; (2) they require finding assignments to decision variables, such that an optimisation criterion or constraint on the random variables is satisfied; (3) they may involve other constraints on the decisions we can take.

The example above has a fourth characteristic: (4) the optimisation criterion requires the calculation of an expectation over the probabilistic variables given an assignment to the decision variables, where the expectations are higher if more nodes are selected. This makes its probability distribution *monotonic* and makes the problem an instance of a *stochastic constraint (optimisation) problem on monotonic distributions* (SCPMD). While (4) seems limiting, problems with this characteristic are plentiful.

SCPs are NP-hard, as they generalise NP-hard constraint satisfaction and optimisation problems. The approach we take is to treat SCPs as problems that combine *weighted model counting* (WMC) with constraint satisfaction and optimisation: we are looking for assignments to decision variables, such that a WMC is either optimised or satisfies a given constraint. Hence, we treat the problem of expectation calculation as a WMC problem.

We evaluate two methods for solving SCPs exactly. They have in common that they use *knowledge compilation* to address the WMC problem. The challenge is how to incorporate the results of the knowledge compilation in a subsequent step that solves the overall SCP.

The first method converts *decision diagram* representations of probability distributions into *arithmetic circuits* (ACs) that are used to compute conditional probabilities through WMC. A constraint on a distribution translates to a constraint on the AC. This constraint is then *decomposed* into a multitude of local constraints, which are solved by an off-the-shelve CP or MIP solver [6]. This *decomposition method* is solver-agnostic and straightforwardly implemented. However, some relations between variables are lost during decomposition, causing the method to prune the search space inadequately. Specifically, it does not guarantee *generalised arc consistency* (GAC).

The second method preserves GAC by introducing a *global* constraint for SCPMDs (whose underlying probability distributions are *monotonic*). We propose and implement a constraint propagator for this *stochastic constraint on monotonic distributions* (SCMD), which

preserves relations between variables. Because the monotonicity of the distribution is crucial to preserving GAC, this *global constraint method* is less general than decomposition.

We followed the paradigm of *programming by optimisation* [3] and implemented alternative *design choices*. We automatically configured the resulting solvers on problem instances from different domains and compared the PAR10 values of the resulting optimised methods. We found that the global method performs comparably and largely complementarily to a MIP-based decomposition method for smaller instances, but outperforms this variant of the decomposition method for larger instances. The optimised global method outperforms the CP-based decomposition method by up to two orders of magnitude.

References

[1] Domingos, P.M., Richardson, M.: Mining the network value of customers. In: KDD. pp. 57–66. ACM (2001)

[2] Dueñas-Osorio, L., Meel, K.S., Paredes, R., Vardi, M.Y.: Counting-based reliability estimation for power-transmission grids. In: AAAI. pp. 4488–4494 (2017)

[3] Hoos, H.H.: Programming by optimization. Commun. ACM **55**(2), 70–80 (2012)

[4] Kempe, D., Kleinberg, J.M., Tardos, É.: Maximizing the spread of influence through a social network. In: KDD. pp. 137–146. ACM (2003)

[5] Latour, A.L.D., Babaki, B., Fokkinga, D., Anastacio, M., Hoos, H.H., Nijssen, S.: Stochastic constraint optimisation with applications in network analysis (extended abstract). In: WMC 2020 (jul 2020)

[6] Latour, A.L.D., Babaki, B., Dries, A., Kimmig, A., Van den Broeck, G., Nijssen, S.: Combining stochastic constraint optimization and probabilistic programming — From knowledge compilation to constraint solving. In: CP. Lecture Notes in Computer Science, vol. 10416, pp. 495–511. Springer (2017)

[7] Latour, A.L.D., Babaki, B., Fokkinga, D., Anastacio, M., Hoos, H.H., Nijssen, S.: Exact stochastic constraint optimisation with applications in network analysis. Artif. Intell. **304**, 103650 (2022)