Abstract: Stochastic Constraint Programming and Reinforcement Learning

S. D. Prestwich[1], R. Rossi[2], S. A. Tarim[3]
[1]Insight Centre for Data Analytics, University College Cork, Ireland
[2]University of Edinburgh Business School, Edinburgh, UK
[3]Department of Management, Cankaya University, Ankara, Turkey
s.prestwich@cs.ucc.ie, roberto.rossi@ed.ac.uk, at@cankaya.edu.tr

Stochastic Constraint Programming (SCP) is designed to model and solve complex problems involving uncertainty and probability. It was first proposed in [3, 20] and developed in [18]. SCP problems are in a higher complexity class than CP problems and can be much harder to solve, but many real-world problems contain elements of uncertainty so SCP is an important problem class.

Many improvements have been made to SCP in recent years. Better search and filtering algorithms have been proposed [2, 6, 9, 15, 19]. Confidence intervals can be applied to control approximations [16]. If a problem has many decision variables metaheuristics can be applied [11, 14, 20]. If there are many scenarios then we can apply scenario reduction by sampling [8], and in special cases a subset can be used without approximation [13]. However, so far SCP has not been applied to very large problems.

In contrast, far larger stochastic (and adversarial) problems have been successfully solved by methods from Reinforcement Learning (RL) [10, 17], and the related fields of Neuro-Dynamic Programming [4] and Approximate Dynamic Programming [12]. In RL researchers are not concerned with sample sizes, confidence intervals or other statistical issues, though they are concerned with convergence. RL methods have been successfully applied to real-world problems in many fields including robotics, control, game playing, trading and human-computer interfaces.

We argue that SCP has been held back by its roots in Stochastic Programming [5] and that a more scalable future direction lies in Machine Learning. As a first step we have implemented the simple TD(0) algorithm [17] in the Eclipse constraint system [1] and applied it to a few problems. In our framework, assigning a decision variable is an RL *action*, assigning a random variable is an RL *environmental response* that moves to a new state, a constraint is an environmental response that restricts future actions, a partial assignment is an RL *state*, a complete assignment is an RL *terminal state* achieved by an RL *episode*, the empty assignment is the RL *initial state*, and the objective function is the RL *reward*. *State aggregation* techniques are used to handle exponential numbers of scenarios.

Initial results are promising: our algorithm was competitive with a specialised Monte Carlo method on a hard 1-stage problem with a billion scenarios, and on a multistage inventory control problem it found a policy within a few percent of optimal. Our approach might also be applied to adversarial problems such as those in [7].

Acknowledgments This publication has emanated from research supported in part by a research grant from Science Foundation Ireland (SFI) under Grant Number SFI/12/RC/2289.

References

1. K. R. Apt, M. Wallace. Constraint Logic Programming Using Eclipse. Cambridge University Press, 2007.
2. T. Balafoutis, K. Stergiou. Algorithms for Stochastic CSPs. *12th International Conference on Principles and Practice of Constraint Programming, Lecture Notes in Computer Science* vol. 4204, Springer, 2006, pp. 44–58.
3. T. Benoist, E. Bourreau, Y. Caseau, B. Rottembourg. Towards Stochastic Constraint Programming: A Study of On-Line Multi-Choice Knapsack with Deadlines. *7th International Conference on Principles and Practice of Constraint Programming, Lecture Notes in Computer Science* vol. 2239, Springer, 2001, pp. 61–76.
4. D. P. Bertsekas, J. N. Tsitsiklis. Neuro-Dynamic Programming. Athena Scientific, 1996.
5. J. Birge, F. Louveaux. Introduction to Stochastic Programming. Springer Series in Operations Research, 1997.
6. L. Bordeaux, H. Samulowitz. On the Stochastic Constraint Satisfaction Framework. *ACM Symposium on Applied Computing*, 2007, pp. 316–320.
7. K. N. Brown, J. Little, P. J. Creed, E. C. Freuder. Adversarial Constraint Satisfaction by Game-tree Search. *Proceedings of ECAI*, 2004, pp. 151–155.
8. J. Dupăcová, N. Gröwe-Kuska, W. Römisch. Scenario Reduction in Stochastic Programming: an Approach Using Probability Metrics. *Mathematical Programming Series A* **95**:493–511, 2003.
9. B. Hnich, R. Rossi, S. A. Tarim, S. D. Prestwich. Filtering Algorithms for Global Chance Constraints. *Artificial Intelligence* **189**:69–94, 2012.
10. L. P. Kaelbling, M. L. Littman, A. W. Moore. Reinforcement Learning: A Survey. *Journal of Artificial Intelligence Research* **4**:237–285, 1996.
11. B. Kawas, M. Laumanns, E. Pratsini, S. Prestwich. Risk Averse Production Planning. *2nd International Conference on Algorithmic Decision Theory, Lecture Notes in Computer Science* vol. 6992, Springer 2011, pp. 108-120.
12. W. B. Powell. Approximate Dynamic Programming. John Wiley & Sons, 2007.
13. S. D. Prestwich, M. Laumanns, B. Kawas. Value Interchangeability in Scenario Generation. *19th International Conference on Principles and Practice of Constraint Programming, Lecture Notes in Computer Science* Vol. 8124, Springer, 2013, pp. 587–595.
14. S. D. Prestwich, S. A. Tarim, R. Rossi, B. Hnich. Hybrid Metaheuristics for Stochastic Constraint Programming. *Constraints*, 2014.
15. R. Rossi, S. A. Tarim, B. Hnich, S. D. Prestwich. Cost-Based Domain Filtering for Stochastic Constraint Programming. *14th International Conference on Principles and Practice of Constraint Programming, Lecture Notes in Computer Science* 5202, Springer, 2008, pp. 235–250.
16. R. Rossi, B. Hnich, S. A. Tarim, S. Prestwich. Finding (α,θ)-Solutions via Sampled SCSP. *22nd International Joint Conference on Artificial Intelligence, AAAI Press/International Joint Conferences on Artificial Intelligence*, 2011, pp. 2172–2177.
17. R. S. Sutton, A. G. Barto. Reinforcement Learning: An Introduction. MIT Press, Cambridge, MA, 1998.
18. S. A. Tarim, S. Manandhar, T. Walsh. Stochastic Constraint Programming: A Scenario-Based Approach. *Constraints* 11(1):1383–7133, 2006.
19. S. A. Tarim, I. Miguel. A Hybrid Bender's Decomposition Method for Solving Stochastic Constraint Programs with Linear Recourse. *Lecture Notes in Computer Science* vol. 3978, Springer, 2006, pp. 133–148.
20. T. Walsh. Stochastic Constraint Programming. *15th European Conference on Artificial Intelligence*, 2002, pp. 111–115.