Cardinality optimisation in constraint-based modelling of metabolism

Hoai Minh Le, Hulda S. Haraldsdottir, Tu Vuong Phan, Ines Thiele, **Ronan M.T. Fleming**
(*ronan.mt.fleming@gmail.com)
[1]Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg.

http://stanford.edu/group/SOL/multiscale/

Project Goals: A growing number of genome-scale biochemical networks are becoming available, especially for metabolism. Although the models can be constructed from omics data, several types of errors can occur that make the quality of the model insufficient. manually testing the accuracy of models exceeding several thousands of metabolites and reactions is impractical. Hence, it is important to develop tools to verify the biological models, or to automatically detect and correct errors. These tools often involve the solution of an optimisation problem with a discrete solution. In general, this problem is computationally intractable at large scale but in specific cases continuous approximation can yield an adequate solution, even for large computational models.

Several biochemical applications can be mathematically formulated as a cardinality optimisation problem [1, 2, 3]. The cardinality optimisation problem refers to an optimisation problem that involves optimising the number of non-zero components of vector, that is, the ℓ_0 norm. In general, this cardinality optimisation problem is computationally intractable at large scale because of the discontinuity of the ℓ_0 norm renders the problem NP-hard. In this work, we investigate a non-convex approximation approach which consists in replacing the ℓ_0 norm by non-convex continuous functions. The resulting problem is non-convex and can be reformulated as a Difference of Convex (DC) program. Based on the theory of DC programming and DC Algorithm (DCA), we present a generic algorithm for solving several cardinality optimisation problems arising in constraint-based modelling of biochemical networks. As application of the proposed algorithm, we study three problems in systems biochemistry: the detection of the stoichiometrically consistent part of a metabolic network, the sparse flux balance analysis and relaxation of constraints to obtain non-zero steady state flux in an otherwise infeasible flux balance analysis problem. Numerical tests on various biochemical networks from a range of species clearly show that our algorithms outperform existing related algorithms.

References

[1] Vlassis N, Pacheco MP, Sauter T (2014) Fast reconstruction of compact context-specific metabolic network models. PLoS Comput Biol 10: e1003424.

[2] Gevorgyan A, Poolman MG, Fell Da (2008) Detection of stoichiometric inconsistencies in biomolecular models. Bioinformatics 24: 2245–51.

[3] Fleming RM, Vlassis N, Thiele I, Saunders MA (2015) Conditions for duality between fluxes and concentrations in biochemical networks. arXiv preprint arXiv:151202690 .

This work was supported by the U.S. Department of Energy, Offices of Advanced Scientific Computing Research and the Biological and Environmental Research as part of the Scientific Discovery Through Advanced Computing program, grant #DE-SC0010429.