An algorithm with improved time complexity for the calculation of elementary flux modes

Axel von Kamp and Stefan Schuster
Friedrich-Schiller-Universität Jena, Biologisch-Pharmazeutische Fakultät
Ernst - Abbe - Platz 1-2, 07743 Jena
kamp@minet.uni-jena.de, schuster@minet.uni-jena.de

seit 1558

1. Introduction

- Elementary (flux) modes serve as a mathematical formalization of metabolic pathways and describe the fluxes through metabolic networks under steady-state conditions [2].
- For the calculation of elementary modes, the metabolites in the network have to be classified as either external (buffered) or internal (balanced by production/consumption).
- The set of elementary modes for a given metabolic system is characterized by the following properties:
  1. The elementary modes are unique up to scaling factors.
  2. Fluxes through irreversible reactions are non-negative.
  3. Elementarity: Let S(m) be the set of reactions that do not occur in elementary mode m. Then for all other modes n≠m it follows that $S(n) \not\subset S(m)$.
- Elementarity entails that no elementary mode is a subset of any other mode.
- Elementary modes can also be used to assess the robustness and versatility of metabolic networks.

2. Objectives

- The calculation of elementary modes in large or complex networks often meets with the problem of combinatorial explosion. This leads to a concurrent growth in computation time and space for the calculation of the modes.
- One of the major factors which contributes to the time complexity is the test for elementarity of the modes during calculation.
- Two different methods for performing the elementarity test are evaluated with respect to their efficiency.

3. Methods

During calculation of the elementary modes, new candidate modes are generated by combining accepted modes from the preceding iteration [3]. The two methods for testing whether the candidate mode is elementary are:

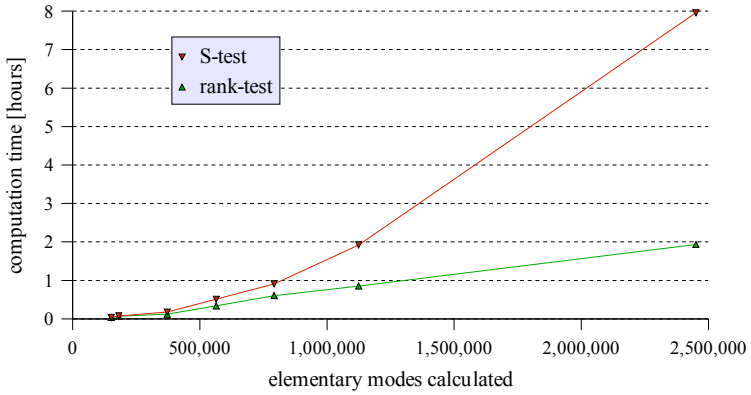
1. *S-test*:
   - Tests each candidate against the already accepted modes.
   - Although it is possible to restrict the subset of modes against which the test has to be made, the time complexity of this test remains correlated to the number of accepted modes.
2. *rank-test*:
   - Takes preprocessed stoichiometric matrix R as a parameter.
   - For each candidate a submatrix of R is created by taking those columns that correspond to the reactions in the candidate mode. The rank of this submatrix determines elementarity.
   - The time complexity of this test scales with the number of reactions in the candidate mode.

5. Conclusions

- The rank-test allows for a more efficient computation of elementary modes than the S-test, but the problem of combinatorial explosion remains.
- Parallelization can further increase performance; for this the rank-test is well suited, because once it is set up with the modified stoichiometric matrix it only needs the candidate mode as input to determine whether it is elementary or not.

4. Results

Figure 1: Computation times for S3 during final iterations



- In total, the computation times of three metabolic systems were measured (table 1). The comparison of system S1 with systems S2/S3 shows that the speed-up with the rank test increases with the total number of modes calculated.
- Figure 1 shows that the difference between the computation times increases as the number of calculated modes increases. This is due to the fact that the time needed to perform the S-test increases with the number of elementary modes. On the other hand, the time complexity of the rank test is limited by the total number of reactions in the system which remains constant during calculation.

Test systems (*E. coli* central metabolism) for benchmarks

system name		S1	S2	S3
substrates		Glc	S1 + Ac, Gly, Succ	S2 + Asp
products		CO_2, Ac, Form, Eth, Lac		S2 + Succ
reactions (reduced)		106 (42)	110 (47)	112 (51)
metabolites (reduced)		89 (25)	89 (26)	89 (28)
elementary modes		27,100	508,632	2,450,787
computation time	S-test	2.0 sec	14 min 33 sec	477 min 27 sec
	rank-test	1.7 sec	3 min 55 sec	116 min 01 sec
speed-up with rank-test		15%	73%	76%

Table 1: Details of the test systems used for calculating the benchmarks. The metabolic systems encompass a significant part of the *E. coli* central metabolism and only differ in their substrates/products (external metabolites); for details see [1]. The computation times were measured on a standard desktop PC with a Pentium 4 processor at 2.4 GHz and 512 MB RAM. Only the binary representations of the modes were calculated. Computation time does not include back transformation to obtain flux values; for details of this procedure see [1].

References

1. Gagneur, J. and Klamt, S.: 'Computation of elementary modes: a unifying framework and the new binary approach.', under review.
2. Schuster, S.; Hilgetag, C.; Woods, J., H.; Fell, D., A.: 'Reaction routes in biochemical reaction systems: algebraic properties, validated calculation procedure and example from nucleotide metabolism.', *J. Math. Biol.*, 2002, **45**, pp. 153-181.
3. Urbanczik, R. and Wagner, C.: 'A novel algorithm for stoichiometric network analysis.', under review.