**Exercise for**

**Data Integration in Cellular Networks**

**11 April 2024**

**Modified pFBA**

It is expected that when cells are growing exponentially, there will be selection for the fastest growers, and then among the fastest growers, there will be a fitness advantage to cells using the least amount of enzyme.

The goal of this exercise is to minimize the weighted sum of enzyme-associated fluxes, subject to optimal biomass. The weight for an enzyme-associated reaction is defined by the minimum number of genes need to be present to have flux through that reaction. The example below shows how this weight can be retrieved.

|  |  |  |  |
| --- | --- | --- | --- |
| **Reaction** | **GPR Rule** | **Essential gene set(s)** | **Weight** |
| A | A | {A} | 1 |
| B | B1 or B2 | {B1},{B2} | 1 |
| C | C1 and C2 | {C1,C2} | 2 |
| D | D1 or (D2 and D3) | {D1},{D2,D3} | 1 |

Use the *E. coli* model ecoli\_core\_model.xml from BiGG database to determine the minimum weighted sum of enzyme-associated fluxes at the optimum biomass in *E. coli*.

**Step by step to implement modified pFBA:**

1. Using the Cobra Toolbox to read the *E. coli* model from the file provided in BiGG database.
2. Use the function convertToIrreversible to split all reversible reactions into two irreversible reactions.
3. Run FBA using the function optimizeCbModel to determine the optimum biomass for the model.
4. Write a MATLAB function that for a given reaction uses its GPR rule to calculate the weight. This function associate the reaction to . Clearly, for a reaction with no associated GPR rule, the weight is zero.
5. Write an appropriate linear programming to solve the following problem:

s.t