Logic and linear programming for seed identification in metabolic networks
Chabname Ghassemi Nedjad, Clémence Frioux, Loïc Paulevé

To cite this version:
Chabname Ghassemi Nedjad, Clémence Frioux, Loïc Paulevé. Logic and linear programming for seed identification in metabolic networks. Biorégul 2023 - Modélisation Formelle de Réseaux de Régulation Biologique, Jun 2023, Porquerolles (Hyères), France. pp.1-1, 2023. hal-04328778

HAL Id: hal-04328778
https://inria.hal.science/hal-04328778
Submitted on 7 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
From Genome-Scale Metabolic Network to behaviour prediction

**Optimization:**
- Genome-scale metabolic networks (GSMNs) [1]
- Productivity modelling: network expansion [2]

**Prediction of behaviour** of an organism in a particular environment:
- Environmental response data
- Mathematical modeling

**Linear programming problems** to solve:
- Flux Balance Analysis (standard)
- Discrete dynamic models

---

Reverce ecology [3]: Identification of environmental nutrients

**What?**
- Seeds: precursor compounds to produce essential metabolites

**Why?**
- Ensure the growth of a species, represented by the biomass reaction

**Purpose?**
- Understand the link between a system and its environment [4]
- Design of culture for uncultivated species

---

**Data and additional approaches**

- 108 GSMNs from database BIGG
- Up to thousands of reactions for a species
- Import Reactions removed
- Validation: fba with cobra.py

**Objectives:**
- **Full Network:** All metabolites are produced (fba test only)
  - only on Biomass Reaction described on GSMNs
- **Target:** Focus on producing only the metabolites needed
  - for the Biomass Reaction described on GSMNs

**Optimization:**

<table>
<thead>
<tr>
<th>P: set of solutions</th>
<th>A,B,C,D</th>
<th>≠ O</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subset Minimal:</td>
<td>A,B ∈ P such as A ∩ B = ∅, S = {A}</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A,C ∈ P such as A ∩ C = ∅, S = {A,C}</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A,D ∈ P such as A ∩ D ≠ ∅ and A ∩ D ≠ A, S = {A,B,C,D}</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A ∈ P such as v∈Y ∈ A, size(y) = size(y) and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B ∈ P such as v∈Y ∈ B, v∈A, size(y) &gt; size(y)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S = {A}</td>
<td></td>
</tr>
</tbody>
</table>

---

**Discussion and Perspective**

**Work in Progress:**
- Improve hybrid method (limited tractability)
- Evaluation of solutions of different methods for steady growth
- Evaluate diversity and size of identified seeds compared to other methods

**Discussion:**
- Improve discrete method to avoid hybrid methods

**Perspective:**
- Optimization for minimize in order to obtain solutions in less than 20 min

---

**References:**


---

**ASP**

**Answer Set Programming**

**Hybrid Method**

**Assumption:**
- Transient dynamic

**Approach:**
- Dynamic Programming

**Result:**
- 10 to 20 subset minimal models found for few species (≤20min)
- 1 minimize model found for 2 species

**Data:**
- 108 GSMNs from database BIGG
- Up to thousands of reactions for a species
- Import Reactions removed
- Validation: fba with cobra.py

**Objectives:**
- **Full Network:** All metabolites are produced (fba test only)
  - only on Biomass Reaction described on GSMNs
- **Target:** Focus on producing only the metabolites needed
  - for the Biomass Reaction described on GSMNs

---

**Discrete Method**

**Assumption:**
- Discrete Programming

**Approach:**
- Can_reach, can_reach

**Result:**
- 20 subset minimal models found for the whole benchmark (≤20min)
- 20 minimize models found for 3 species

---

**Discussion:**
- Improve discrete method to avoid hybrid methods

---

**References:**