ILP approaches on the DNA recombination problem

Timeline

May, 22

Preliminary explanations.

July, 3

Formal stage start. Problem definition.

July, 6

ILP: first constraints and properties.

July, 27

Review of the ILP formulation: corrections, objective function (min MDSs quantity), functions definitions (equivalence, inverse, MIC/MAC substring notation). LP solver: *GUROBI.*

Preprocessing phase (properties directly verifiable on the instance):

- Generate all possibile MDSs annotations
- Compute values of some variables
- Python or Ruby

August, 21

- Constraints description or brief explanation;
- Group similar constraints;
- Correct every non-compliant constraint (e.g. `if` cannot be used, if not in the preprocessing phase);
- Correct `Cov_{MIC}(i,j)`;
- Remove useless and redunant constraints;
- Necessary variable and constraints: proof of correctness;
- Missing constraints (e.g. MDS_End > MDS_Start);
- Preprocessing: use `find` or Python's `re.search` on every MIC's substring to find if they exist in the MAC (Consider length > 3 as minimum for IESs and MDSs);
- Thesis: should be understandable to CS undergraduates. 35-40 pages. Main elements:
    1. Introduction
    2. Prerequisites
    3. What I've learnt during the stage experience
    4. What I've done during the stage

Proof of correctness:

Being I an instance of the problem, P the correspondent ILP formulation, A any solution of P:

1. Show how to use a solution of P (computed by Gurobi) to build a solution of the starting problem;
2. Show that (1) is always possible.

Resources

- *DNA recombination through assembly graphs* - Angela Angeleska, Nataša Jonoska, Masahico Saito
- `mds ies db`: *a database of ciliate genome rearrangements* - Jonathan Burns, Denys Kukushkin, Kelsi Lindblad, Xiao Chen, Natasa Jonoska and Laura F. Landweber
- MDS and IES annotation algorithm (Python) used in `mds ies db`
  - BLAST - *Basic Local Alignment Search Tool*
- *Programmed genome rearrangements in the ciliate Oxytricha* - V. Talya Yerlici, Laura F. Landweber
- *RNA-guided DNA assembly* - Angela Angeleska, Natasa Jonoska
- Gusfield ILP tutorial