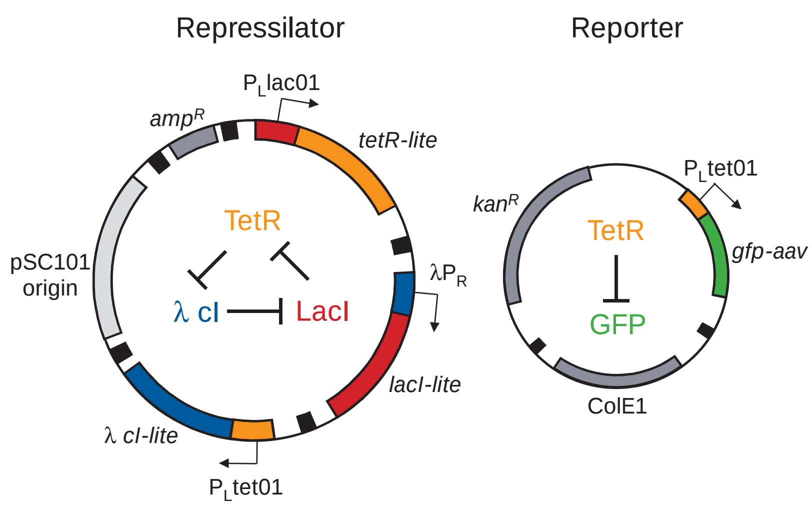
Models and Languages for Computational Systems Biology
Session 2009–2010, Semester 2

Assignment 1

This is the first of two pieces of credit-bearing coursework for MLCSB, each marked out of 100 and each contributing 15% to the final course grade. Please submit your written coursework to the collection box outside the Informatics Teaching Organisation office on level 4 of the Appleton Tower. This work will be marked anonymously: please use the cover sheet provided by the ITO, writing on it the title of the course, and your name in the corner where the ITO can conceal it for anonymisation. Do not mark your name, student number, or exam number anywhere else.

Due date: 4pm Thursday 25 February 2010 (2010-02-25 16:00Z) *Ian Stark*

The *repressilator* is a landmark synthetic regulatory network, demonstrating a feedback loop among three genes.



M. B. Elowitz and S. Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–338, January 2000. DOI:10.1038/35002125

The following questions concern the Petri net in Figure 1. This gives a highly simplified model of the repressilator network, omitting the fluorescent reporter and eliding details of transcription/translation. All the net records is whether a gene is expressed or inhibited, and all transitions are purely qualitative.

1. Write out the pre-state (flow-in) and post-state (flow-out) matrices for the Petri net in Figure 1. Write out the corresponding vector form of its initial marking. Calculate the activity (reaction, stoichiometry, incidence) matrix for the net. [10 marks]

2. Find the minimal place and transitions invariants of this system. Describe what each one represents in terms of the original regulatory network. [10 marks]

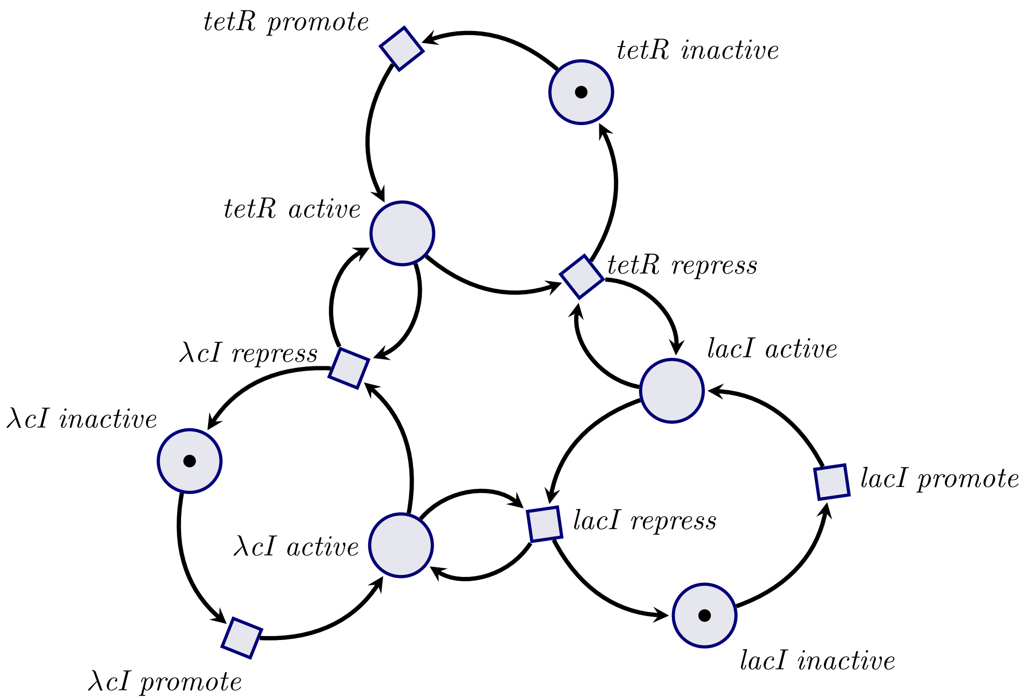


Figure 1: A simplified Petri net model of the repressilator.

3. Draw the graph of reachable states from this initial marking, labelling each node and each arc to show how they correspond to the net. [10 marks]

4. For each of the following LTL formulae: explain in words what property it expresses; state whether it holds for all runs, some runs, or no runs of the net given; and explain, with examples as appropriate, why this is so.

   (a) $\mathbf{G}(\lambda cI\text{-}active \vee tetR\text{-}active \vee lacI\text{-}active)$

   (b) $\mathbf{FG}(\lambda cI\text{-}active \vee tetR\text{-}active \vee lacI\text{-}active)$

   (c) $\mathbf{GF}(\lambda cI\text{-}active)$

   (d) $\mathbf{F}(lacI\text{-}active \wedge ((\neg \lambda cI\text{-}active)\ \mathbf{U}\ lacI\text{-}inactive))$

[10 marks]

5. For each of the following CTL formulae: describe in words what property it expresses of a state; construct a valuation table for the appropriate subformulae at all reachable states; and deduce for which states the formula holds.

   (a) $\mathbf{AG}(\lambda cI\text{-}active \vee tetR\text{-}active \vee lacI\text{-}active)$

   (b) $\mathbf{AF}(tetR\text{-}inactive)$

   (c) $\mathbf{AG}\,\mathbf{EF}(tetR\text{-}inactive)$

   (d) $lacI\text{-}active \wedge \mathbf{E}(tetR\text{-}inactive\ \mathbf{U}\ lacI\text{-}inactive)$

[10 marks]

6. For each of the following HML formulae: explain in words what property of a state it describes; construct a valuation table for all subformulae at all reachable states; and deduce for which states the formula holds.

   (a) $\mathit{tetR\text{-}active} \Rightarrow \langle \lambda \mathit{cI\text{-}repress} \rangle \mathit{True}$
   (b) $\langle \mathit{tetR\text{-}repress} \rangle \langle \mathit{tetR\text{-}promote} \rangle \mathit{tetR\text{-}active}$
   (c) $[\lambda \mathit{cI\text{-}promote}] \langle \mathit{lacI\text{-}repress} \rangle \mathit{True}$
   (d) $[\mathit{lacI\text{-}repress}][\mathit{tetR\text{-}repress}] \mathit{False}$

[10 marks]

The next questions involve an *E. coli* pathway, as represented in the EcoCyc database at `http://ecocyc.org`. Specifically the portion of the enterobacterial common antigen (ECA) biosynthesis pathway where UDP-N-acetyl-D-glucosamine is converted to undecaprenyl N-acetyl-glucosaminyl-N-acetyl-mannosaminurate pyrophosphate. (ECA is a component of the cell wall outer membrane of Gram-negative bacteria, and this portion of the path represents the synthesis one of its two repeating parts.) The EcoCyc database includes a diagrammatic representation of this pathway — you want the left-hand side in the EcoCyc diagram, not the whole thing — in various degrees of detail.

7. Construct a Petri net to represent this portion (down to undecaprenyl N-acetyl-glucosaminyl-N-acetyl-mannosaminurate pyrophosphate). You should include the undecaprenyl phosphate precursor and the uridine-5′-phosphate/-diphosphate products, but can ignore H_2O, H^+, NADH, NAD^+ and all of the enzymes. You don't need to include any arrival or departure transitions, nor the reverse reactions. [10 marks]

8. Give the matrix representation of your net, as pre-, post- and activity matrices. Identify its minimal place and transition invariants. For each of these draw the corresponding subnet, and describe in words what it represents biochemically. [10 marks]

The final questions are based on a reference pathway from the Kegg database at `http://www.genome.jp`. The *Entner-Doudoroff* pathway is included as module M00008 within the Kegg map of the pentose phosphate pathway. In the later portion of the Entner-Doudoroff pathway 6-phospho-D-gluconate is dehydrated and cleaved into pyruvate and glyceraldehyde-3-phosphate.

9. Draw a Petri net to represent this segment of the Entner-Doudoroff pathway. Include the enzymes listed on Kegg; the water molecule involved; and entry and exit transitions for all the non-enzyme constituents. You can leave out the reverse reactions. [10 marks]

10. Write CTL expressions for the following statements about the system.

    - Phosphogluconate dehydratase is essential for the eventual production of glyceraldehyde-3-phosphate.
    - Conversion of 6-phospho-D-gluconate into pyruvate and glyceraldehyde-3-phosphate always involves the production of a water molecule.

[10 marks]