Automatic design of digital synthetic gene circuits

Mario A. Marchisio and Joerg Stelling
Department of Biosystems Science and Engineering
ETH Zurich

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Anaheim, 15/06/10

SIB
Swiss Institute of Bioinformatics

Summary

- Automatic gene circuit design: the problem.
- The Karnaugh map method in biology.
- Comparison with a different design.
- Circuit complexity and performance.
- Conclusion and future work.

Automatic gene circuit design: previous approaches.

- Given the output, how to derive the corresponding circuit (structure and parameter values?)
- Brute force optimization via evolutionary algorithm (François and Hakim, *PNAS* **101**, 580, 2004)
- Similar implementations: OptCircuit (Dasika and Maranas, *BMC Syst. Bio.* **2**, 24, 2008); Genetdes (Rodrigo *et al*., *Bioinformatics* **23**, 1857, 2007).

Problems:

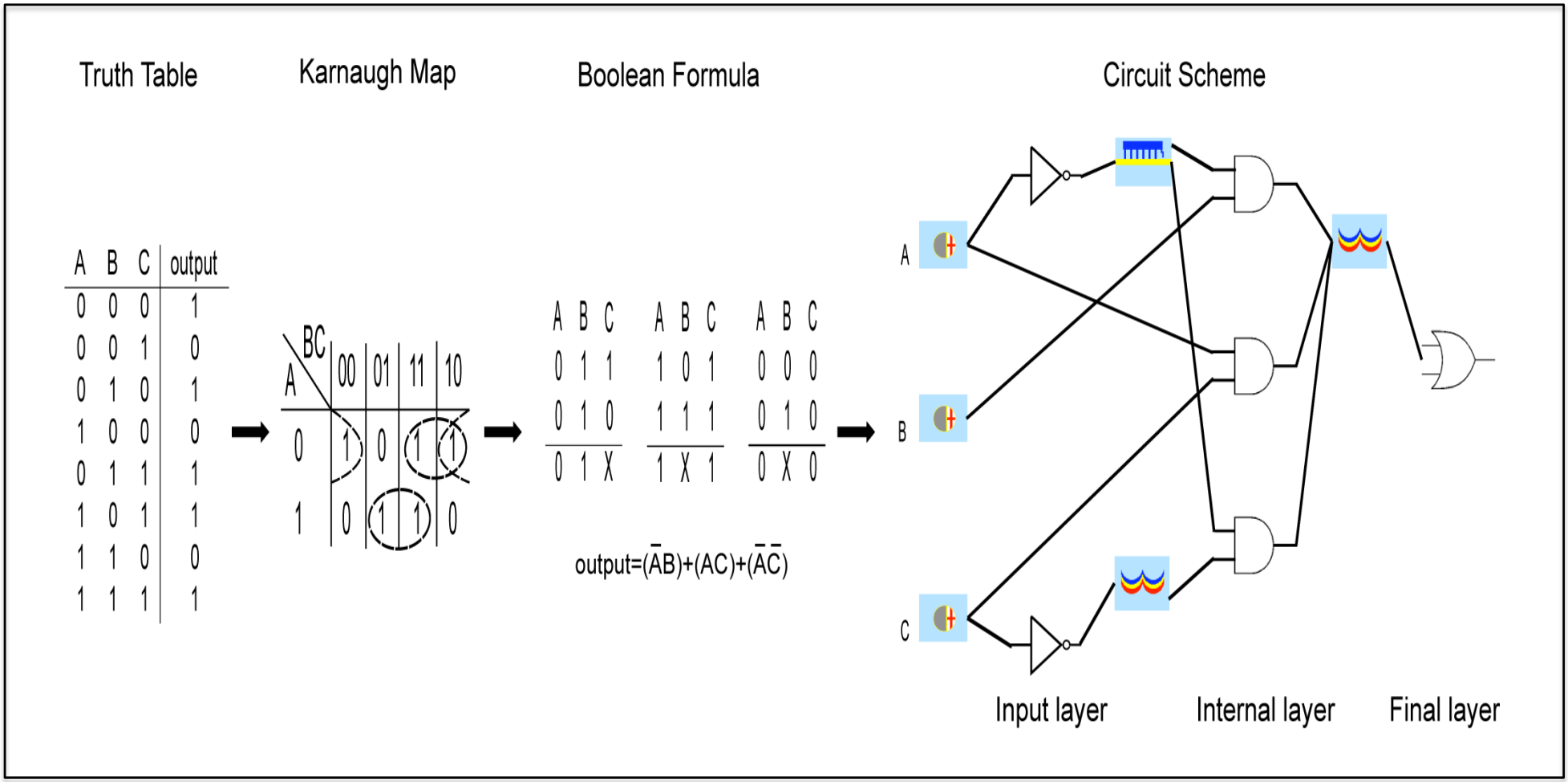
- Transcription units as bio-bricks (instead of parts).
- Limited model (translation as single-step event).
- Double optimization procedure: long computational time.

Looking for a different strategy.

Digital gene circuits

- Input/Output relation fully described by a truth table.
- The Karnaugh map method converts a truth table into a circuit scheme – no optimization required.
- Boolean gates due to promoter and RBS regulation mechanisms.
- Important application as biosensors.

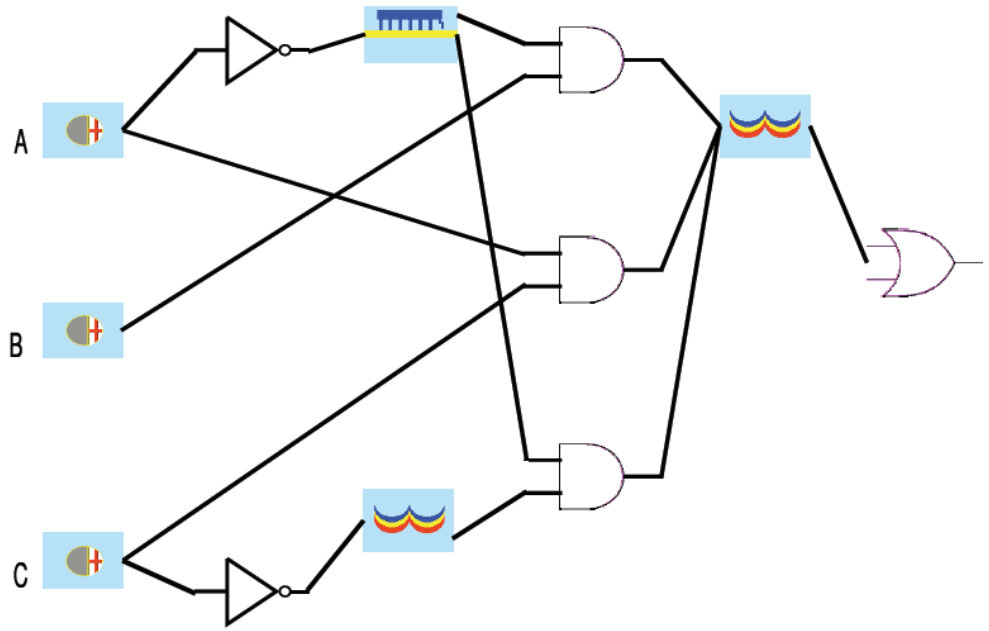
The Karnaugh map method



Circuit **structure** in three layers – **No optimization** required

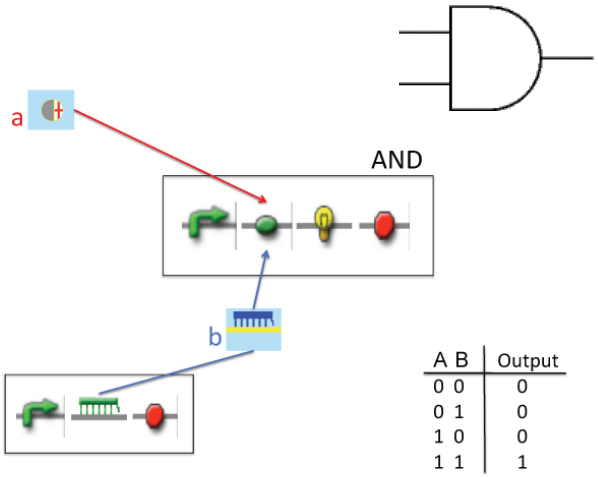
Circuit characteristics

- Activated/Repressed Promoters and RBSs (Bintu *et al.*, *Curr. Opin. Genet. Dev.* **15**, 125, 2005; Isaacs et al., *Nat. Biotech.* **22**, 841, 2004).
- **Pools** of transcription factors, sRNAs, and chemicals (M.A. Marchisio and J. Stelling, "Computational design of synthetic gene circuits with composable parts." *Bioinformatics*, **24**, 1903, 2008).
- A circuit takes up to four inputs (chemicals) and produces a single output (fluorescent protein).

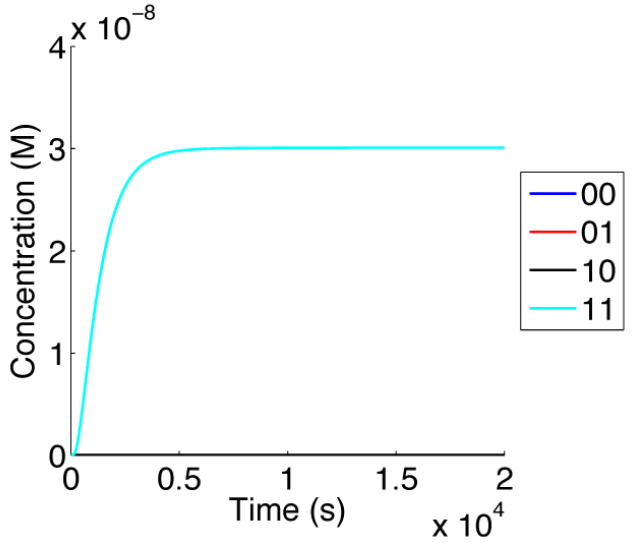


Gate structure and new designs.

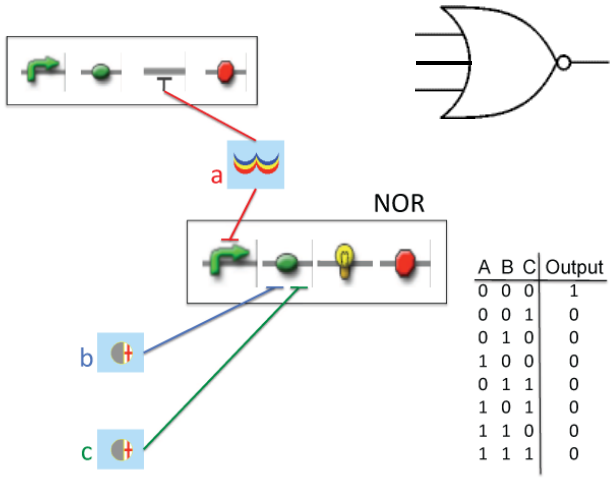
- Riboswitches + sRNA on the RBS.



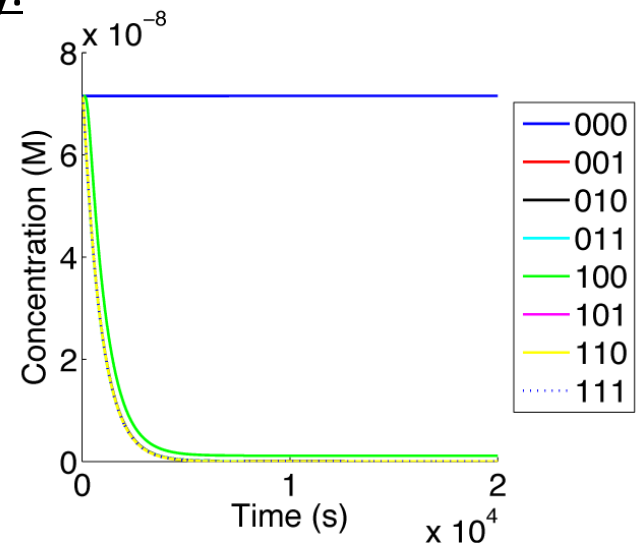
A	B	Output
0	0	0
0	1	0
1	0	0
1	1	1



- Promoters and RBS are controlled simultaneously.



A	B	C	Output
0	0	0	1
0	0	1	0
0	1	0	0
1	0	0	0
0	1	1	0
1	0	1	0
1	1	0	0
1	1	1	0



Comparison with electronics

- Every truth table corresponds to two Boolean formulas: CNF (POS) and DNF (SOP).
- In electronics the minimal circuit is given by the formula with the lowest number of clauses of NOT operations.
- In biology several circuit schemes arise from the same Boolean formula.
- How to define a **minimal** circuit in biology?

The complexity score

- **Regulatory factors** matter more than gene number.
- Only a handful of repressors and activators is currently used.
- Engineering new proteins is more difficult than synthesizing antisense small RNAs.
- Riboswitches simplify the structure of a gate.
- We define as minimal the circuit with the lowest complexity score defined as

$$S = 2^{R-1} + 2^{A-1} + n$$

where: R, repressor number (>= 1); A, activator number (>=1),n antisense sRNA number

- A circuit should avoid to re-use the same kind of transcription factors and prefer RBS controls to the promoter ones.
- Riboswitches do not increase the circuit complexity.

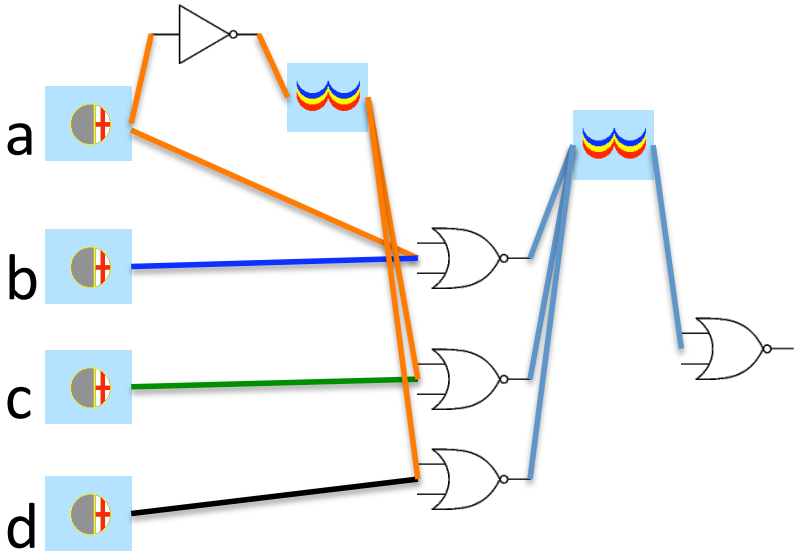
Our tool

- The truth table is the only input.
- All the schemes compatible with POS and SOP formulas are computed (less than 1s up to 8s).
- They are ranked according to their complexity score.
- The user can choose a solution: this is built by parts, pools, and device composition and encoded in MDL (Model Definition Language) to be visualized in ProMoT (http://www.mpimagdeburg.mpg.de/projects/promot/).

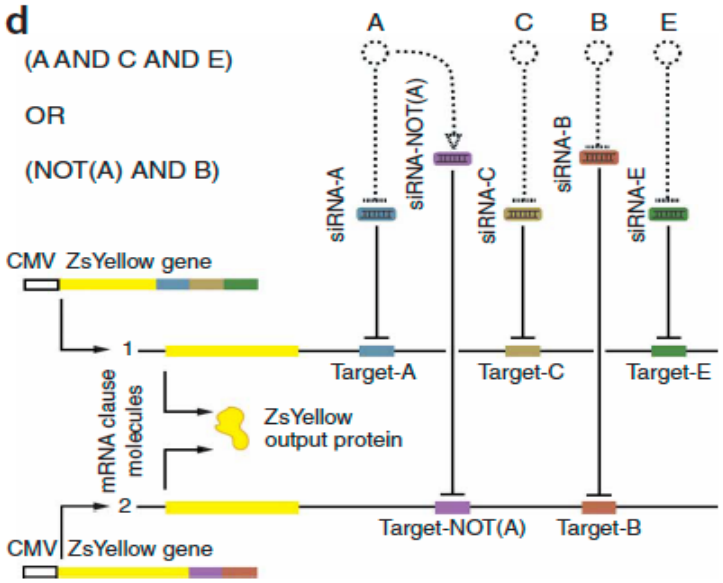
Comparison with RNAi-based design

(Rinaudo *et al.*, *Nat. Biotech.*, **25**, 795, 2007)

Our best solution
(a+b)(A+c)(A+d)



Rinaudo's solution
(acd)+(Ab)

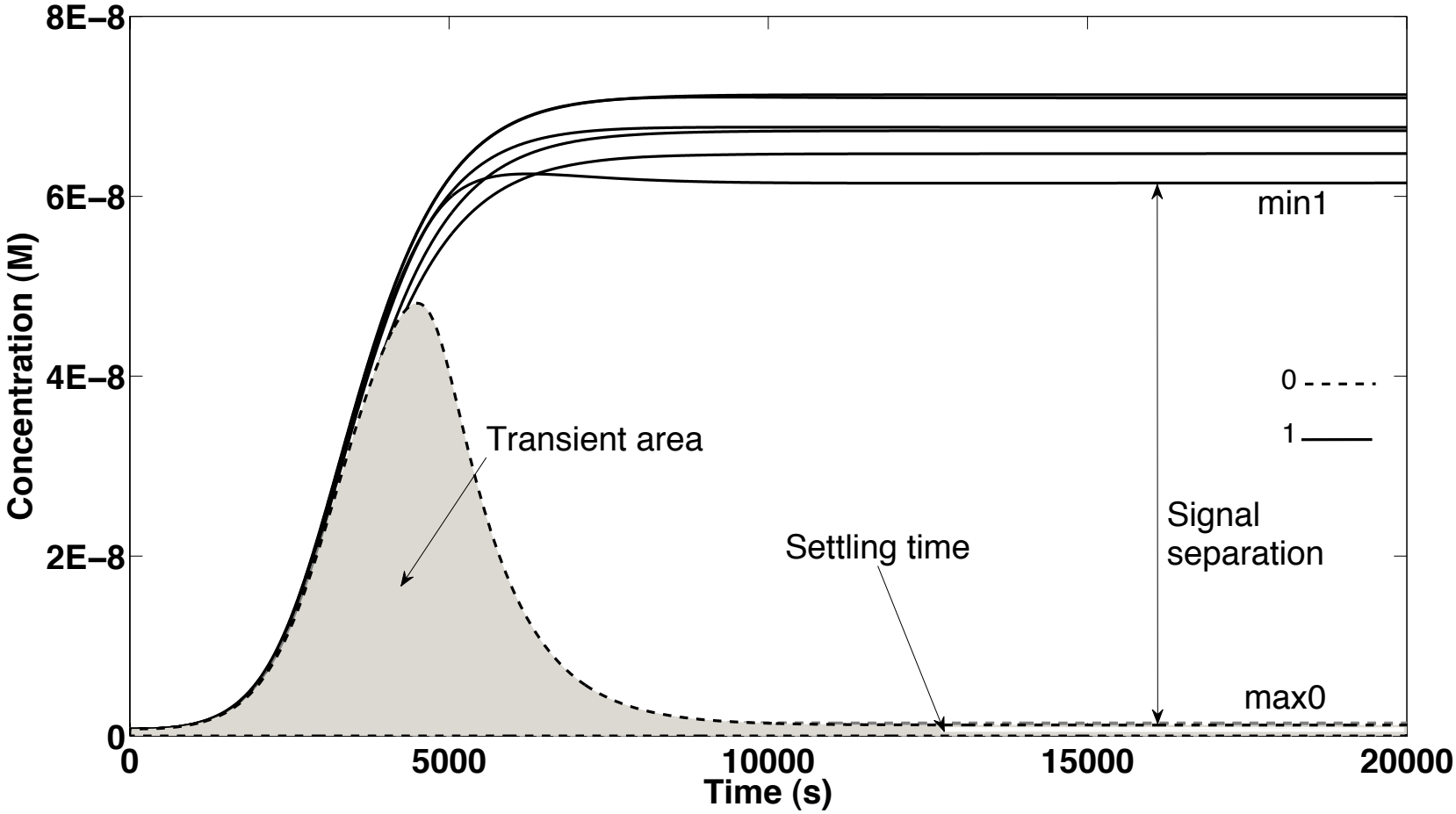


Circuit	Score	A	R	RNA
Our best POS	2	0	2	0
Our best SOP	4	2	1	1
Rinaudo	5	0	0	5

Our tool found 15 designs with complexity lower than 5.

How do these circuits work?

- Circuit **performance** is estimated through **signal separation** and **transient** calculation and depends both on structure and parameter values.

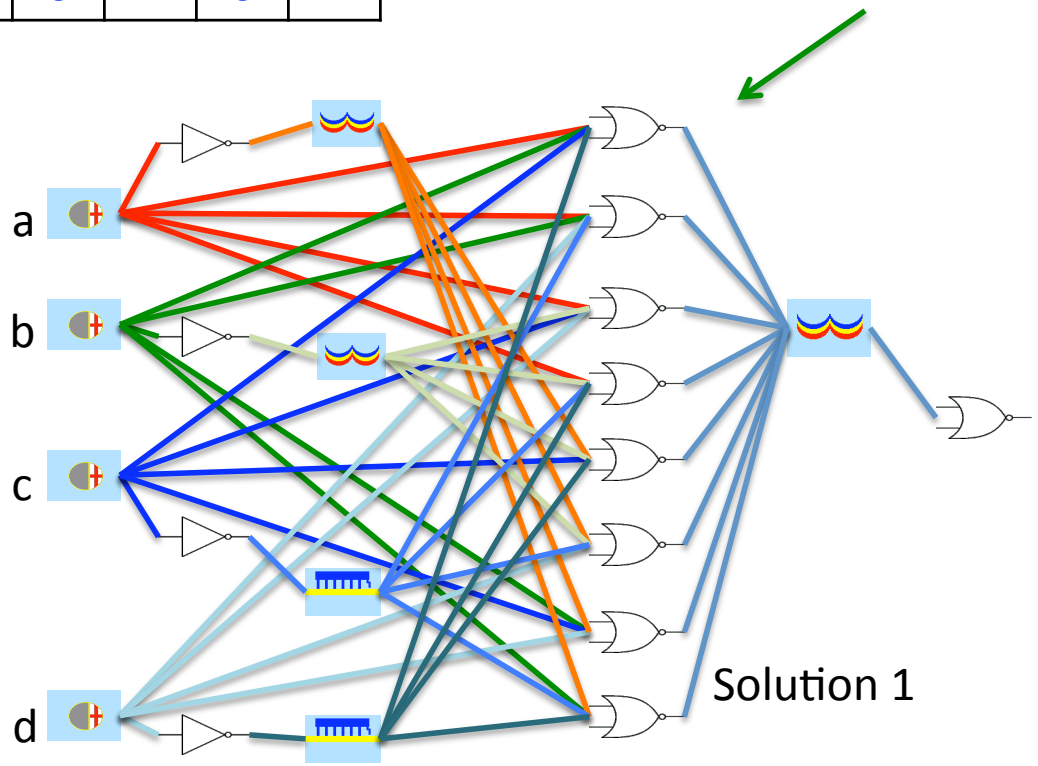


A benchmark

a	b	c	d	O
0	0	0	0	1
0	0	0	1	0
0	0	1	0	0
0	1	0	0	0
1	0	0	0	0
0	0	1	1	1
0	1	0	1	1
1	0	0	1	1
0	1	1	0	1
1	0	1	0	1
1	1	0	0	1
0	1	1	1	0
1	0	1	1	0
1	1	0	1	0
1	1	1	0	0
1	1	1	1	1

ab \ cd	00	01	11	10
00	1	0	1	0
01	0	1	0	1
11	1	0	1	0
10	0	1	0	1

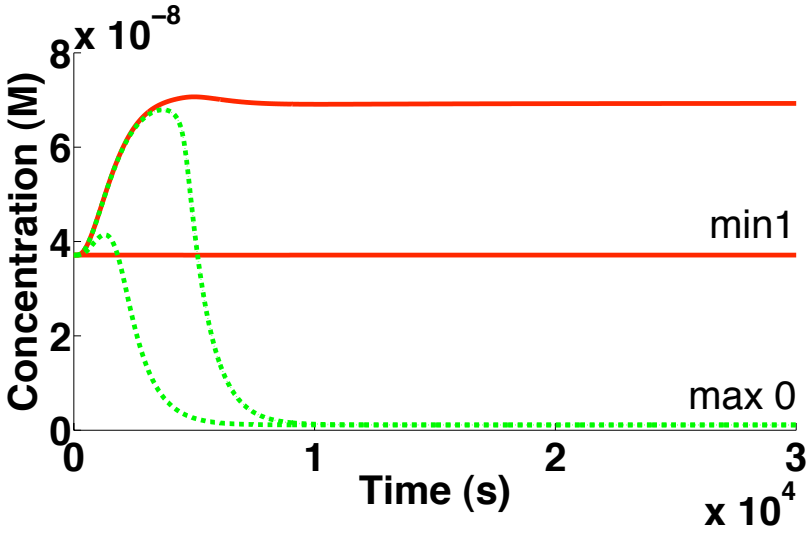
- POS and SOP formulas have 8 clauses each of which contains 4 inputs.
- 48 possible schemes with S varying from 20 to 2062.



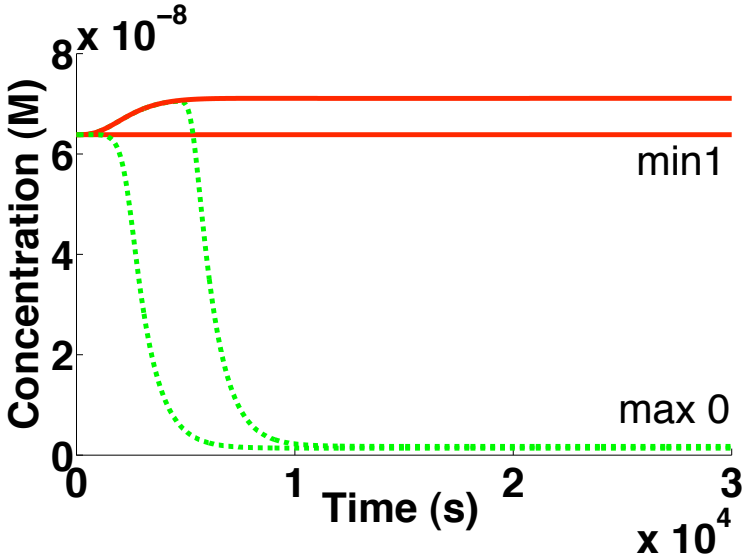
Solution 1

Comparison of two possible solutions

Solution 1



Solution 4

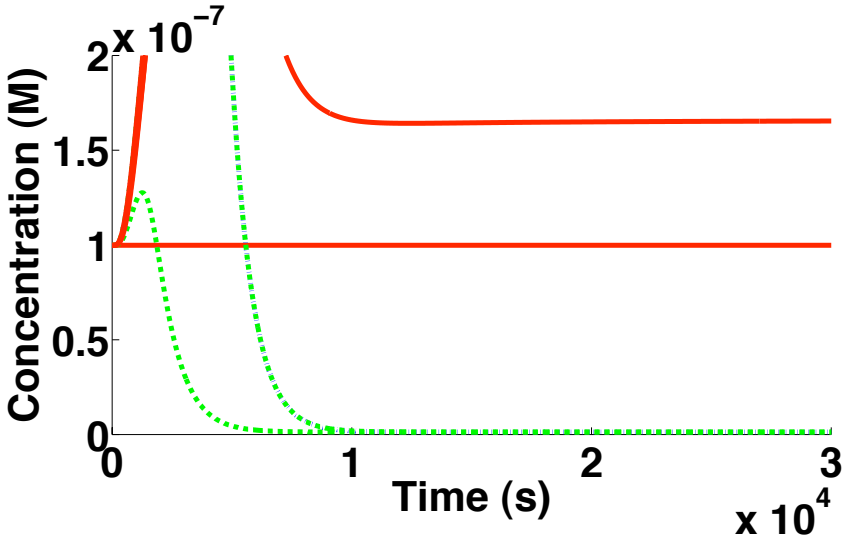


Solution	Rank	Score	A	R	RNA	Gene	Separation	Transient
1	1	20	2	5	2	17	36 nM	3709
4	25	548	10	6	4	21	62.1 nM	5112

Higher complexity seems to guarantee better performance.

Improving the performance

- The signal separation is mostly influenced by parameters belonging to the **final gate**.
- Tuning only **one** parameter (the strength of the promoter in the final gate) the signal separation can be drastically amplified.



- Stochastic algorithms can be avoided but a good set of default parameter values is required.

Conclusion and future work

- Automatic design of digital synthetic gene circuits via the Karnaugh map method.
- Circuit structure calculation does not require any optimization procedure.
- Theoretical new design of Boolean gates where promoter and RBS are simultaneously.
- Computer simulations show an unequivocal signal separation between 0/1 outputs with our choice of default parameter values.

- Insertion of other translation regulation mechanisms.
- Extension to eukaryotic cells.
- **MAIN GOAL**: Wet-lab implementation of single gates and more complex circuits.