Modern Engineering as Translational Science

Edward R. Dougherty

Department of Electrical and Computer Engineering
Center for Bioinformatics and Genomic Systems Engineering
Texas A&M University

Reading

- Dougherty, E. R., "Translational Science: Epistemology and the Investigative Process," *Current Genomics*, Vol. 10, No. 2, 102-109, April, 2009.
- Dougherty, E. R., Pal, R., Qian, X., Bittner, M. L., and A. Datta, "Stationary and Structural Control in Gene Regulatory Networks: Basic Concepts," *International Journal of Systems Science*, Vol. 41, No. 1, 5-16, January, 2010.

Classical Engineering

- Trial and error based on hunches (intuition).
  - The end product may or may not be scientifically understood.
- Can work well when problems are simple.
  - Rarely works optimally for complex systems, for instance, drug development with complex disease.

Texas A&M University Genomic Signal Processing Laboratory

Scientific Theory

- *Mathematical model* consisting of variables and relations between the variables.
- *Operational definitions* relating the variables to observable (and measurable) phenomena.
- *Experimental design* to test predictions made by the model.
- A scientific theory is validated to the extent that predictions derived from it agree with experimental observations.

Why Scientific Knowledge Is Mathematical

- Four reasons:
  - Scientific knowledge is based on quantitative measurements.
  - Scientific knowledge concerns relations and mathematics provides the formal structure for relations.
  - Validity depends on predictions. This requires a quantitative structure from which to generate predictions and a theory of probability in which to quantify the goodness of predictions.
  - Mathematics provides a formal language sufficiently simple so that both the constituting theory and the experimental protocols for prediction are inter-subjective.

1/23/2017

Modern Engineering

- Modern engineering is translational science: it transforms a mathematical model, whose purpose is to provide a predictive conceptualization of some portion of the physical world, into a model characterizing human intervention (action) in the physical world.
- Scientific knowledge is translated into practical knowledge by expanding a scientific system to include inputs that can be adjusted to affect the behavior of the system and outputs that can be used to monitor the effect of the external inputs and feed back information on how to adjust the inputs.

Science and Action

- **Arturo Rosenblueth and Norbert Wiener**: “The intention and the result of a scientific inquiry is to obtain an understanding and a control of some part of the universe.”
  - For them, science and translational science are inextricably linked, the ultimate purpose of acquiring scientific knowledge being to translate that knowledge into action.

1/23/2017

Analysis

- Given a system and an operator, what can be said about the properties of the output system in terms of the properties of the input system?
- It might be mathematically difficult to characterize completely the output system given the complete input system or we may only know certain properties of the input system, so that the best we can hope for is to characterize related properties of the output system.

Synthesis

- Given a system, we would like to design an operator to transform the system in some desirable manner.
  - Synthesis forms the existential basis of engineering.
  - Trial and error is groping in the dark, not translational science
- Synthesis starts with a mathematical theory constituting the relevant scientific knowledge and the theory be utilized to arrive at an optimal (or close to optimal) operator for accomplishing the desired transformation under the constraints imposed by the circumstances.
- Synthesis does not guarantee a physical transformation.

Synthesis Protocol

- Four steps.
  - Construct the mathematical model for the system.
  - Identify the class of operations from which to choose.
  - Define the optimization problem.
  - Solve the optimization problem.
- Level of abstraction.
  - Sufficiently complex to formulate problem sufficiently.
  - Sufficiently simple that the translational problem is not obscured by too much structure, the necessary parameters can be estimated, and the optimization is tractable.

Historical Turning Point

- Translational scientific synthesis, which is synonymous with modern engineering, begins with optimal time series filtering in the classic work of Andrei Kolmogorov [1941] and Norbert Wiener [1942].
  - The scientific model is a random signal and the translational problem is to linearly operate on the signal so as to transform it to be more like some ideal (desired) signal.
  - The synthesis problem is to find an optimal weighting function and the goodness criterion is the mean-square difference between the ideal and filtered signals.

Optimal Linear Filtering

- Filter model
  - Original signal: s
  - Blurred signal: $B(s)$
  - Noisy signal, blur plus point noise: $B(s) + n$
  - Filtered noisy signal: $\Psi[B(s) + n]$
- Use a linear integral operator.
- Optimization problem:
  - Find Ψ to minimize distance between $\Psi[B(s) + n]$ and s.
- Solve optimization: Involves stochastic processes.

Linear Image Filtering

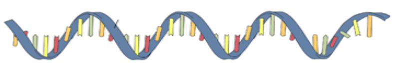
- Image filtering:
  - Original, blur + noise, filtered image



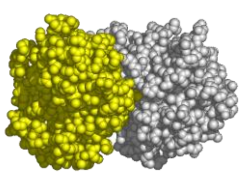
Gene Regulation

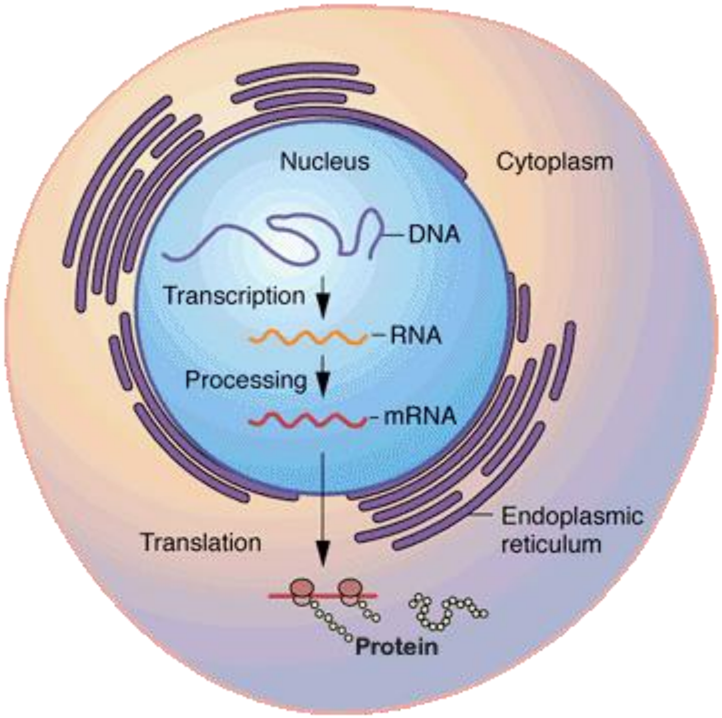
DNA 

Transcription

RNA 

Translation

Protein 



Multivariate Regulation

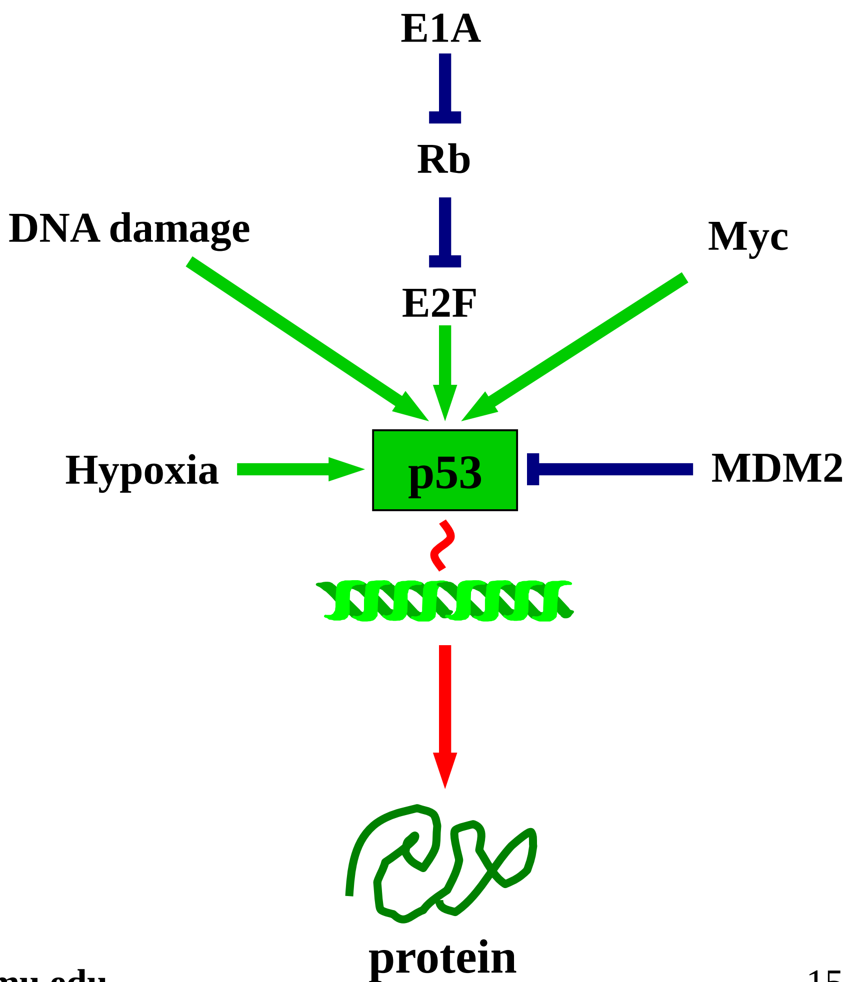
Gene regulatory controls

transcription

translation

Gene expression

the process by which gene products (proteins) are made

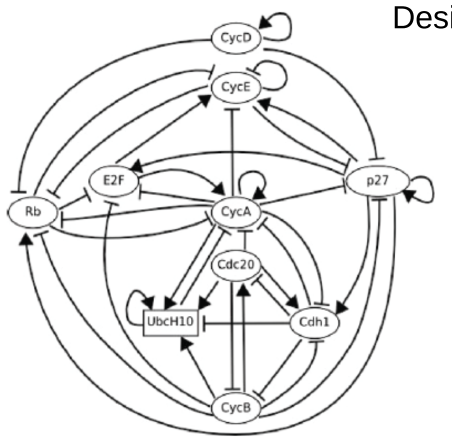


Texas A&M University Genomic Signal Processing Laboratory

Intervention

- A key goal of network modeling is to determine intervention targets (genes) such that the network can be “persuaded” to transition into desired states.
- We desire genes that are the best potential “lever points” in the sense of having the greatest possible impact on desired network behavior.

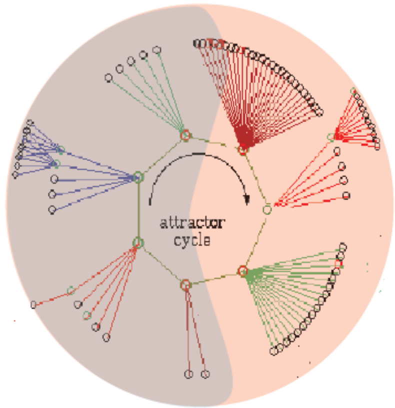
Intervention Paradigm



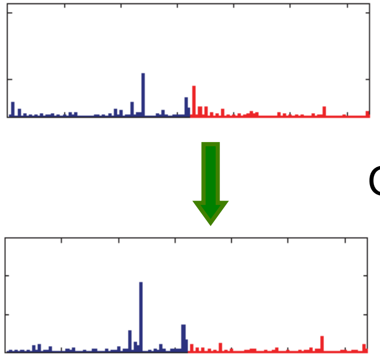
Regulatory networks

Desirable states

Undesirable states



State space



Control

Steady-state distribution

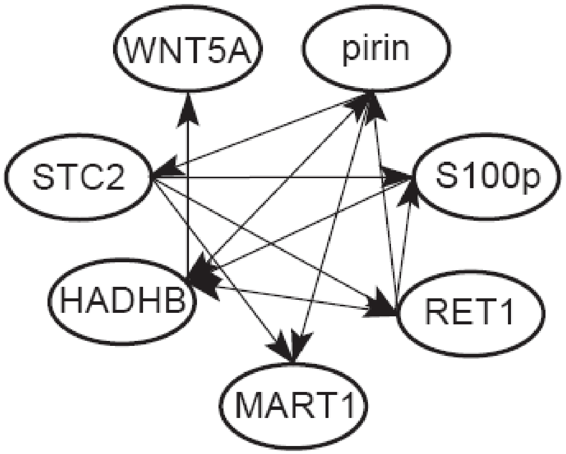
Optimal Control

- Key Objective : Optimally manipulate external controls to move the gene activity profile (GAP) from an undesirable pattern to a desirable pattern.
- Use available information, e.g., phenotypic responses, tumor size, etc.
- Requires a paradigm for modeling the evolution of the GAP under different controls.
- Often a Markov chain.

External Control

- Consider an external control variable and a cost function depending on state desirability and cost of action.
- Minimize the cost function by a sequence of control actions over time – control policy.
- Application: Design optimal treatment regime to drive the system away from undesirable states.

WNT5A Boolean Network

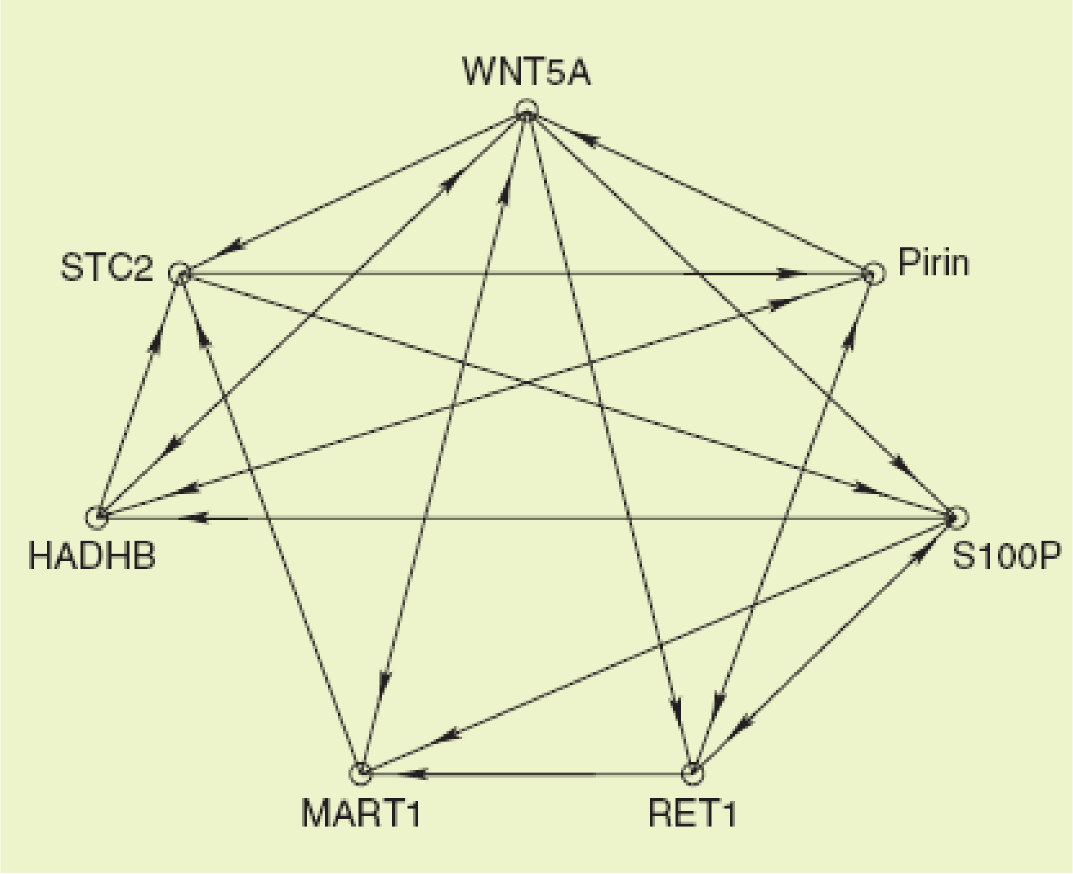


1. WNT5A
2. Pirin
3. S100P
4. RET1
5. MART1
6. HADHB
7. STC3

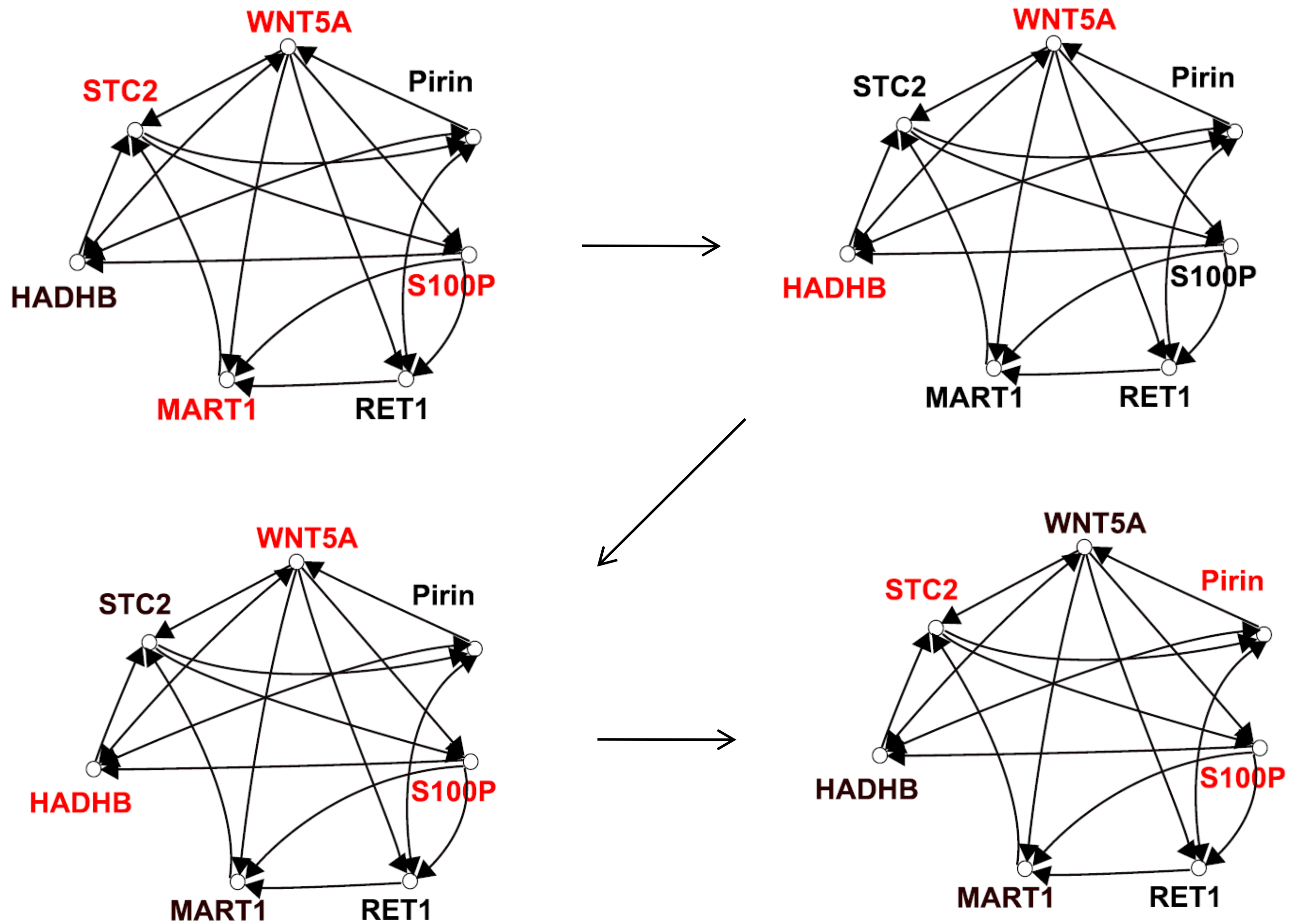
Function	Input variables	Output
f_1	x_6	10
f_2	x_2, x_4, x_6	00010111
f_3	x_3, x_4, x_7	10101010
f_4	x_4, x_6, x_7	00001111
f_5	x_2, x_5, x_7	10101111
f_6	x_2, x_3, x_4	01110111
f_7	x_2, x_7	1101

WNT5A Control

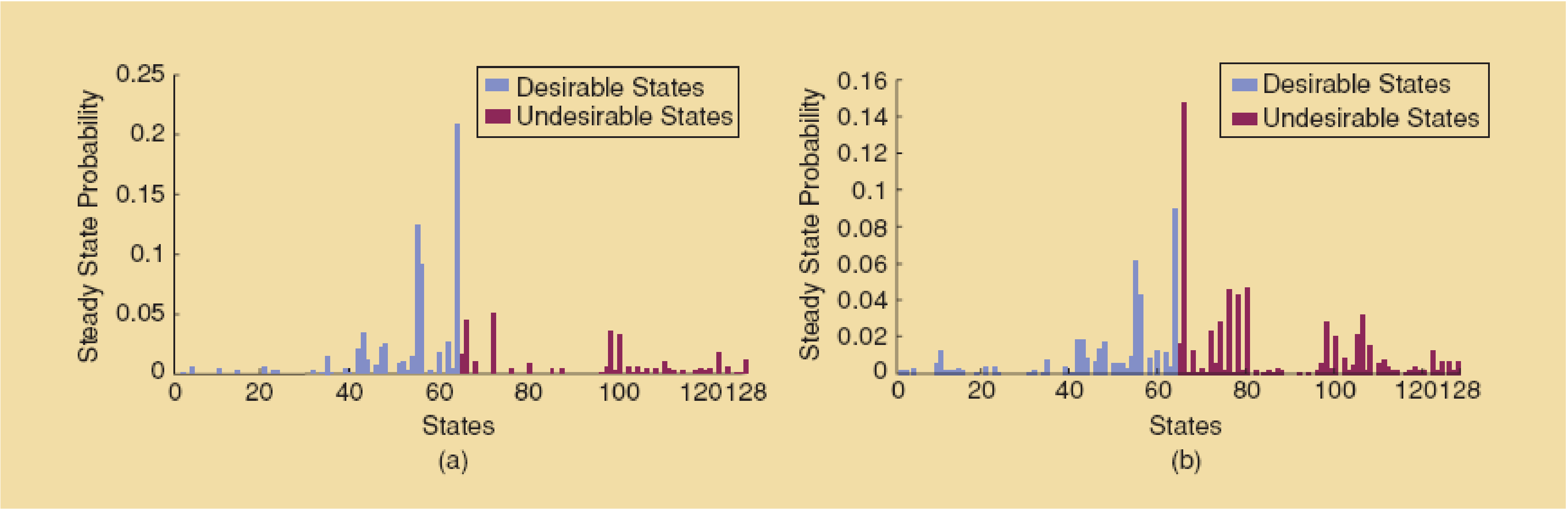
- Up-regulated WNT5A associated with increased metastasis.
- Cost function penalizes WNT5A being up-regulated.
- Optimal control policy with Pirin as control gene.



Sample Trajectory



Shift of Steady-State Distribution



- Optimal (infinite horizon) control with pirin has shifted the steady-state distribution to states with WNT5A down-regulated: (a) with control; (b) without control.

Optimal Structural Intervention

- Find the one-bit change in the rule structure that optimally reduces the undesirable steady-state mass.
  - Based on analytic formulation of change in the steady state.
  - Iterative procedure can be used for multiple-bit changes.
- Analytic procedure allows optimality to be constrained.
  - Allow only biologically implementable changes.
  - Impose constraints on new steady state – no new attractors.
- Based on the fundamental matrix of a Markov chain.

– Qian, X., and E. R. Dougherty, "**Effect of Function Perturbation on the Steady-State Distribution of Genetic Regulatory Networks: Optimal Structural Intervention**," *IEEE Trans. Signal Processing*, **56**(10), Part 1, 4966-4975, 2008.

Perturb Gene Logic

- Input network has probability transition matrix *P* and there is a single flip in the truth table defining predictor functions. Find SSD for output network.
- Theorem: If $P^{\sim} = P + E$, where $E = ab^T$ is a rank-one perturbation matrix, then the new SSD $\pi^{\sim}$ is given by

$$\tilde{\pi}^T = \pi^T + \frac{\pi^T a}{1 - b^T Z a} b^T Z$$

where *Z* is the fundamental matrix of original network.

- More complicated perturbations are handled by proceeding iteratively.

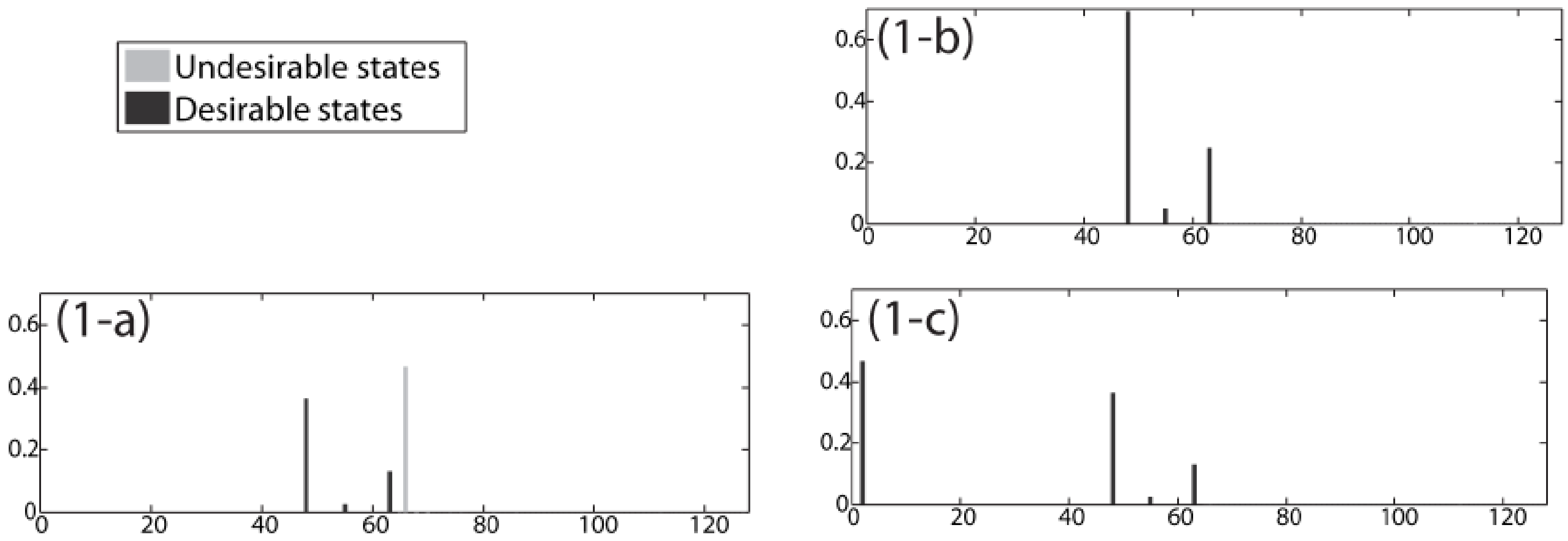
Optimal 1-bit Perturbations

- Table gives the optimal one bit perturbation for each predictor function and the resulting SSD shift of mass for *p* = 0.001.

Function	Input variables	Input values for optimal perturbation	Optimal perturbation	$\sum_{x_1=1} \tilde{\pi}(\mathbf{x})$
f_1	x_6	0	$1 \rightarrow 0$	0.0010
f_2	x_2, x_4, x_6	100	$0 \rightarrow 1$	0.3777
f_3	x_3, x_4, x_7	101	$0 \rightarrow 1$	0.2084
f_4	x_4, x_6, x_7	011	$0 \rightarrow 1$	0.0035
f_5	x_2, x_5, x_7	100	$1 \rightarrow 0$	0.4728
f_6	x_2, x_3, x_4	000	$0 \rightarrow 1$	0.0038
f_7	x_2, x_7	01	$1 \rightarrow 0$	0.2305

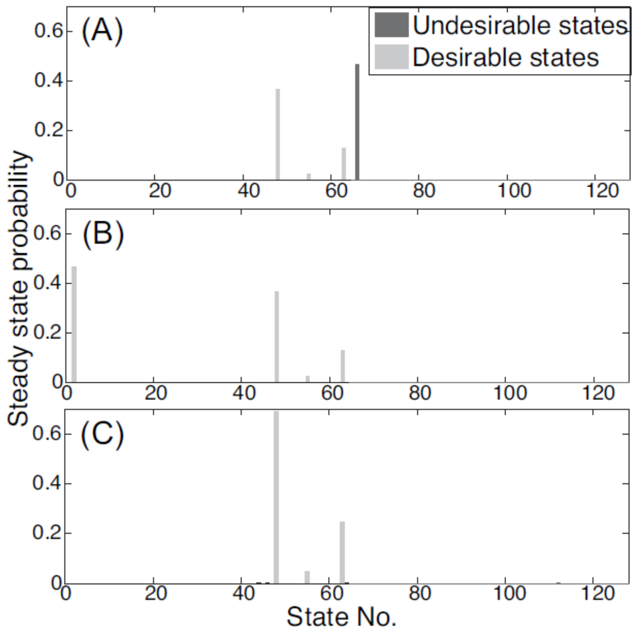
What is the Best Perturbation?

- Optimal 1-bit perturbations for $p = 0.001$ in the table.
  - f_1 is optimal but introduces new attracter.
- SSD: (a) original, (b) perturb f_4, (c) perturb f_1.



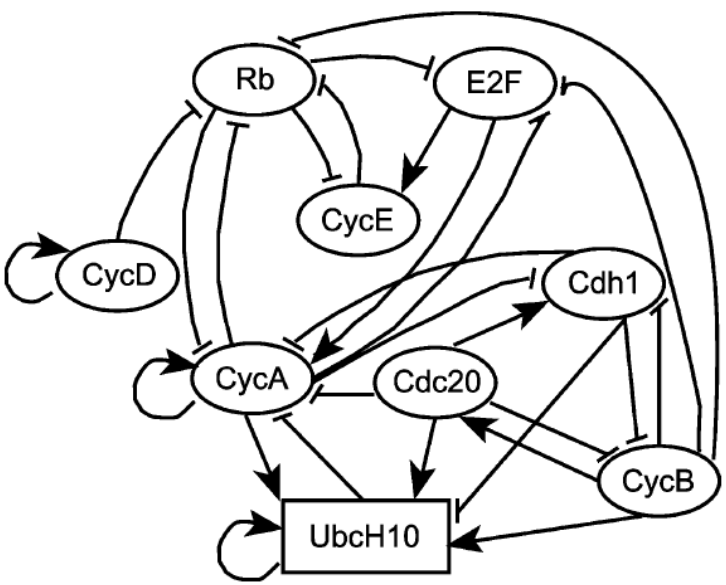
Phenotypically Constrained Intervention

- For the previously considered melanoma network, we found (b) to provide the optimal structural intervention.
- But although WNT5A = 0 for STC2 (leftmost state), STC2 is known to be associated with carcinogenesis.
- Thus, we constrain interventions to those only allowing tiny new mass with original tiny mass (c).



Mutated Mammalian Cell Cycle PBN

- CycD is independent of the cell's content and reflects the state of the growth factor, which is not part of the cell and is determined by surrounding cells.
- CycD = 0 and CycD = 1 determines two BNs for PBN.



Product	Predictors
CycD	Input
Rb	$(\overline{CycD} \wedge \overline{CycE} \wedge \overline{CycA} \wedge \overline{CycB})$
E2F	$(\overline{Rb} \wedge \overline{CycA} \wedge \overline{CycB})$
CycE	$(E2F \wedge \overline{Rb})$
CycA	$(E2F \wedge \overline{Rb} \wedge \overline{Cdc20} \wedge \overline{(Cdh1 \wedge Ubc)}) \vee (CycA \wedge \overline{Rb} \wedge \overline{Cdc20} \wedge \overline{(Cdh1 \wedge Ubc)})$
Cdc20	CycB
Cdh1	$(\overline{CycA} \wedge \overline{CycB}) \vee (Cdc20)$
Ubc	$(\overline{Cdh1}) \vee (Cdh1 \wedge Ubc \wedge (Cdc20 \vee CycA \vee CycB))$
CycB	$(\overline{Cdc20} \wedge \overline{Cdh1})$

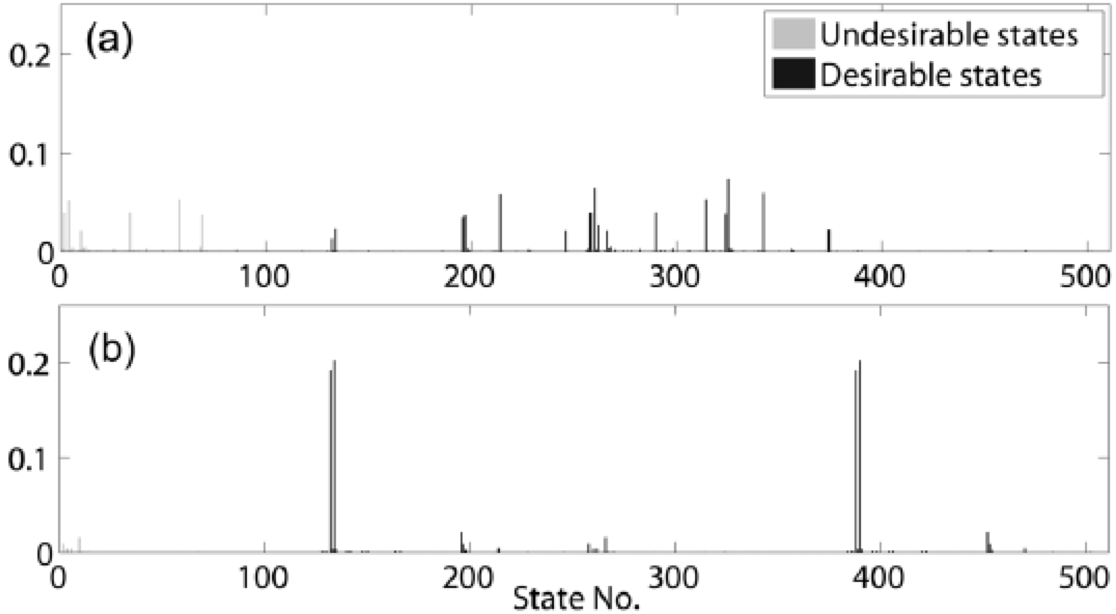
Optimal Perturbations

- Undesirable SSD mass: Rb = CycD = 0.
  - Cell cycles in the absence of any growth factor.
- Optimal 1-bit perturbations for $p = 0.001$ in the table.
  - Maximally reduce undesirable steady-state mass
  - Perturbing predictor function *Rb* in BN_2 provides maximal reduction.

Gene	Rb	E2F	CycE	CycA	Cdc20	Cdh1	UbcH10	CycB
BN_1	0.1901	0.2903	0.2534	0.2484	0.2071	0.2576	0.2587	0.2532
BN_2	0.0413	0.2198	0.2529	0.2543	0.2568	0.2576	0.2587	0.2550

Shift in Steady-State Distribution

- Steady-state distributions: (a) before, (b) after.
  - Maximally reduce undesirable steady-state mass
  - Optimal intervention with *Rb* in *BN*2.
  - Note introduction of new attractors.



Translational Protocol in GRN Intervention

- The four translational steps in structural intervention for the mammalian cell cycle network.
  - (1) The interaction among key genes in the cell cycle is modeled – via a probabilistic Boolean network.
  - (2) A class of actions is specified – one-bit perturbations in the regulatory rules.
  - (3) An optimization criterion is posited – minimize the undesirable probability mass in the steady state.
  - (4) mathematical methods are used to find an optimal action – via Markov chain perturbation theory

Benefits of a Translational System

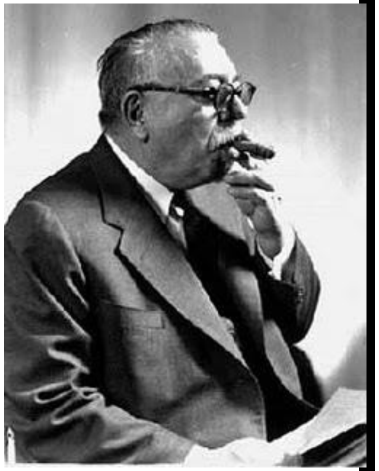
- A translational mathematical system provides guides.
  - Guide the scientist in building a fruitfully applicable model
  - Guide the engineer in studying costs and benefits of action
  - Guide the technologist in devising devices or treatments.
- In a properly functioning relationship, the scientist does not hand the engineer a set of data and ask the engineer to find something in it; instead, assuming a translational goal, the enterprise should be guided by the goal and this goal should already have led a carefully designed experiment.

Implications for Research

- **Norbert Wiener**: “If the difficulty of a physiological problem is mathematical in essence, ten physiologists ignorant of mathematics will get precisely as far as one physiologist ignorant of mathematics. If a physiologist who knows no mathematics works together with a mathematician who knows no physiology, the one will be unable to state his problem in terms that the other can manipulate, and the second will be unable to put the answers in any form that the first can understand.”

Implications for Research

- **Norbert Wiener**: "The mathematician need not have the skill to conduct a physiological experiment, but he must have the skill to understand one, to criticize one, and to suggest one. The physiologist need not be able to prove a certain mathematical theorem, but he must be able to grasp its physiological significance and tell the mathematician for what he should look for."



1/23/2017

Implications for Education

- The mathematician cannot be a scientific dilettante carrying his toolbox around searching for some problem that directly fits some existing theory.
- The biologist need not know the formal mathematical structure or be able to prove theorems but should see how the theory manifests itself in the physical world.
- In essence, Wiener posits two requirements.
  - Mathematician and experimentalist are experts in their fields
  - Their domain of intersection is large enough for interaction
- Education should aim to achieve these goals.

Final Comment

- **Norbert Wiener** (*Cybernetics*, 1948): As far back as four years ago, the group of scientists about Dr. Rosenblueth and myself had already become aware of the essential unity of the set of problems centering about communication, control, and statistical mechanics, whether in the machine or in living tissue.