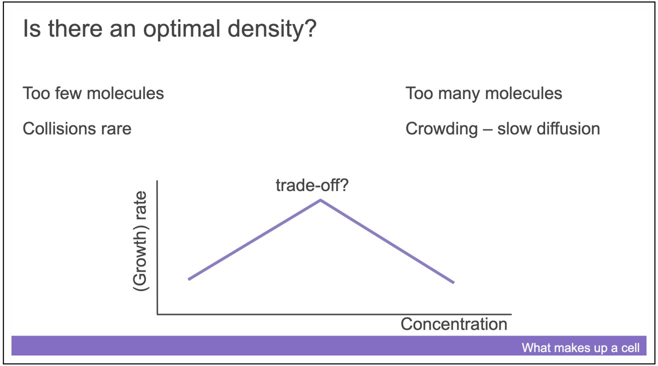
Optimality in Biology

Wolfram Liebermeister & **Markus Köbis**

… what is this talk about?

from the previous talk(?)



Proclaimer

Economic Principles in **Cell Physiology**

Goals of this talk

- Motivate (a bit more) what is about to follow, i.e. the economy-of-the-cell analogy
- Establish ‘our’ view on ‘economic principles’, and ‘our’ subset of ‘cell physiology’
- Set the “mathematical background” (mostly: optimization) and cell modeling via the constraint-based framework, some intro to algorithms

Outline

Book chapter "OPT"

1. Optimality principles in biology
2. History of mathematical optimality problems and their applications
3. Mathematical optimality problems
4. Multi-objective problems
5. Examples of optimality problems in cells
6. Constraints and trade-offs in models: relation to empirical knowledge, mechanisms, and optimality
7. Discussion: beyond optimality thinking

Outline of this lecture

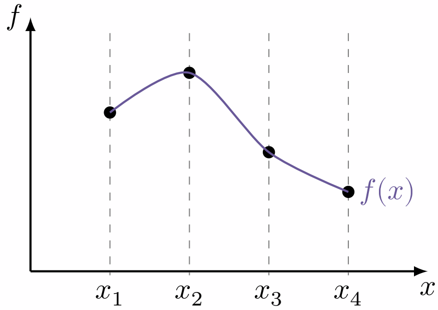
1. Primer on Optimization
2. Connection to 'Economy' and 'Evolution'
3. 'Chapter 1': Primer on metabolic networks
4. Examples
5. Some historical notes (if time permits)
6. Discussion

A Primer on Optimization

Optimization Problems (I)

Optimization problems (f. ex. 1-D)

Choose x such that some value $f(x)$ becomes maximal/minimal.



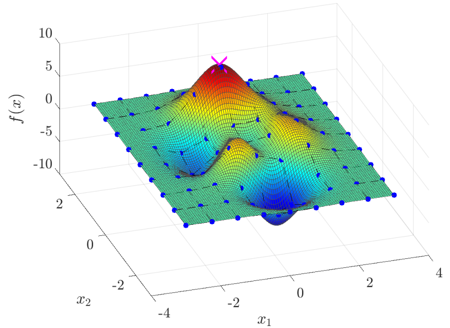
- 'Choose 'the best' out of possible decisions.'
- 'Find 'the best' possible configuration.'
- 'Pick 'the best', according to your preferences and/or quantifiable criteria.'

$$\min_{x \in S} f(x) \text{ or } \max_{x \in S} f(x)$$

Optimization Problems (II)

Optimization problems (f. ex. 2-D)

Choose x such that some value $f(x)$ becomes maximal/minimal.



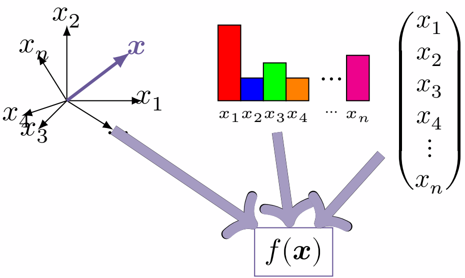
- 'Choose 'the best' out of possible decisions.'
- 'Find 'the best' possible configuration.'
- 'Pick 'the best', according to your preferences and/or quantifiable criteria.'

$$\min_{x \in S} f(x) \text{ or } \max_{x \in S} f(x)$$

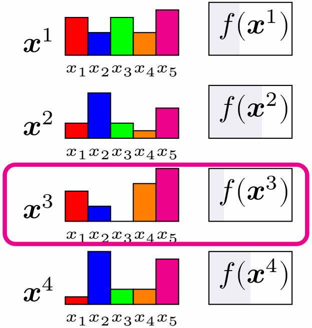
Optimization Problems (III)

$$\min_{\boldsymbol{x} \in S} f(\boldsymbol{x}) \text{ or } \max_{\boldsymbol{x} \in S} f(\boldsymbol{x})$$

Optimization Problems (f. ex. n-D)



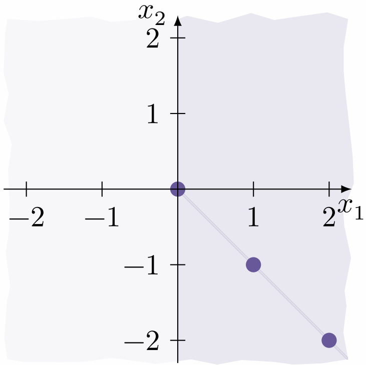
'Picking' the optimum



Optimization Problems (IV)

$$\min_{\boldsymbol{x} \in S} f(\boldsymbol{x}) \text{ or } \max_{\boldsymbol{x} \in S} f(\boldsymbol{x})$$

Constraining the Feasible Points



Constraint by means of

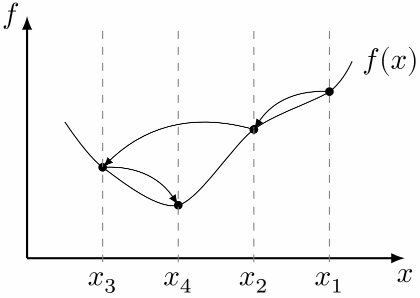
- (non-) linear inequalities
  $x_1 \geq 0$
- (non-) linear equalities
  $x_1 + x_2 = 0$
- set inclusions
  $x_2 \in \mathbb{Z}$

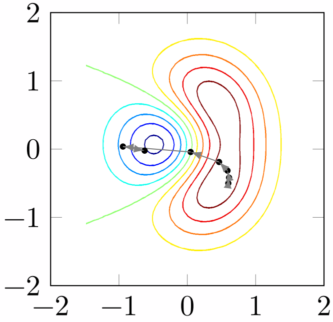
Optimization Algorithms: Some theory

Simplest Case: Gradient Descent for $\min_{\boldsymbol{x}} f(\boldsymbol{x})$

Iteration:

$$\boldsymbol{x}^{n+1} := \boldsymbol{x}^n - \nabla_{\boldsymbol{x}} f(\boldsymbol{x}^n)$$





If necessary: Gradient approximation:

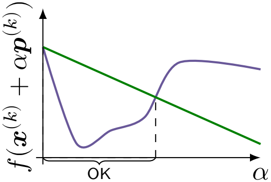
$$\frac{\partial f}{\partial x_i} \approx \frac{f(\boldsymbol{x} + h\boldsymbol{e}_i) - f(\boldsymbol{x})}{h}$$

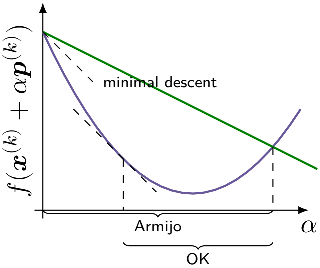
Optimization Algorithms: Some theory (cont.)

Armijo condition

Sufficient descent (with $c_1 \in (0,1)$)

$$\underbrace{f(\boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{p}^{(k)})}_{\Phi_k(\alpha_k)} \leq \underbrace{f(\boldsymbol{x}^{(k)}) + c_1 \cdot \alpha_k \cdot \nabla f(\boldsymbol{x}^{(k)})^\top \cdot \boldsymbol{p}^{(k)}}_{\Phi_k(0)+\alpha_k \cdot c_1 \cdot \Phi_k'(0) =: l(\alpha_k)}$$





Wolfe condition(s)

Avoid too small steps (with $c_2 \in (c_1, 1)$)

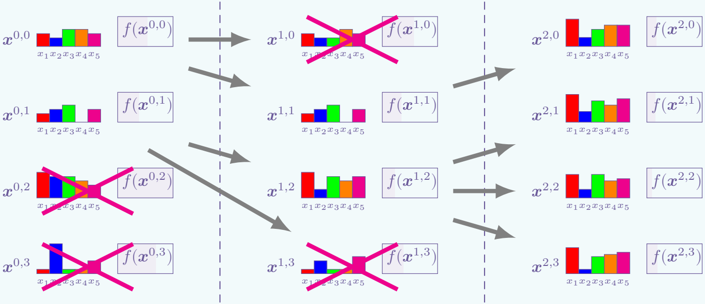
$$\underbrace{\nabla f(\boldsymbol{x}^{(k)} + \alpha_k \cdot \boldsymbol{p}^{(k)})^\top \cdot \boldsymbol{p}^{(k)}}_{\Phi_k'(\alpha_k)} \geq \underbrace{c_2 \cdot \nabla f(\boldsymbol{x}^{(k)})^\top \cdot \boldsymbol{p}^{(k)}}_{c_2 \cdot \Phi_k'(0)}$$

Optimization Algorithms: Some theory (cont.)

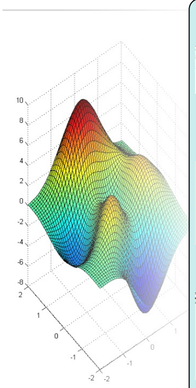
$$\min_{\boldsymbol{x} \in S} f(\boldsymbol{x}) \text{ or } \max_{\boldsymbol{x} \in S} f(\boldsymbol{x})$$

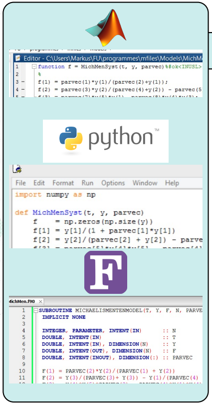
So-called 'genetic' (global optimization) algorithms

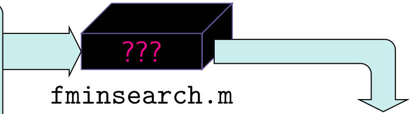
0. Pick a sample of initial guesses $\boldsymbol{x}^{0,0}$, $\boldsymbol{x}^{0,1}$, ...
1. Calculate function values
2. Sort out the worst cases, adapt/mix the best performers, goto 1.



Optimization Algorithms: ‘Practice’



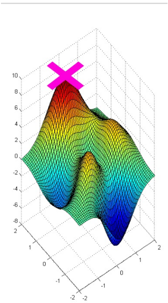




`fminsearch.m`
`linprog.m`
`fmincon.m`

`scipy.optimize`
`cvxopt.py`
`APMonitor`

`CPLEX`
`GUROBI`
`IPOPT`
`NLopt`
⋮



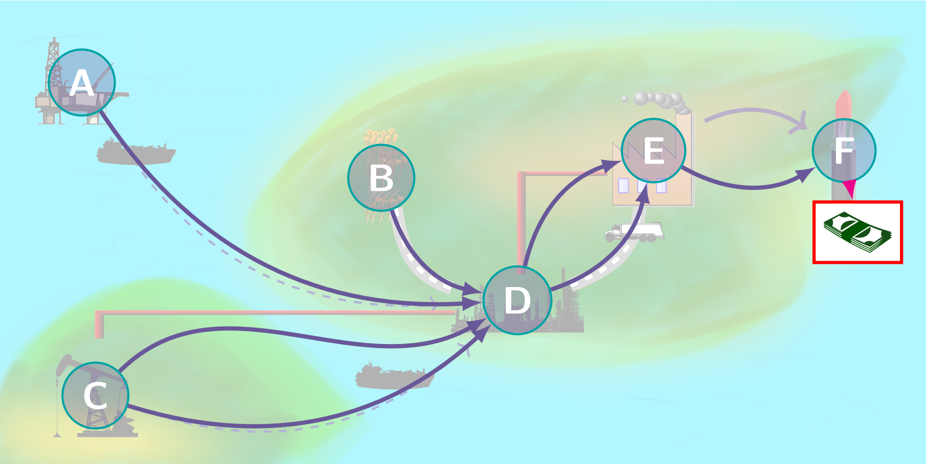
So what is a "good algorithm"?

Properties of ´good' algorithms (Why isn't there a best one?)

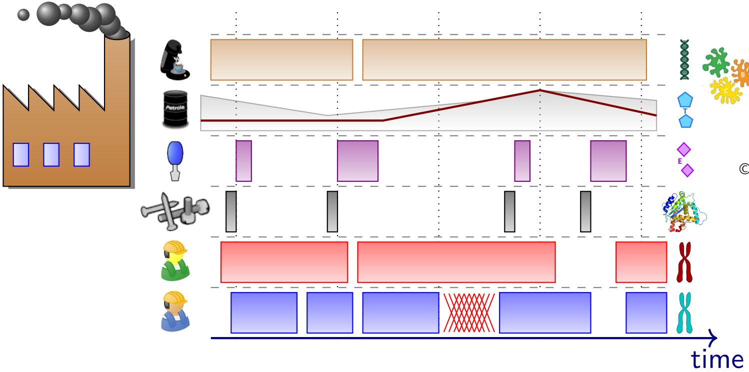
- Accuracy
- Efficiency (also: Scalability)
  - Speed
    - Total execution time (avoid NP if possible, parallel vs. sequential)
    - Number of computer operations
      - Core function/ Linear algebra/GPU-routine-calls
      - Floating point operations (plus, times, roots, exp./trig, ...), bit operations
      - Number of $f/\nabla f/\nabla^2 f$ calls
  - Memory
- Robustness
  - Reliably find all solutions for a variety of problems
  - Feedback if algorithm unsuccessful, reproducibility
- User-friendliness, ability to include expert knowledge, interactivity
- Easily extendable and maintainable
- Cheap, trusted, supported, ...

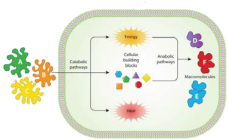
Connection to ‘Economy’ and ‘Evolution’

Our understanding of ‘economy’

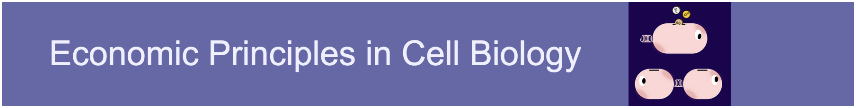


Our understanding of ‘economy’



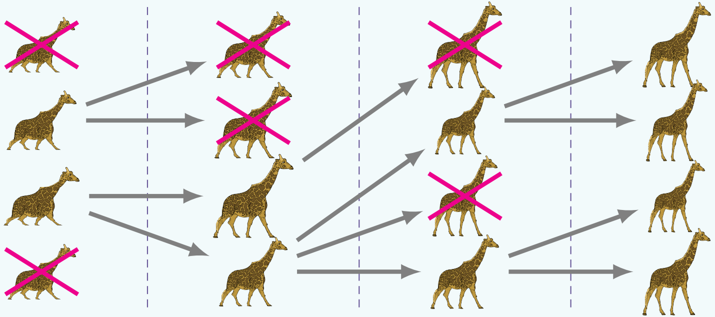


© Nature Education, 2010



Evolution

"Case study 'Plant eater' " ;-)



Quintessentially

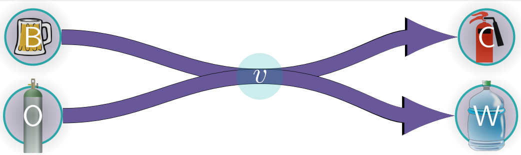
- Even if "Cells don't optimize", they have been optimized by evolution.
- "Optimization is the last religion in science."

Primer on Metabolic Network Modeling

(Hopefully) no spoilers for the next talk :-)

Metabolic Network Models

$$C_2H_5OH + 2\,O_2 \longrightarrow 2\,CO_2 + 3\,H_2O$$



Input-Output

$$\begin{array}{lc} \text{B} & 1 \text{ out} \\ & 2 \text{ out} \\ & 2 \text{ in} \\ & 3 \text{ in} \end{array} \rightsquigarrow \begin{pmatrix} -1 \\ -2 \\ 2 \\ 3 \end{pmatrix} = \mathbf{S}$$

(Time-) Dynamical System

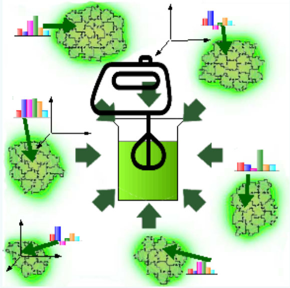
Description as an ODE/IVP

$$\dot{\boldsymbol{y}} = \mathbf{S} \cdot \boldsymbol{v}$$
$$\boldsymbol{y}(0) = \boldsymbol{y}_0$$

plus rate laws, enzymes, genes, regulation, etc.

Metabolic Network Models

The "well-stirred" metabolism



Dynamics

$$\boxed{\dot{\boldsymbol{y}}(t) = \mathbf{S} \cdot \boldsymbol{v}(t)}$$

$$\boxed{0 = \mathbf{S} \cdot \boldsymbol{v}(t)}$$

Flow dependent on quota

$$3A + B \leftrightarrow C$$

Simple rate laws

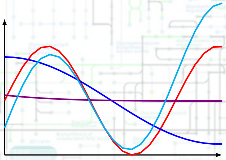
- Mass action: $f_i(t) \propto y_A^3(t) \cdot y_B(t)$ 3 A B C
- Michaelis-Menten: $f_i(t) \propto \frac{y_A(t)}{y_A(t)+K_M}$ A
- Hill-function (act.): $f_i = \tilde{f}_i \cdot \frac{y_E^\alpha}{K^\alpha + y_E^\alpha}$ E
- Hill-function (inh.): $f_i = \tilde{f}_i \cdot \frac{K^\alpha}{K^\alpha + y_E^\alpha}$ E

(Quasi) Steady State Approximation

- Some reactions orders of magnitude faster than others
- Model assumption:
  Always at equilibrium.

Constraint-based Modeling

Continuous Description

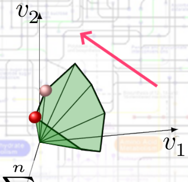


$$\frac{\mathrm{d}y_i}{\mathrm{d}t} = \sum_{j=1}^{n} S_{ij} v_j$$

$$\Rightarrow \quad \dot{\boldsymbol{y}} = \mathbf{S} \cdot \boldsymbol{v}(\boldsymbol{y}, \boldsymbol{p})$$

+ Very accurate
- Computationally expensive
- Not sufficient information

'Constraint-Based'

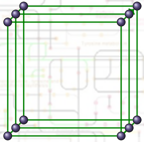


$$0 = \sum_{j=1}^{n} S_{ij} v_j$$

$$\Rightarrow \quad \mathbf{S} \cdot \boldsymbol{v} = 0, \boldsymbol{lb} \leq \boldsymbol{v} \leq \boldsymbol{ub}$$

- Consider just the data 'you know'
- Add (consecutively) as constraints
- hard/soft constraints

Discrete Description



$$Y_i^{(n+1)} = \mathsf{IF}(Y_j^n \wedge \neg Y_i^n \vee \ldots)$$

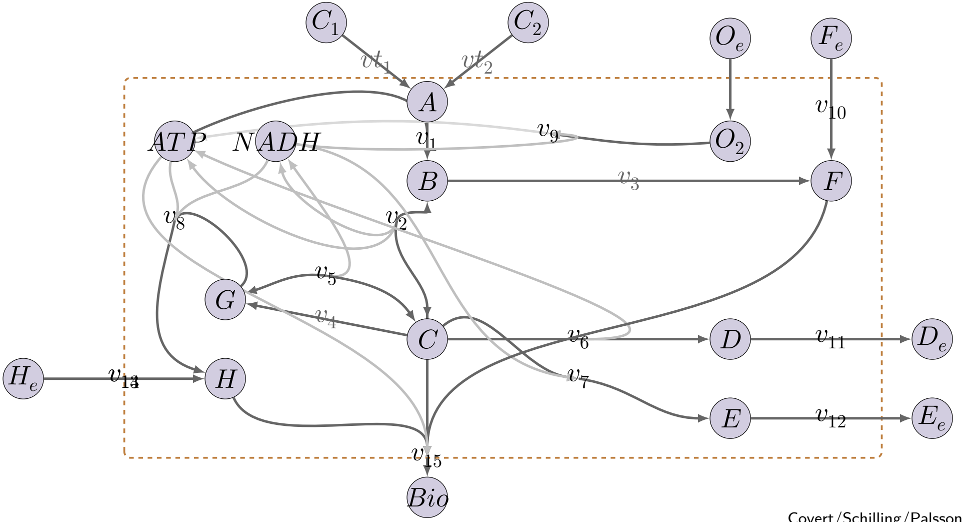
$$\Rightarrow \quad \boldsymbol{Y}^{(n+1)} = \boldsymbol{\Phi}(\boldsymbol{Y}^{(n)}; \boldsymbol{p})$$

+ Single simulations easy
- Very crude
- Just rough qualitative understanding

Examples

Examples (I)

Later today: Flux Balance Analysis



Covert/Schilling/Palsson, 2001

Examples (I)

Flux Balance Analysis

- Collect what you know (stoichiometrics, QSSA, plus lower/upper flux bounds) $\rightsquigarrow$ constraint-based modeling
- Find flux distribution from *linear optimization*

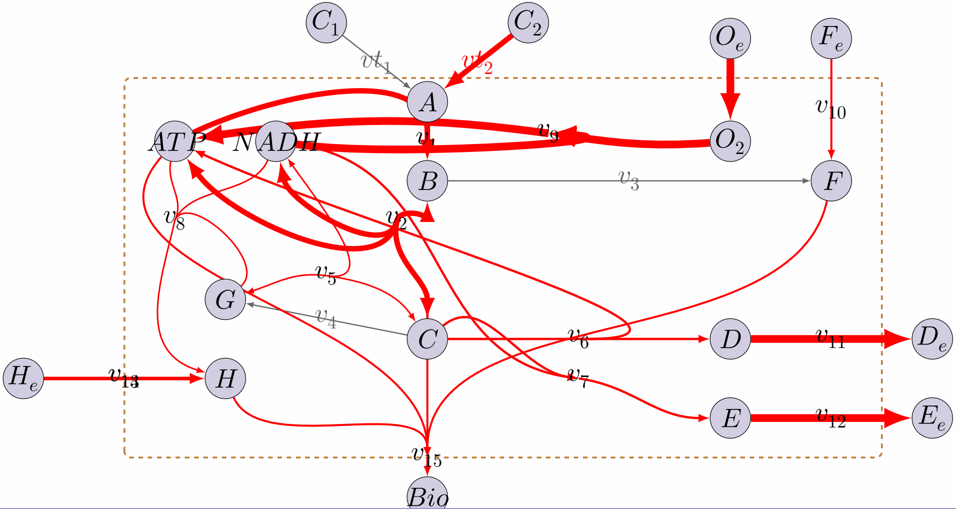
$$\max_{\boldsymbol{v}} f(\boldsymbol{v}) = \boldsymbol{b}^\top \cdot \boldsymbol{v}$$

$$\text{s.t. } 0 = \mathbf{S} \cdot \boldsymbol{v}$$

$$\boldsymbol{lb} \leq \boldsymbol{v} \leq \boldsymbol{ub}$$

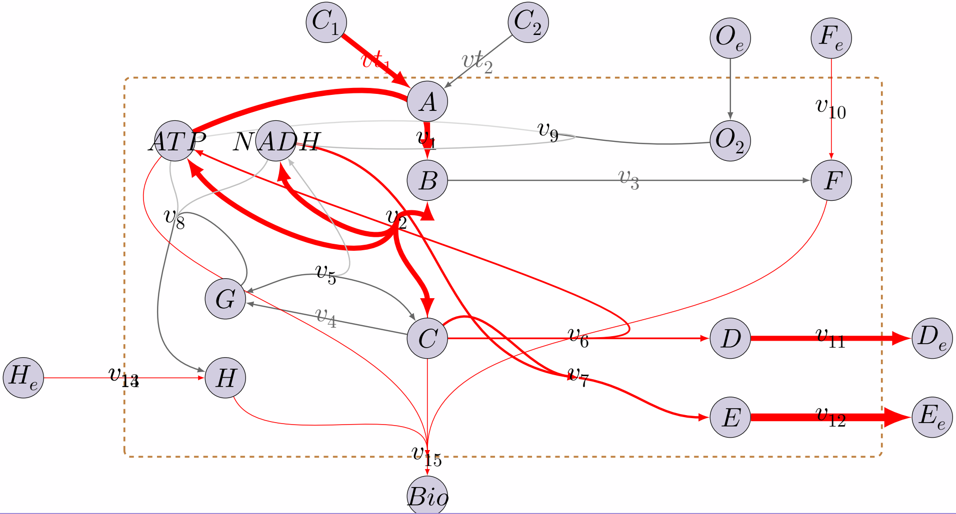
Examples (I)

Growth (i.e. Bio) maximization



Examples (I)

“E” maximization



Examples (II)

Optimal enzyme levels

A linear reaction chain

$$N \xrightarrow[v_1]{E_1,\,k_1} S_1 \xrightarrow[v_2]{E_2,\,k_2} S_2 \xrightarrow[v_3]{E_3,\,k_3} Bio$$

Klipp et al. 2002

Network's constraints: $\mathbf{S} \cdot \boldsymbol{v} = 0,\ 0 \leq \boldsymbol{v}$

$$\boldsymbol{v} \leq \mathrm{diag}(k_1, k_2, k_3) \cdot \begin{pmatrix} E_1 \\ E_2 \\ E_3 \end{pmatrix}$$

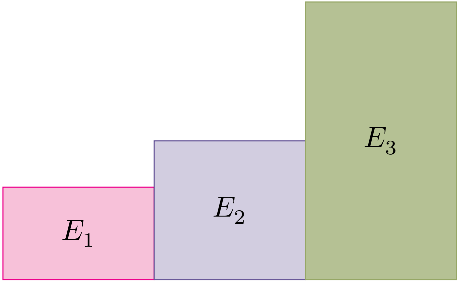
- Goal: Maximize growth reaction v_3
- 'Costs': Enzyme activation $E_1^2 + E_2^2 + E_3^2$

Examples (II)

Optimal enzyme levels

(Toy-) example: $k_1 = 3$, $k_2 = 2$, $k_3 = 1$.

$$J = \quad (E_1^2 + E_2^2 + E_3^2) - \quad v_3$$



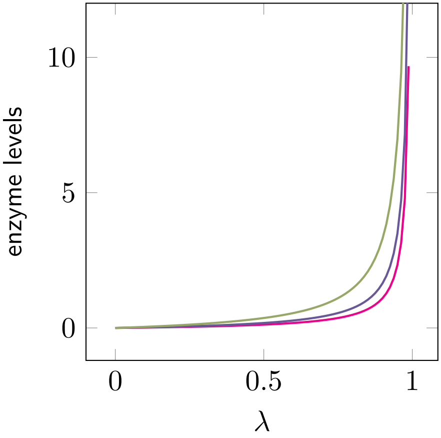
Gain:
0.3673
Cost:
0.3673

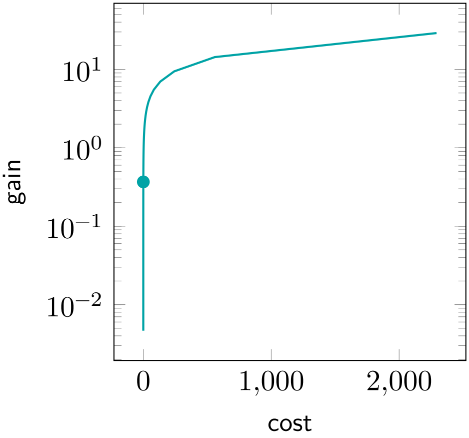
Examples (II)

Optimal enzyme levels

(Toy-) example: $k_1 = 3$, $k_2 = 2$, $k_3 = 1$.

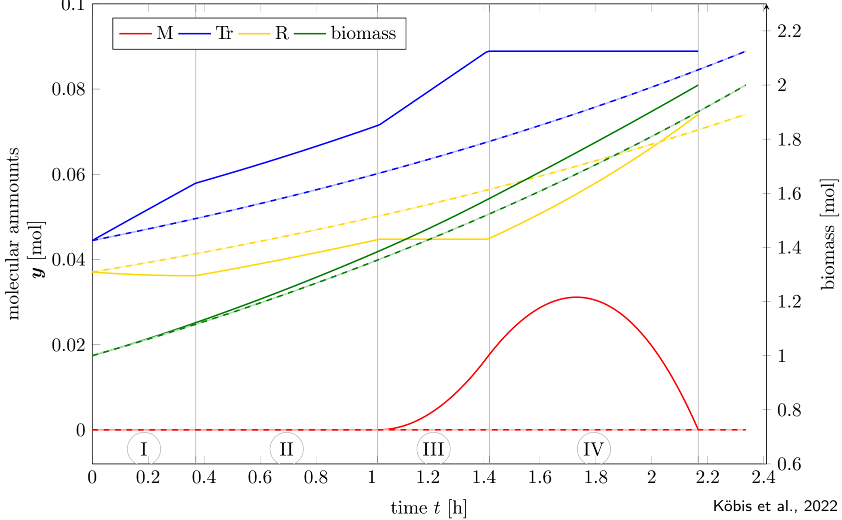
$$J = (1-\lambda)\cdot(E_1^2 + E_2^2 + E_3^2) - \lambda \cdot v_3$$



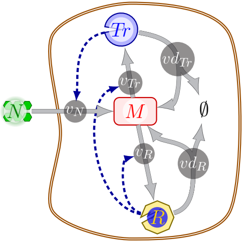


Examples (III)

"Optimality in Time" $\rightarrow$ infinite-dimensional optimization

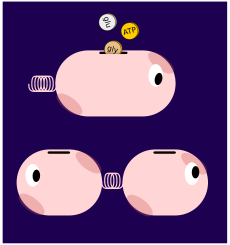


Köbis et al., 2022



Self-replicator

$J = t_{\mathrm{end}}$



Potential Objective Functions

Given: stoichiometrics, flux bounds, some "dynamic data", ...

Biologically inspired optimization principles

1. Cell efficiency: "Minimize fluxes" ("Tikhonov regularization")

$$J := \int \|\boldsymbol{v}(\cdot)\|_*^2 \mathrm{d}t \qquad \text{plus minimum growth conditions}$$

2. Growth (a): Maximize biomass/macro molecule production of the cells

$$J := -\int \|\boldsymbol{w}_{\mathrm{obj}}(t)^\top \cdot \boldsymbol{y}(t)\|_*^2 \mathrm{d}t$$

3. Growth (b): Maximize flux through biomass reaction(s)

$$J := -\int \|\boldsymbol{V}_{\boldsymbol{y}_{\mathrm{growth}}}(t)\|_*^2 \mathrm{d}t$$

4. Robustness (a, b): Maximize survival time, minimize response times

$$J = -t_{\mathrm{end}} = -\int 1 \mathrm{d}t \quad \text{and cell survival}$$

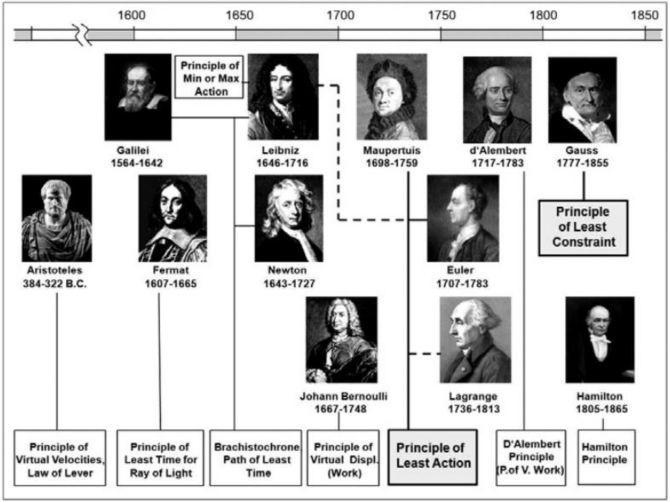
5. Robustness (c): Maximize nutrient uptake

$$J = -\int \|\boldsymbol{w}_{\mathrm{obj}}(t)^\top \cdot \boldsymbol{v}(t)\|_*^2 \mathrm{d}t$$

Historical Notes

Historical Notes (I)

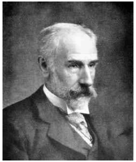
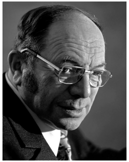
Optimization Methods in the Natural Sciences

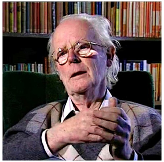


Historical Notes (II)

Optimization Methods in Economy

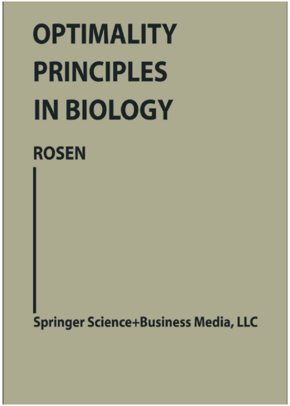
- 1881: Edgeworth, *Mathematical Psychics*
- 1939: Production Planning using linear optimization (Kantorovich)
- 1939–1945: World War II (Operations Planning)
- 1944: von Neumann, Morgenstern, *Theory of games and economic behavior*
- 1947: Simplex Algorithm (Dantzig)
- 1954: Markowitz (quadratic programming, portfolio analysis, later: risk measures)
- 1973: Maynard, Price *The logic of animal conflict*

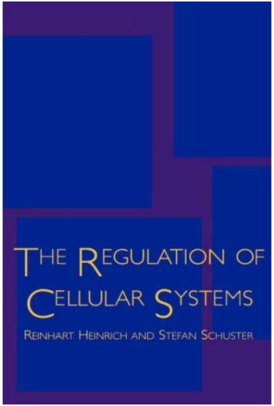



Historical Notes (III)

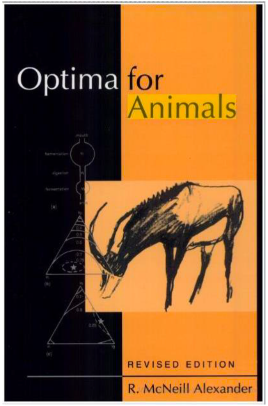
Optimization Methods in (Systems) Biology



1967



1996



1996

- Flux balance analysis: 1990
- (Modern) systems biology fully established: 2000–2010

Concluding Remarks/Discussion

Where do we go from here?

Not covered in this talk

- parameter fitting
- artificial intelligence
- touched upon: multi-objective optimization, optimal control
- handling of uncertainty, robustness ($\rightsquigarrow$ talk tomorrow), regularization
- (mixed-) integer optimization problems (e.g. optimal network design)
- stochastic optimization
- game theory, inverse optimality
- FBA $\rightsquigarrow$ RBA $\rightsquigarrow$ ...
- algorithms, their "optimization"
- mathematics: Duality, optimality conditions, cones, constraint qualification, ...

Conclusion (I)

Central issue: Lack of first principles in (systems) biology

Optimization in Constraint-based Modeling

- Optimization itself not necessarily driving force but often as a proxy based on
  - the viewpoint of cells as 'economic actors'
  - "cells do not optimize", BUT "cells have been optimized by evolution"
- "Sometimes, things *look* optimal." (The end justifies the means.)

Conclusion (II)

'Essentially, all models are wrong, but some are useful.'

George Box, Norman Draper, Empirical Model-Building and Response Surfaces (1987) CC Wikipedia

Thank You!

Image sources

Screenshot "Is there an optimal density?": Talk about Cell components from 2023 summer school (Pranas & Diana)
matlab logo: https://se.mathworks.com/content/mathworks/se/en/company/newsletters/articles/the-mathworks-logo-is-an-eigenfunction-of-the-wave-equation/_jcr_content/mainParsys/image_2.adapt.full.medium.gif/1469941373397.gif
Fortran logo: https://github.com/fortran-lang/fortran-lang.org/blob/master/assets/img/fortran-logo.svg
tanker: https://openclipart.org/detail/318334/tanker-silhouette
offshore rig: https://openclipart.org/detail/323036/an-offshore-oil-rig
oil pump: https://openclipart.org/detail/310626/simple-oil-pump
oil refinery: https://openclipart.org/detail/279473/oil-refinery-silhouette
rape flower: https://openclipart.org/detail/238177/rapeseed-low-resolution
lipstick: https://openclipart.org/detail/311193/full-lipstick
money: https://openclipart.org/detail/222589/money
truck: https://openclipart.org/detail/182107/oil-and-gas-tanker-truck
Coffee machine: https://openclipart.org/detail/17995/coffee-machine
Oil barrel: https://openclipart.org/detail/18090/oil-barrel-baril-de-petrole
Screws: https://openclipart.org/detail/4816/screw
Screw driver: https://openclipart.org/detail/6166/screwdriver
Protein structure: https://commons.wikimedia.org/wiki/File:Spombe_Pop2p_protein_structure_rainbow.png
Giraffe: https://openclipart.org/detail/6958/giraffe
Beer glass: https://openclipart.org/detail/17276/fatty-matty-brewing-beer-mug-icon
Oxygen tank: https://openclipart.org/detail/188627/oxygen-tank
Fire-extinguisher: https://openclipart.org/detail/281430/fire-extinguisher-carbon-dioxide
Water bottle: https://openclipart.org/detail/181115/water-bottle
Mixer: https://findicons.com/icon/568663/mixer
Edgeworth: https://en.wikipedia.org/wiki/Francis_Ysidro_Edgeworth#/media/File:Edgeworth.jpeg
Kantorovich: https://en.wikipedia.org/wiki/Leonid_Kantorovich#/media/File:Leonid_Kantorovich_1975.jpg
von Neumann: https://en.wikipedia.org/wiki/John_von_Neumann#/media/File:JohnvonNeumann-LosAlamos.gif
Dantzig: https://en.wikipedia.org/wiki/George_Dantzig#/media/File:George_B._Dantzig_at_National_Medal_of_Science_Awards_Ceremony,_1976.jpg
Maynard Smith: https://en.wikipedia.org/wiki/John_Maynard_Smith#/media/File:John_Maynard_Smith.jpg
George Box: https://en.wikipedia.org/wiki/George_E._P._Box#/media/File:GeorgeEPBox.jpg