Constraint-Based Workshops

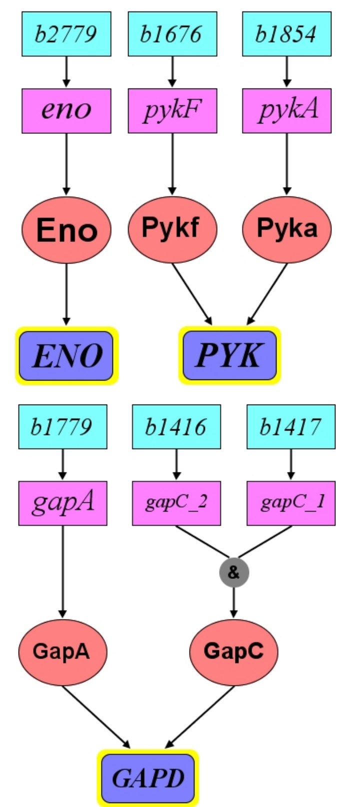
3. Model Construction
December 13th, 2007

Network Assembly and Representation

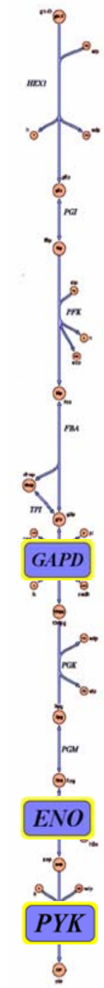
Reconstruction of Glycolytic Pathway

Abbr.	Glycolytic Reactions	Genes
HEX1	[c]glc +atp ⟶ g6p + adp	glk
PGI	[c]g6p ⟷ f6p	pgi
PFK	[c]atp + f6p ⟶ adp + fdp + h	pfkA,pfkB
FBA	[c]fdp ⟷ dhap + g3p	fbaA,fbaB
TPI	[c]dhap ⟷ g3p	tpiA
GAPD	[c]g3p + nad + pi ⟷ 13dpg + h + nadh	gapA,gapC_1,gapC_2
PGK	[c]13dpg + adp ⟷ 3pg + atp	pgk
PGM	[c]3pg ⟷ 2pg	gpmA,gpmB
ENO	[c]2pg ⟷ h2o + pep	eno
PYK	[c]adp + h + pep ⟶ atp + pyr	pykA,pykF

	HEX1	PGI	PFK	FBA	TPI	GAPD	PGK	PGM	ENO	PYK
atp	-1	0	-1	0	0	0	1	0	0	1
glc	-1	0	0	0	0	0	0	0	0	0
adp	1	0	1	0	0	0	-1	0	0	-1
g6p	1	-1	0	0	0	0	0	0	0	0
h	1	0	1	0	0	1	0	0	0	-1
f6p	0	1	-1	0	0	0	0	0	0	0
fdp	0	0	1	-1	0	0	0	0	0	0
dhap	0	0	0	1	-1	0	0	0	0	0
g3p	0	0	0	1	1	-1	0	0	0	0
nad	0	0	0	0	0	-1	0	0	0	0
pi	0	0	0	0	0	-1	0	0	0	0
13dpg	0	0	0	0	0	1	-1	0	0	0
nadh	0	0	0	0	0	1	0	0	0	0
3pg	0	0	0	0	0	0	1	-1	0	0
2pg	0	0	0	0	0	0	0	1	-1	0
pep	0	0	0	0	0	0	0	0	1	-1
h2o	0	0	0	0	0	0	0	0	1	0
pyr	0	0	0	0	0	0	0	0	0	1



PYK: IF *pykA* OR *pykF*
ENO: IF *eno*
GAPD: IF *gapA* OR (*gapC_1* AND *gapC_2*)



Outline

- Formulate simple S and draw network and metabolite maps for a toy system
- Optimization & Flux Balance Analysis
- Implement model into matlab and solve for maximizing a single flux.
- Run analysis of central metabolic model.

Matrix of Stoichiometric Coefficients

Integral numbers
Universal biochemical constants
Constants: time-invariant

chemical reaction:

$$aA + cC \xrightarrow{v_i} eE + hH$$

Representation as a column in a matrix:

compounds ↓

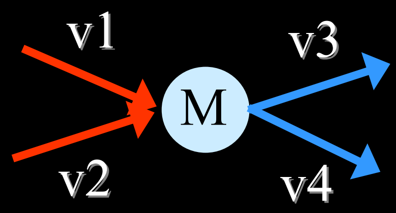
$$
\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \\ F \\ G \\ H \end{array}
\begin{array}{c} v_i \\ \left[\begin{array}{ccccccccc}
\bullet & \bullet & \bullet & \bullet & -a & \bullet & \bullet & \bullet & \bullet \\
 & & & & 0 & & & & \\
 & & & & -c & & & & \\
 & & & & 0 & & & & \\
 & & & & +e & & & & \\
 & & & & 0 & & & & \\
 & & & & 0 & & & & \\
\bullet & \bullet & \bullet & \bullet & +h & \bullet & \bullet & \bullet & \bullet
\end{array}\right] \end{array}
$$

Metabolite Connectivity Maps

Connectivity Map for Metabolite M

Fluxes that produce metabolite M

Fluxes that consume metabolite M



Example

Reaction List

v_1: A → B

v_2: A → C

b_1: → A

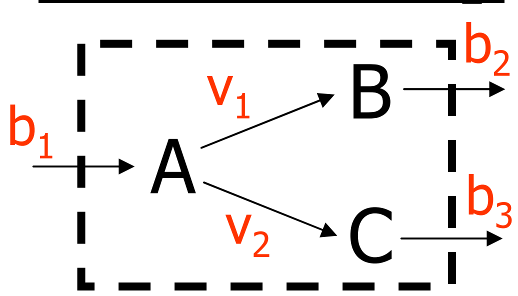
b_2: B →

b_3: C →

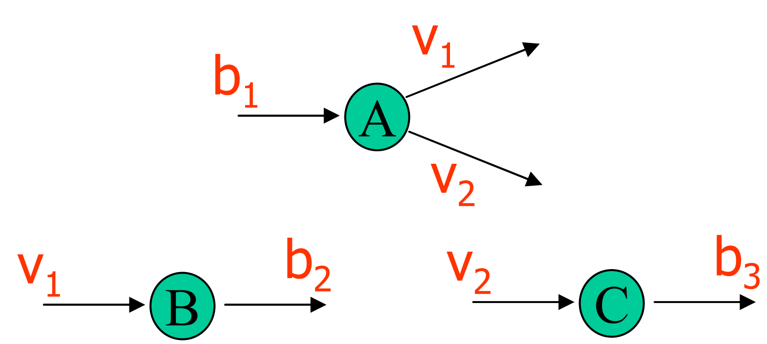
S Matrix

	v_1	v_2	b_1	b_2	b_3
A	-1	-1	1	0	0
B	1	0	0	-1	0
C	0	1	0	0	-1

Metabolic Map



Metabolite Connectivity Maps



Toy Network

Use the reaction list for the toy network to:

1. Write the stoichiometric matrix for the network
2. Draw a pathway map for the network
3. Draw the metabolite connectivity maps

Name	Reaction
v1	$A \rightarrow B$
v2	$2B \rightarrow C + byp$
v3	$2B + cof \rightarrow D + byp$
v4	$D \rightarrow E + cof$
v5	$C + cof \rightarrow D$
v6	$C \rightarrow E$
b1	$\rightarrow A$
b2	$E \rightarrow$
b3	$byp \rightarrow$

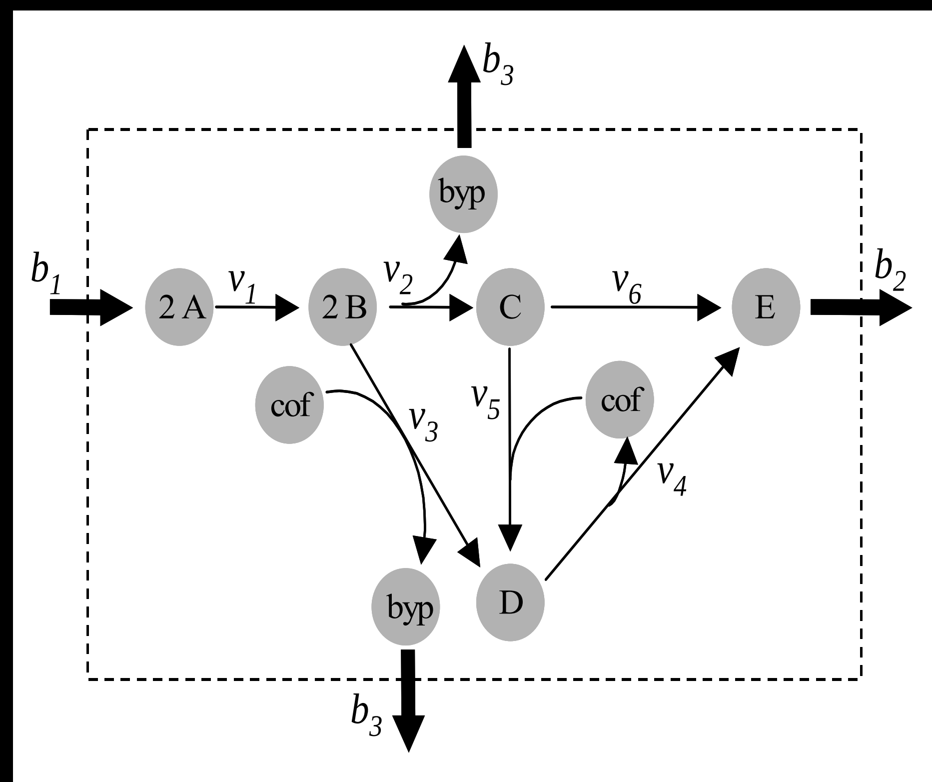
Stoichiometric Matrix and Pathway Map for Toy Network

Stoichiometric Matrix

S =

	v_1	v_2	v_3	v_4	v_5	v_6	b_1	b_2	b_3
A	−1	0	0	0	0	0	+1	0	0
B	+1	−2	−2	0	0	0	0	0	0
C	0	+1	0	0	−1	−1	0	0	0
D	0	0	1	−1	+1	0	0	0	0
E	0	0	0	+1	0	+1	0	−1	0
byp	0	+1	+1	0	0	0	0	0	−1
cof	0	0	−1	+1	−1	0	0	0	0

Pathway Map



Optimization:

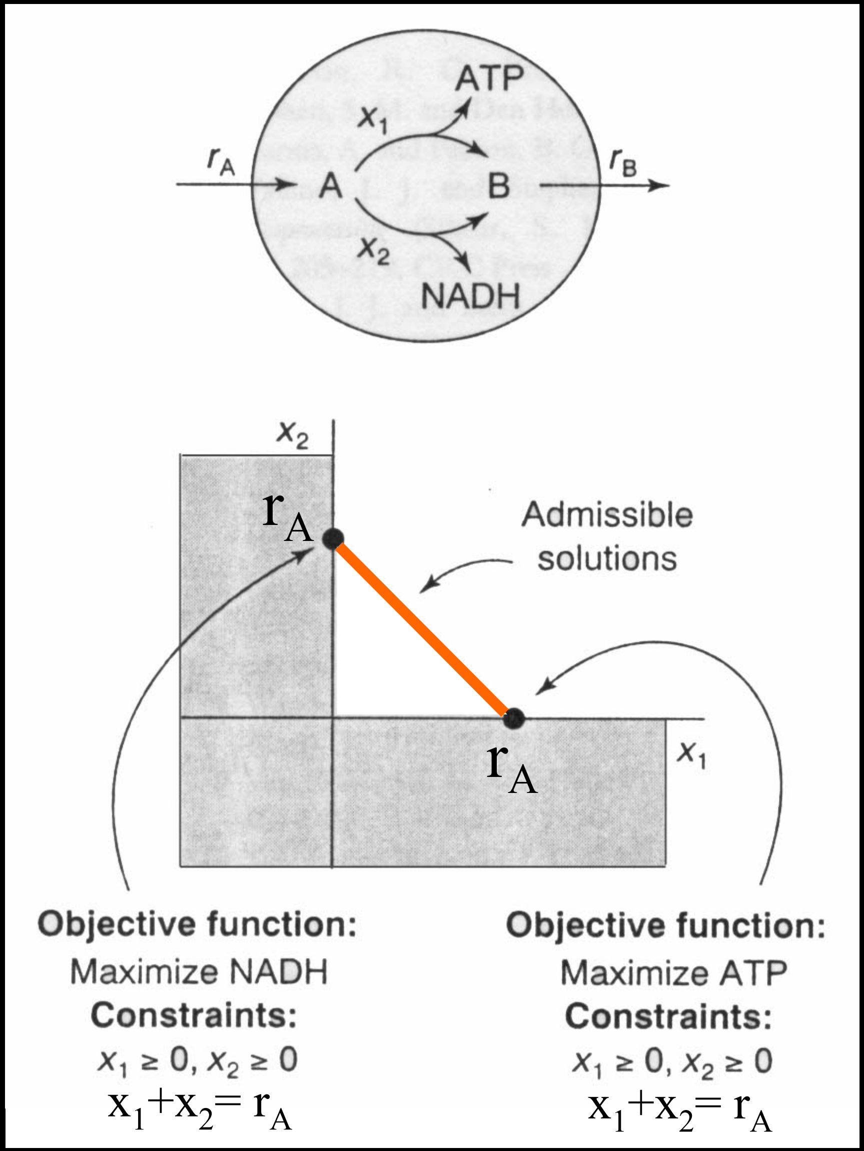
Used when multiple solutions
satisfy your equations.

How does LP work?
A 1D example

The solution space is the line of admissible in the positive orthant.

If we maximize ATP production the solution lies on the x-axis where all the flux would be through reaction x_1. Conversely, maximizing NADH production would give the point at the y-axis, where only reaction x_2 is active.

Note that the optimal solutions lie at the boundary of the admissible space.



Bonarius,et al TIBTECH vol 15:308 (1997)

An Illustrative Example

Consider two variables A and B, which are the amount of toy cars and trucks you can produce.

Do to resource limitations you can make no more than 60 cars a day and no more than 50 trucks a day.

$$0 < A < 60$$

$$0 < B < 50$$

You are also limited by shipping such that the number of cars plus twice the number of trucks must be less than 150.

$$A + 2B < 150$$

You can sell the toys at \$20/car and \$30/truck your earnings (Z) are given by: $Z = 20A + 30B$

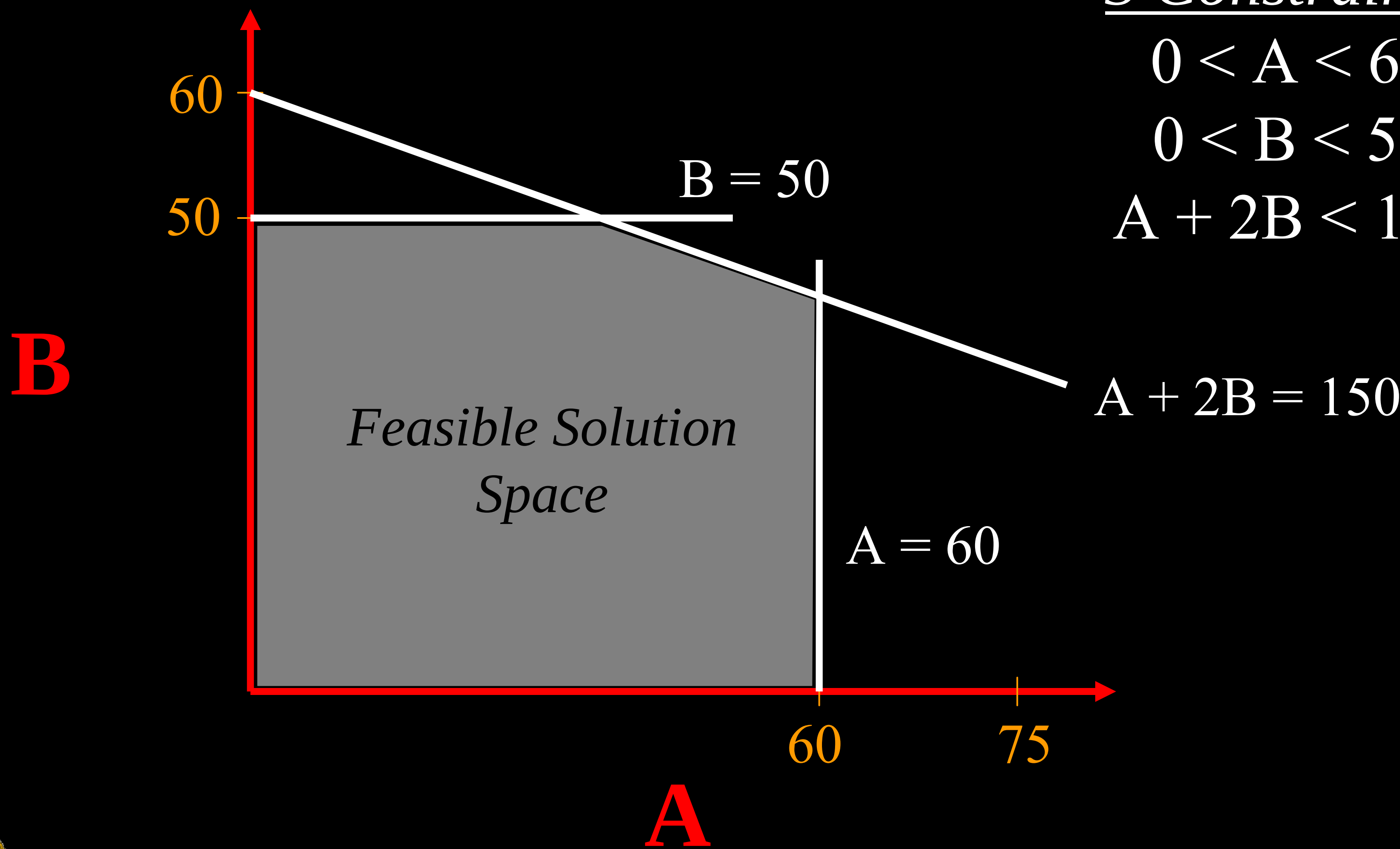
Graphical Representation of Feasible Solution Space

3 Constraints:

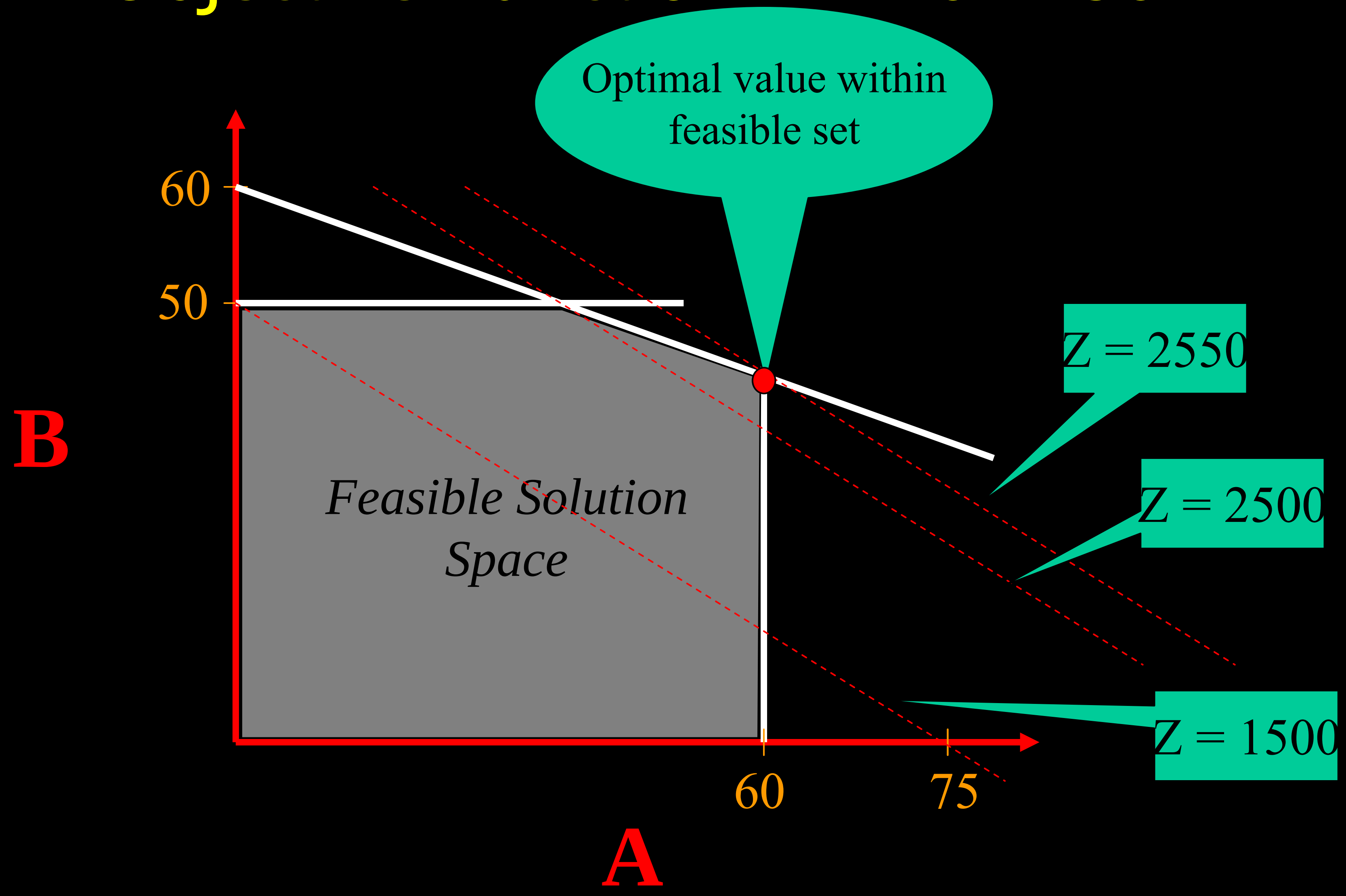
0 < A < 60

0 < B < 50

A + 2B < 150

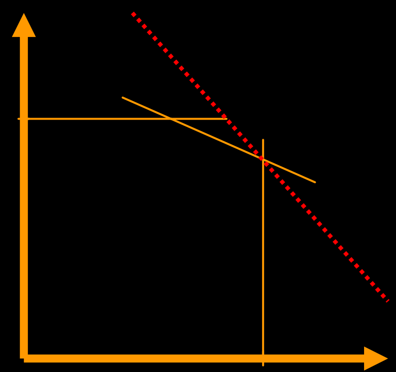


Graphical Representation of the Objective Function: Z=20A+30B



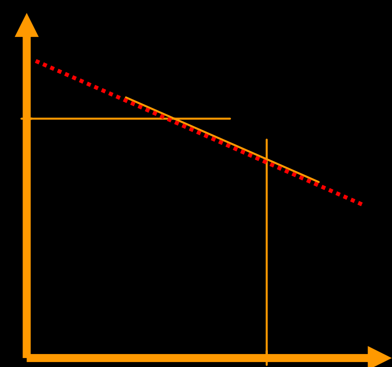
Types of solutions:
the impact of the objective function

Single solution



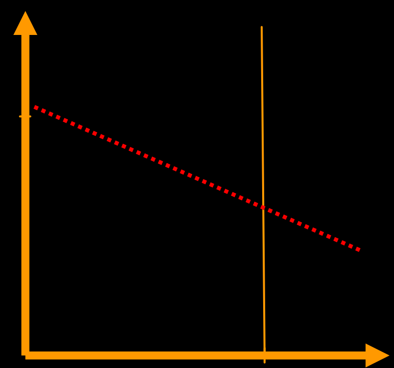
Optimal solution in a corner

Degenerate solution



Optimal solution along an edge

No solution



Optimal solution not found--region unbounded

Lines of constant Z

Flux Balance Analysis

- Used to identify a single flux distribution that that maximizes or minimizes some linear combination of fluxes.

  – eg. find a flux distribution that maximizes ATP production

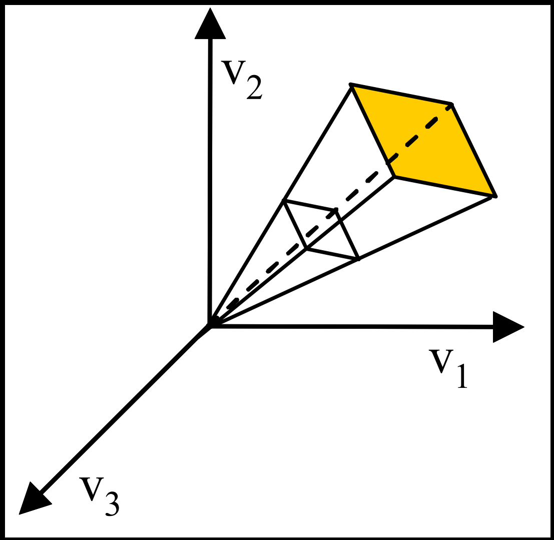
FBA Optimization Problem Statement

- Objective Function: A function that is maximized or minimized to identify optimal solutions
- Constraints: Place limits on the allowable values the solutions can take on.

Maximize: $c \cdot v$

Such that $S \cdot v = 0$

$ub \geq v \geq lb$



COBRA Toolbox

- For detailed directions on how to install COBRA toolbox read: Becker et al. Nature Protocols 2(3):727-738 (2007).

- You can download the matlab toolbox and a number of models you might be interested in from Bernhard Palsson's website:
  - http://gcrg.ucsd.edu/

- To convert SBML files to matlab models you will also need to download the SBML toolbox (I have already converted one model for you).

Directions for Preparing Matlab

- Download and unzip GLPK solver (this is a free optimization solver that Matlab can access)
- Download and unzip Cobra Toolbox.
- In Matlab, add the folders and subfolders in the path for the toolbox and glpk files (→File →Set Path)
- Download CoreTextbookModel.mat file

Directions Cont.

- Change the directory in Matlab to wherever you put the CoreTextbookModel.mat file.
- At matlab command prompt, type:
  - initCobraToolbox
- To load model, type:
  - load COBRA_central_model
- To look at reactions, type:
  - printRxnFormula(model,model.rxns)

FBA Calculations

1. What is the maximum growth rate for glucose aerobic growth (max. glucose uptake rate of 5)?
2. What is the maximum growth rate for glucose anaerobic (no oxygen uptake) growth (max. glucose uptake rate of 5)?
3. What are the by-products that are secreted during maximal glucose anaerobic growth?
4. Can *E. coli* grow anaerobically on acetate ?
   (hint: to get a feasible solution set lowerlimit to -5 for EX_ac_e and upperlimit to 0)

- <u>If you have extra time</u>, what is the maximum amount of ethanol you can produce under aerobic conditions from glucose (uptake rate =5)? (ans = 8.5)