When is complex too complex?

Graph energy and its role in proactive complexity management of cyber-physical systems

Prof. Olivier de Weck
Massachusetts Institute of Technology
deweck@mit.edu

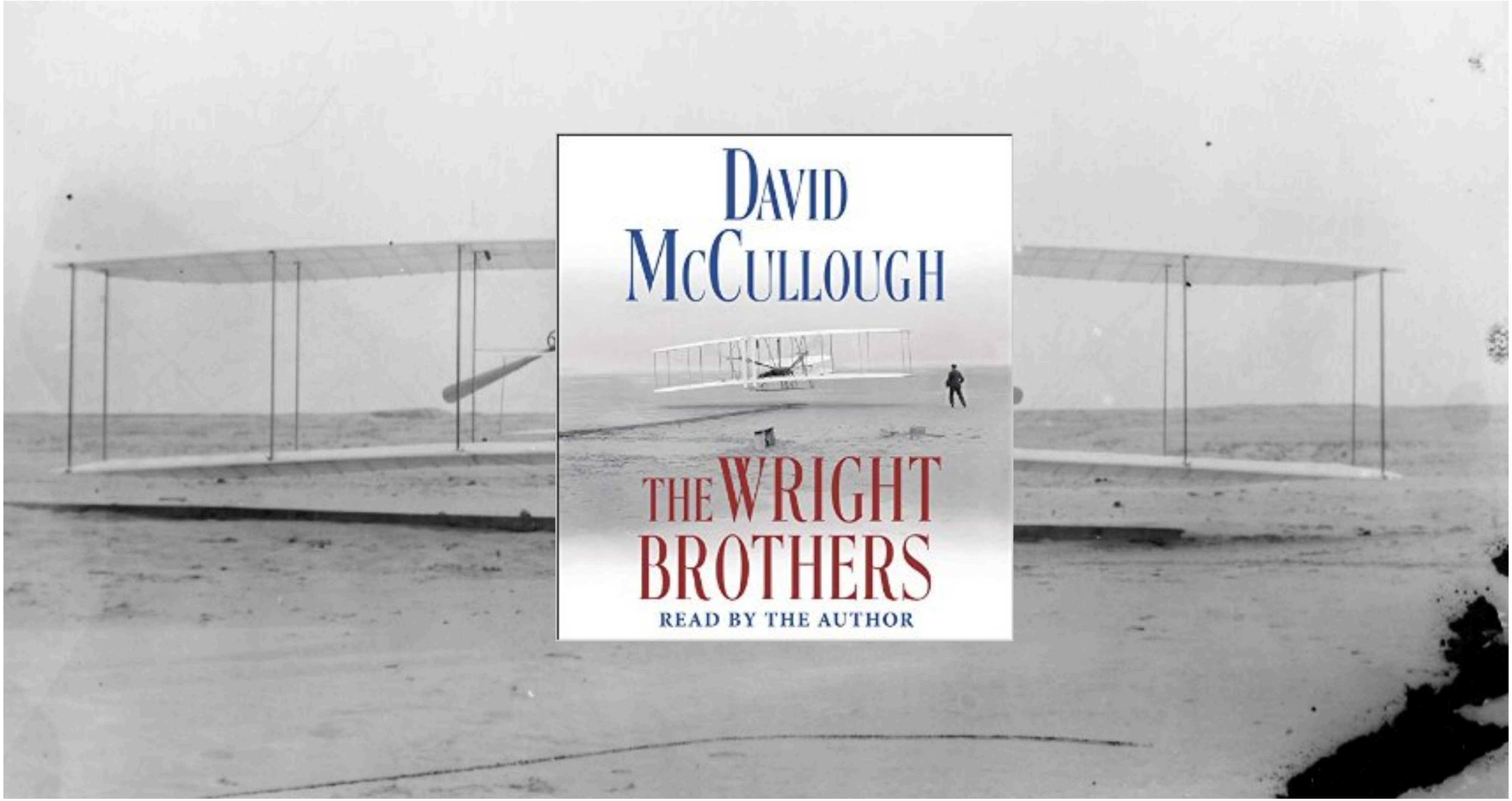
Joint work with Dr. Kaushik Sinha and Narek Shougarian

Why should we care about complexity?

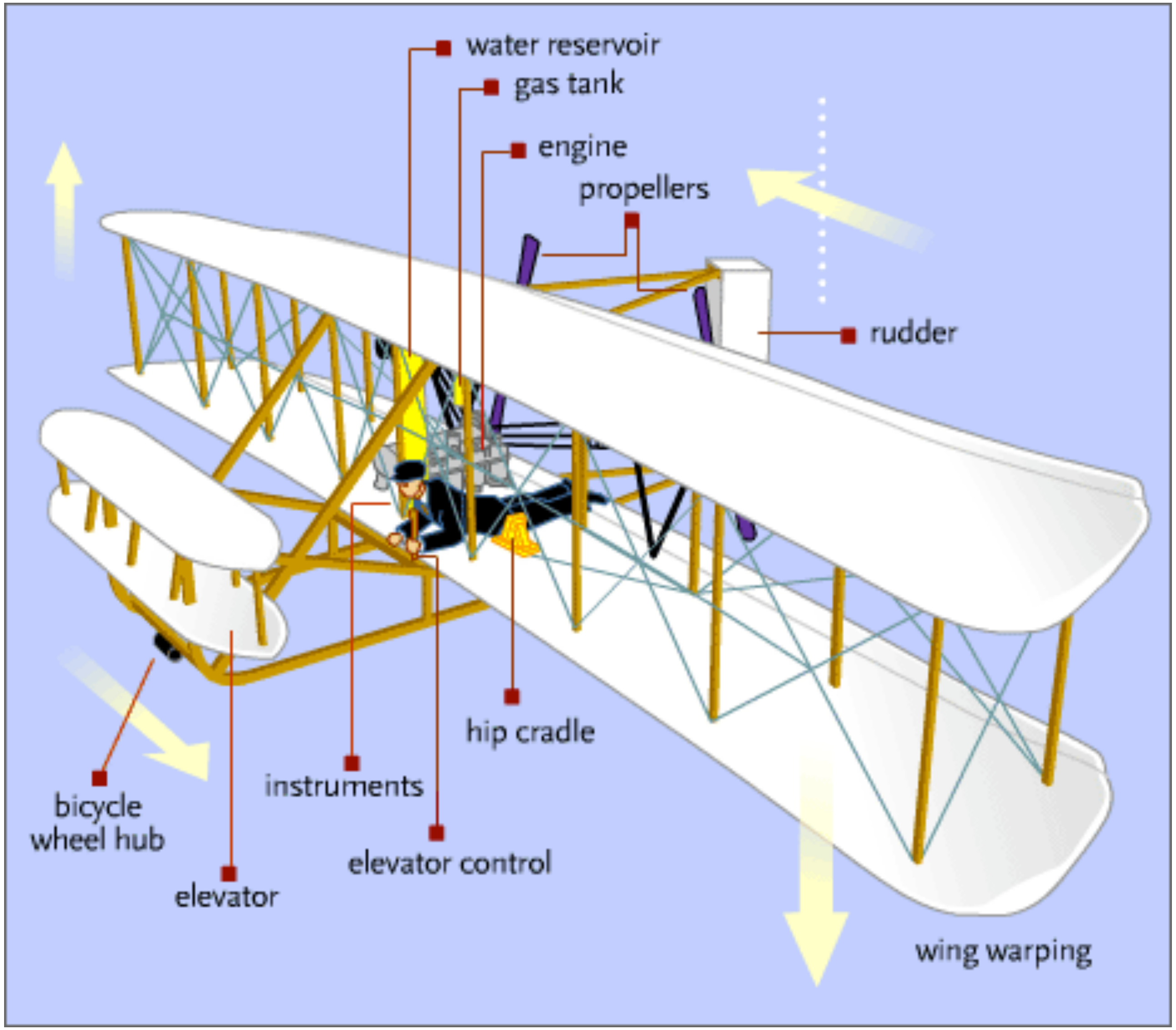
How do we quantify complexity?

How to better manage complexity?

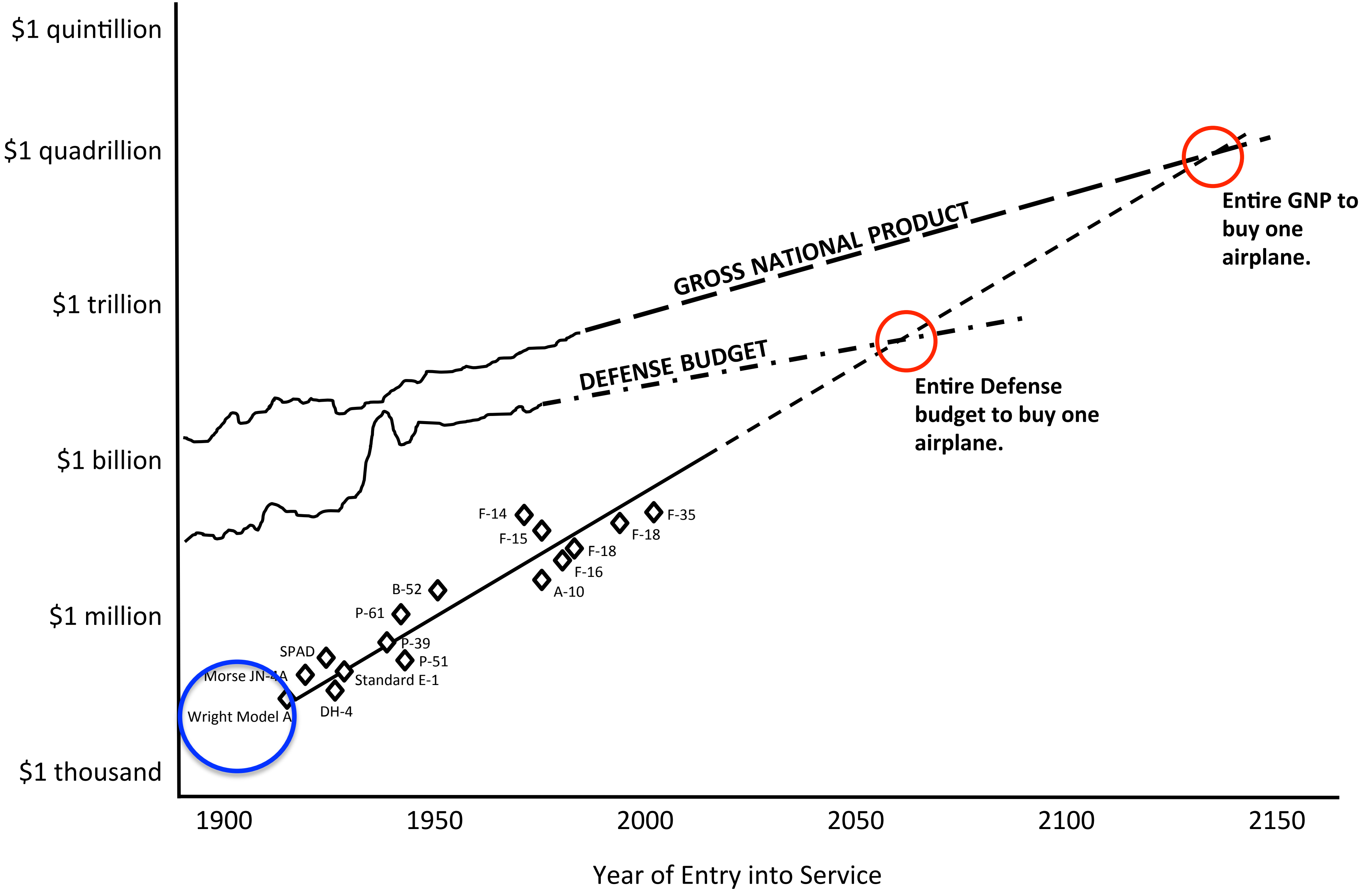
What have you been reading lately?



The Wright Flyer

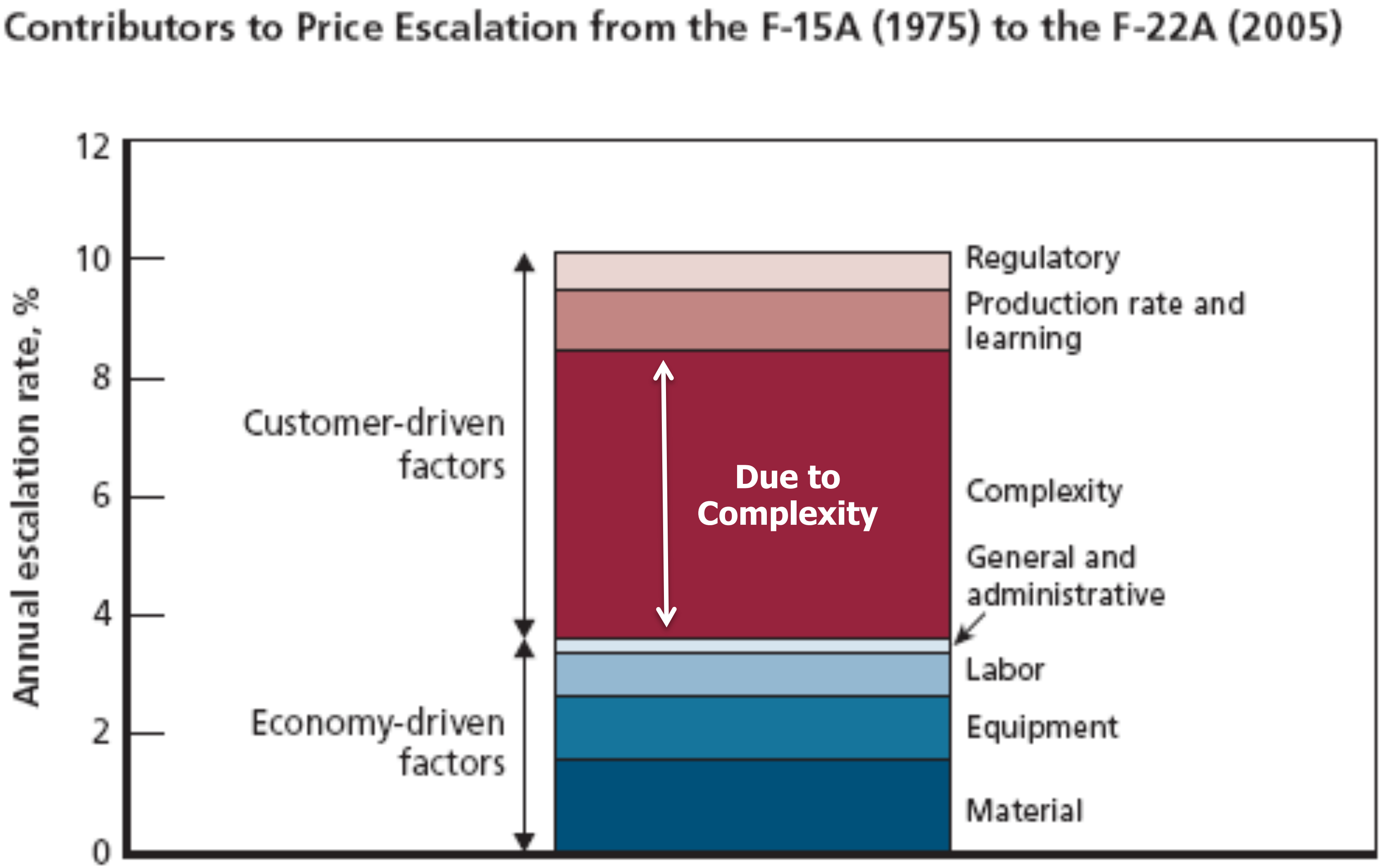


Augustine's 16th Law



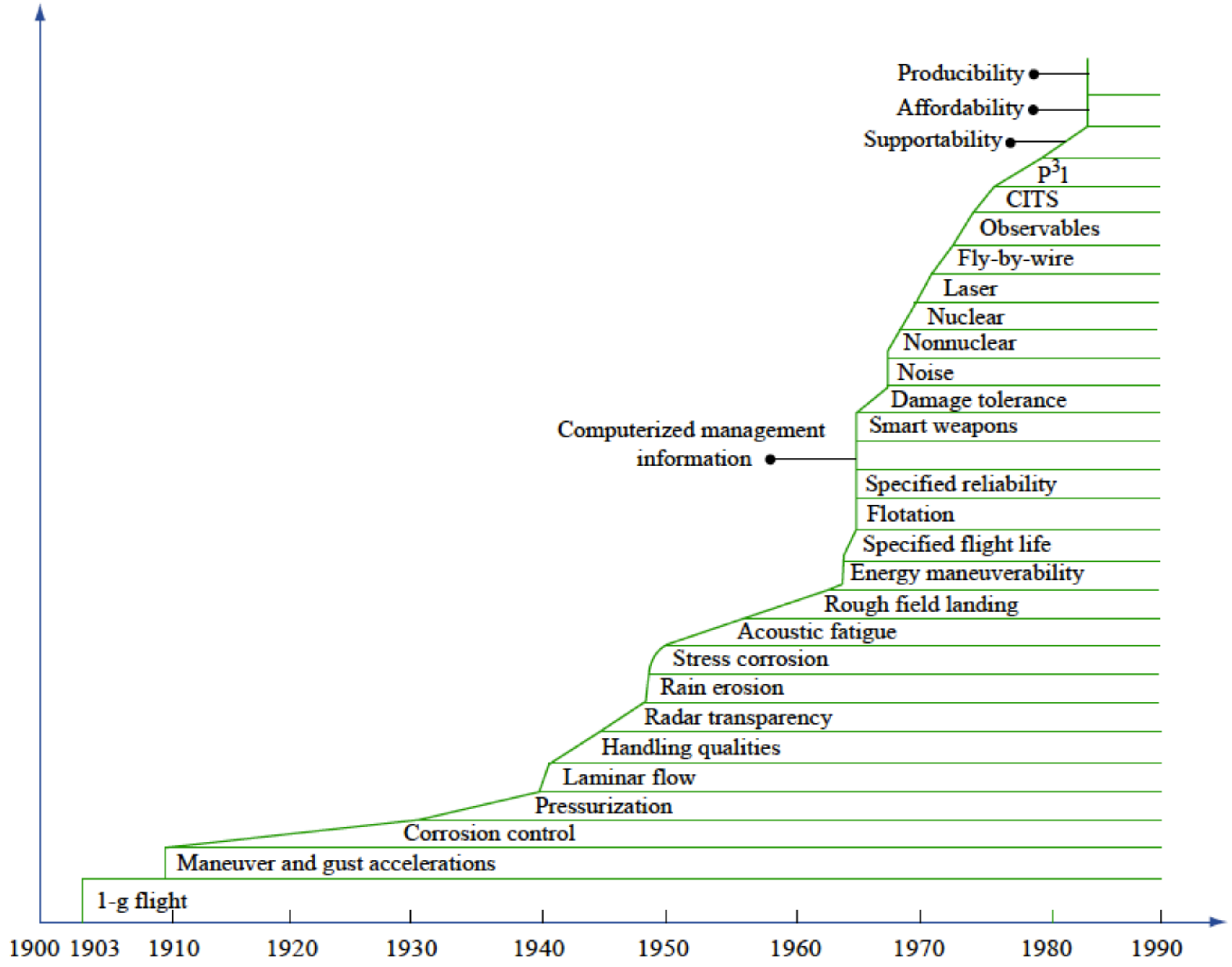
Norm Augustine, *Augustine's Laws*, 6th Edition, AIAA Press, 1997.

What is driving this escalation of cost?



Source: DARPA TTO (2008)

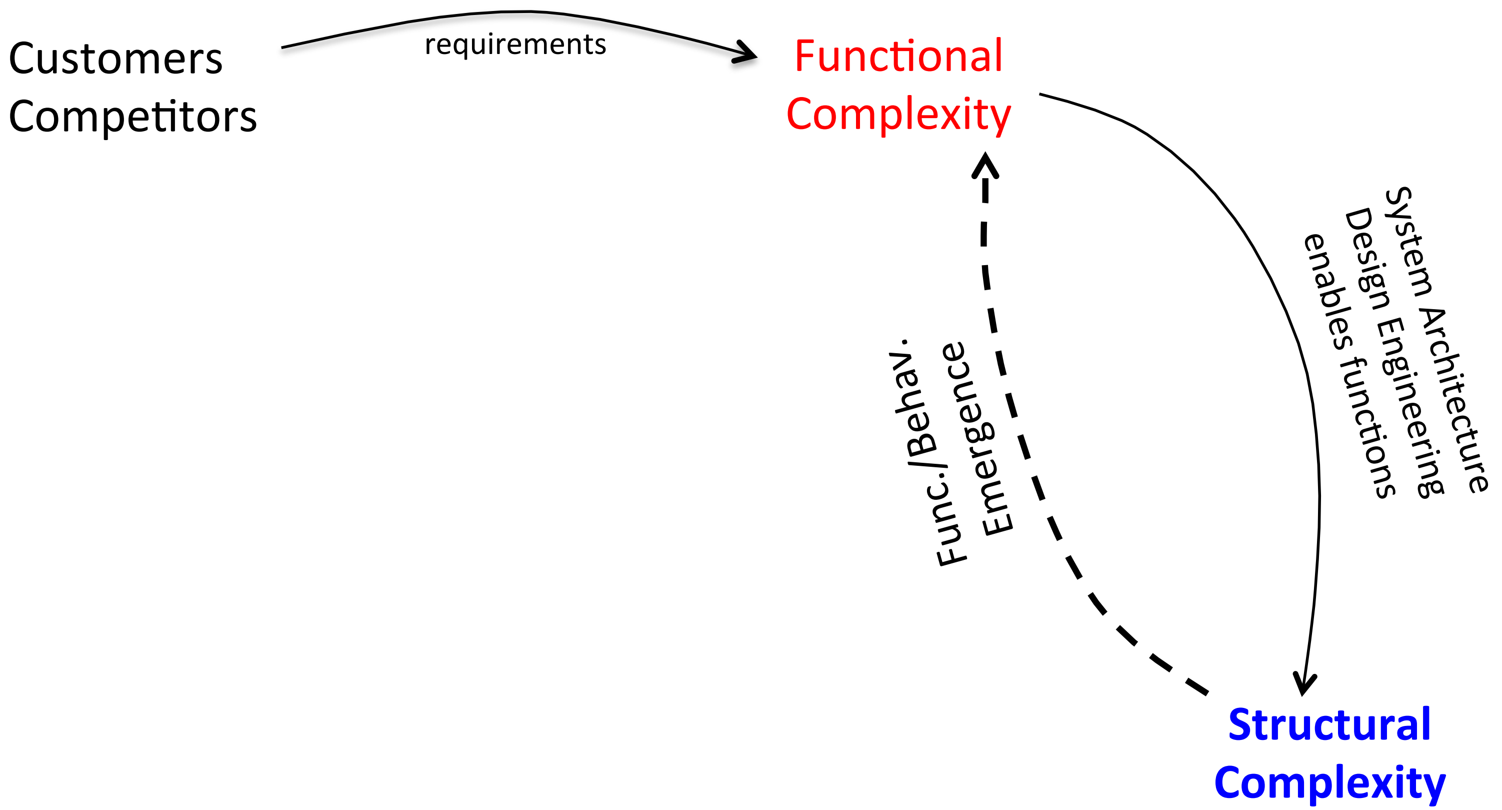
Functional Requirements Explosion



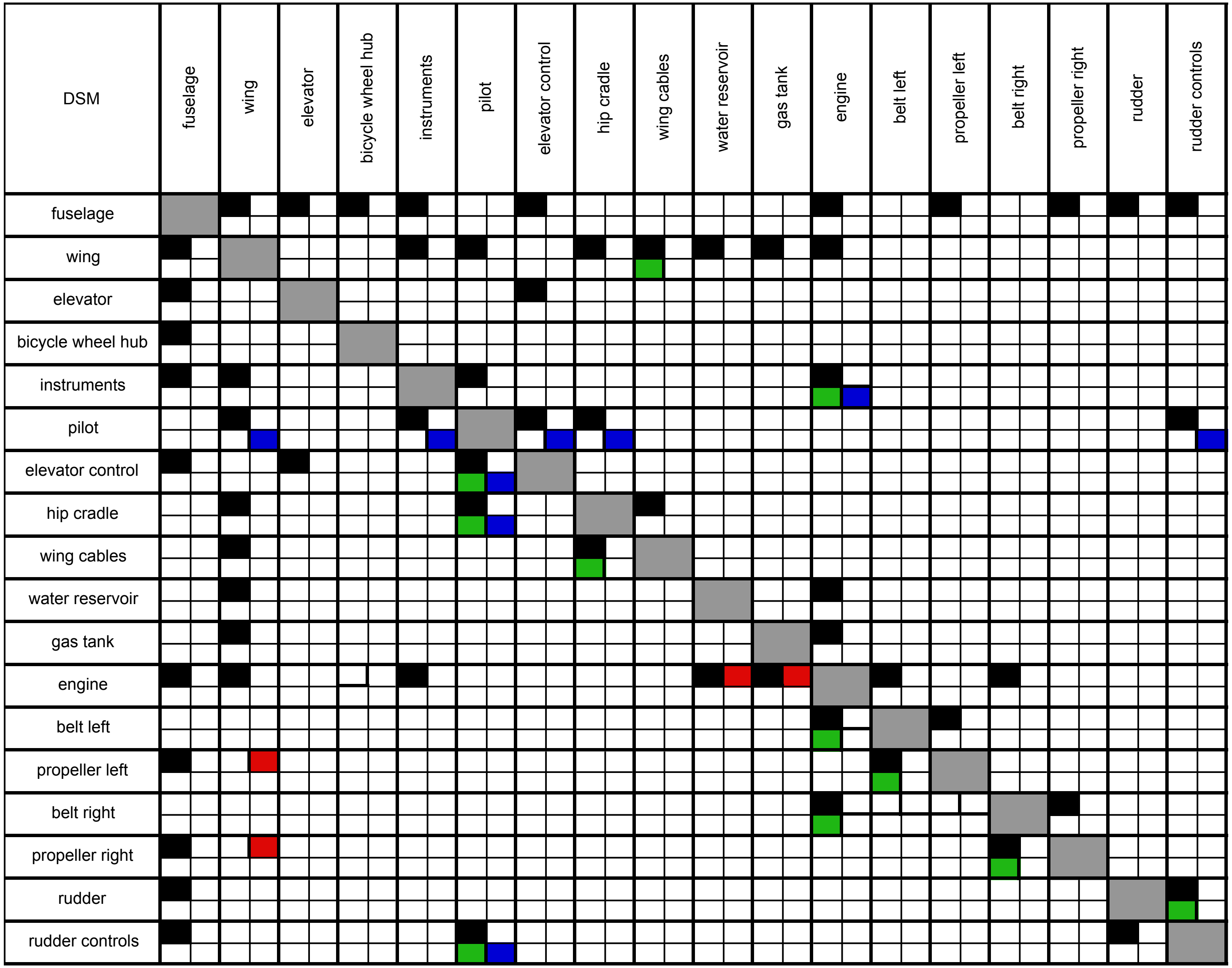
Design requirements growth for aerospace vehicles.

Image by MIT OpenCourseWare.

Two Dimensions of Complexity



Structural DSM of Wright Flyer



Legend
- Physical connection
- Mass flow
- Energy flow
- Information flow

DSM 18x18

Connections

62 Physical

4 Mass Flow

11 Energy Flow

9 Info Flow

Total: 86

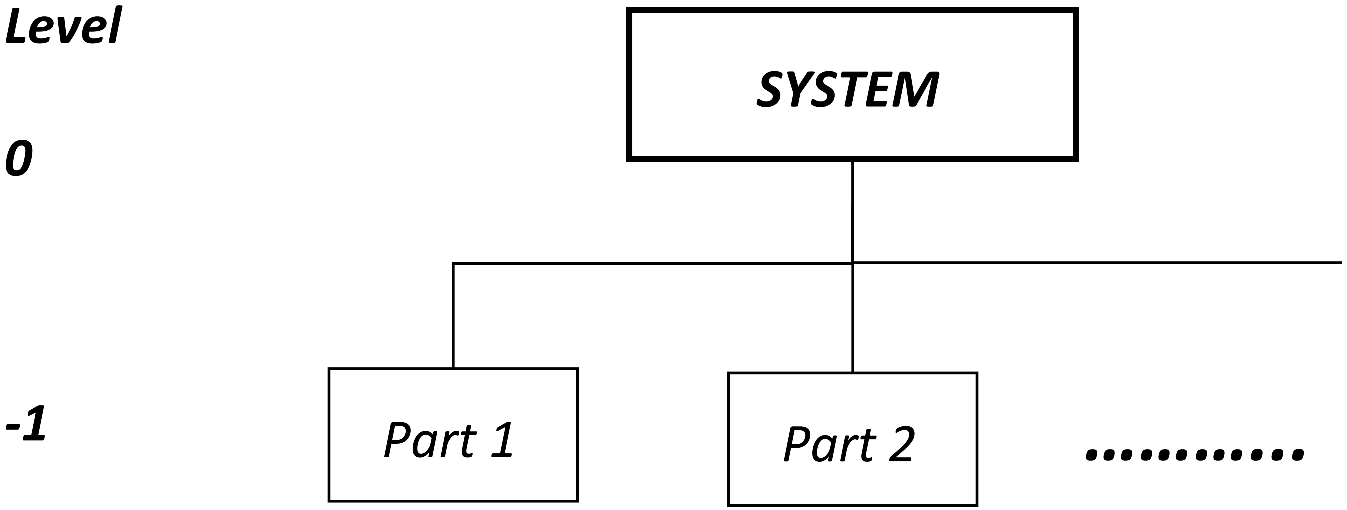
NZF = 86/1,224

= **7% density**

<k>=~5

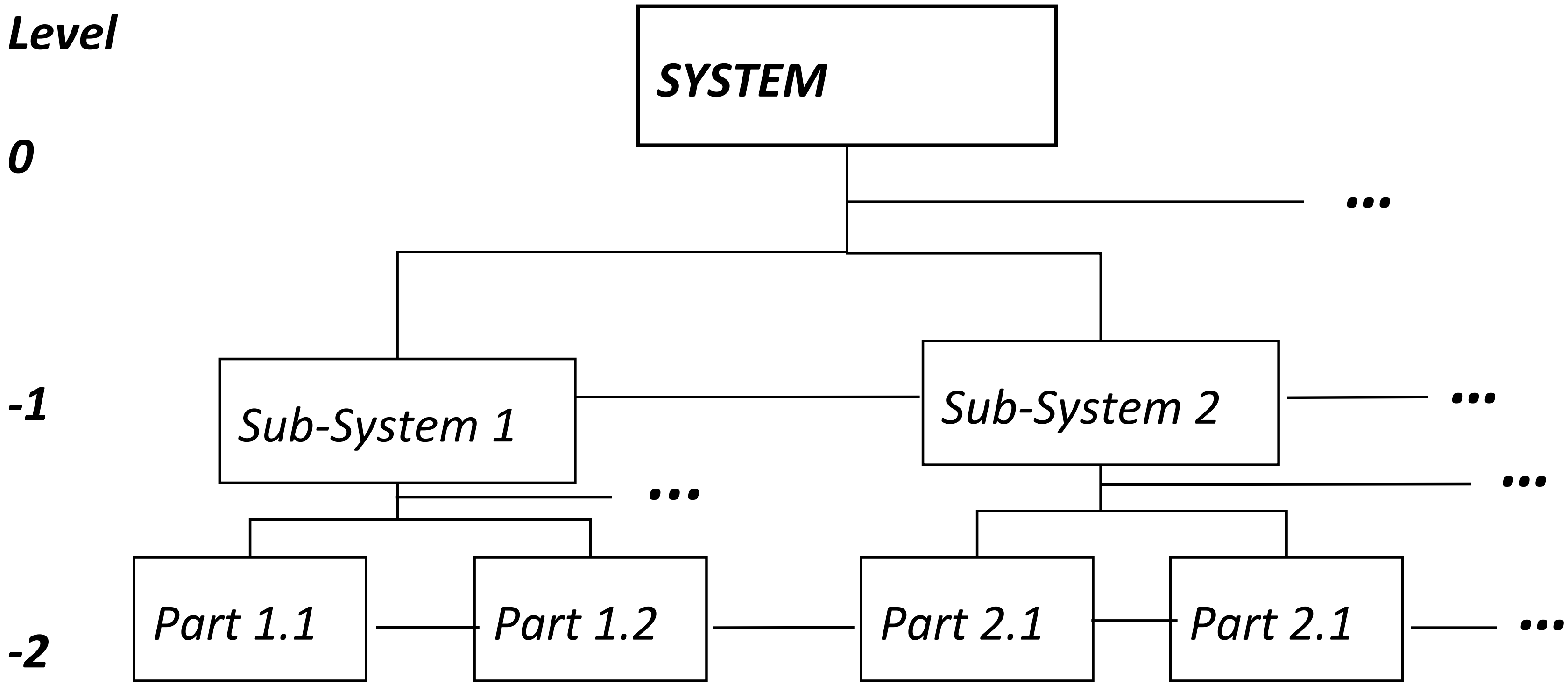
Design Structure Matrix (DSM) – captures structure of elements of form

Form of a Simple System



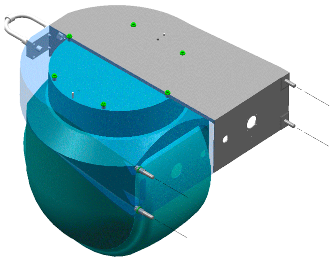
- Generally 5-9 parts (7+/- 2)
- At level -1 we encounter real or **<u>atomic</u>** parts
    - ❑ A part cannot be taken **a-part** without loosing its functionality or integrity
    - ❑ Definition of what is a part is not always unambiguous
- Tree structure is symbolic

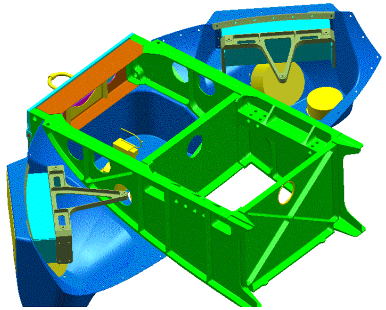
Form of a ‘Medium’ System



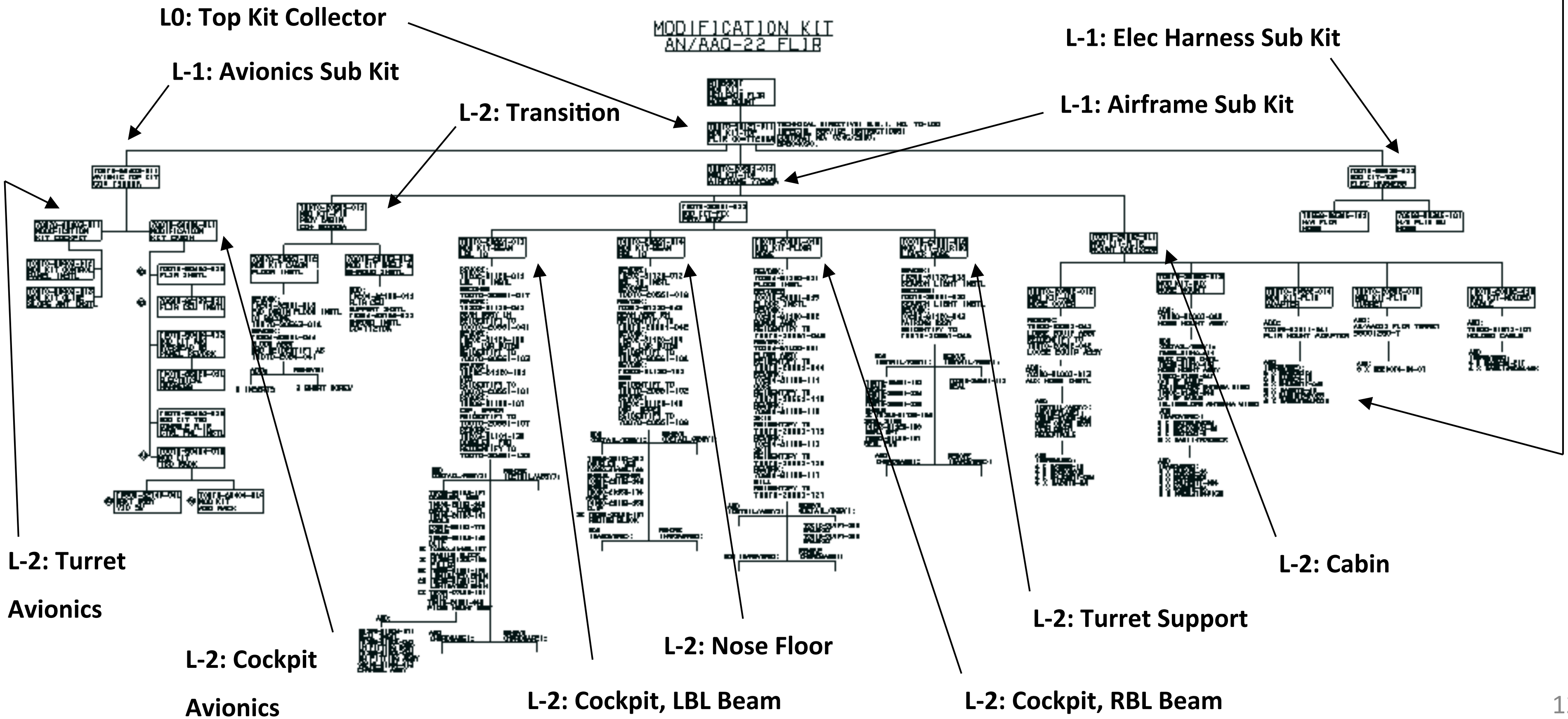
Medium systems typically need 2-3 layers of decomposition

Example : FLIR System Decomposition





L-3: Adds/Removes Hardware & Details



Why do we need system decomposition?

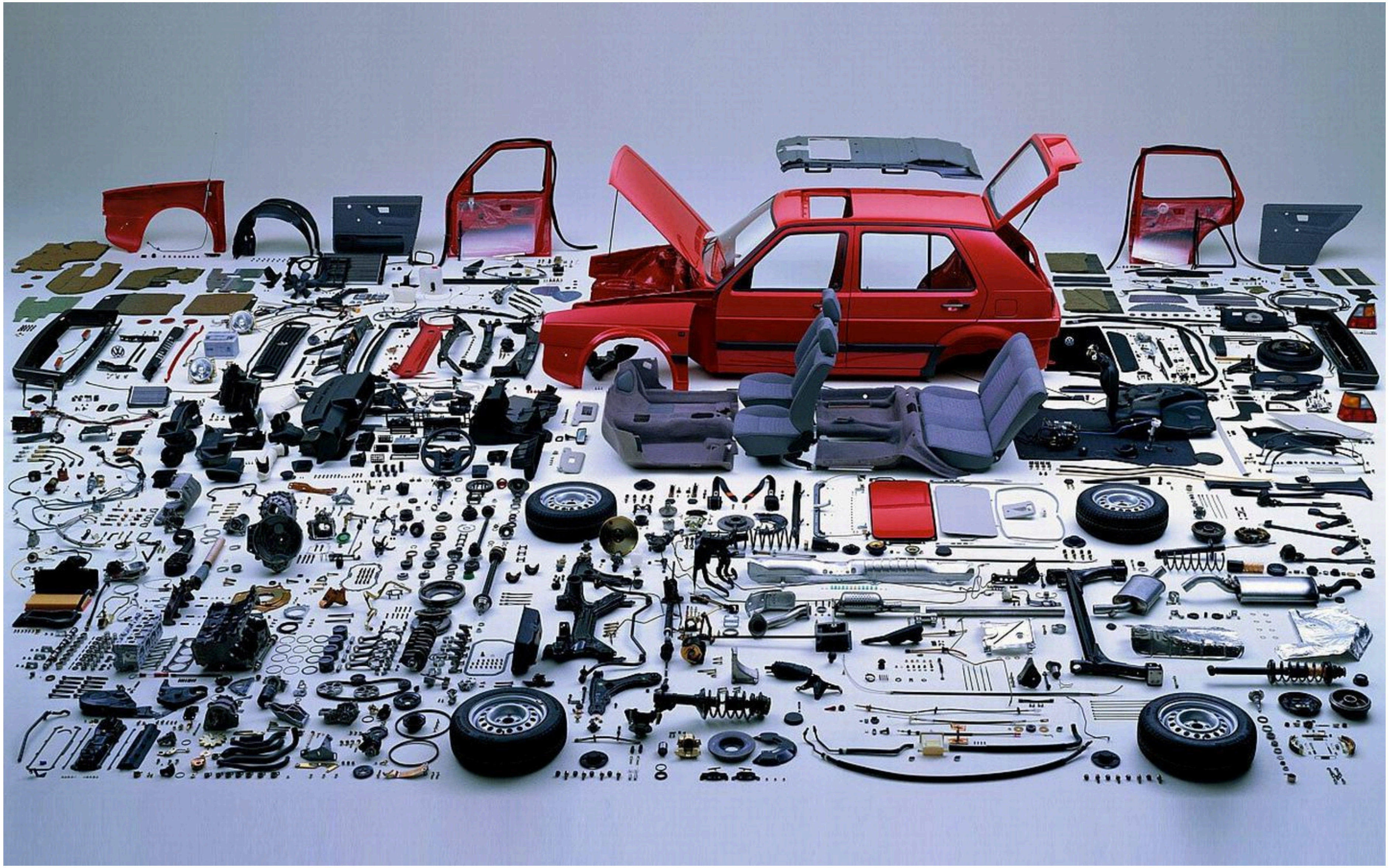


Image Source: http://www.teslamotorsclub.com/showthread.php/29692-How-many-moving-parts-in-a-Model-S/page3

Here is a question for you …

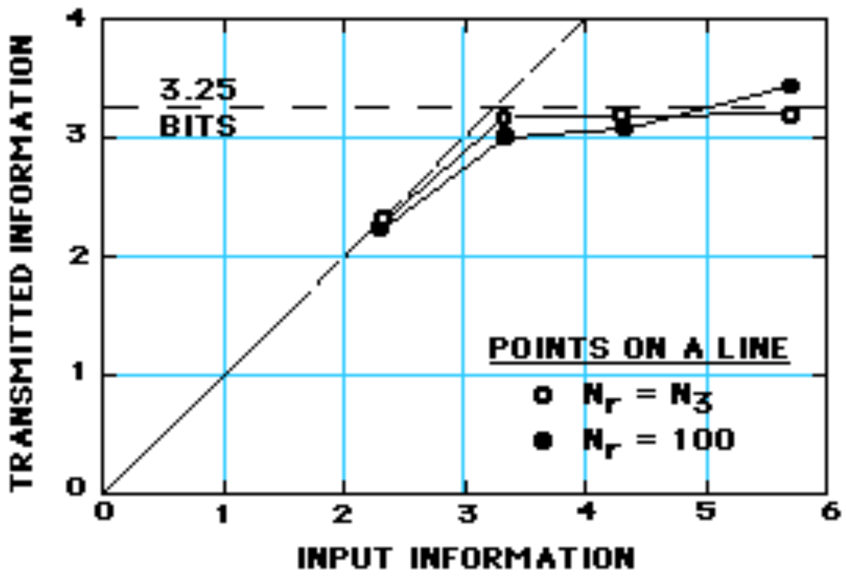
- How many levels of decomposition (depth of drawing tree) do we need to describe the car shown in the previous picture?
  - 1
  - 2
  - 3
  - 4
  - 5
  - 6
  - >6
  - Question is unclear to me

http://tiny.cc/decomp

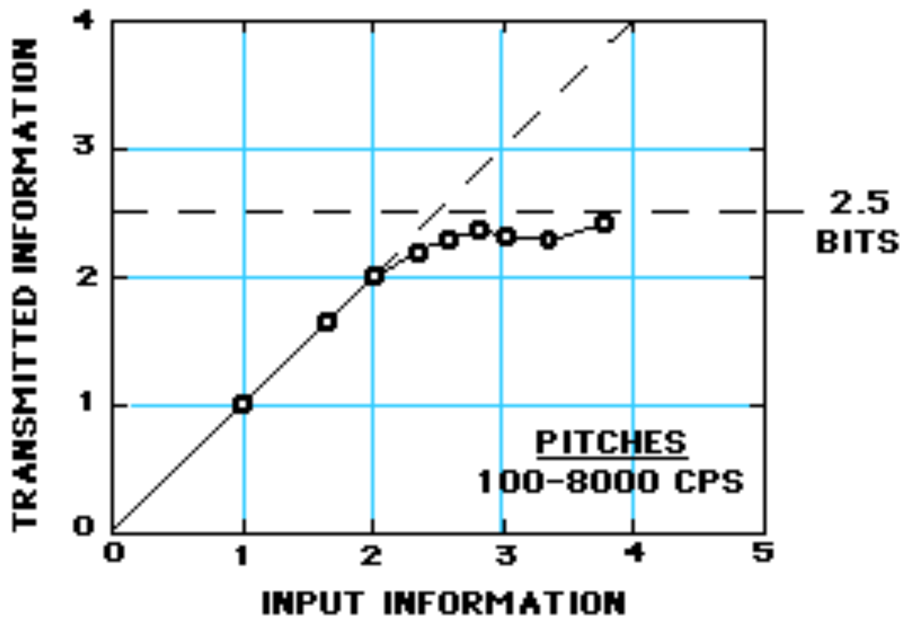
Magic Number 7+/-2

- Human Cognitive Limits for Processing Information
- George Miller (1956)
- http://www.musanim.com/miller1956/

Position of a Pointer
on Linear Interval



Auditory Pitch
Experiments



How many levels of decomposition?

Assume 7-tree

How many levels in drawing tree?

$$\#levels = \left\lceil \frac{\log(\#\,parts)}{\log(7)} \right\rceil$$

	~ #parts	#levels
• Screwdriver (B&D)	3	1
• Roller Blades (Bauer)	30	2
• Inkjet Printer (HP)	300	3
• Copy Machine (Xerox)	2,000	4
• Automobile (GM, VW …)	10,000	5
• Airliner (Boeing)	100,000	6

simple

complex

Source: Ulrich, K.T., Eppinger S.D. , Product Design and Development Second Edition, McGraw Hill, 2nd edition, 2000, Exhibit 1-3

Why should we care about complexity?

How do we quantify complexity?

How to better manage complexity?

Elaine Weyuker's (1998) criteria

Any valid metric for complexity should demonstrate the following broad characteristics (i.e., they act as *necessary conditions* or as *axioms*):

1) Invariant to relabeling (i.e., isomorphism).
2) Possible to have different system architectures have the same complexity level.
3) Differentiate between system architectures.
4) System structure at least partially determines complexity of functionally equivalent systems.
5) Changes in internal architectural patterns , without changes in system size, impact the level of structural complexity.
6) Changing subsystem interfacing patterns impact structural complexity.
7) A system is structurally more complex than the sum of complexities of its constituent subsystems. [whole is larger than the sum of parts]

The Structural Complexity Metric

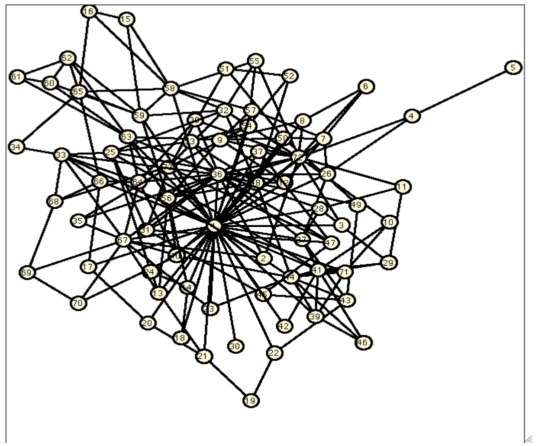
Structural Complexity, $C = C_1 + C_2 . C_3$

[This functional form inspired by the solution of the *steady-state Schrodinger equation* of organic molecular systems [Gutman 1978, 2000].]

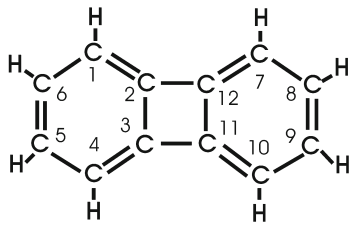
Complexity due to topological formation (a scaling factor) – due to dependency structure

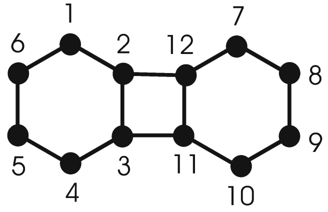
Complexity due to components alone (number and heterogeneity of components)

Complexity due to pair-wise component interactions (number and heterogeneity of interactions)



System Hamiltonian and Complexity





$$[\mathbf{H}]_{ij} = \begin{cases} \alpha & \text{if } i = j \\ \beta & \text{if the atoms } i \text{ and } j \text{ are chemically bonded} \\ 0 & \text{if there is no chemical bond between the atoms } i \text{ and } j. \end{cases}$$

$$\mathbf{H} = \alpha \begin{bmatrix} 1&0&0&0&0&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0&0&0&0&0 \\ 0&0&0&0&1&0&0&0&0&0&0&0 \\ 0&0&0&0&0&1&0&0&0&0&0&0 \\ 0&0&0&0&0&0&1&0&0&0&0&0 \\ 0&0&0&0&0&0&0&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&0&0&0&0&1 \end{bmatrix} + \beta \begin{bmatrix} 0&1&0&0&0&1&0&0&0&0&0&0 \\ 1&0&1&0&0&0&0&0&0&0&0&1 \\ 0&1&0&1&0&0&0&0&0&0&1&0 \\ 0&0&1&0&1&0&0&0&0&0&0&0 \\ 0&0&0&1&0&1&0&0&0&0&0&0 \\ 1&0&0&0&1&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&1&0&0&0&1 \\ 0&0&0&0&0&0&1&0&1&0&0&0 \\ 0&0&0&0&0&0&0&1&0&1&0&0 \\ 0&0&0&0&0&0&0&0&1&0&1&0 \\ 0&0&1&0&0&0&0&0&0&1&0&1 \\ 0&1&0&0&0&0&1&0&0&0&1&0 \end{bmatrix}.$$

$$H = \alpha I_n + \beta A(G)$$

$$H\psi = \varepsilon\psi$$

$$|\varepsilon_i| = \alpha + \beta\sigma_i; \; \varepsilon_\pi = \sum_{i=1}^{n} h_i |\varepsilon_i|$$

$$\varepsilon_\pi = n\alpha + \beta \sum_{i=1}^{n} h_i \sigma_i \le n\alpha + \beta \underbrace{\left(\sum_{i=1}^{n} h_i\right)}_{n} \underbrace{\left(\sum_{i=1}^{n} \sigma_i\right)}_{E(A)}$$

$$\therefore \varepsilon_\pi \le n\alpha + n^2\beta\left(\frac{E(A)}{n}\right)$$

Introduce a notion of of *configuration energy*:

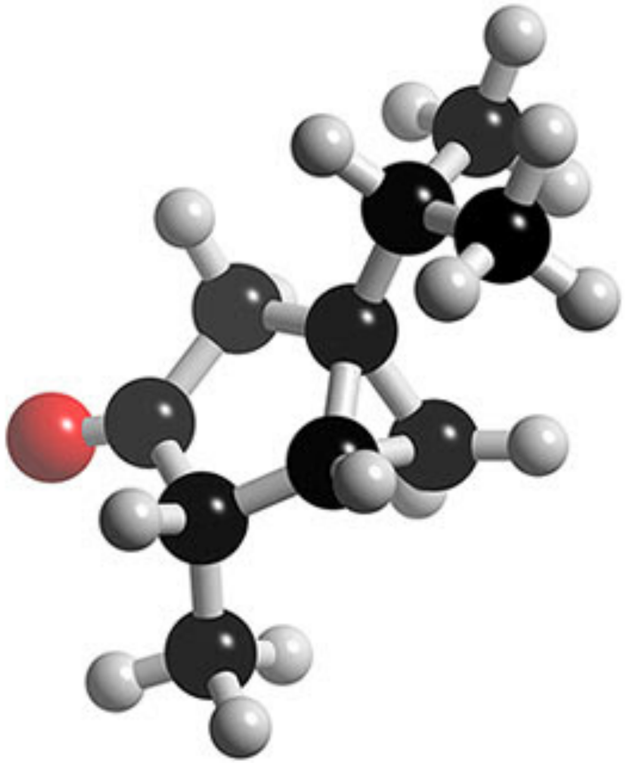
$$\Xi := \underbrace{n\hat{\alpha}}_{C_1} + \underbrace{m\hat{\beta}}_{C_2} \underbrace{\left(\frac{E(A)}{n}\right)}_{C_3} = C_1 + C_2 C_3$$

Use the above functional form to measure the complexity associated to the system structure – ***Structural Complexity*** of the system where α's stand for component complexity while β's stand for interface complexity:

$$C = C_1 + C_2 C_3$$

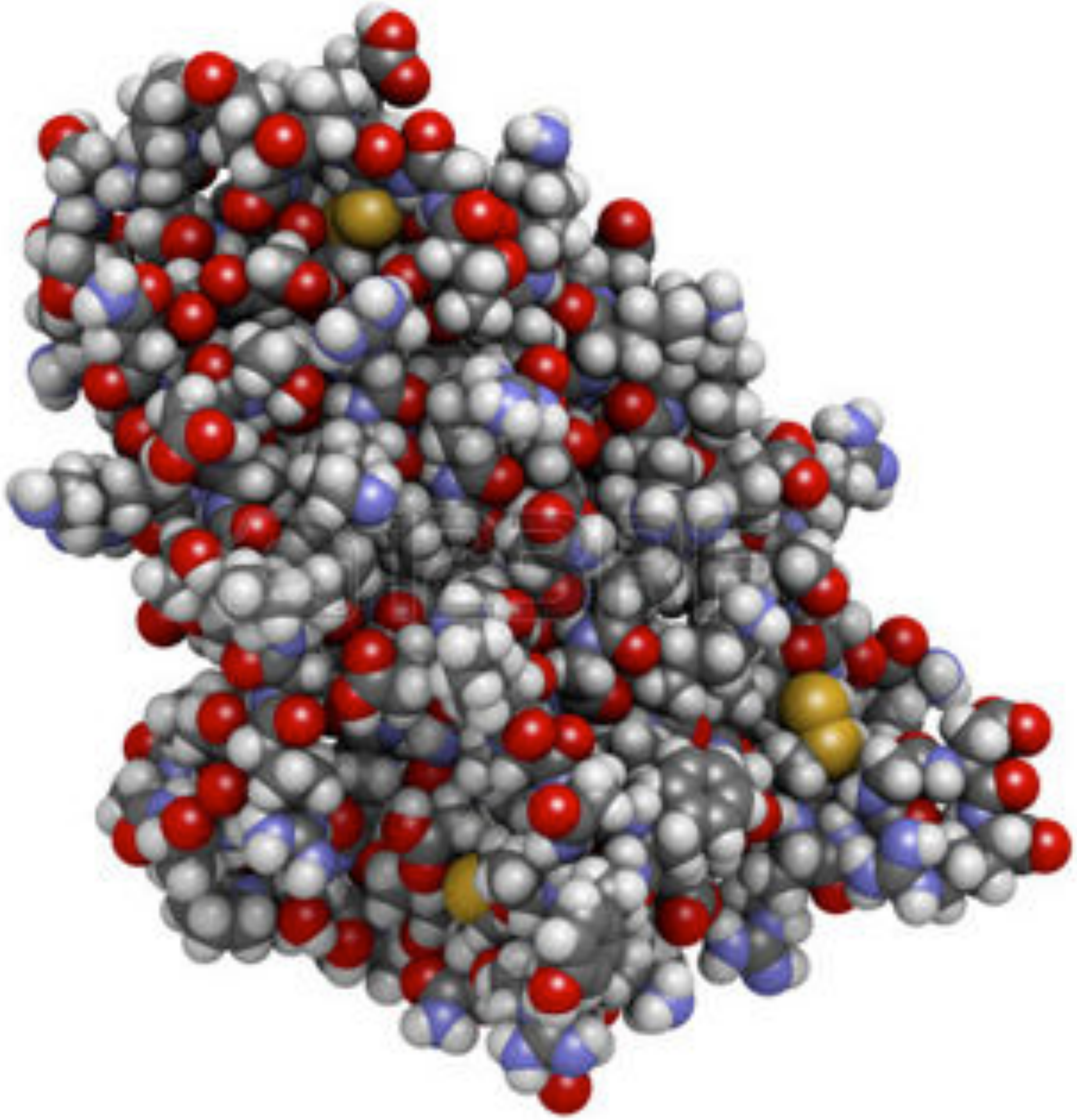
$$= \sum_{i=1}^{n} \alpha_i + \left(\sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij}\right)\left(\frac{E(A)}{n}\right) = \sum_{i=1}^{n} \alpha_i + \left(\sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij}\right) \gamma E(A)$$

Simple Molecule



Molecule #10

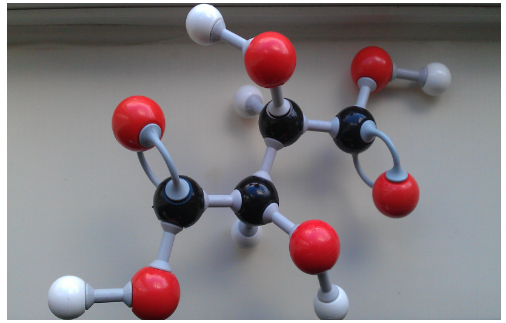
Complex Molecule



https://en.wikipedia.org/wiki/Erythropoietin

Validation using Human Experiments

- Empirical validation of the structural complexity metric
  - Recruited volunteer test subjects.
  - Provided: (a) ball and stick chemistry toolkit;
    (b) a set of pictures of molecules to be built.
  - Task: Assemble the depicted architecture.



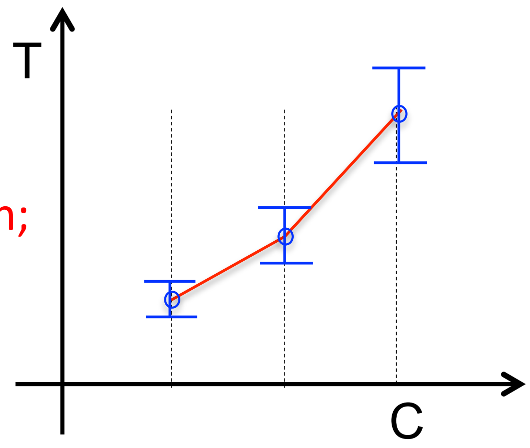
- Record for each model (for each individual)
  - **C** = computed structural complexity
  - **T** = [time to build, including rework if any]

\# Molcules: 12
\# Subjects: 17

Hypothesis:
High Structural Complexity leads to measurably …

→ **Slower Progress** (Cognition; Schedule)

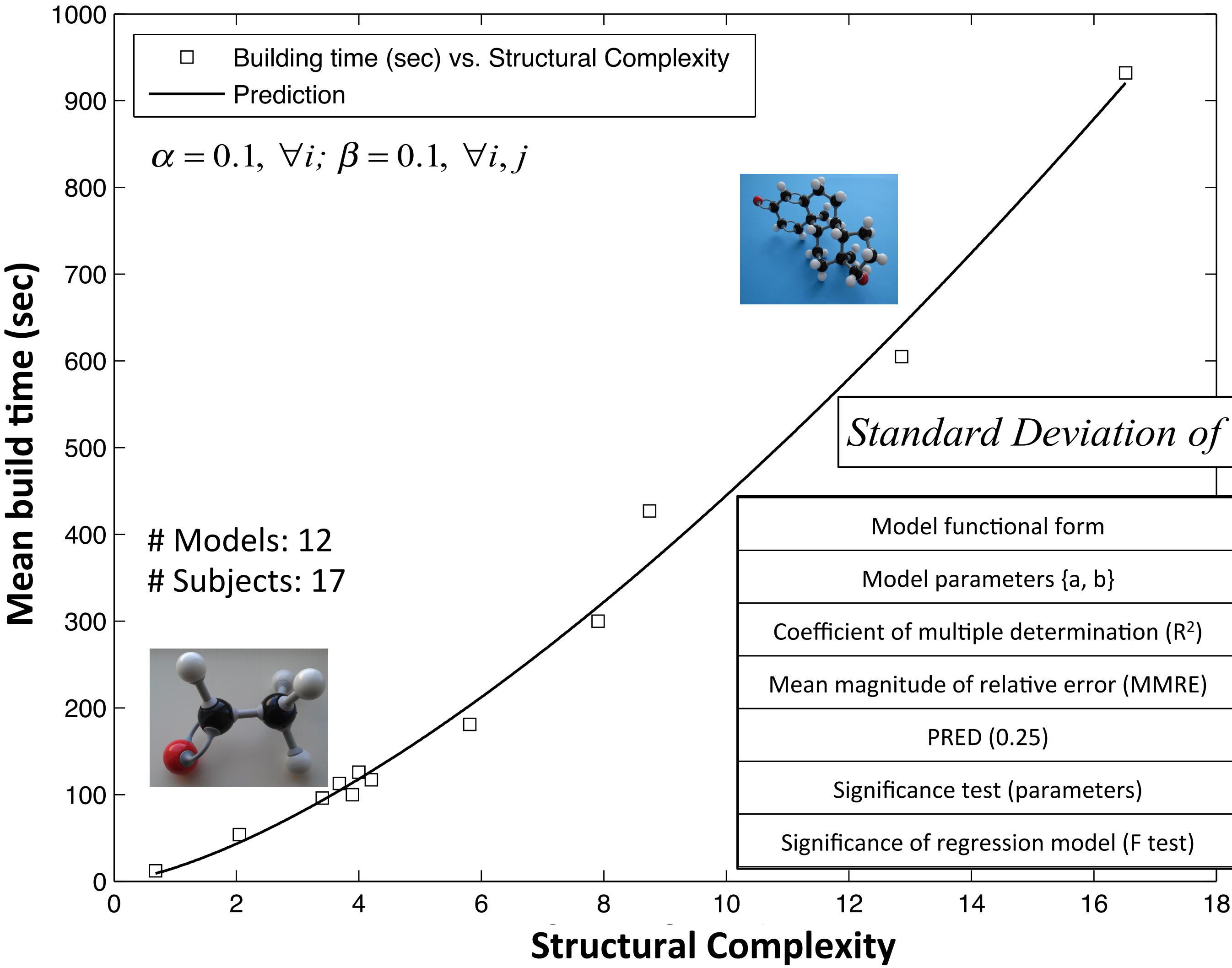
→ **Higher Error-rate** (Rework)



Experimental Design (12 molecules)

Molecule No.	n	m	C1	C2	C3= E(A)/n	C2*C3	SC = C1 + C2*C3
1	3	4	0.3	0.4	0.94	0.38	0.68
2	7	12	0.7	1.2	1.13	1.35	2.05
3	12	22	1.2	2.2	1.13	2.48	3.68
4	12	22	1.2	2.2	1.00	2.20	3.40
5	12	22	1.2	2.2	1.27	2.80	4.00
6	14	26	1.4	2.6	0.96	2.50	3.90
7	15	28	1.5	2.8	0.97	2.70	4.20
8	16	30	1.6	3	1.40	4.21	5.81
9	19	38	1.9	3.8	1.58	6.00	7.90
10	27	56	2.7	5.6	1.08	6.05	8.75
11	39	80	3.9	8	1.12	8.96	12.86
12	46	100	4.6	10	1.19	11.92	16.52

Experimental Results are super-linear



$$Standard\ Deviation\ of\ build\ time,\ Y_{sd} = 8.48 X_*^{1.3}$$

Model functional form	$Y = aX^b$
Model parameters {a, b}	{ 14.68, 1.4775}
Coefficient of multiple determination (R^2)	0.992
Mean magnitude of relative error (MMRE)	0.107
PRED (0.25)	0.9167
Significance test (parameters)	$t_a = 28.2$, $t_b = 30.67$ ($>t_0= 2.131$)
Significance of regression model (F test)	$f = 124 > f_{0.05,1,10} = 4.54$

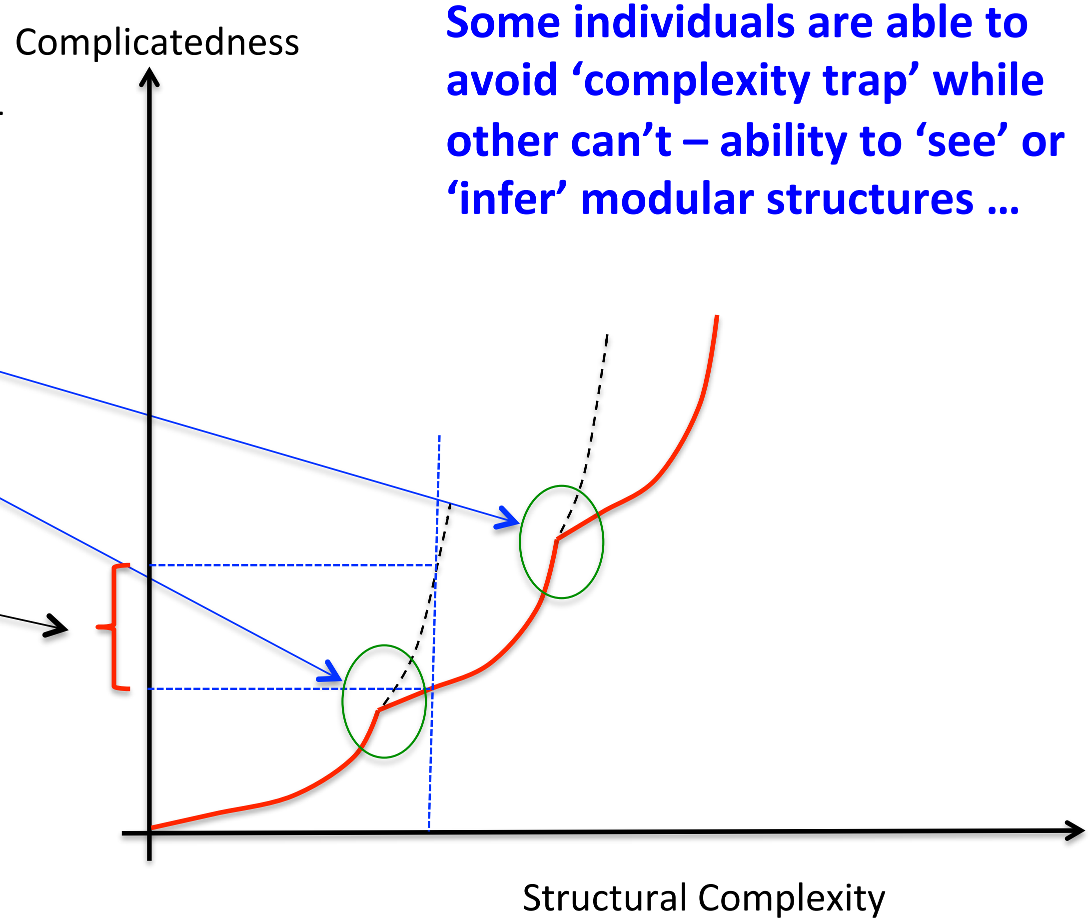
$$Structural\ Complexity,\ C = O(n^{1.08}) \leftarrow mild\ super\text{-}linearity$$

$$Average\ build\ time,\ t = O(C^{1.48}) \leftarrow strong\ super\text{-}linearity$$

Empirical Observation about Modularity

- Avoid '*complexity trap*' by understanding higher level patterns - individual cognitive ability!

- Significant reduction in *perceptive complexity* or *complicatedness* -



Some individuals are able to avoid 'complexity trap' while other can't – ability to 'see' or 'infer' modular structures …

Structural Complexity Metric

IDSS
MIT INSTITUTE FOR DATA, SYSTEMS, AND SOCIETY

$$C = C_1 + C_2 . C_3$$

$$C(n,m,A) = \underbrace{\boxed{\sum_{i=1}^{n} \alpha_i}}_{\text{Components}} + \underbrace{\left(\sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} A_{ij} \right)}_{\text{Interfaces}} \underbrace{\gamma E(A)}_{\text{Architecture}}$$

Normalization factor (1/n)

$$E(A) = \sum_{i=1}^{n} \sigma_i$$

Graph energy

Binary (0/1) conn. matrix

# of components

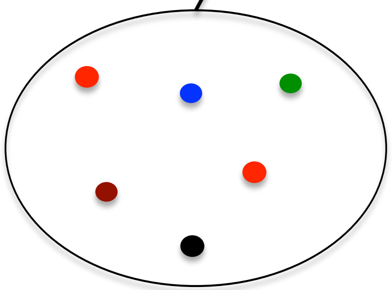
#interfaces

DSM

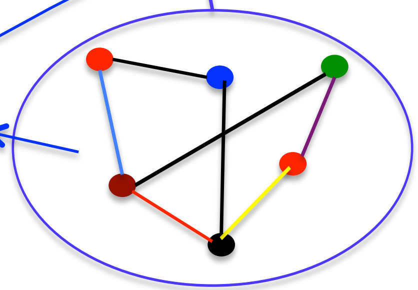
Related to component engineering

Related to interface design and mgmt.

Related to System Integration Effort
(topological complexity)



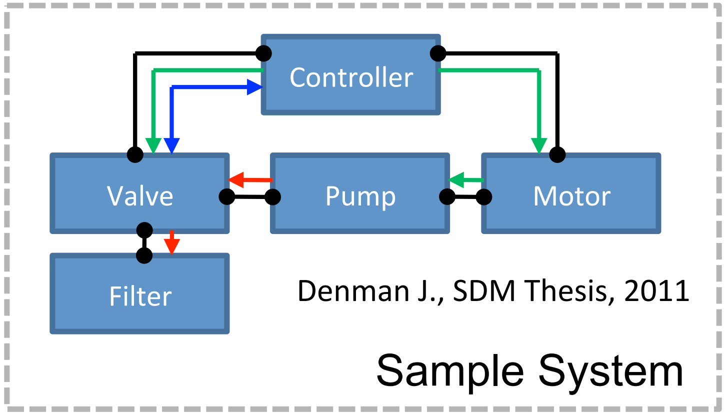
$$A = \begin{bmatrix} 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$



$$A = U \Sigma V^T$$

$$\Rightarrow A = \sum_{i=1}^{n} \sigma_i \underbrace{u_i v_i^T}_{R_i}$$

Example: Cyber-Physical System

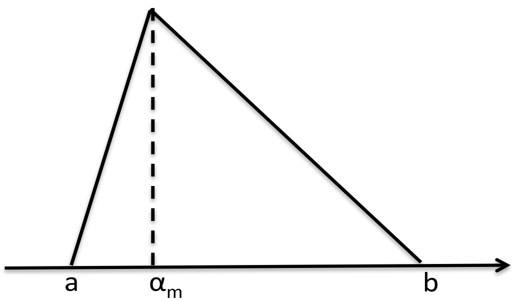


aggregation

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Component	ID	Complexity
Controller	1	1.5
Pump	2	1.0
Valve	3	0.3
Filter	4	0.3
Motor	5	1.2

Comp. 1	Comp. 2	$1/c^{(k)}$
1	3	0.05
1	3	0.10
1	3	0.15
1	5	0.05
1	5	0.10
2	3	0.05
2	3	0.10
		0.05
		0.15
		0.05
		0.10



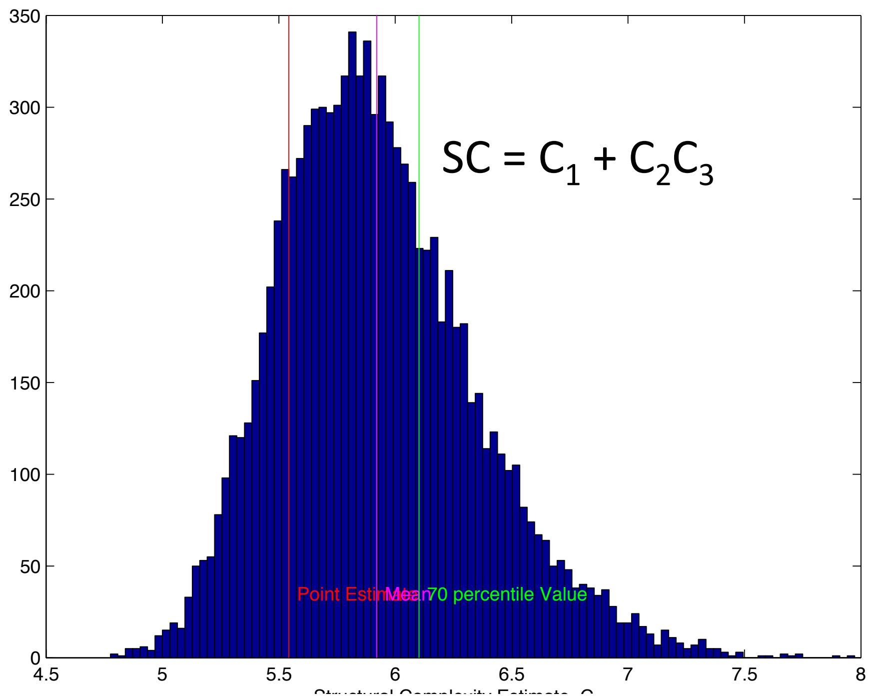
$(b-\alpha_m) = p(\alpha_m-a)$

$$p \in [1.0;3.0]$$
$$a \in [0.8\alpha_m;0.9\alpha_m]$$
$$b \in [1.1\alpha_m;1.6\alpha_m]$$

$$\beta_{ij}^{(k)} = g(\alpha_i,\alpha_j,c^{(k)})$$

$$\beta_{ij}^{(k)} = \frac{\max(\alpha_i,\alpha_j)}{c^{(k)}},$$

$$\forall \alpha_i,\alpha_j \neq 0,\ k \text{ is the interface type}$$



Structural Complexity Estimate, C

Construct Validity: Weyuker's Criteria

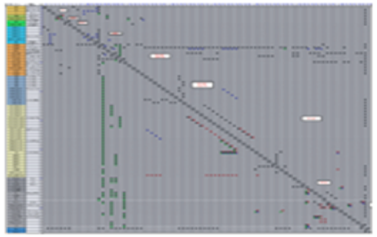
- Graph Energy stands out as both computable and satisfies Weyuker's criteria and establishes itself as a theoretically valid measure (i.e., construct validity) of complexity.

Complexity Measure	Computability	Aspect emphasized	Weyuker's Criteria
Number of components [Bralla, 1986]	✔	Component development (count-based measure)	✗
Number of interactions [Pahl and Beitz, 1996]	✔	Interface development (count-based measure)	✗
Whitney Index [Whitney *et al.*, 1999]	✔	Components and interface developments	✗
Number of loops, and their distribution []	✗	Feedback effects	✗
Nesting depth [Kerimeyer and Lindemann, 2011]	✗	Extent of hierarchy	✗
Graph Planarity [Kortler *et al.*, 2009]	✔	Information transfer efficiency	✗
CoBRA Complexity Index [Bearden, 2000]	✔	Empirical correlation in similar systems	✗
Automorphism-based Entropic Measures [Dehmer *et al.*, 2009]	✗	Heterogeneity of network structure, graph reconfigurability	✔
Matrix Energy / Graph Energy	✔	Graph Reconstructabality	✔

Complexity should be abstraction-Invariant

Digital Printing Press (Xerox) Example

Size: 50x50



Size: 91x91

DSM attribute	Coarse Representation	Finer representation
System size, N	50	91
C_3	1.3534	1.3597

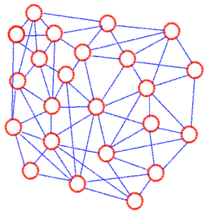
Functional Area	Coarse DSM (50x50)	Fine DSM (91x91)
ROS Assembly	4	10
Marking elements	16	38
Paper Path	7	12

Why should we care about complexity?

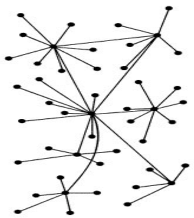
How do we quantify complexity?

How to better manage complexity?

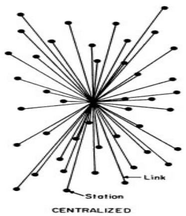
Topological Complexity: Important Properties



"Distributed" Architecture



"Hierarchical" Architecture



Centralized architecture

Simple components / constituents / building blocks with intricate connectivity structure

Higher system integration effort

Increasing Topological Complexity (C_3)

Complex components / constituents / building blocks with simple connectivity structure

Lower system integration effort

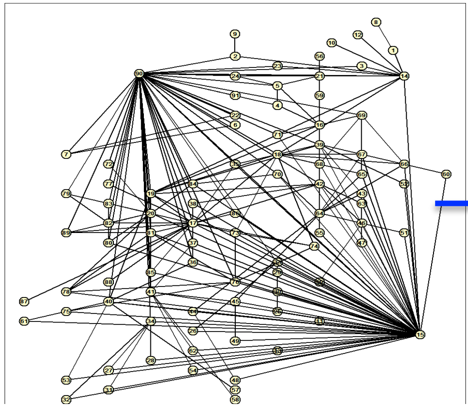
$$Centralized\ Architecture \rightarrow hypoenergetic,\ C_3 < 1$$

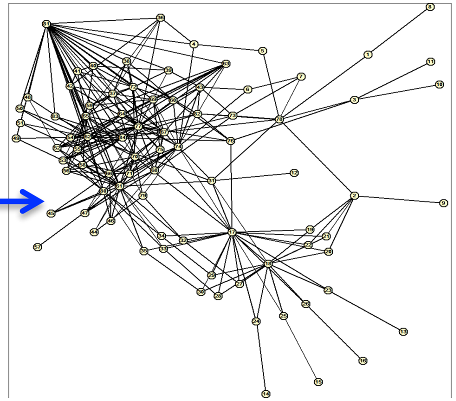
$$Hierarchical\ /\ layered\ Architecture \rightarrow transitional,\ 1 \leq C_3 < 2$$

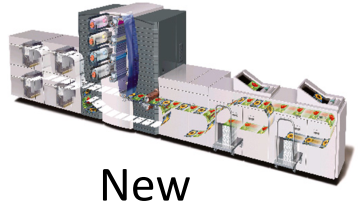
$$Distributed\ Architecture \rightarrow hyperenergetic,\ C_3 \geq 2$$

Case Study 1: Printing Engines

Old







New

Complexity = 186

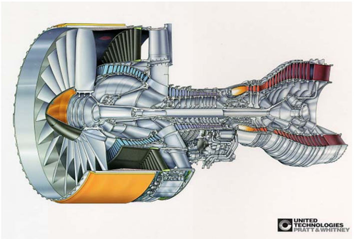
Complexity = 354

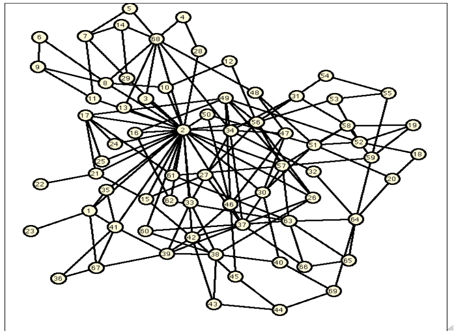
Complexity increase +90%

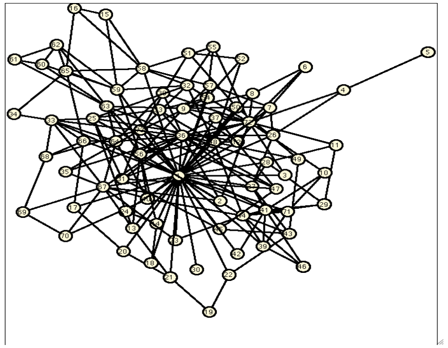
	C_1		C_2		C_3		C		C_{New} $/C_{Old}$
	Old	**New**	**Old**	**New**	**Old**	**New**	**Old**	**New**	
Most Likely	110.2	169	55.68	102.78	1.36	1.804	185.93	354.42	1.9062
Mean	125.62	213.6	63.29	130.6	1.36	1.804	211.69	449.2	2.122
Median	124.47	211.84	62.46	128.62	1.36	1.804	209.42	443.88	2.12
70 percentile	127	219	65.82	134.2	1.36	1.804	216.2	461.1	2.133

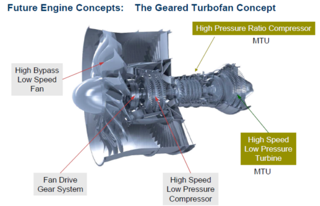
- Trend towards more distributed architecture with higher structural complexity and significantly higher development cost*

Case Study 2: Aircraft Engines









Old

New

Complexity = 351

Complexity = 499

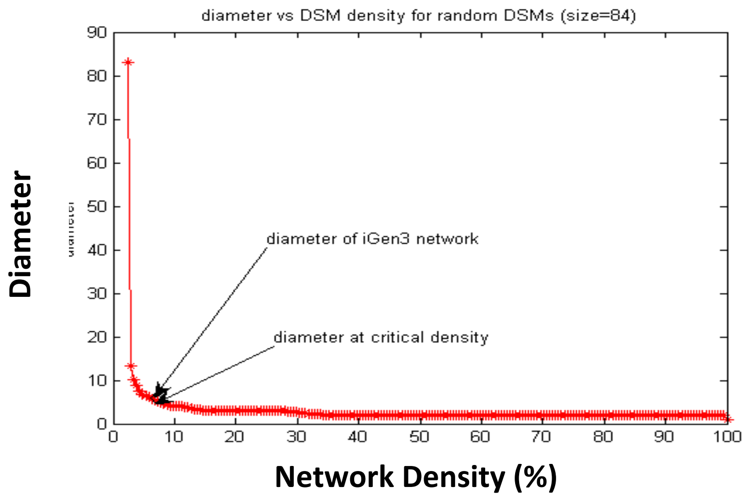
Complexity increase +42%

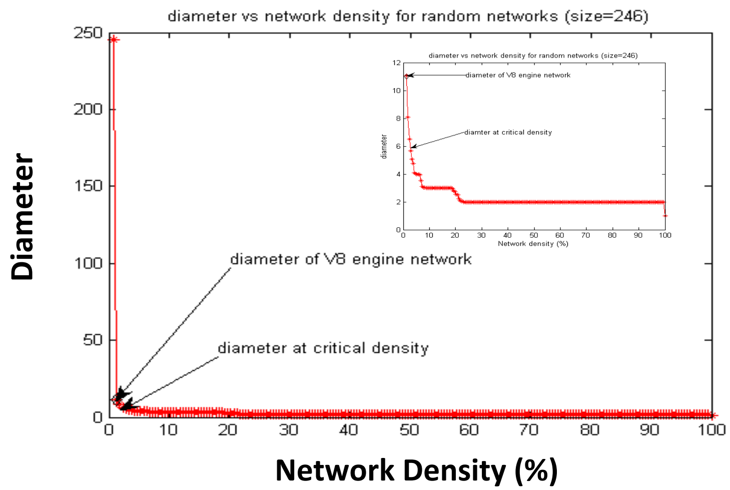
	C_1		C_2		C_3		C		C/C_{ML}		C_{new}/C_{old}
	Old	**New**	**Old**	**New**	**Old**	**New**	**Old**	**New**	**Old**	**New**	
Most Likely	161	188	126	184	1.51	1.69	351	499	1	1	1.42
Mean	179	244	141	240.4	1.51	1.69	392	650.3	1.12	1.30	1.65
Median	178	242	139	238.9	1.51	1.69	388	646.8	1.10	1.29	1.66
70 percentile	181	247.9	145	246.2	1.51	1.69	399.6	663.94	1.14	1.33	1.66

- Trend towards more distributed architecture with higher structural complexity and significantly higher development cost*. Similar trend was observed in Printing Systems.

P point – complexity phase transition

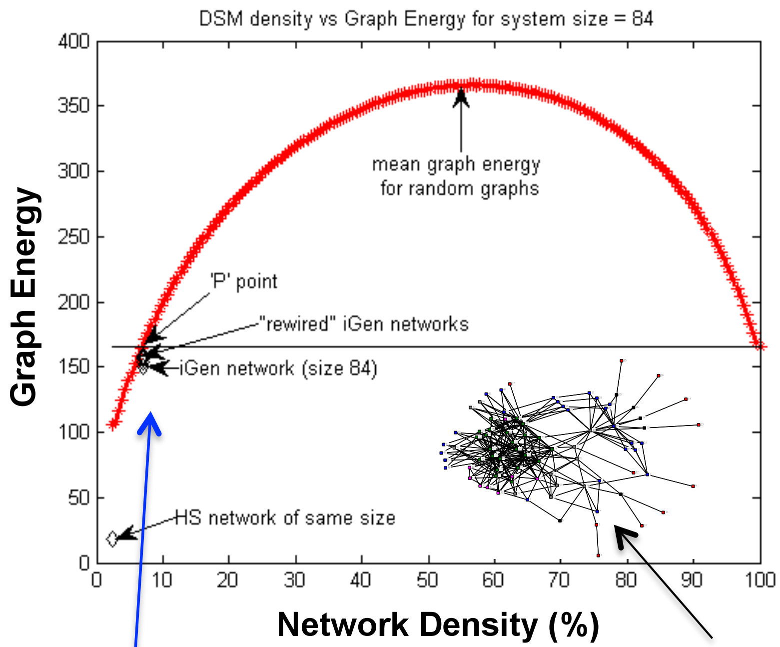
- The ***P* point** on graph energy – density plot: Phase transition for complxity
- At densities higher than ***P*** point, structural complexity increases but that does not buy much improvement in terms of performance measures (e.g., network diameter)

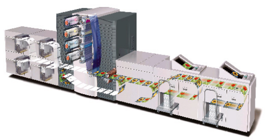




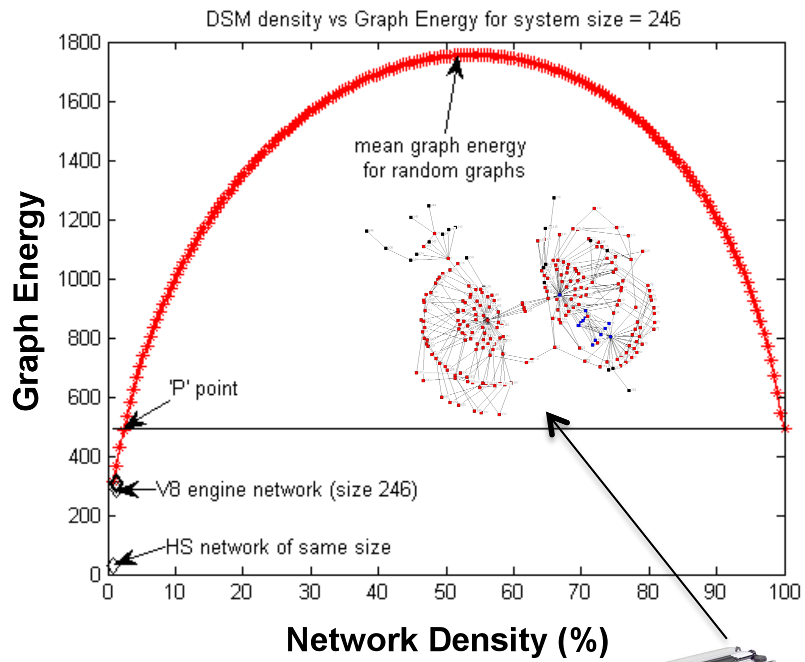
- Use equivalent random networks (Erdős–Rényi) as background.
- P-point has E(A) equivalent to fully connected system, and architectures become rank-dense beyond this point (critical for design).

Real Product Design and P-Point Complexity

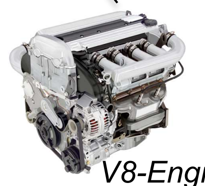




P-point is critical, because
here DSM reaches full rank



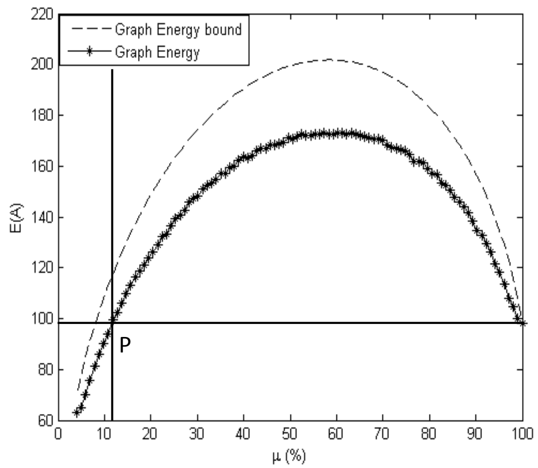
iGen3 (digital printing system)
Xerox

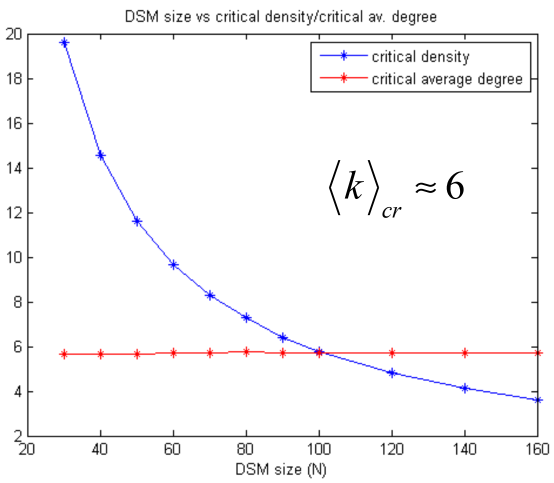


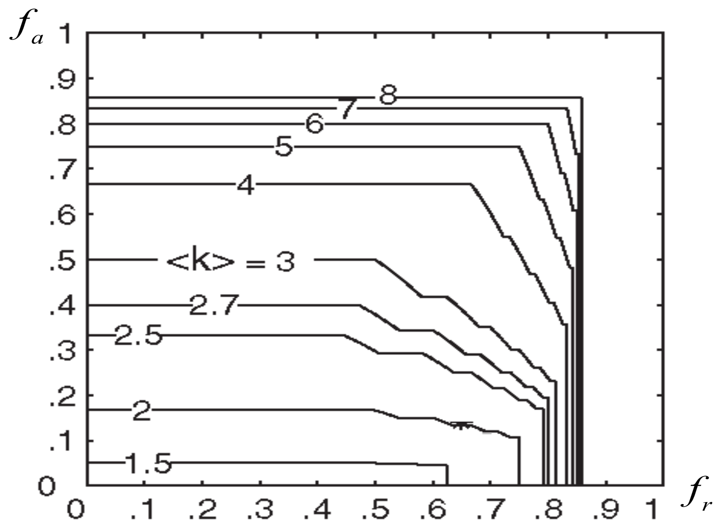
V8-Engine

- Can compare systems at same level of abstraction in this space
- Use equivalent random networks (Erdős–Rényi) as background (red curve)
- P-point has E(A) equivalent to fully connected system, critical for design
- If we go beyond the P-point in System Design will have diminishing returns

Critical Nodal Degree <k>$_{cr}$ =~ 6

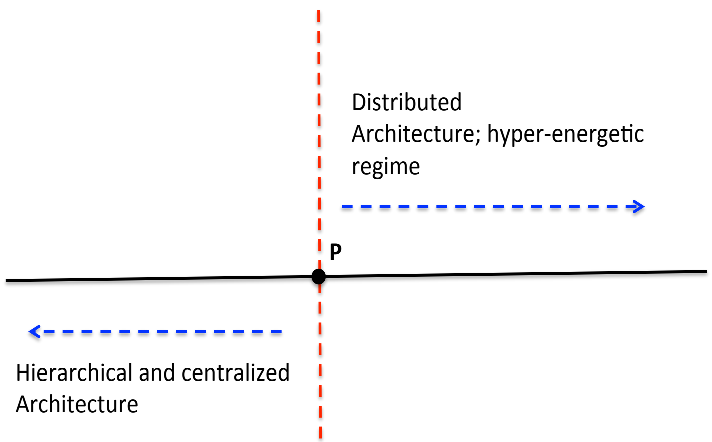


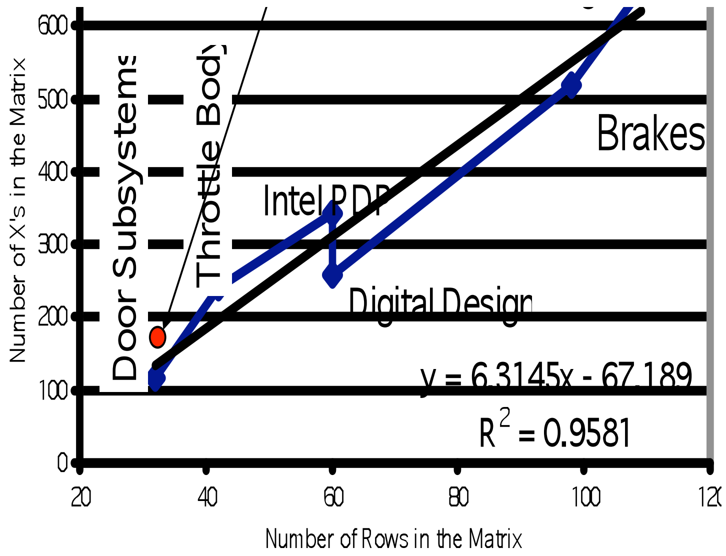




Network resilience contour (f_r vs. f_a)
[Valente et al., 2004]

$$\mu_{cr} \geq \frac{4}{n};\ \langle k \rangle_{cr} \geq 4(1-\frac{1}{n}) \text{ and } m_{cr} \geq 2(n-1)$$

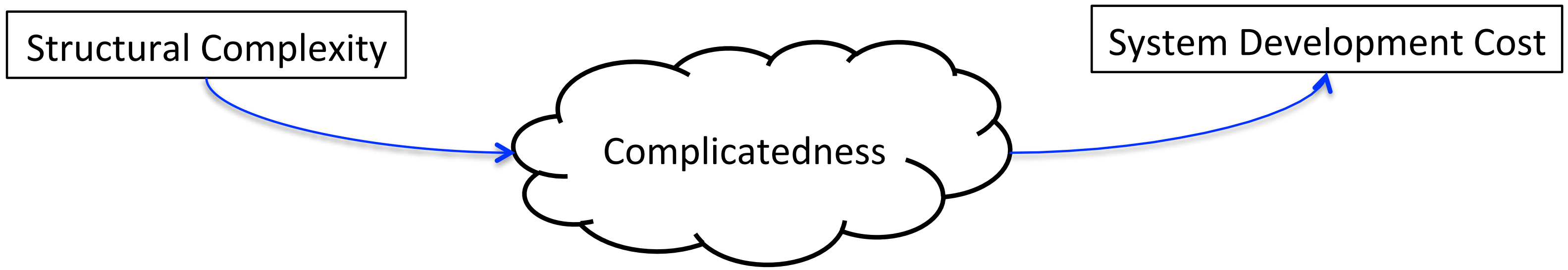




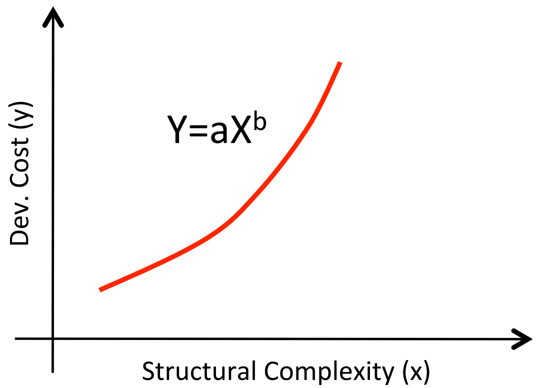
Use of *P* point as a system architecting guideline – entering regime of diminishing returns

[Whitney et al., 1999]

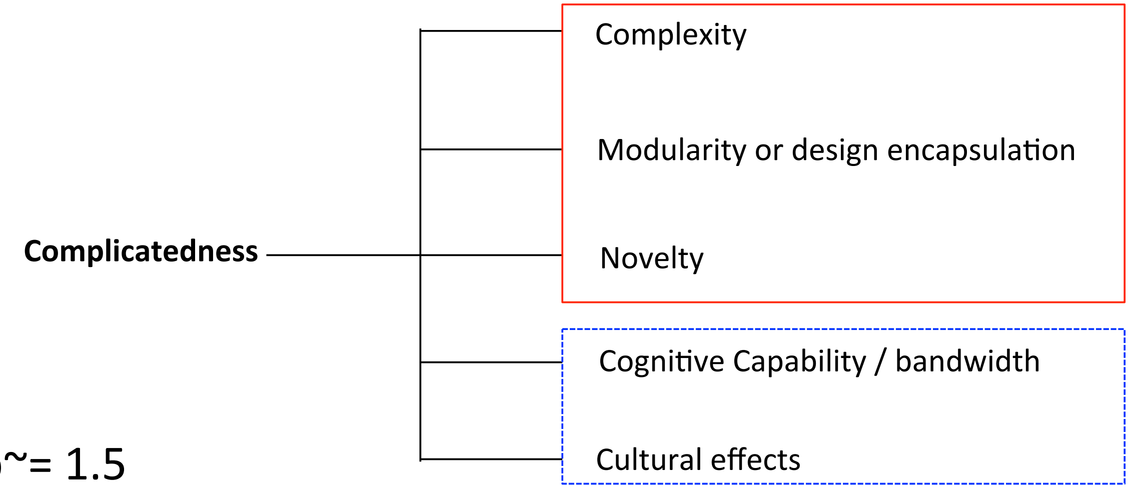
Complicatedness vs. Complexity



- Complicatedness, $b = g(complexity, modularity, novelty, cognitive\ bandwidth,\ \ldots)$



Experiments suggested b~= 1.5

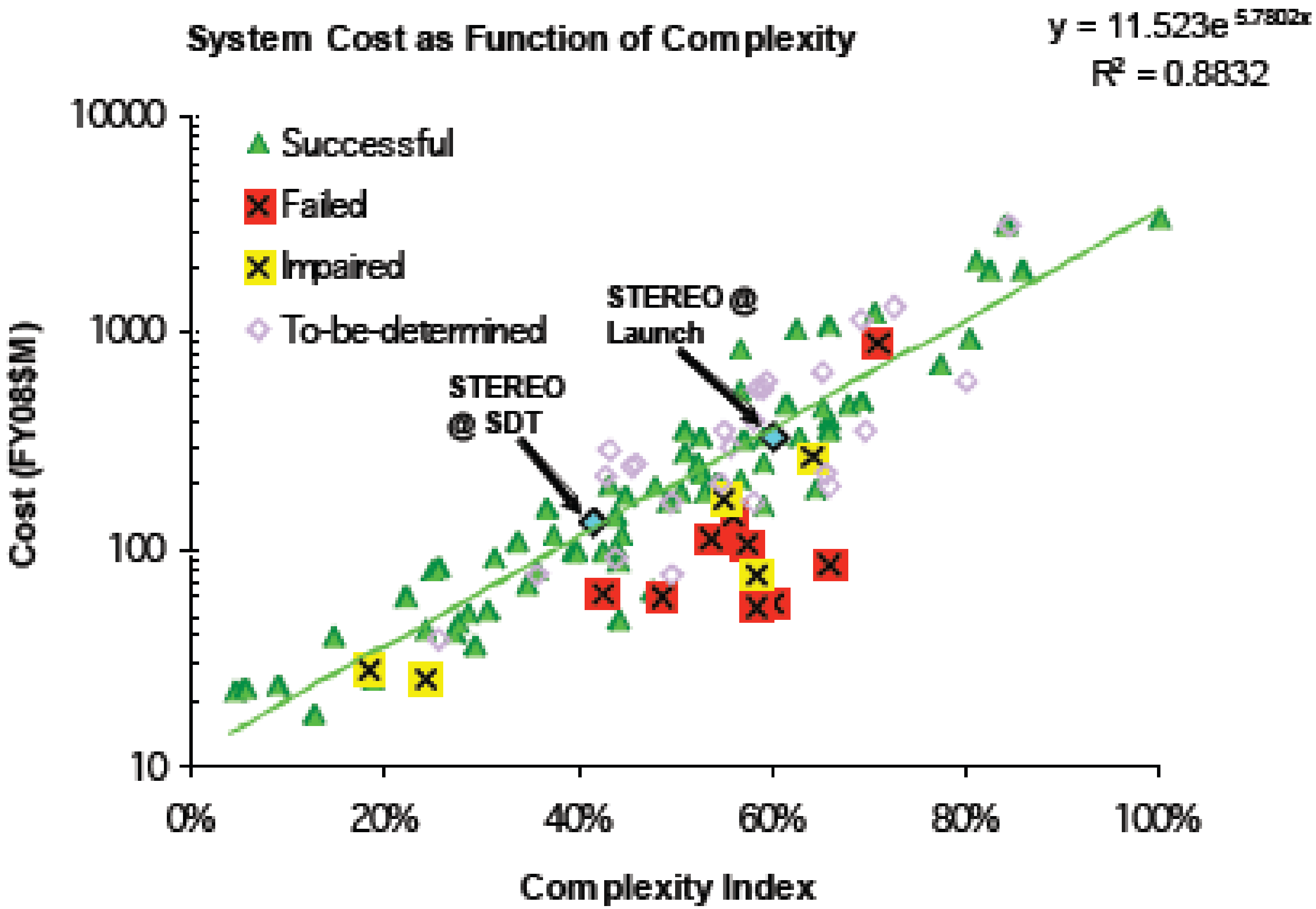


Implication: A 42% increase in complexity
Will lead to a 69% increase in R&D cost

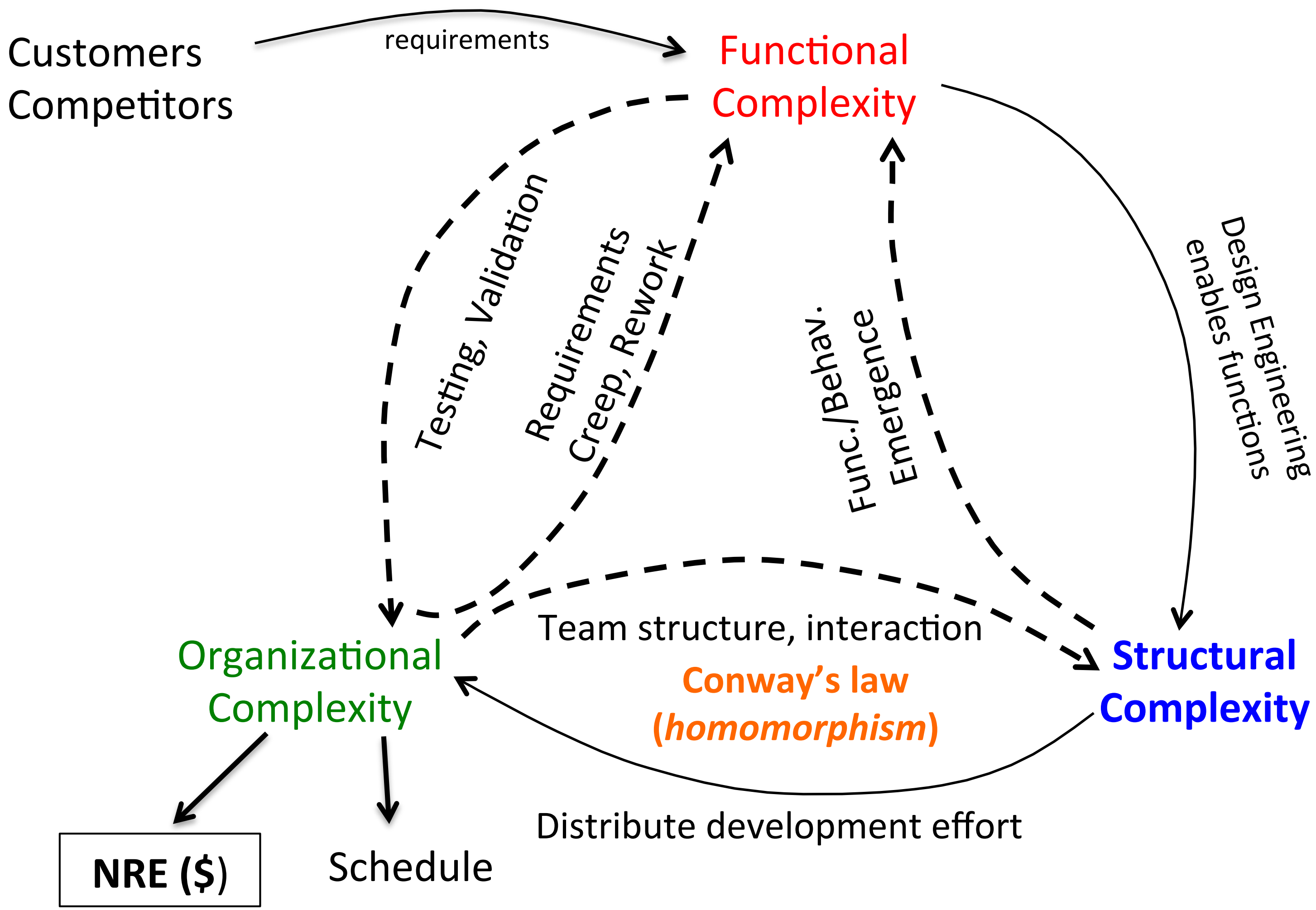
Ramasesh and. Browning, 2012 (preprint)

Development Cost and "Complexity"

- **CoBRA (Aerospace Corp., 2008) – Complexity Index based on analysis of historical data.**
- **Projects that were highly complex but tried to cut development cost had high failure rates**



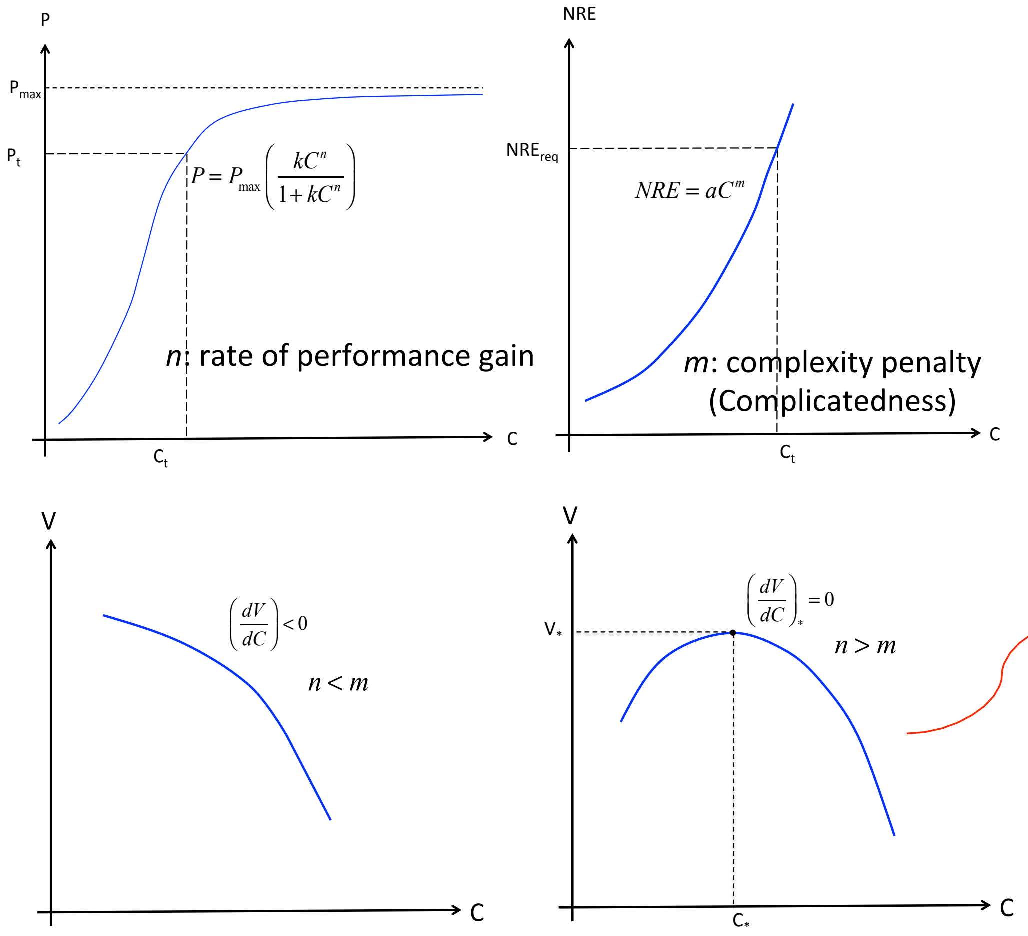
Three Dimensions of Complexity



NRE Cost – Non-Recurrent Engineering Cost

We need to do Complexity Budgeting

Complexity budget is the level of complexity that maximizes Value !



$$P = P_{max}\left(\frac{kC^n}{1+kC^n}\right)$$

$$NRE = aC^m$$

$$\boxed{V = \frac{P}{NRE} = P_{max}\left(\frac{k}{a}\right)\left[\frac{C^{(n-m)}}{1+kC^n}\right] = S\left[\frac{C^{(n-m)}}{1+kC^n}\right]}$$

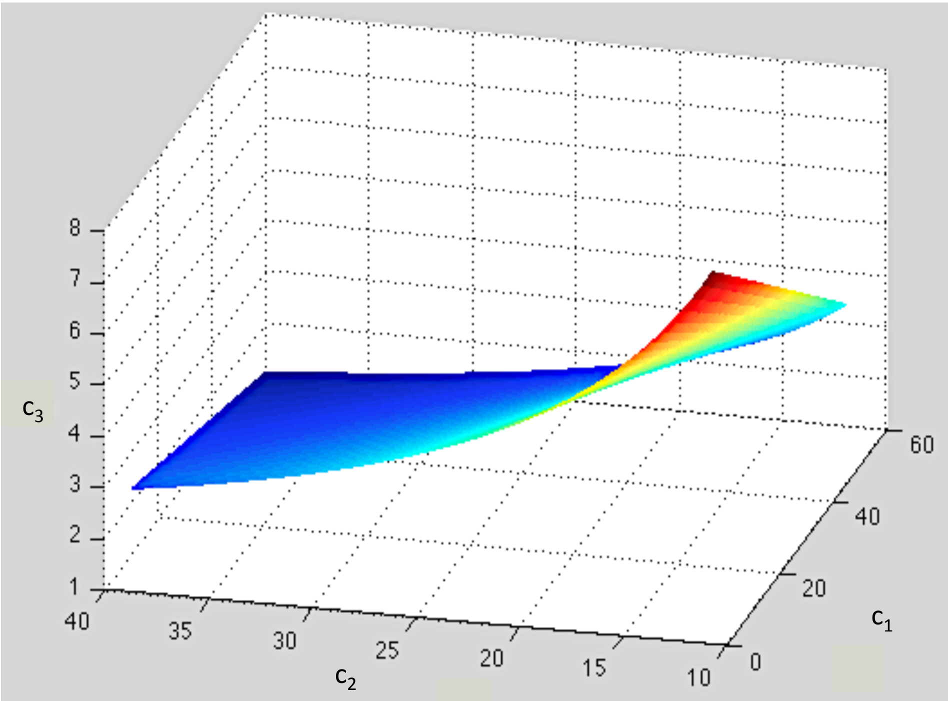
Value function as the complexity price for performance gain – Maximize V:

$$C_*^n = \frac{\left(\frac{n}{m}\right)-1}{k};\; P_* = P_{max}\left(1-\frac{m}{n}\right)$$

$$NRE_* = a\left[\frac{\left(\frac{n}{m}\right)-1}{k}\right]^{\frac{m}{n}};\; V_* = S\left(\frac{m}{n}\right)\left[\frac{\left(\frac{n}{m}\right)-1}{k}\right]^{\left(1-\frac{m}{n}\right)}$$

Iso-Complexity → how to allocate C?

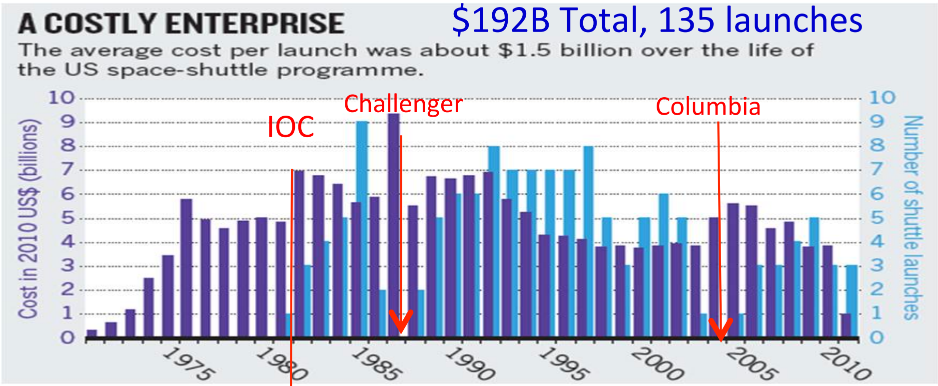
- Once we set a complexity budget, there are different ways to distribute this total structural complexity, C into its three components $\{C_1, C_2, C_3\}$: *IsoComplexity Surface*



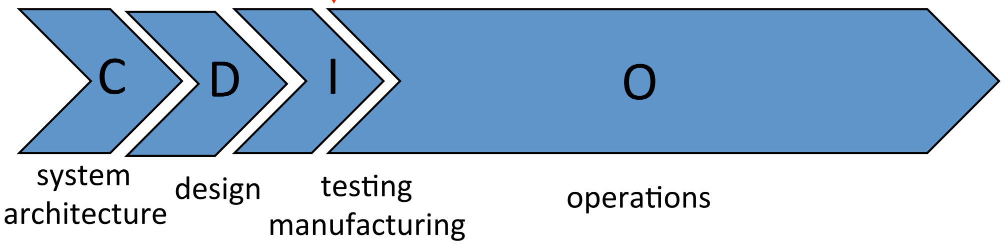
Iso-complexity surface: n = 20 components, assuming, c_1 in [10,60]; c_2 in [12,40] and C = 100.

- Tradeoff between (i) complex components and simple architecture, or (ii) simpler components and more complex architecture.
- Choice can be made depending on complexity handling capabilities of the development organization. E.g.
  - Excellent component designers
  - Systems integrators

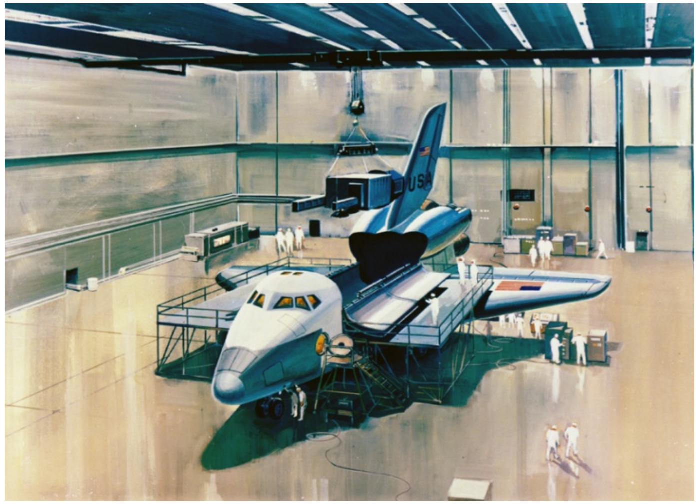
Space Shuttle Lifetime Cost (1971-2011)



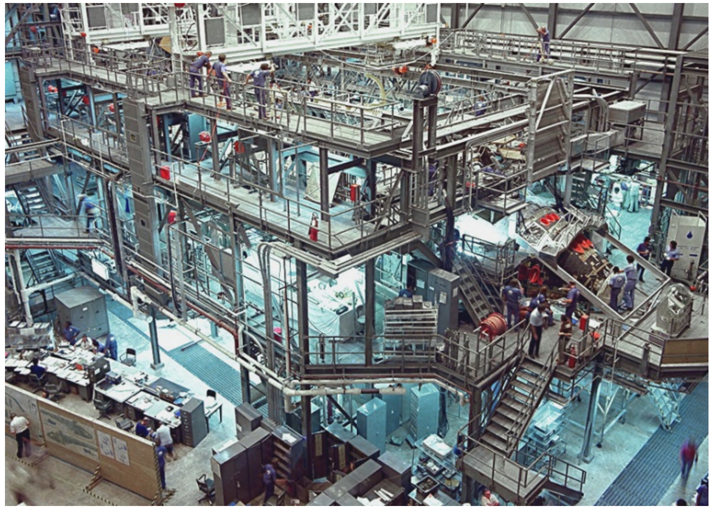
Roger Pielke Jr & Radford Byerly, Shuttle programme lifetime cost, *Nature* 472, 38 (07 April 2011)



What we wanted



What we got



- Vision: partially reusable space vehicle with quick turnaround and high flight rate
- Actual: complex and fragile vehicle with average cost of about $1.5B/flight (20,000 workforce)
- Why?
  - Congress capped RDT&E at $B5.15 (1971)
  - Did not do complexity budgeting

Why should we care about complexity?

How do we quantify complexity?

How to better manage complexity?

Summary of key points

- Structural complexity of cyber-physical systems has been increasing steadily since industrial revolution
- Driven by customer needs and competition → functional complexity → structural complexity → organizational complexity
- Due to human cognitive bandwidth limitation (magic 7+/-2) → Complicatedness drives super-linear cost in effort (b ~ 1.5)
  - Abstraction layers and decomposition into modules
- A rigorous measure of complexity
  - Satisfies Weyuker's criteria (1998)
  - C = C1+ C2*C3; Graph Energy is a measure of topological complexity
- **Better complexity-based management**
  - P-Point is a critical transition point
  - Critical nodal degree <k>cr=6
  - Iso-complexity based budgeting with clear targets