When is complex too complex?

Graph Energy, Proactive Complexity Management, and the First Law of Systems Engineering

UTC-IASE Distinguished Lecture: October 6, 2016

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?

IDSS
MIT INSTITUTE FOR DATA, SYSTEMS, AND SOCIETY

In the news

New Geared Turbofan Engine Challenges

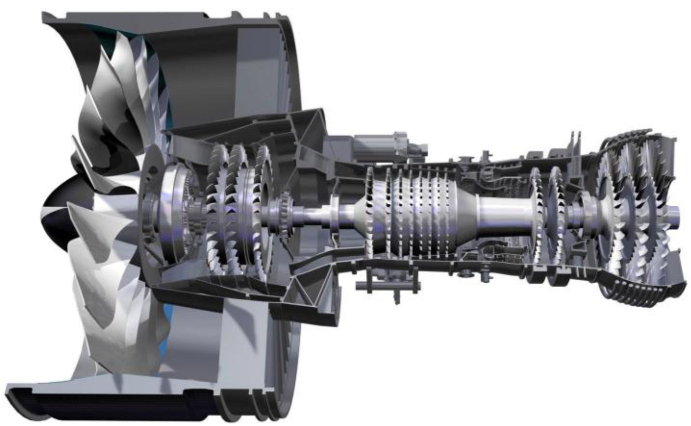
"Pratt **invested $10 billion** over the past two decades to develop the engine, which promises dramatic cuts in fuel use, emissions and noise over previous-generation models."


Pratt & Whitney Cuts Jet-Engine Delivery Plan as Output Lags

Bloomberg — Sep 16, 2016

EMAIL | in SHARE | Tweet | G+1 | Recommend 4 | COMMENTS 0

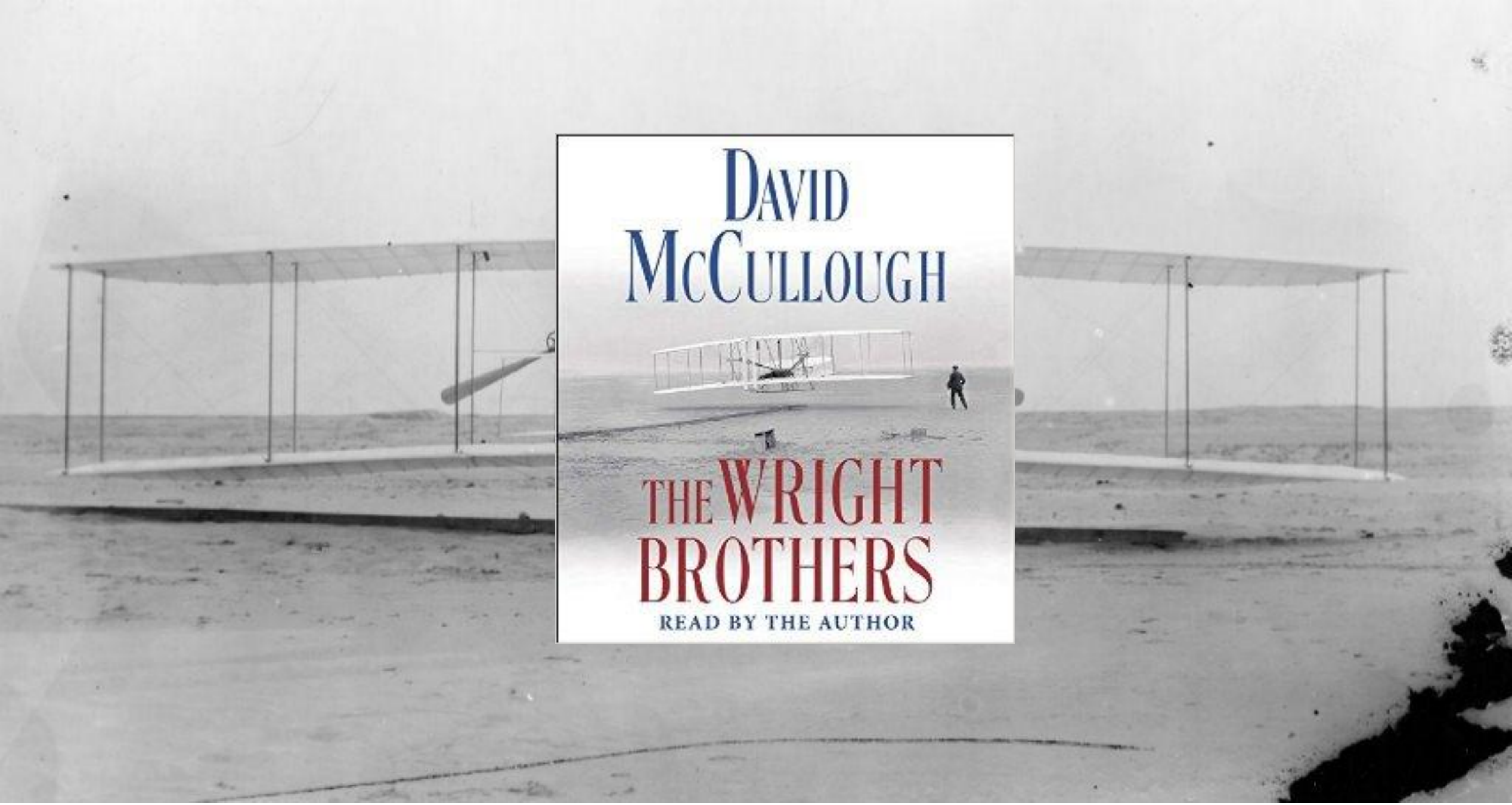
The engine maker will hand over about 150 units this year, down from the 200 it previously expected.



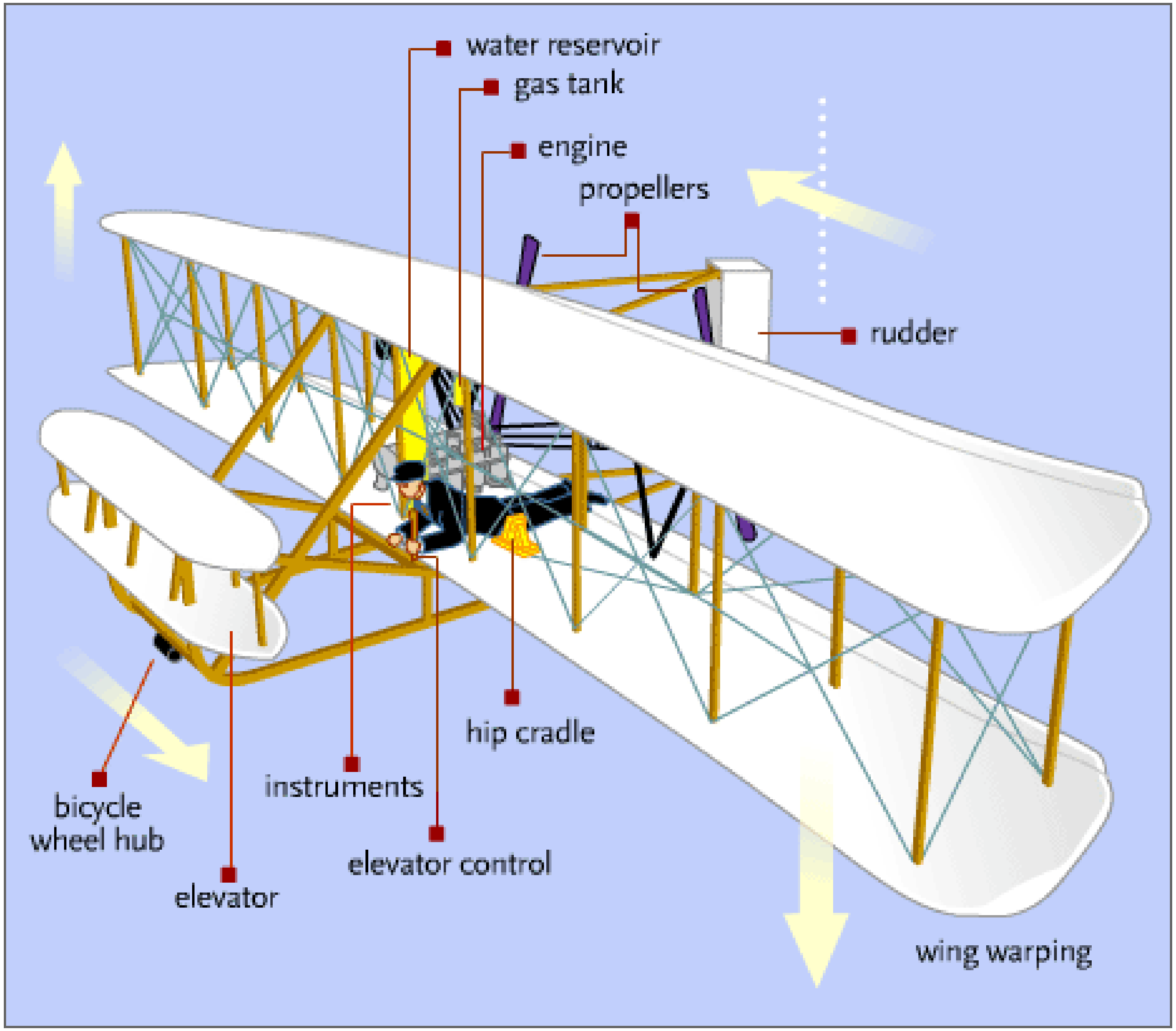
"**The engine has about 800 parts**, and difficulties with about five of those are causing problems that have slowed down production, Hayes said. In particular, the fan blades now take about 60 days to produce **because of the complex technology**, but that should take about 30 days, he said."

http://www.industryweek.com/operations/pratt-whitney-cuts-jet-engine-delivery-plan-output-lags

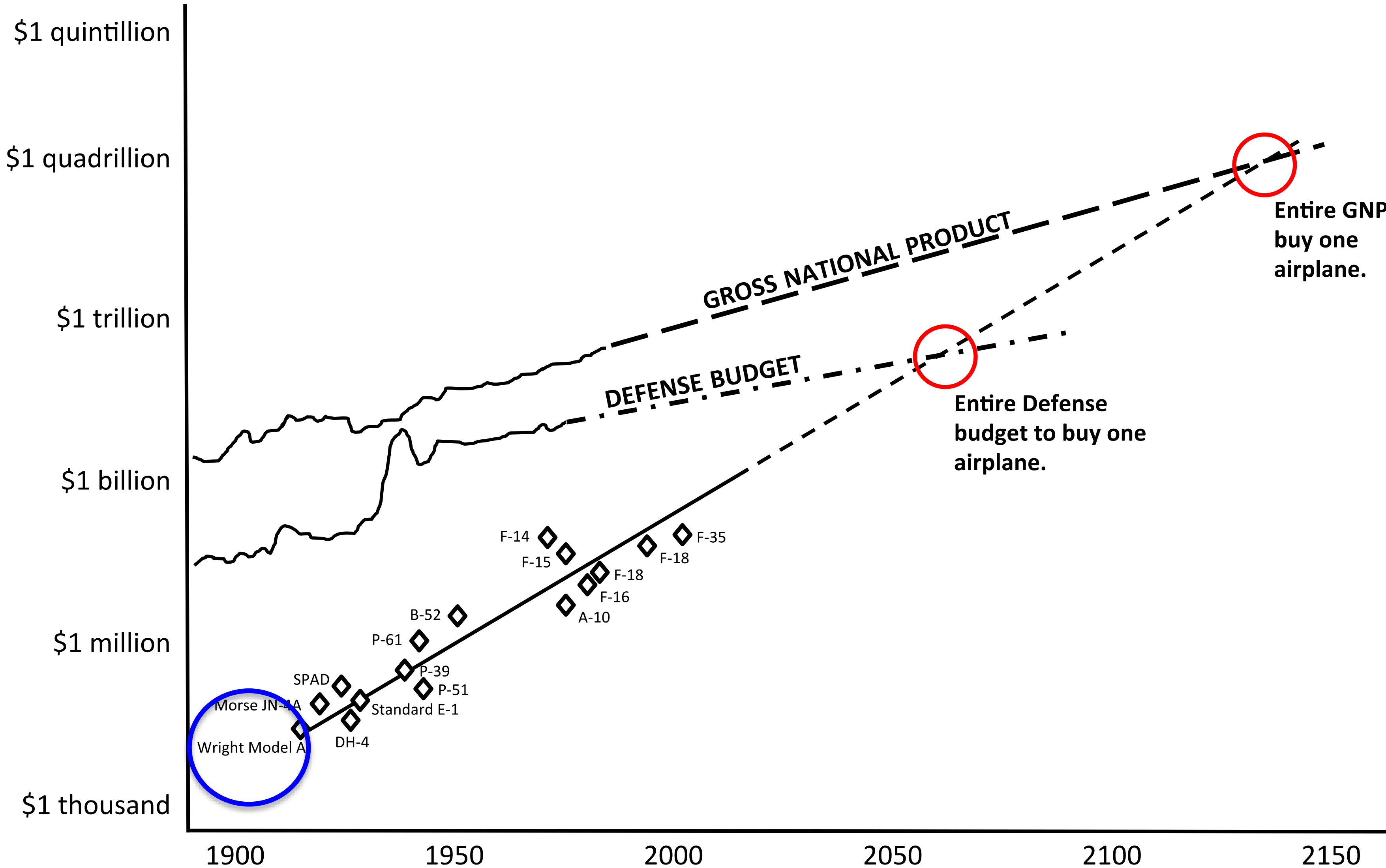
What have you been reading lately?



The Wright Flyer

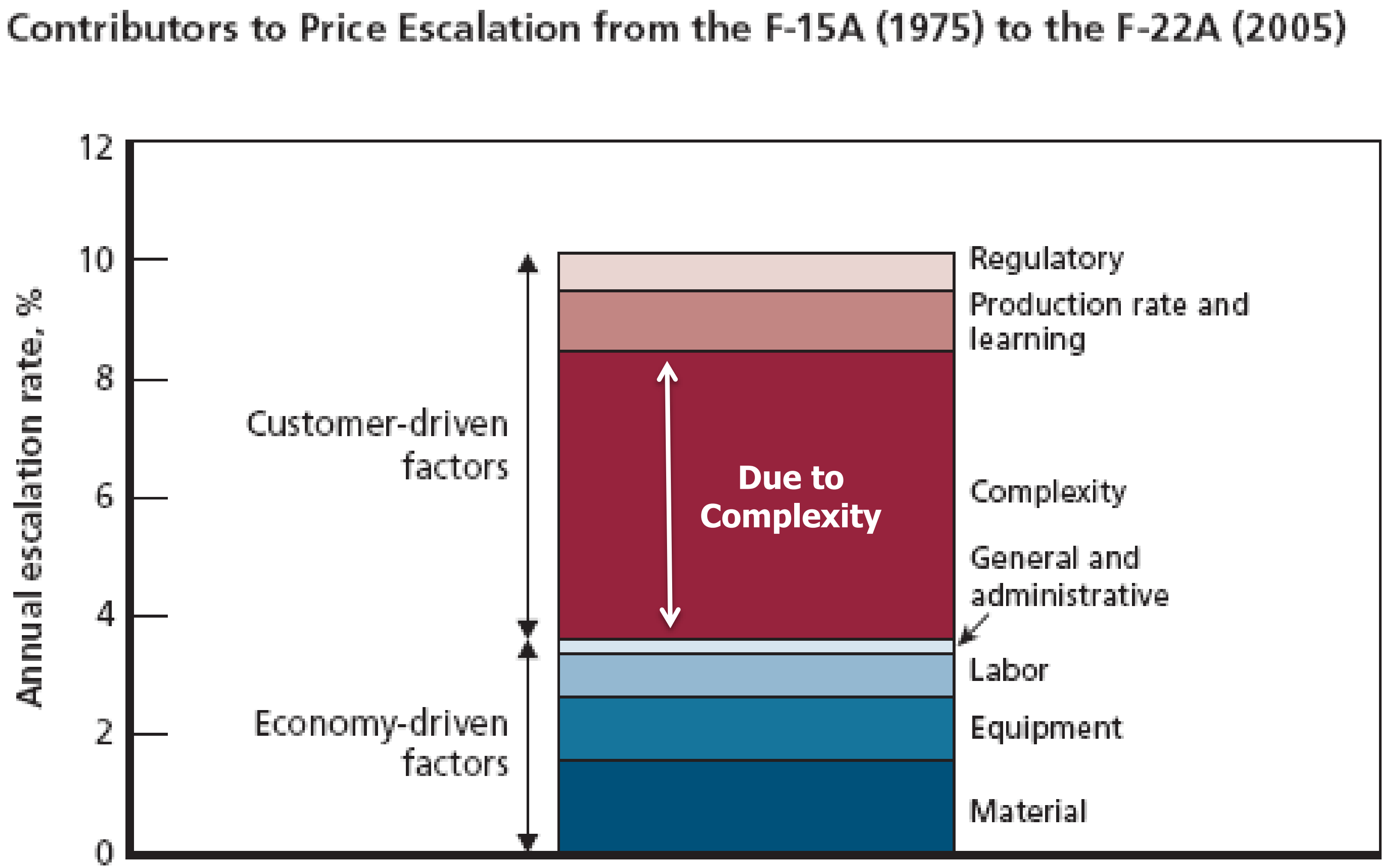


Augustine's 16th Law



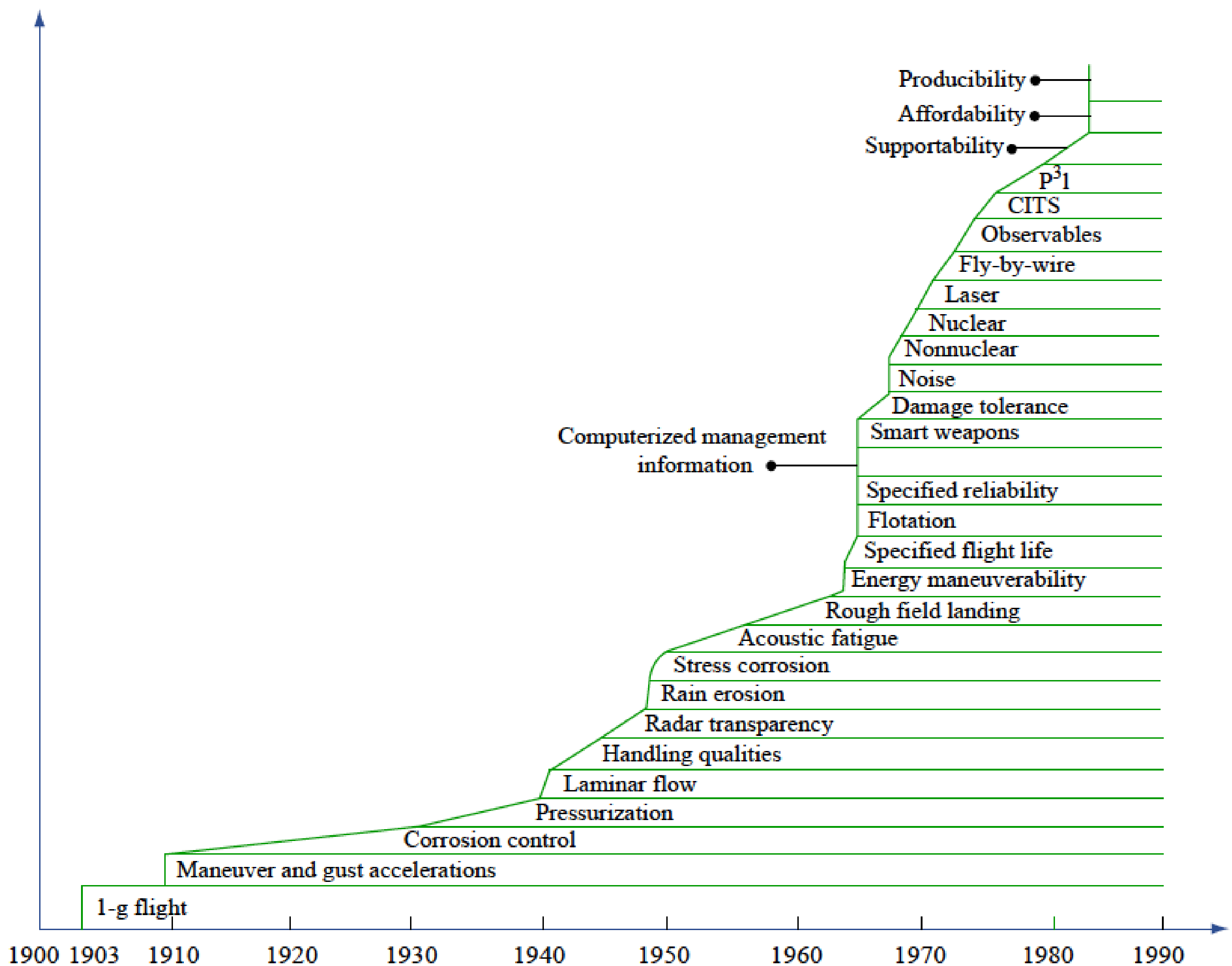
Norm Augustine, *Augustine's Laws*, 6th Edition, AIAA Press, 1997.

What is driving this escalation of cost?



Source: DARPA TTO (2008)

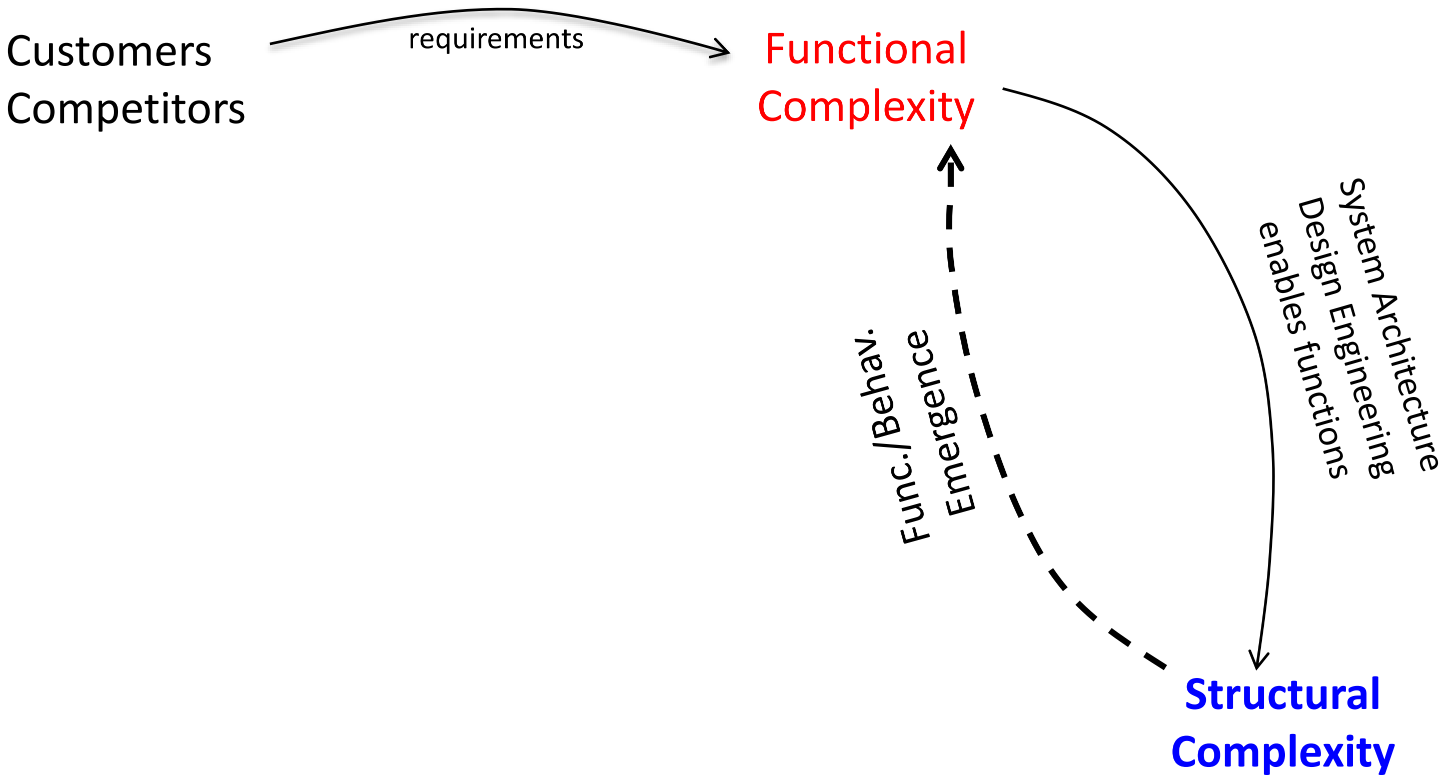
Functional Requirements Explosion



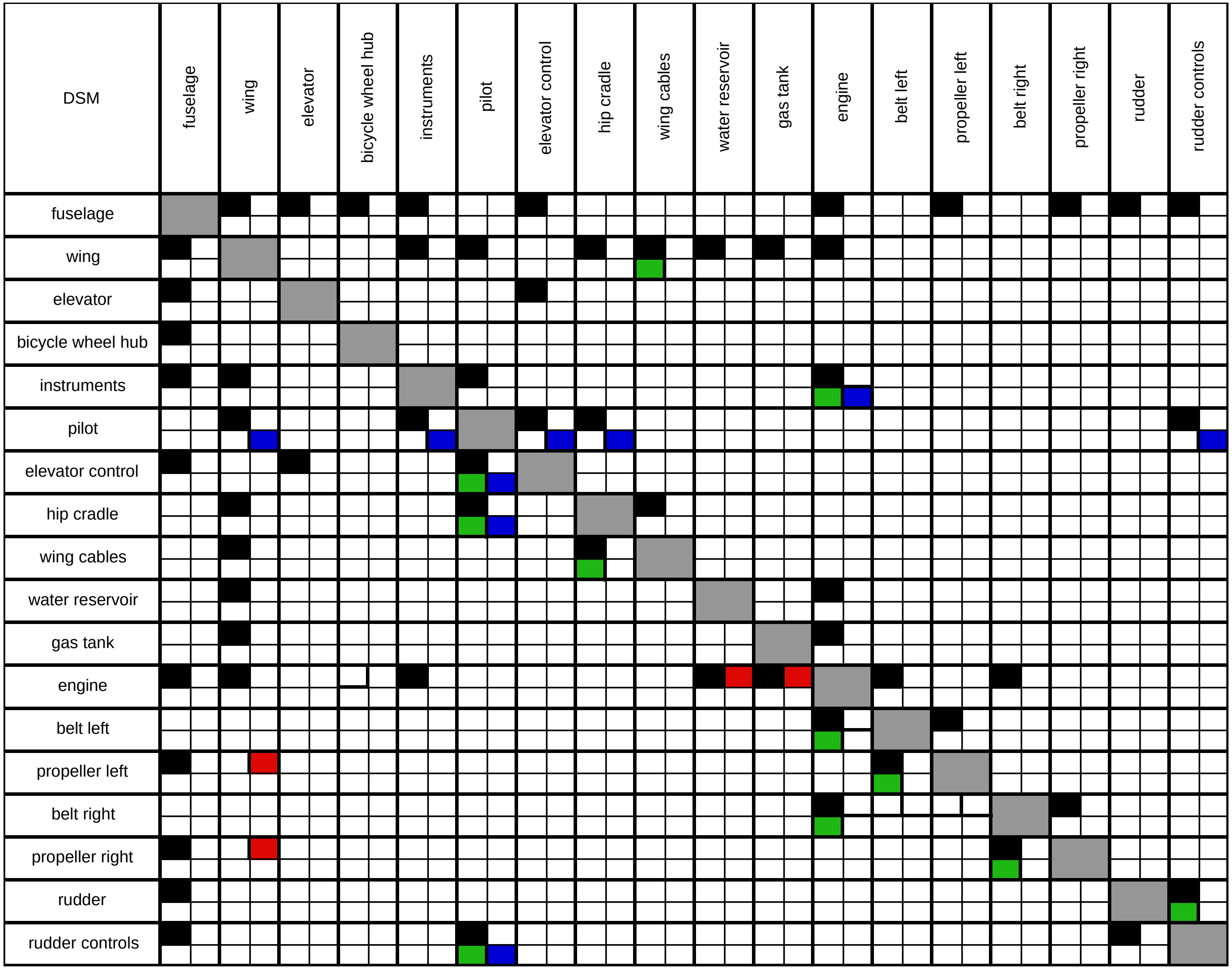
Design requirements growth for aerospace vehicles.

Image by MIT OpenCourseWare.

Two Dimensions of Complexity



Structural DSM of Wright Flyer



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

DSM 18x18

<u>Connections</u>

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

Why should we care about complexity?

How do we quantify complexity?

How to better manage complexity?

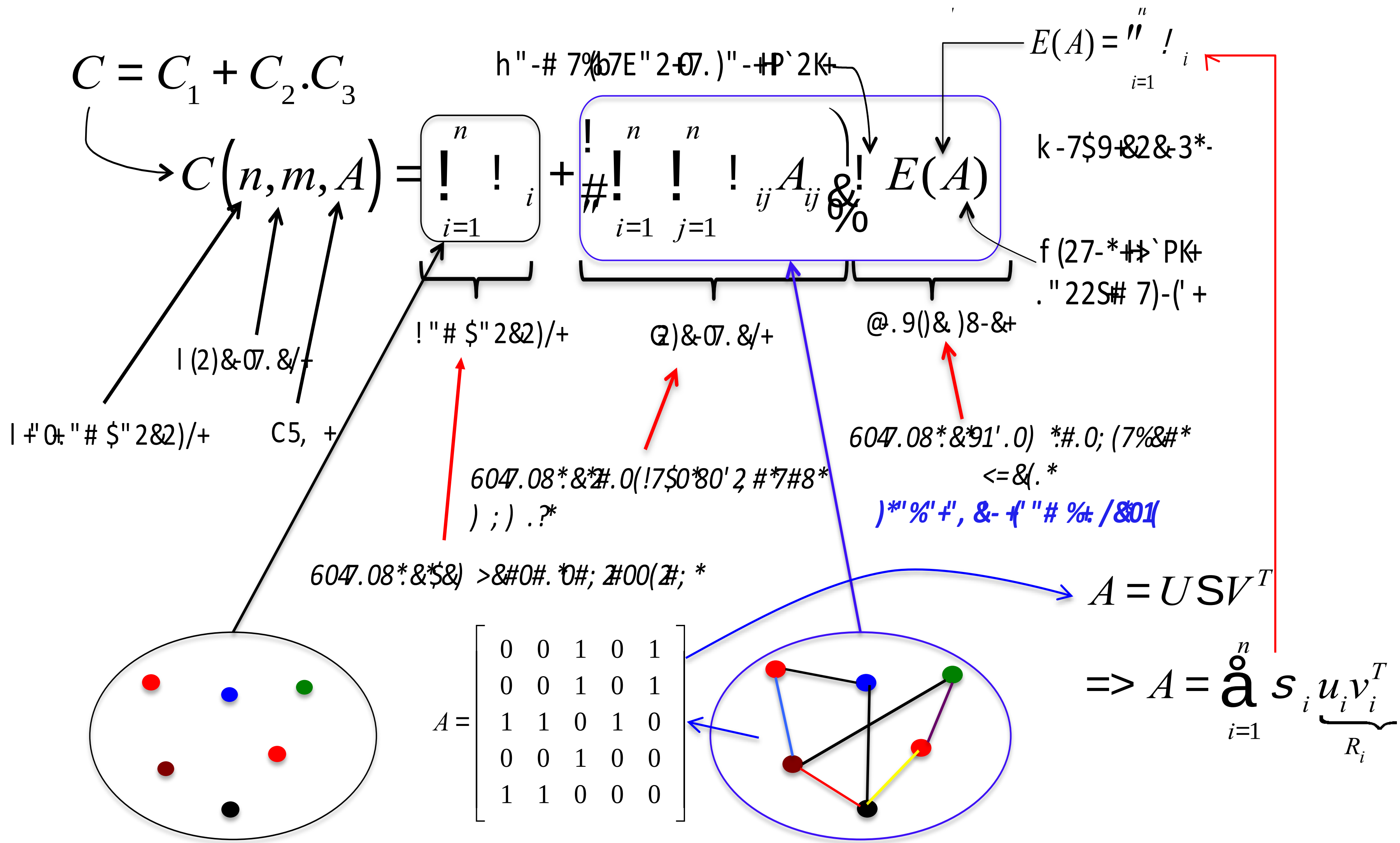
Elaine Weyuker's (1988) criteria

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

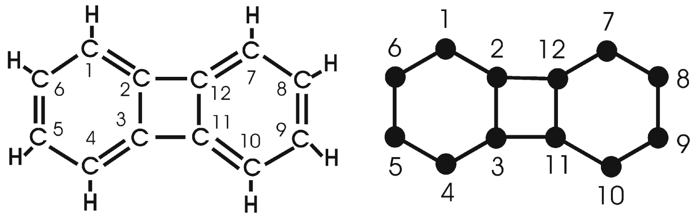
1) Invariant to relabeling (i.e., isomorphism).
2) Possible to have different system architectures have the same complexity level.
3) Differentiate between system architectures.
4) System structure at least partially determines complexity of functionally equivalent systems.
5) Changes in internal architectural 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]

Weyuker, E. J. (1988). Evaluating software complexity measures. *IEEE transactions on Software Engineering*, *14*(9), 1357-1365.

Structural Complexity Metric



System Hamiltonian and Complexity



$$e_p = na + b\sum_{i=1}^{n} h_i s_i \le na + b\underbrace{\left(\sum_{i=1}^{n} h_i\right)}_{n}\underbrace{\left(\sum_{i=1}^{n} s_i\right)}_{E(A)}$$

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

$$\backslash \quad e_p \le na + n^2 b\left(\frac{E(A)}{n}\right)$$

Introduce a notion of of *configuration energy*:

$$X := \underbrace{n\hat{a}}_{C_1} + \underbrace{m\hat{b}}_{C_2}\underbrace{\left(\frac{E(A)}{n}\right)}_{C_3} = C_1 + C_2 C_3$$

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

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:

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

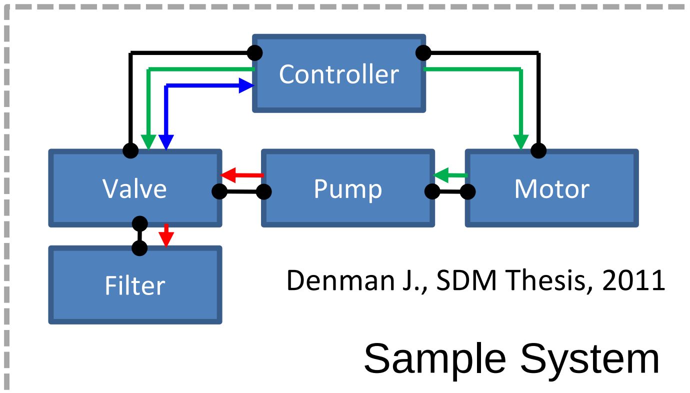
$$C = C_1 + C_2 C_3$$

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

$$= \sum_{i=1}^{n} a_i + \left(\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}\right)\left(\frac{E(A)}{n}\right) = \sum_{i=1}^{n} a_i + \left(\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}\right) gE(A)$$

$$|e_i| = a + bs_i; \; e_p = \sum_{i=1}^{n} h_i |e_i|$$

Example: Cyber-Physical System

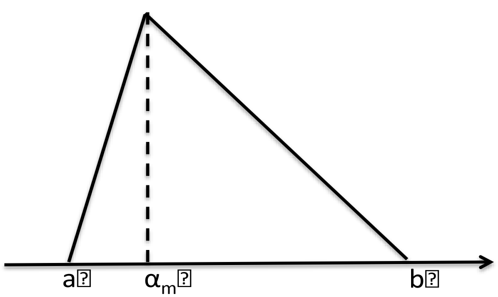


aggregation

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

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
		0.10
		0.05
		0.15
		0.05
		0.10



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

p Î [1.0;3.0]

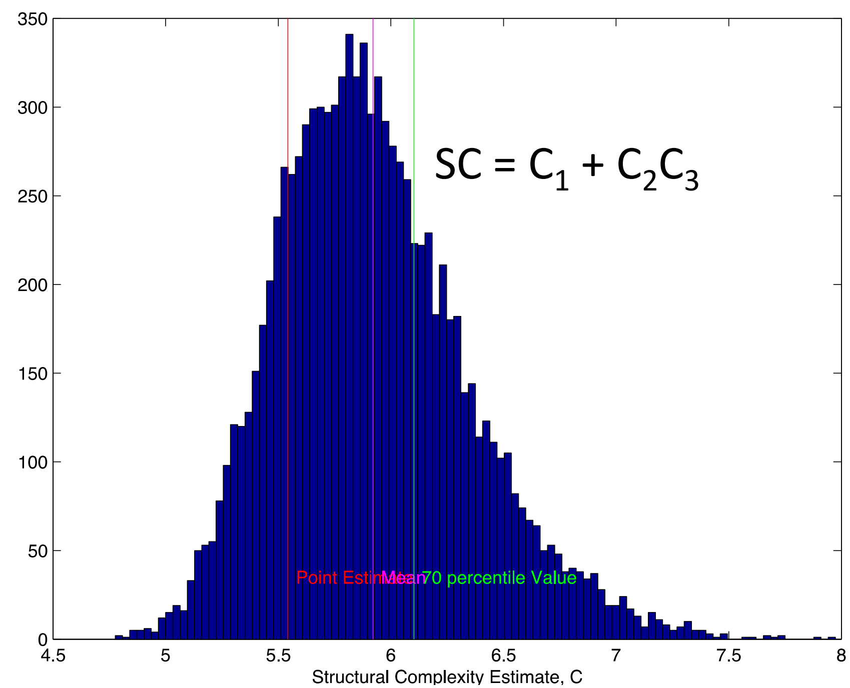
a Î $[0.8a_m; 0.9a_m]$

b Î $[1.1a_m; 1.6a_m]$

$$b_{ij}^{(k)} = g(a_i, a_j, c^{(k)})$$

$$b_{ij}^{(k)} = \frac{\max(a_i, a_j)}{c^{(k)}},$$

" a_i, a_j ¹ 0, *k is the interface type*

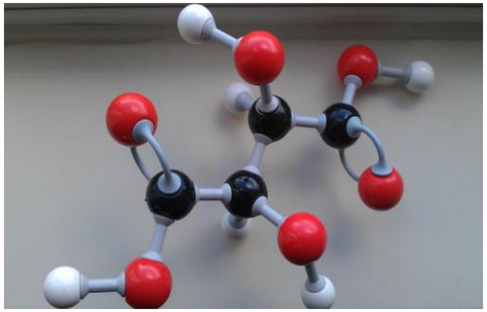


$SC = C_1 + C_2 C_3$

Structural Complexity Estimate, C

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.



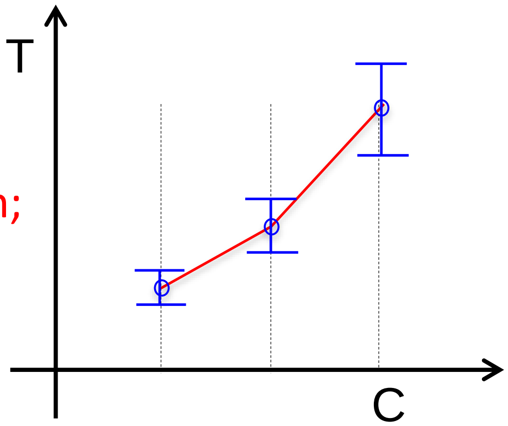
- Record for each model (for each individual)
  - **C** = computed structural complexity
  - **T** = [time to build, including rework if any]

Molcules: 12
Subjects: 17

Hypothesis:
High Structural Complexity leads to measurably …

Slower Progress (Cognition; Schedule)

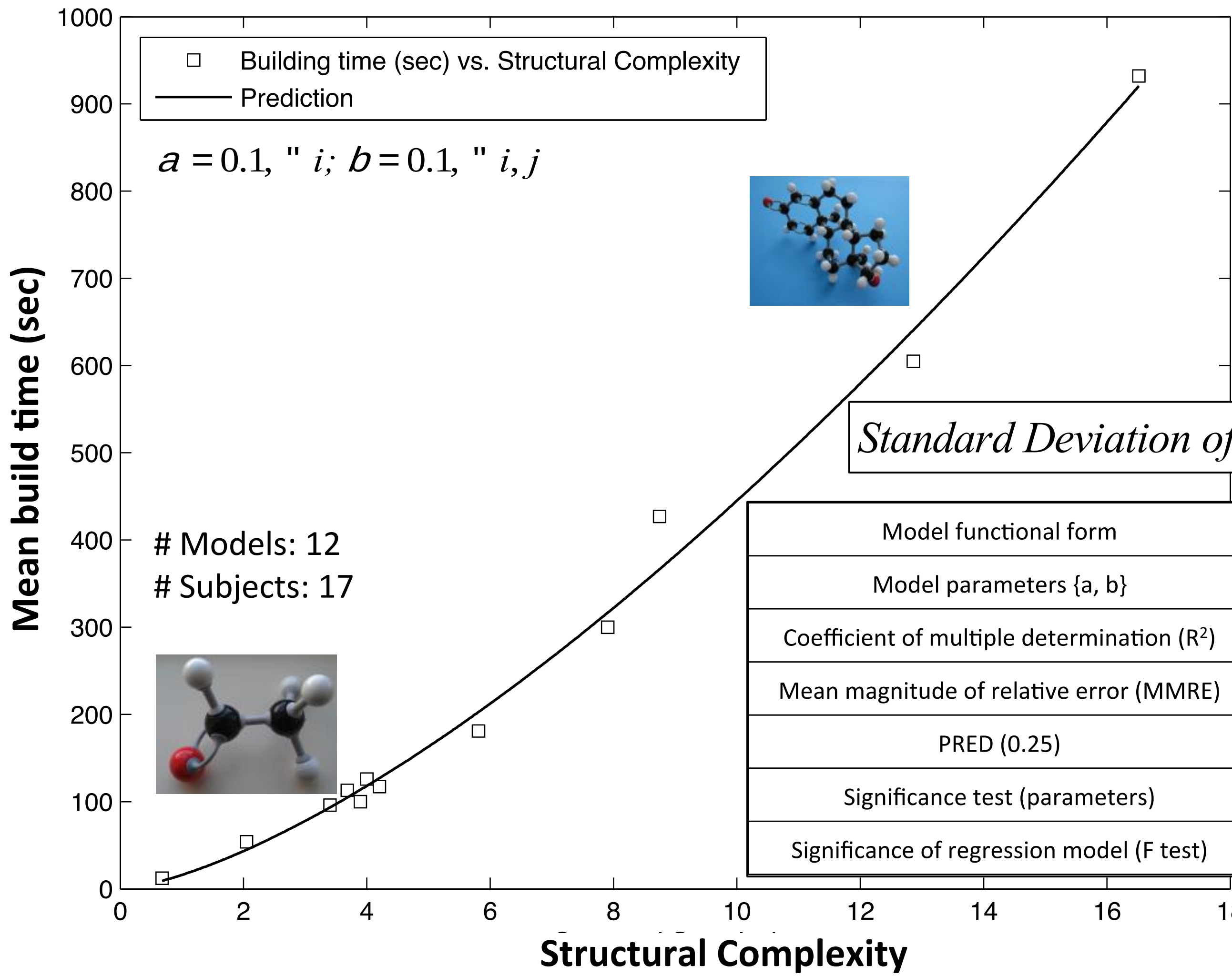
Higher Error-rate (Rework)



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

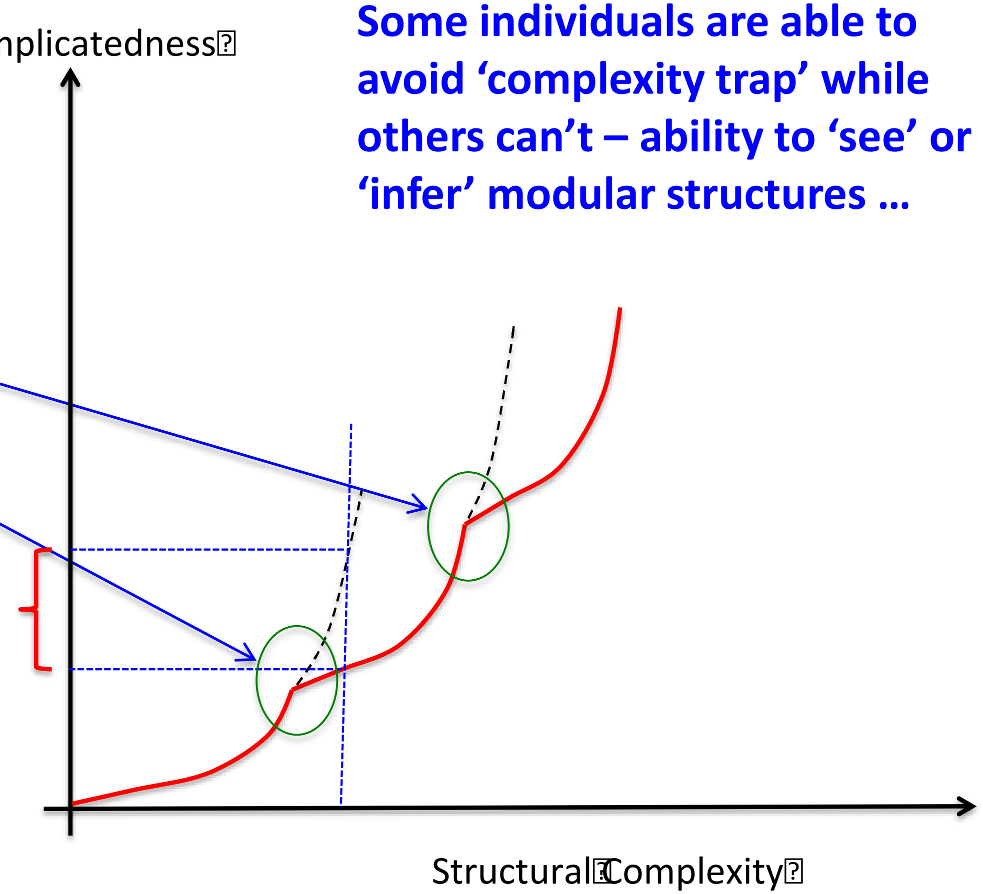
$$Structural\ Complexity,\ C = O(n^{1.08}) \neg\ \ mild\ super\text{-}linearity$$

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

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 others can't – ability to 'see' or 'infer' modular structures …



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.

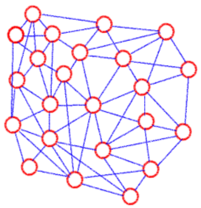
Complexity Measure	Computability	Aspect emphasized	Weyuker's Criteria
Number of components [Bralla, 1986]	✔	Component development (count-based measure)	✗
Number of interactions [Pahl and Beitz, 1996]	✔	Interface development (count-based measure)	✗
Whitney Index [Whitney *et al.*, 1999]	✔	Components and 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 Reconstructability	✔

Why should we care about complexity?

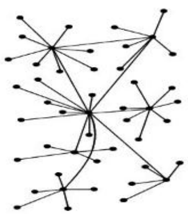
How do we quantify complexity?

How to better manage complexity?

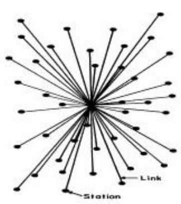
Topological Complexity: Important Properties



“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 ! *hypoenergetic,* $C_3 < 1$!!

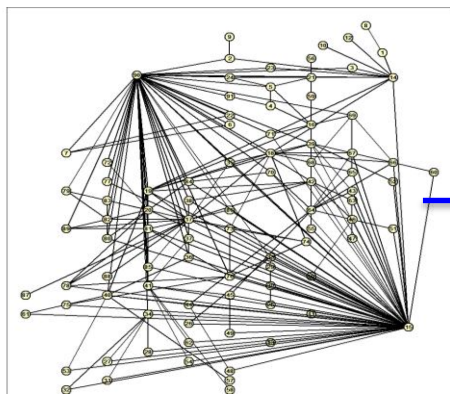
Hierarchical / layered Architecture ! *transitional,* $1'' \; C_3 < 2$!!

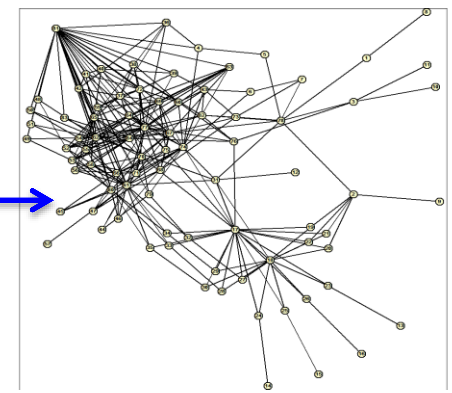
Distributed Architecture ! *hyperenergetic,* $C_3 '' \; 2$!!

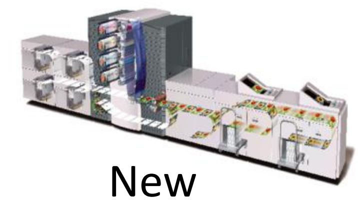
Case Study 1: Printing Engines

Old

Complexity = 186







New

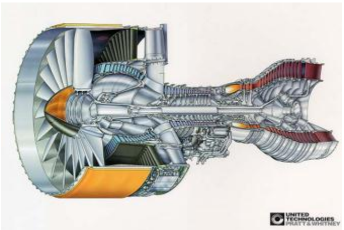
Complexity = 354

Complexity increase +90%

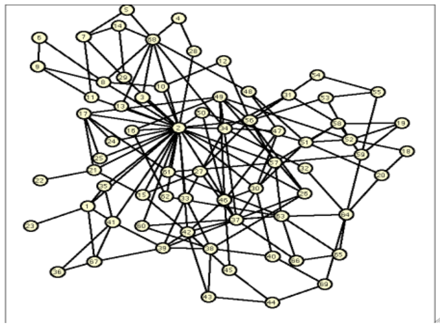
	C_1		C_2		C_3		C		C_{New} $/C_{Old}$
	Old	**New**	**Old**	**New**	**Old**	**New**	**Old**	**New**	
Most Likely	110.2	169	55.68	102.78	1.36	1.804	185.93	354.42	1.9062
Mean	125.62	213.6	63.29	130.6	1.36	1.804	211.69	449.2	2.122
Median	124.47	211.84	62.46	128.62	1.36	1.804	209.42	443.88	2.12
70 percentile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

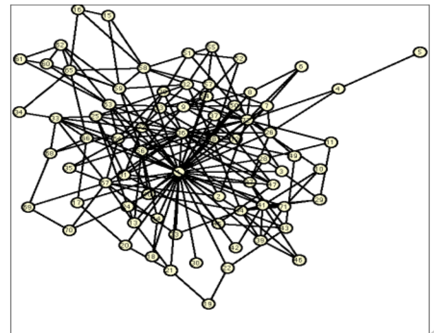
Case Study 2: Aircraft Engines

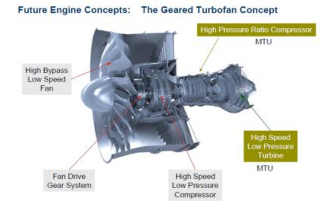


Old



Complexity = 351





New

Complexity = 499

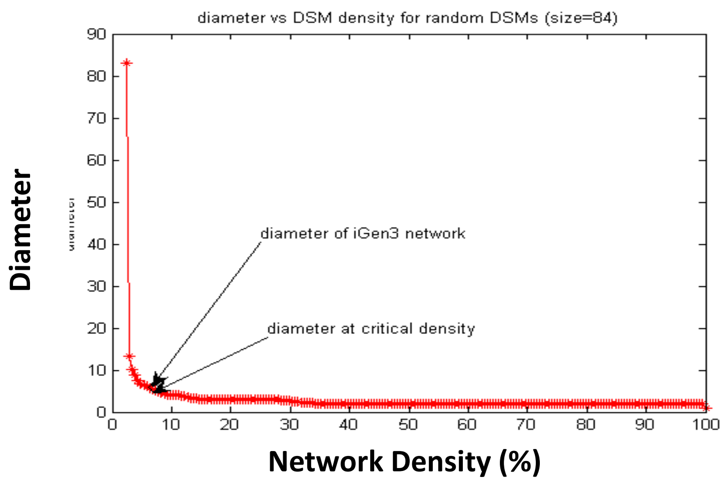
Complexity increase +42%

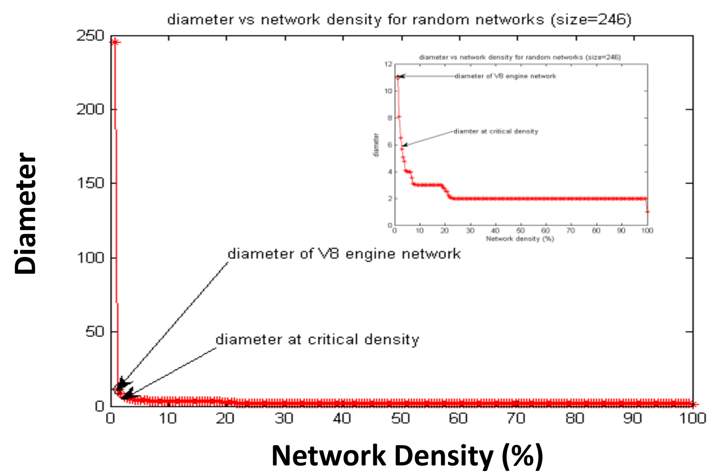
	C_1		C_2		C_3		C		C/C_{ML}		C_{new}/C_{old}
	Old	**New**	**Old**	**New**	**Old**	**New**	**Old**	**New**	**Old**	**New**	
Most Likely	161	188	126	184	1.51	1.69	351	499	1	1	1.42
Mean	179	244	141	240.4	1.51	1.69	392	650.3	1.12	1.30	1.65
Median	178	242	139	238.9	1.51	1.69	388	646.8	1.10	1.29	1.66
70 percentile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

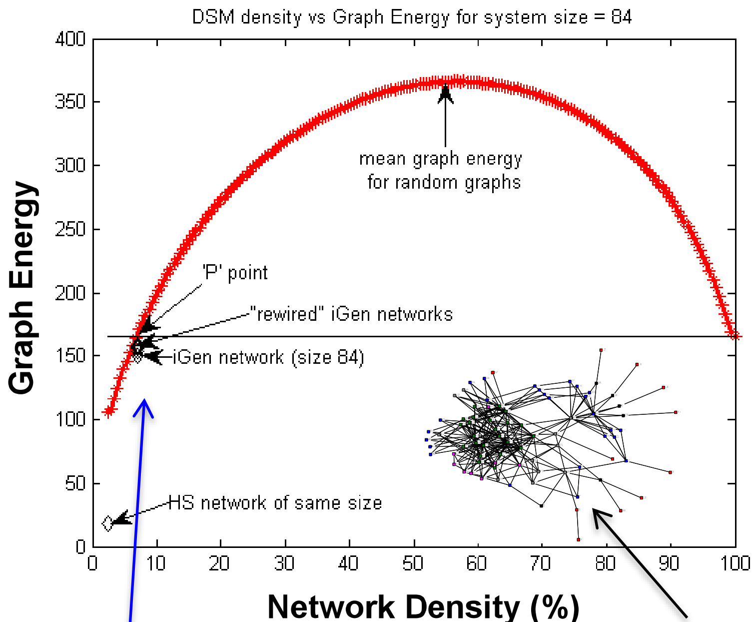
- The ***P* point** on graph energy – density plot: Phase transition for complxity
- At densities higher than ***P*** point, structural complexity increases but that does not buy much improvement in terms of performance measures (e.g., network diameter)

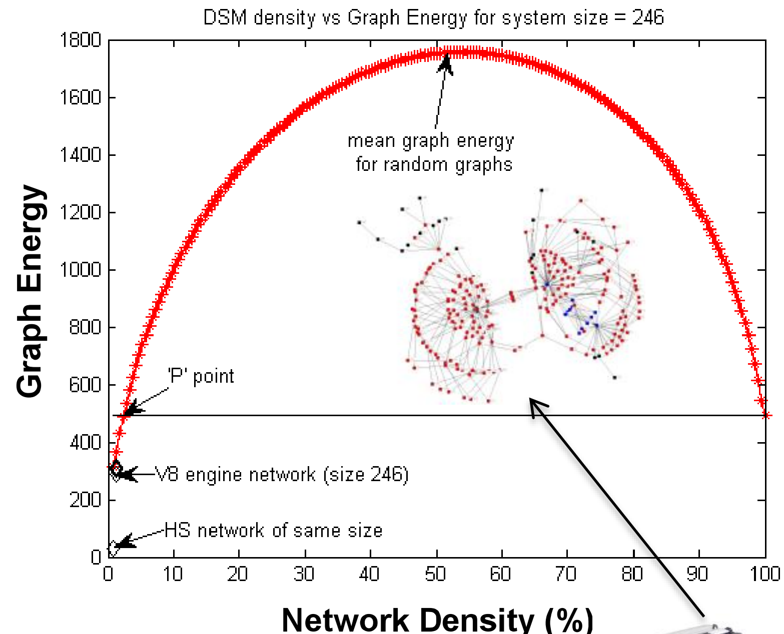




- Use equivalent random networks (Erdős–Rényi) as background.
- P-point has E(A) equivalent to fully connected system, and architectures become rank-dense beyond this point (critical for design).

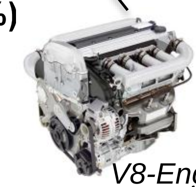
Real Product Design and P-Point Complexity





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

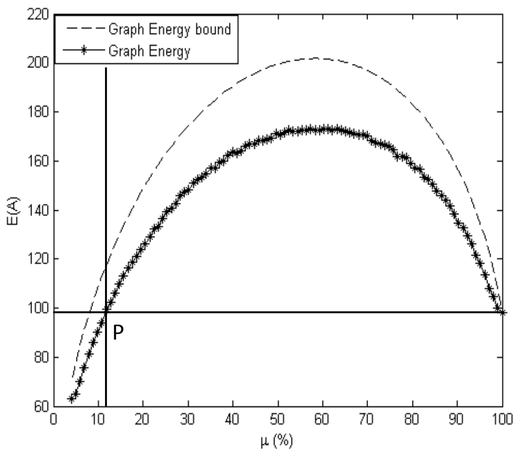
iGen3 (digital printing system)
Xerox

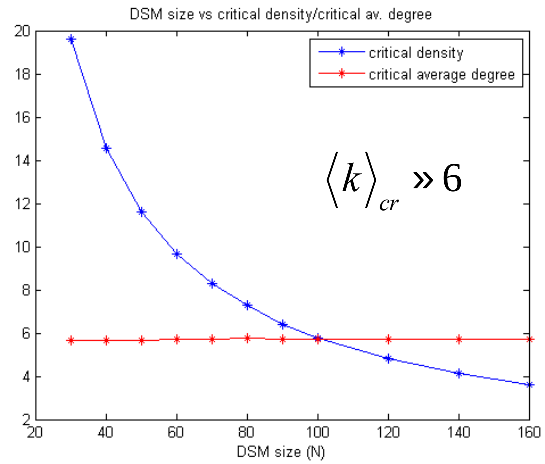


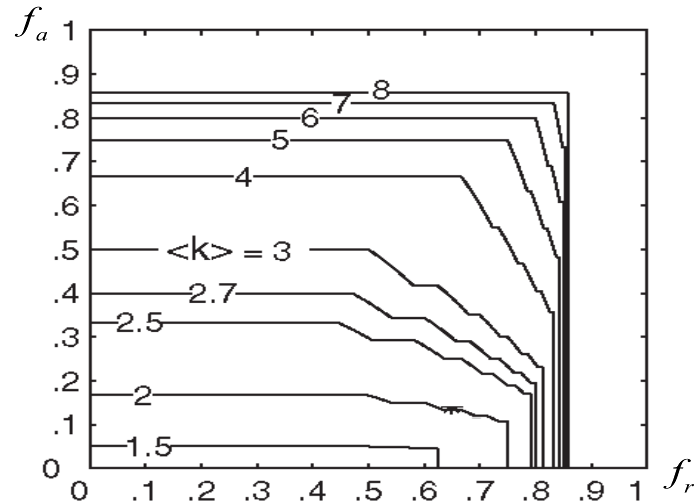
V8-Engine

- 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 $\langle k \rangle_{cr} \approx 6$

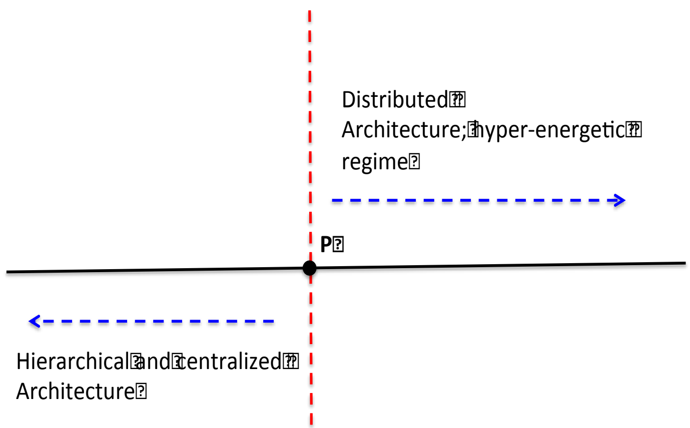


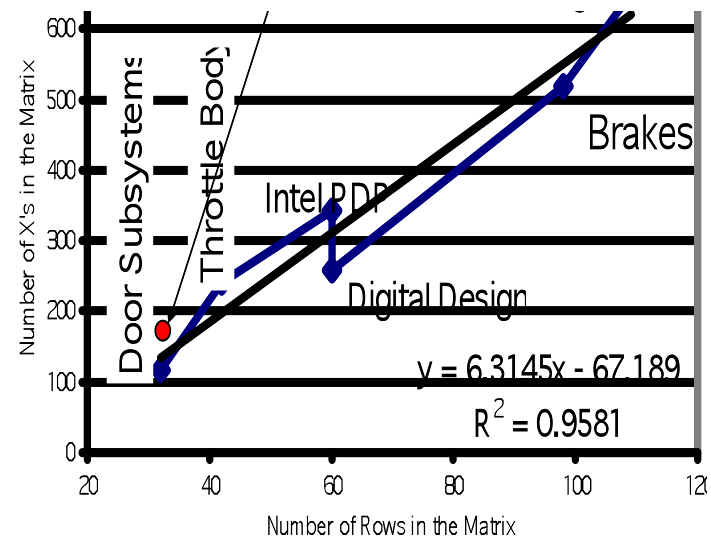




Network resilience contour (f_r vs. f_a)
[Valente et al., 2004]

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

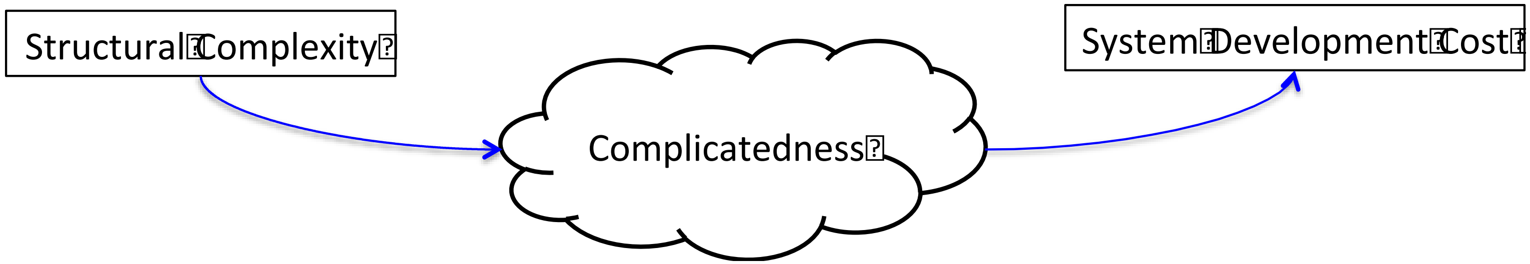




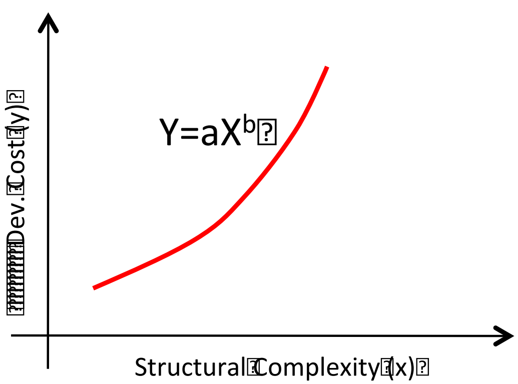
[Whitney et al., 1999]

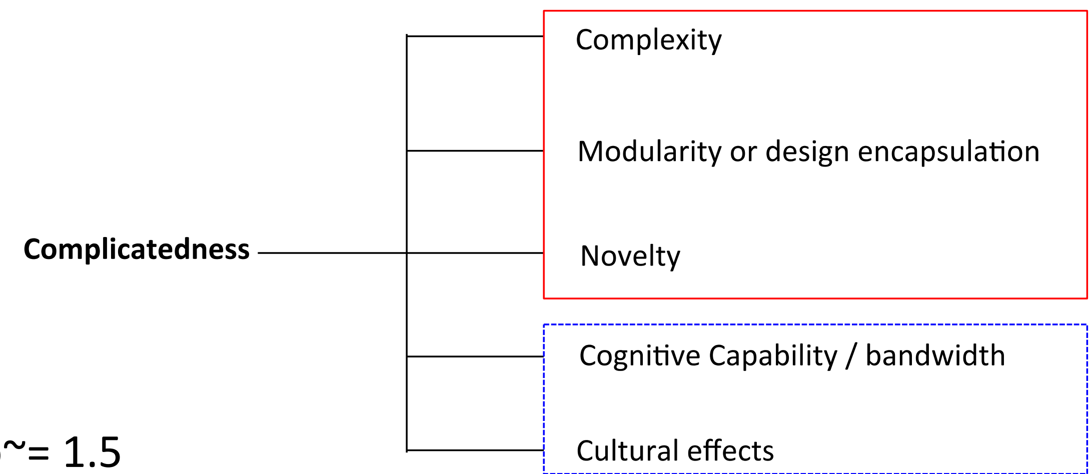
Use of *P* point as a system architecting guideline – entering regime of diminishing returns

Complicatedness vs. Complexity



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



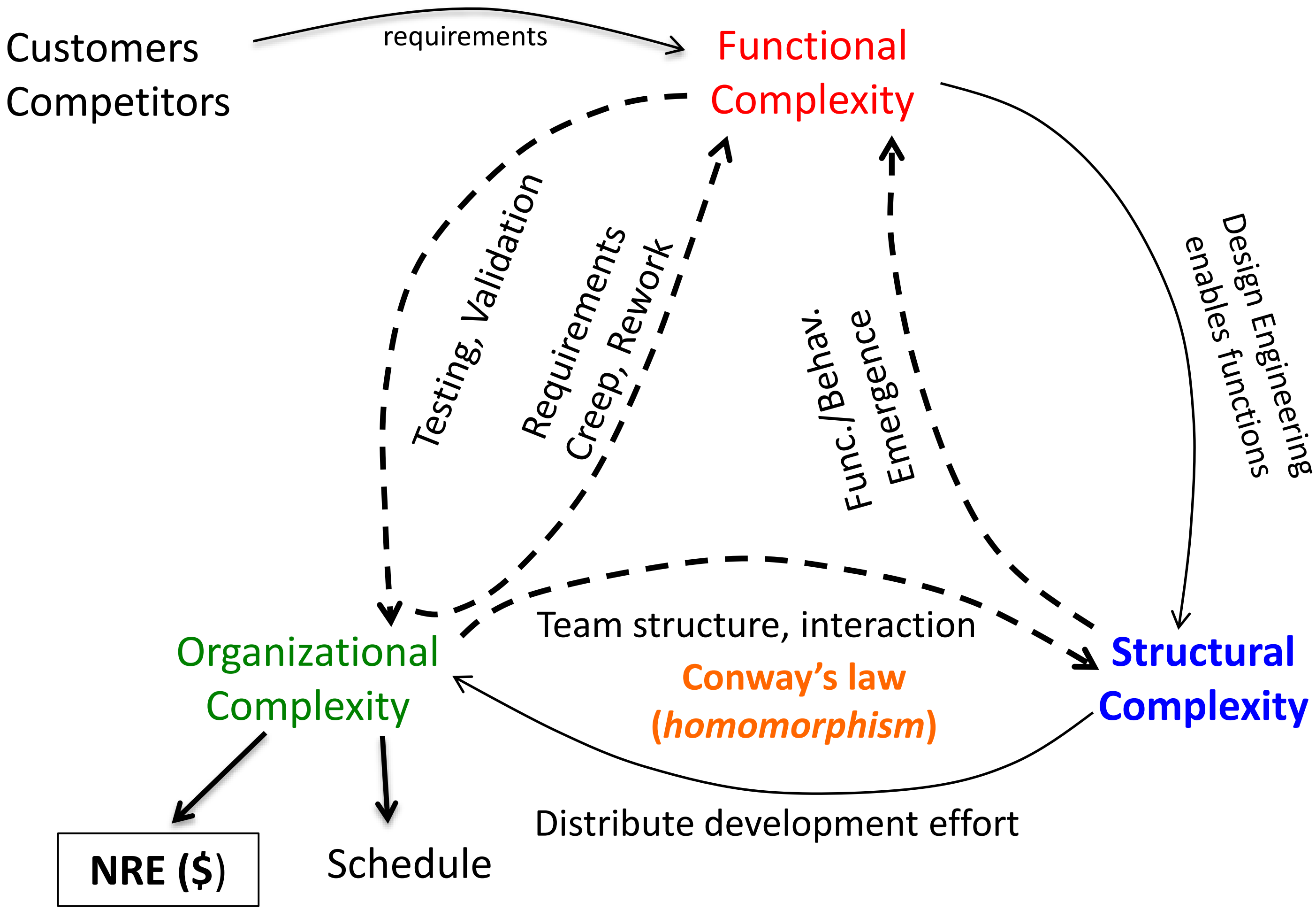


Experiments suggested b~= 1.5

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

Ramasesh and. Browning, 2012 (preprint)

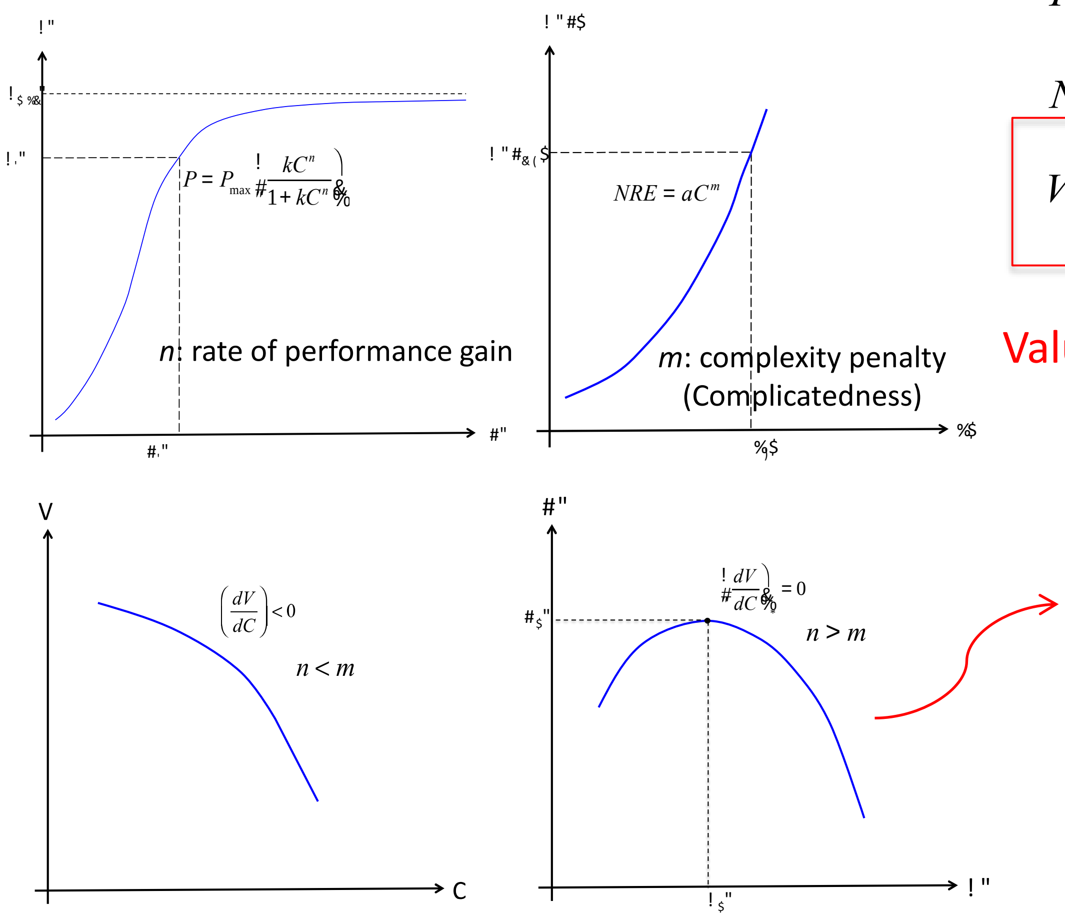
Three Dimensions of Complexity



NRE Cost – Non-Recurrent Engineering Cost

We need to do Complexity Budgeting

Complexity budget is the level of complexity that maximizes Value !



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

$$NRE = aC^m$$

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

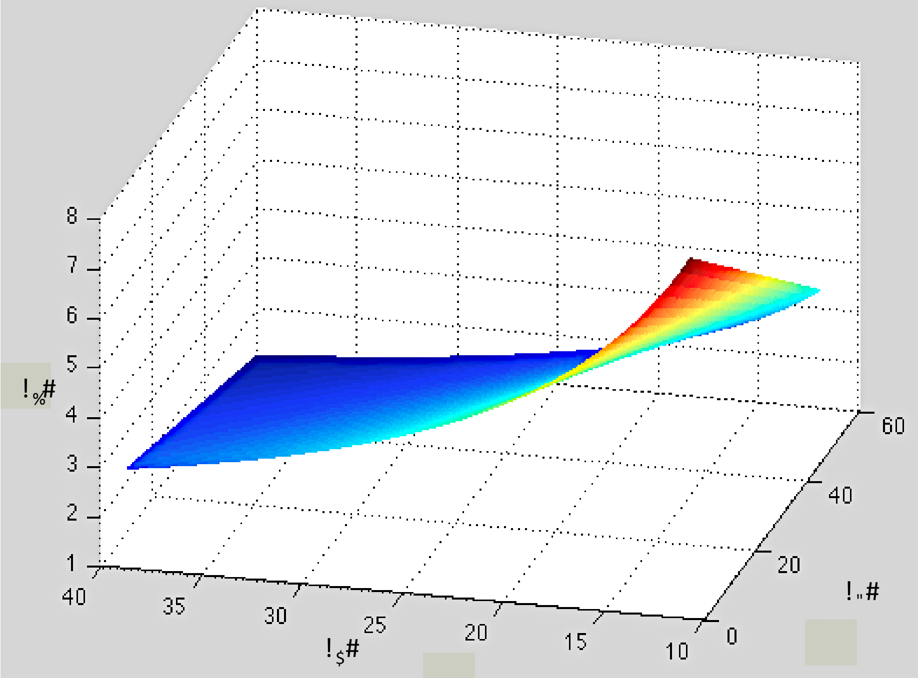
Value function as the complexity price for performance gain – Maximize V:

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

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

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.

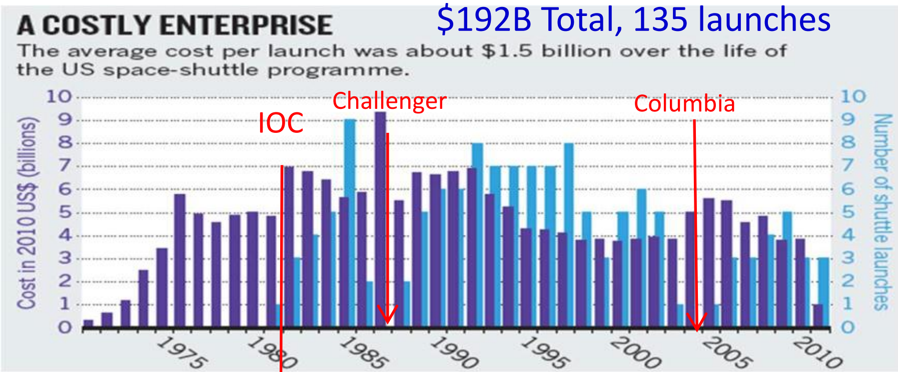
- Tradeoff between (i) complex components and simple architecture, or (ii) simpler components and more complex architecture.
- Choice can be made depending on complexity handling capabilities of the development organization. E.g.
  - Excellent component designers
  - Systems integrators

The First Law

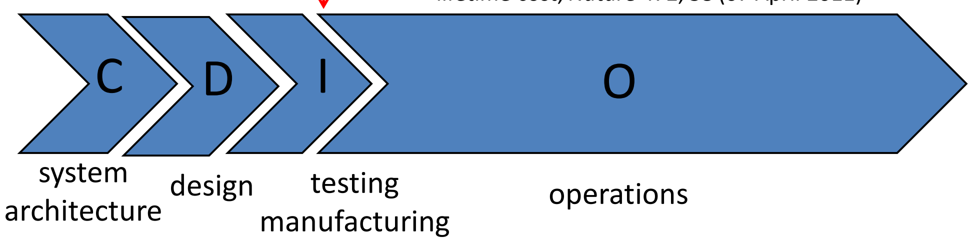
- First Law of Thermodynamics (ca. 1850):
  - The law of conservation of energy states that the total energy of an isolated system is constant; energy can be transformed from one form to another, but cannot be created or destroyed.
    - $E_{tot} = E_{kin} + E_{pot} + U$

- The First Law of Systems Engineering (ca. 2015)
  - Given a fixed set of functional requirements and associated performance levels, the total complexity of a system is conserved; complexity can be traded between its components and its interfaces and topology but cannot be decreased beyond a minimum level.
    - $C_{tot} = C_1 + C_2*C_3 + W$

Space Shuttle Lifetime Cost (1971-2011)



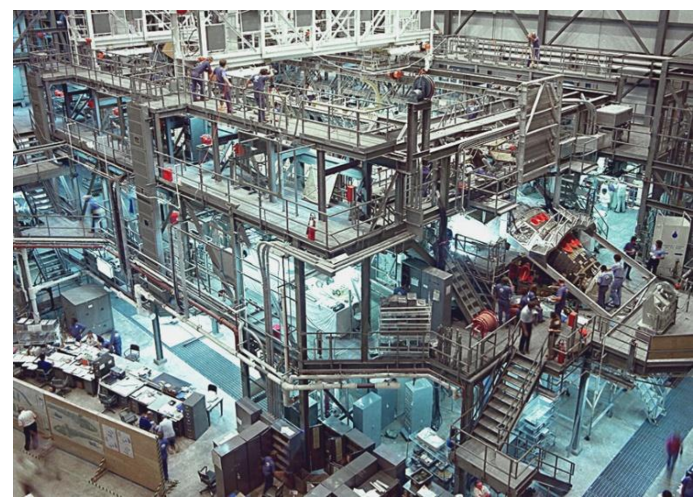
Roger Pielke Jr & Radford Byerly, Shuttle programme lifetime cost, *Nature* 472, 38 (07 April 2011)



- 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

What we wanted

What we got



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 is based on Graph Energy
  - Satisfies Weyuker's criteria (1998)
  - C = C1+ C2*C3; C3: Graph Energy is a measure of topological complexity
- **Better complexity-based management**
  - P-Point is a critical transition point
  - Critical nodal degree $\langle k \rangle_{cr}=6$
  - Iso-complexity based budgeting with clear targets
- First Law of Systems Engineering → Conservation of Complexity