The Structure and Value of Modularity in Software Design

Kevin J. Sullivan,
William G. Griswold

1

Outline

2

Introduction

The Goals of Software Design:

- Increase product quality
- Maximize return on investment

The value of a design can be measured relative to advancing these goals

3

Introduction

Modularity in Software Design:

The modularization chosen has a considerable impact on the development process over time

4

Introduction

Difficult Questions:

- Which of the available modularizations is best?
- At what point should one be selected in the face of uncertainty about the first question?

5

Introduction

Design for Changeability:

- Delay difficult decisions
- The changeable design creates options
- These options have the potential to create value

6

Background

Baldwin and Clark's Theory:

- Modularity adds value in the form of *real options*
- An option provides the right to make an investment in the future
- Example: stock options

  The right to purchase company stock at a fixed price within a specified time period

7

Background

Electronic Data Systems EDS (NYSE)
Last Price: $19.95
(Data as of 09/07/2004)

($19.95 - $60.00) x 1000 = **-$40,050**

8

Background

- There is no need to exercise negative options
- Each option represents either a positive change or no change

9

Model

Baldwin and Clark's Model:

1. Design Structure Matrix
2. Design Rules and Design Evolution
3. Net Options Value

10

Model

1. **Design Structure Matrix**

- DSM represents the design space in terms of design parameters
- These design parameters represent choices

  Example: algorithm, data structures, interface, etc.
- These choices offer value

11

Model

1. **Design Structure Matrix**

   Dependencies:
   B depends on A
   B depends on C
   C depends on B

	A	B	C
A	.		
B	X	.	X
C		X	.

12

Model

1. **Design Structure Matrix**
   Hierarchically dependent - one parameter choice must precede another
   Example: choice of component standard typically precedes choice about component design (B-A)

	A	B	C
A	.		
B	X	.	X
C		X	.

13

Model

1. **Design Structure Matrix**
   Interdependent parameters – parameter choices must be mutually consistent
   Example: algorithm and data structure choices go together (B-C) & (C,B)

	A	B	C
A	.		
B	X	.	X
C		X	.

14

Model

1. **Design Structure Matrix**
   Independent parameters – those parameters that can be changed without coordination
   Example: truly hidden data representation choices

	A	B	C
A	.		
B	X	.	X
C		X	.

15

Model

2. **Design Rules and Evolution**
   A group of *interdependent* design parameters is clustered into a *proto-module* to show that the decisions are managed collectively as a single task.

	A	B	C
A	.		
B	X	.	X
C		X	.

16

Model

2. **Design Rules and Evolution**

- To obtain a modular design dependencies across proto-modules must be broken
- Breaking them requires the use of the modular operator called *splitting*

	I	A	B	C
I	.			
A	X	.		
B	X	X	.	X
C			X	.

17

Model

2. **Design Rules and Evolution**

- A design parameter (I) that decouples interdependent proto-modules is a *design rule*
- Design rules impose constraints that other parameters must respect

	I	A	B	C
I	.			
A	X	.		
B	X		.	X
C			X	.

18

Model

3. Net Options Value

- Modularity in design multiples and decentralizes real options
- Non-modular designs can only be replaced whole
- In finance a portfolio of options is worth more than an option on a portfolio
- This is also how modularity adds value

19

Model

3. Net Options Value

Real Options Formula:

System value = base value + options values

$$\mathbf{V = S_0 + NOV_1 + NOV_2 + \ldots + NOV_n}$$

20

Model

3. Net Options Value

Real Options Formula:

Option value (one per module) = **max**$_k$ {expected value of best of *k* substitutes - cost of creating them - cost of ripple effects}

$$\mathbf{NOV_i = max_{k_i} \{\sigma_i n_i^{1/2} Q(k_i) - c_i n_i k_i - \Sigma\, c_i n_j\}}$$

21

Approach

- Use Parnas's KWIC example to test the potential of DSMs and NOV
- Parnas's modularizations: the traditional *strawman* based design (functional or flowchart-based) and newer information hiding approach
- Create DSMs for both designs
- Create DSM for a proto-modular design

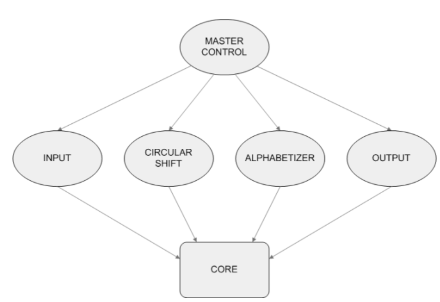
22

Approach

- Create DSM for a proto-modular design
- In the proto-modular design data structures are clustered with algorithms

23

Strawman Design:

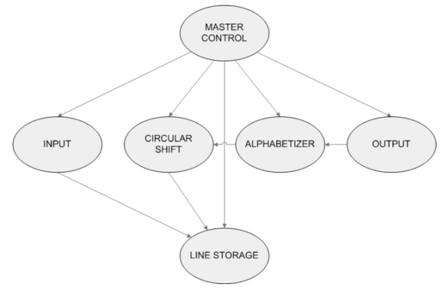


24

DSM for Strawman:

	A	D	G	J	B	E	H	K	C	F	I	L	M
A - Input Type	.												
D - Circ Type		.											
G - Alph Type			.										
J - Out Type				.									
B - In Data					.	X	X						
E - Circ Data					X	.	X						
H - Alph Data					X	X	.						
K - Out Data								.					
C - Input Alg	X				X				.				
F - Circ Alg		X			X	X				.			
I - Alph Alg			X		X	X	X				.		
L - Out Alg				X	X		X	X				.	
M - Master	X	X	X	X									.

Information Hiding Design:



DSM for Information Hiding:

	N	A	D	G	J	O	P	B	C	E	F	H	I	K	L	M
N - Line Type	.															
A - In Type		.														
D - Circ Type			.													
G - Alph Type				.												
J - Out Type					.											
O - Line Data	X					.	X									
P - Line Alg	X					X	.									
B - In Data		X						.	X							
C - In Alg	X	X						X	.							
E - Circ Data	X		X							.	X					
F - Circ Alg	X		X							X	.					
H - Alph Data	X			X								.	X			
I - Alph Alg	X			X								X	.			
K - Out Data					X									.	X	
L - Out Alg	X				X									X	.	
M - Master	X	X	X	X	X											.

DSM for Proto-Modular:

	A	B	C	D	E	F	G	H	I	J	K	L	M
A - In Type	.												X
B - In Data	X	.	X		X	X		X	X			X	
C - In Alg	X	X	.										
D - Circ Type				.									X
E - Circ Data		X		X	.	X		X	X				
F - Circ Alg		X		X	X	.							
G - Alph Type							.						X
H - Alph Data		X			X		X	.	X			X	
I - Alph Alg		X			X		X	X	.				
J - Out Type										.			X
K - Out Data										X	.	X	
L - Out Alg		X						X		X	X	.	
M - Master	X			X			X			X			.

Extension

- DSM does not represent the impact of environment on design parameters
- Extend DSM to include *environment parameters*
- Environment parameters are not controlled by the designers
- Create an EDSM for each design

EDSM for Strawman:

	X	Y	Z	A	D	G	J	B	E	H	K	C	F	I	L	M
X - Computer	.															
Y - Corpus	X	.	X													
Z - User	X		.													
A - In Type				.												
D - Circ Type					.											
G - Alph Type						.										
J - Out Type							.									
B -In Data	X	X						.	X	X						
E - Circ Data	X	X						X	.	X						
H - Alph Data	X	X						X	X	.						
K - Out Data	X	X									.					
C - In Alg	X	X		X				X				.				
F - Circ Alg		X	X		X			X	X				.			
I - Alph Alg	X	X	X			X		X	X	X				.		
L - Out Alg	X	X					X	X		X	X				.	
M - Master			X	X	X	X	X									.

EDSM for Information Hiding:

	X	Y	Z	N	A	D	G	J	O	P	B	C	E	F	H	I	K	L	M
X - Computer	.																		
Y - Corpus	X	.	X																
Z - User	X		.																
N - Line Type				.															
A - In Type					.														
D - Circ Type						.													
G - Alph Type							.												
J - Out Type								.											
O - Line Data	X	X		X					.	X									
P - Line Alg	X	X		X					X	.									
B - Input Data	X	X			X						.	X							
C - Input Alg		X	X	X	X						X	.							
E - Circ Data	X	X		X		X							.	X					
F - Circ Alg		X	X	X		X							X	.					
H - Alph Data	X	X		X			X								.	X			
I - Alph Alg	X	X	X	X			X								X	.			
K - Out Data	X	X						X									.	X	
L - Out Alg	X	X		X				X									X	.	
M - Master			X	X	X	X	X	X											.

EDSM for Proto-Modular:

	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
X - Computer	.															
Y - Corpus	X	.	X													
Z - User	X		.													
A - In Type				.												X
B - Input Data	X	X		X	.	X		X	X		X	X			X	
C - Input Alg	X	X		X	X	.										
D - Cir Type							.									X
E - Circ Data	X	X			X		X	.	X		X	X				
F - Circ Alg		X	X		X		X	X	.							
G - Alph Type										.						X
H - Alph Data	X	X			X			X		X	.	X			X	
I - Alph Alg	X	X	X		X			X		X	X	.				
J - Out Type													.			X
K - Out Data	X	X											X	.	X	
L - Out Alg	X	X			X						X		X	X	.	
M - Master			X	X			X			X			X			.

Analysis

NOV-Based Analysis of KWIC:

- N is the number of design parameters in a given design
- For proto-modular and strawman N=13
- For information hiding N=16
- Given a module of P parameters it has a complexity of p/N
- The value of one experiment on an unmodularized design, $\boldsymbol{\sigma N^{1/2}Q(1)=1}$, is the value of the original system

Analysis

- The design cost **c=1/N** of each design parameter is the same, and the cost to redesign the whole system is **cN=1**
- The visibility cost of module **i** of size **N** is $\boldsymbol{Z_i = \Sigma_{i \text{ sees } i} Cn}$
- One experiment on an unmodularized system breaks even: $\boldsymbol{\sigma N^{1/2}Q(1)=0}$

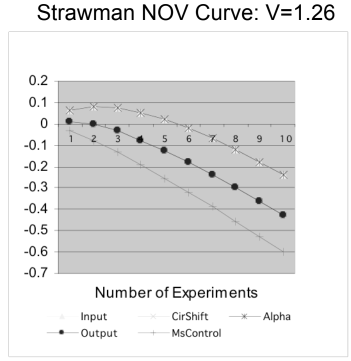
Analysis

- The values used for σ and **Z**

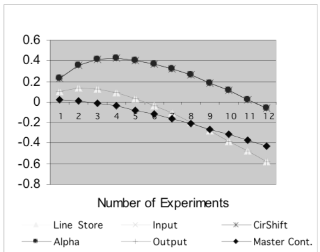
  (This represents the assumed technical potential and visibility)

	Strawman		*Info Hiding*	
Module Name	***sigma***	***Z***	***sigma***	***Z***
Design Rules	2.5	1	0	1
Line Storage	NA	NA	1.6	0
Input	1.25	0	2.5	0
Circular Shift	1.25	0	2.5	0
Alphabetizing	1.25	0	2.5	0
Output	0.8	0	1.6	0

Strawman NOV Curve: V=1.26



Information Hiding NOV Curve: V=2.56



37

Analysis

The results of the NOV calculations:

- Proto-Modular: 1 (base)
- Strawman: 1.26
- Information Hiding: 2.56

38

Conclusion

- The results agree with Parnas's conclusion
- The environment parameters extension to Baldwin and Clark's model proves useful
- The estimates for sigma could use additional justification

39

Conclusion

Future Work:

- Determine what proxies can support estimates for technical potential
- Produce evaluations using real projects to give realistic scale and access to data for estimating NOV
- Develop and evaluate a tool for large hierarchical EDSMs
- Explore a generalized concept of interface comprising all of the dependencies that cross module boundaries

40

Resources

- Sullivan, Kevin. "Strategic Software Design." http://sunset.usc.edu/events/2001/dod/Event/dod_presentations/4.5-Sullivan.ppt
- Cai, Yuanfang. "The Structure and Value of Modularity in Design." http://www.cs.virginia.edu/posters/cai.ppt
- Everett, Jim. "The Structure and Value of Modularity in Design." http://www.cs.wm.edu/~coppit/csci780-fall2003/presentations/07-structure-and-value.pdf

41