Socio-Technical Coordination

James Herbsleb
jdh@cs.cmu.edu
Mars Climate Orbiter
“evidence of inadequate communications between the project elements, including the development and operations teams, the operations and navigation teams, the project management and technical teams, and the project and technical line management.”

-- Report of the Mishap Investigation Board
Shared Memory

Component 1

Development Teams
8,000 Miles

Component 2
Coordination Failures

Managing dependencies among tasks.
-- Malone and Crowston (1994)
A Brief, Selective, and Biased History of Coordination

How do we coordinate work in software engineering?

We’ve applied an astonishing variety of techniques.
Paleozoic Era
“The ultimate method for managing . . . activity with a small group of 10 or 20 people is 10 hours of meetings a day. And then you go work 5 hours.”

What the Data Showed

“There are seven unique personal contacts per day on average, representing continuing interactions.”

“Direct observation showed us that developers spend about 75 minutes per day in unplanned interpersonal interactions.”

Figure 7. Messages sent and received across four media types. The figure shows the number of messages sent and received, by media type and according to whether they were received (r) or initiated (s). We applied a square-root transformation to stabilize the variance. Each box contains data on all seven study subjects across five days of observation per subject.
Not Just Meetings . . .
There are also methods to the madness
The Waterfall Model

Requirements

Design

Implementation

Code-and-fix

Final product

Surprise!
Not Just Methods . . . Processes!
Software Process

SOFTWARE PROCESSES ARE SOFTWARE TOO

Leon Osterweil

University of Colorado Boulder, Colorado USA


The major theme of this meeting is the exploration of the importance of a process as a vehicle for improving both the quality of software products and the way in which we describe a class or set of objects related to each other by virtue of the fact that they are all activities which follow the dictated behavior. We shall have reason to return to this point later in this presentation. For now we should return to our consideration of the intui-
Maturity Framework 1988

- Initial: Basic management control
- Repeatable: Process definition
- Defined: Process measurement
- Managed: Process control
- Optimizing: Optimal software development practices

Not Just Process . . .

• Don’t forget product structure!
Modularity and information hiding

Mesozoic Era
Empirical Studies

Models of Development
How to distribute work across global sites.

Best Practices
- Planning Travel
- Establishing Liaisons
- Building Trust
- Communication Etiquette
- Preventing Delay
- Using Commercial Tools

Research Team

Global Development Solutions

New Products

Team

Codes

Test

Design

Tools
- TeamPortal
- Rear View Mirror
- CalendarBot
- Experience Browser

Planning Travel
Establishing Liaisons
Building Trust
Communication Etiquette
Preventing Delay
Using Commercial Tools
Expertise Browser


Organizational Models

Subsystem

LMTH

TMNH

ROI

IOP

Requirements

Requirements

Process

Design

Code

Test

Release

N

N-1

N-2

N-3

Cenozoic Era
Meeting Innovation
So Many Techniques . . .

“Sometimes the magic works and sometimes it doesn’t.”
– Little Big Man

This is the history of technology and the evolution of useful knowledge.
THE GIFTS OF ATHENA

HISTORICAL ORIGINS OF THE KNOWLEDGE ECONOMY

JOEL MOKYR
## History of Useful Knowledge

<table>
<thead>
<tr>
<th>Technique</th>
<th>Epistemic Base</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Make iron (from 2000 BC)</td>
<td>• Metallurgy</td>
</tr>
<tr>
<td>• Mix ore, charcoal</td>
<td>• Eliminate phosphorus</td>
</tr>
<tr>
<td>• Apply heat</td>
<td>• Add carbon at right time</td>
</tr>
<tr>
<td>• Pour when ready</td>
<td>• Reduce oxygen</td>
</tr>
<tr>
<td></td>
<td>• Siemens Martin process (1865)</td>
</tr>
</tbody>
</table>
History of Useful Knowledge

Technique
- Analgesic (1763)
  - Ingest willow bark
  - Pain relief
  - Side effects

Epistemic Base
- Chemistry
  - Salicin
  - Explore related compounds
  - Salicylic acid (1835)
Future Useful Knowledge

Techniques
• Meetings
• Communication tools
• Processes
• Tweets
• Etc.

Epistemic Base
• Theory of Coordination
Where to Start?

- Product modularity, task modularity, mirroring
  - Baldwin (2000); Conway (1968); Parnas (1974); Sosa & Eppinger (2004); Colfer & Baldwin (2010)
- Collaboration over distance
  - Olson & Olson (2000); Olson, Malone, & Smith (2001); Olson & Teasley (1996)
- Implicit and explicit coordination
- Interdisciplinary theory of coordination
  - Malone and Crowston (1994)
- Social network analysis
  - Krackhardt & Carley (1998)
Socio-Technical Coordination

Network of decisions . . .

. . . establishes a coordination problem . . .

. . . that the organization must solve.
Expressed More Formally . . .


Example: Modularity and Teams
Example: Modularity and Teams
Socio-Technical Coordination

Network of decisions . . .

. . . establishes a coordination problem . . .

. . . that the organization must solve.
Socio-Technical Coordination

Network of decisions . . .

. . . establishes a coordination problem . . .

. . . that the organization must solve.

Decisions and Constraints

Prior work history
Socio-Technical Coordination

Network of decisions . . .

. . . establishes a coordination problem . . .

. . . that the organization must solve.

Create communication channels
Research Challenges

Measure structure of network

Decision network structure

Measure coordination techniques

Coordination techniques

Congruence between decision network and coordination techniques

Compute congruence

Bugginess

Productivity

Measure effects

+ -
Decision Constraint Matrix (DC)

D1  D2  D3
File 1

D4  D5  D6
File 2

D7  D8  D9
File 3

...  Dn

D1  1  1
1  1
1

D2  1
1
1

D3  1
1
1

D4  1
1
1

D5  1
1
1

D6  1
1
1

D7  1
1
1

D8  1
1
1

D9  1
1
1

Dn  1
1
1

Decisions and Constraints

<table>
<thead>
<tr>
<th></th>
<th>$F_1$</th>
<th>$F_2$</th>
<th>$F_3$</th>
<th>$F_4$</th>
<th>$F_5$</th>
<th>$F_6$</th>
<th>$F_7$</th>
<th>$F_8$</th>
<th>$F_9$</th>
<th>...</th>
<th>$F_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F_2$</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F_3$</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F_4$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F_5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F_6$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F_n$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
### Decision Assignment Matrix (DA)

<table>
<thead>
<tr>
<th>File 1</th>
<th>File 2</th>
<th>File 3</th>
<th>...</th>
<th>D&lt;sub&gt;n&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>D&lt;sub&gt;1&lt;/sub&gt;</td>
<td>D&lt;sub&gt;2&lt;/sub&gt;</td>
<td>D&lt;sub&gt;3&lt;/sub&gt;</td>
<td>D&lt;sub&gt;4&lt;/sub&gt;</td>
<td>D&lt;sub&gt;5&lt;/sub&gt;</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

**Decision Assignment Matrix (DA)**

<table>
<thead>
<tr>
<th>F₁</th>
<th>F₂</th>
<th>F₃</th>
<th>F₄</th>
<th>F₅</th>
<th>F₆</th>
<th>F₇</th>
<th>F₈</th>
<th>F₉</th>
<th>...</th>
<th>Fₙ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Carnegie Mellon University

**DA**

<table>
<thead>
<tr>
<th></th>
<th>$F_1$</th>
<th>$F_2$</th>
<th>$F_3$</th>
<th>$F_4$</th>
<th>$F_5$</th>
<th>$F_6$</th>
<th>$F_7$</th>
<th>$F_8$</th>
<th>$F_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>F₁</td>
<td>F₂</td>
<td>F₃</td>
<td>F₄</td>
<td>F₅</td>
<td>F₆</td>
<td>F₇</td>
<td>F₈</td>
<td>F₉</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
$C_R$ Coordination Requirements

\[
DA \times DC \times DA^T = C_R
\]
Coordination Activities and Congruence

<table>
<thead>
<tr>
<th>Coordination Requirements ($C_R$)</th>
<th>Actual Coordination ($C_A$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Team structure</td>
<td>1 1</td>
</tr>
<tr>
<td>Geographic location</td>
<td>1</td>
</tr>
<tr>
<td>Use of chat</td>
<td>1 1</td>
</tr>
<tr>
<td>On-line discussion</td>
<td>1</td>
</tr>
</tbody>
</table>

Congruence = proportion of nonzero cells in $C_R$ that are also nonzero in $C_A$

### Impact on Productivity

**Table 2: Results from OLS Regression of Effects on Task Performance († p < 0.10, * p < 0.05, ** p < 0.01).**

<table>
<thead>
<tr>
<th></th>
<th>Model I</th>
<th>Model II</th>
<th>Model III</th>
<th>Model IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>2.987**</td>
<td>3.631**</td>
<td>1.572*</td>
<td>1.751*</td>
</tr>
<tr>
<td>Dependency</td>
<td>0.897*</td>
<td>0.653*</td>
<td>0.784*</td>
<td>0.712*</td>
</tr>
<tr>
<td>Priority</td>
<td>-0.741*</td>
<td>-0.681*</td>
<td>-0.702*</td>
<td>-0.712*</td>
</tr>
<tr>
<td>Re-assignment</td>
<td>0.423*</td>
<td>0.487*</td>
<td>0.304*</td>
<td>0.324*</td>
</tr>
<tr>
<td>Customer MR</td>
<td>-0.730</td>
<td>-0.821</td>
<td>-0.932</td>
<td>-0.903</td>
</tr>
<tr>
<td>Release</td>
<td>-0.154*</td>
<td>-0.137*</td>
<td>-0.109*</td>
<td>-0.098*</td>
</tr>
<tr>
<td>Change Size (log)</td>
<td>1.542*</td>
<td>1.591*</td>
<td>1.428*</td>
<td>1.692*</td>
</tr>
<tr>
<td>Team Load</td>
<td>0.307*</td>
<td>0.317*</td>
<td>0.356*</td>
<td>0.374*</td>
</tr>
<tr>
<td>Programming Experience</td>
<td>-0.062*</td>
<td>-0.162*</td>
<td>-0.117*</td>
<td>-0.103*</td>
</tr>
<tr>
<td>Tenure</td>
<td>-0.269*</td>
<td>-0.265*</td>
<td>-0.239*</td>
<td>-0.248*</td>
</tr>
<tr>
<td>Component Experience (log)</td>
<td>-0.143*</td>
<td>-0.143*</td>
<td>-0.195*</td>
<td>-0.213*</td>
</tr>
<tr>
<td><strong>Structural Congruence</strong></td>
<td><strong>-0.526</strong></td>
<td><strong>-0.483</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Geographical Congruence</strong></td>
<td><strong>-0.317</strong></td>
<td><strong>-0.312</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MR Congruence</strong></td>
<td><strong>-0.189</strong></td>
<td><strong>-0.129</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>IRC Congruence</strong></td>
<td><strong>-0.196</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interaction: ReleaseX Structural Congruence</td>
<td>0.007</td>
<td>0.009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interaction: ReleaseXGeographical Congruence</td>
<td>-0.013</td>
<td>-0.017</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interaction: Release X MR Congruence</td>
<td>-0.009†</td>
<td>-0.011†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interaction: Release X IRC Congruence</td>
<td>-0.017*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>809</td>
<td>809</td>
<td>1983</td>
<td>1983</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.787</td>
<td>0.872</td>
<td>0.756</td>
<td>0.854</td>
</tr>
</tbody>
</table>

## Impact on Bugginess

<table>
<thead>
<tr>
<th></th>
<th>Model I</th>
<th>Model II</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOC (log)</td>
<td>1.125**</td>
<td>1.136**</td>
</tr>
<tr>
<td>Avg. Lines Changed (log)</td>
<td>1.128**</td>
<td>1.121**</td>
</tr>
<tr>
<td>Number Logical Dep. (log)</td>
<td>2.219**</td>
<td>2.109**</td>
</tr>
<tr>
<td>Clustering Logical Dep. (log)</td>
<td>0.012**</td>
<td>0.012**</td>
</tr>
<tr>
<td>Coordination Req. Dep. (log)</td>
<td>2.187**</td>
<td>1.962**</td>
</tr>
<tr>
<td>Structural Congruence</td>
<td></td>
<td>0.281*</td>
</tr>
<tr>
<td>Geographical Congruence</td>
<td></td>
<td>0.317</td>
</tr>
<tr>
<td>MR Congruence</td>
<td></td>
<td>0.209**</td>
</tr>
<tr>
<td>IRC Congruence</td>
<td></td>
<td>0.271**</td>
</tr>
</tbody>
</table>

**Model Fit**

<table>
<thead>
<tr>
<th></th>
<th>Model I</th>
<th>Model II</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>3980</td>
<td>3980</td>
</tr>
<tr>
<td>Model $\chi^2$</td>
<td>1663**</td>
<td>1859**</td>
</tr>
<tr>
<td>df</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Deviance Explained</td>
<td>0.302</td>
<td>0.335</td>
</tr>
<tr>
<td>Model Comparison $\chi^2$</td>
<td>--</td>
<td>196.24**</td>
</tr>
</tbody>
</table>

(+ $p < 0.10$; * $p < 0.05$; ** $p < 0.01$)
Research Challenges

Measure structure of network

- Decision network structure
- Coordination techniques

Congruence between decision network and coordination techniques

Compute congruence

Measure effects

Bugginess

Productivity

Measure coordination techniques
Selected Work on Congruence in Software Engineering

- Kwan, Schröter, & Damian (2011)
  - Examined the relationship of congruence to build success.
- Kwan & Damian (2011)
  - Developed an aggregated congruence measure based on multiple awareness mechanisms
- Avritzer, Paulish, Cai, & Sethi (2010)
  - DSMs to represent architectural dependencies and social communication networks, compute congruence
- Kwan, Schröter, & Damian (2009)
  - Developed a weighted congruence measure
- Sarma et al (2009)
  - Designed Tesseract for visualizing social networks, dependency networks, and congruence
  - Used socio-technical network measures to predict failure-prone components
- Bolici, Howison, & Crowston (2009)
  - Examined stygmergy as a mechanisms for establishing congruence in open source projects
- Valetto, Chulani, & Williams (2009)
  - Analyzed costs and risks of different approaches to close congruence gaps
  - Develop a graph-theoretic algorithm for computing congruence
Theory and Social Coding

Repositories: clumps of decisions
Decision owners ≅ git access
Constraints ≅ “uses” relation

... establishes a coordination problem ...

This problem is typical of open source
Social Coding

Repositories: clumps of decisions
Decision owners ≅ git access
Constraints ≅ “uses” relation

. . . establishes a coordination problem . . .

. . . that the organization must solve.

Tool Affordances
Adjustment & intervention
Power asymmetries
Hard power, soft power
Audience and accountability
We’re at the Beginning

- Not just code!
- Popular frameworks, libraries, APIs
- Temporal order, pace of decisions
- Predict early, use in planning

- Match decision networks with techniques
- How to plan, correct, adjust

- What is the full set of techniques?
- Substitute, complement, compose?
- Role of new, social and transparent media?
Takeaways

• We have a great many coordination techniques, what we need is a theory
  • We have made a start – we are at the beginning

• A good theory will incorporate the social and the technical
  • Either alone is “one hand clapping”
An Observation

- Human dimension increasingly taken into account in software engineering, e.g.
  - API usability
  - End user programming
  - Tools and environments
Another Observation

- Practical value versus enduring value
  - We need a portfolio of research
  - Validation of a technique can have immediate value
  - Theory development will not yield immediate practical results, yet
- In the long run, “There is nothing so practical as a good theory.”
  - Kurt Lewin (1959)
Yet Another Observation

- Software engineering research is based in behavioral science as strongly as it is based in computer science.
Questions?

Collaborators

- Kathleen Carley
- Marcelo Cataldo
- Laura Dabbish
- Audris Mockus
- Anita Sarma
- Colleen Stuart
- Jason Tsay
- Patrick Wagstrom