Socio-Technical Coordination

James Herbsleb
jdh@cs.cmu.edu
“evidence of inadequate communications between the project elements, including the development and operations teams, the operations navigation and operations 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 Build

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 software 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
  - Optimizing

Not Just Process . . .

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

Mesozoic Era
Bell Labs Collaboratory

New Products

Global Development Solutions

Research Team

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

Tools

- TeamPortal
- Rear View Mirror
- CalendarBot
- Experience Browser

Design → Code → Test

Best Practices

- Establishing Liaisons
- Building Trust
- Communication Etiquette
- Preventing Delay
- Using Commercial Tools
Instant Messaging
Rear View Mirror


Cenozoic Era
Meeting Innovation
Today
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

**Technique**
- Make iron (from 2000 BC)
  - Mix ore, charcoal
  - Apply heat
  - Pour when ready

**Epistemic Base**
- Metallurgy
  - Eliminate phosphorus
  - Add carbon at right time
  - Reduce oxygen
  - Siemens Martin process (1865)
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

$M_1 \hspace{1cm} M_2 \hspace{1cm} M_3$
Example: Modularity and Teams

$M_1$, $M_2$, $M_3$, $T_1$, $T_2$, $T_3$
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.
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

Compute congruence

- Congruence between decision network and coordination techniques
- Measure effects

Measure coordination techniques

- Productivity
- Bugginess

Coordination techniques
<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>\ldots</th>
<th>$F_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_1$</td>
<td></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>$F_2$</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F_3$</td>
<td></td>
<td>1</td>
<td></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>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F_5$</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$F_6$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>\ldots</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></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></td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

**Carnegie Mellon University**
Carnegie Mellon University

Decision Assignment Matrix (DA)

<table>
<thead>
<tr>
<th></th>
<th>File 1</th>
<th>File 2</th>
<th>File 3</th>
<th>...</th>
<th>D_n</th>
</tr>
</thead>
<tbody>
<tr>
<td>D_1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D_2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D_3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D_4</td>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>D_5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D_6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D_7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D_8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D_9</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>
</tr>
<tr>
<td>D_n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### DA

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>F_1</td>
<td>F_2</td>
<td>F_3</td>
<td>F_4</td>
<td>F_5</td>
<td>F_6</td>
<td>F_7</td>
<td>F_8</td>
<td>F_9</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>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>1</td>
<td>1</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>1</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>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</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></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

Carnegie Mellon University
\[ \text{DA} \times \text{DC} \times \text{DA}^T = \mathbf{C}_R \]

\[ \mathbf{C}_R \] Coordination Requirements
Coordination Activities and Congruence

**Coordination Requirements** ($C_R$)

- Team structure
- Geographic location
- Use of chat
- On-line discussion

**Actual Coordination** ($C_A$)

$Congruence = \text{proportion of nonzero cells in } C_R \text{ that are also nonzero in } C_A$
## 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>Model</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>
</tbody>
</table>

**Structural Congruence** -0.526* -0.483*

**Geographical Congruence** -0.317* -0.312*

**MR Congruence** -0.189* -0.129*

**IRC Congruence** -0.196* __--__

Interaction: ReleaseX Structural Congruence 0.007 0.009

Interaction: ReleaseX Geographical Congruence -0.013 -0.017

Interaction: Release X MR Congruence -0.009+ -0.011+

Interaction: Release X IRC Congruence -0.017* __--__

<table>
<thead>
<tr>
<th>N</th>
<th>809</th>
<th>809</th>
<th>1983</th>
<th>1983</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjusted $R^2$</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>
</tbody>
</table>

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Structural Congruence</td>
<td>0.281*</td>
<td></td>
</tr>
<tr>
<td>Geographical Congruence</td>
<td>0.317</td>
<td></td>
</tr>
<tr>
<td>MR Congruence</td>
<td>0.209**</td>
<td></td>
</tr>
<tr>
<td>IRC Congruence</td>
<td>0.271**</td>
<td></td>
</tr>
</tbody>
</table>

<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 χ²</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 χ²</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

Compute congruence

Congruence between decision network and coordination techniques

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
Social Coding

Repositories: clumps of decisions
Decision owners $\approx$ git access
Constraints
“uses” relation
accept code contribution

. . . establishes a coordination problem . . .

. . . that the organization must solve.

Tool Affordances
Signals
Power asymmetries
Negotiation
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”
- 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