Aligning Coordination Behavior with Coordination Needs:

Congruence in Software Development

James D. Herbsleb
School of Computer Science
Carnegie Mellon University

Overview

- Coordination requirements and congruence
- In search of a theory: Distributed Constraint Satisfaction
- Initial test of theory: partial confirmation
- Implications
  - For tools
  - For coordination research

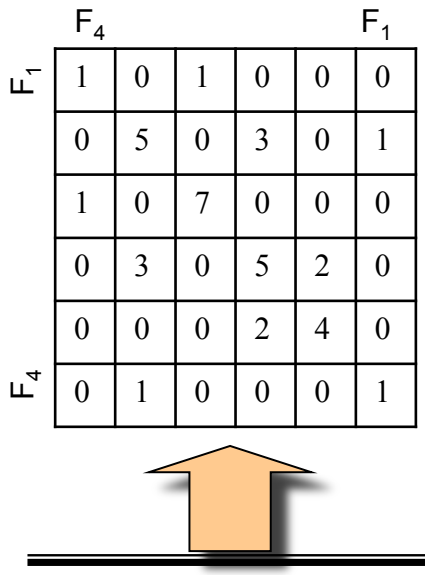
Coordination

- Coordination is managing dependencies among tasks (Malone & Crowston)
- Coordination is a central concern in software engineering, e.g.,
  - Modularity
  - Architecture
- Coordination is central concern of work collaboration more generally
- Generally assume modularizing the product design modularizes the tasks

Measuring Coordination Requirements

- Dependencies between files
- Number of times the two files were modified in same Modification Request

Files Modified Together



Measuring Coordination Requirements

- Task assignment
- Number of times each developer modified each file for some unit of work

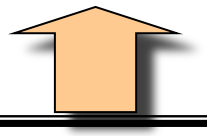
Files Modified Together

	F_4					F_1
F_1	1	0	1	0	0	0
	0	5	0	3	0	1
	1	0	7	0	0	0
	0	3	0	5	2	0
	0	0	0	2	4	0
F_4	0	1	0	0	0	1

X

Developers Modified Files

	F_4					F_1
D_1	1	0	0	1	0	0
	0	1	0	0	0	0
	0	1	0	0	1	0
D_4	0	0	0	1	1	1



Measuring Coordination Requirements

- Transpose of task assignment

Files Modified Together

	F_4					F_1
F_1	1	0	1	0	0	0
	0	5	0	3	0	1
	1	0	7	0	0	0
	0	3	0	5	2	0
	0	0	0	2	4	0
F_4	0	1	0	0	0	1

X

Developers Modified Files

	F_4					F_1
D_1	1	0	0	1	0	0
	0	1	0	0	0	0
	0	1	0	0	1	0
D_4	0	0	0	1	1	1

X

Developers Modified Files (Transpose)

	D_4			D_1
F_1	1	0	0	0
	0	1	1	0
	0	0	0	0
	1	0	0	1
	0	0	1	1
F_4	0	0	0	1

Measuring Coordination Requirements

- Coordination requirements
- Extent to which two developers worked on interdependent files

Files Modified Together

	F_4					F_1
F_1	1	0	1	0	0	0
	0	5	0	3	0	1
	1	0	7	0	0	0
	0	3	0	5	2	0
	0	0	0	2	4	0
F_4	0	1	0	0	0	1

X

Developers Modified Files

	F_4					F_1
D_1	1	0	0	1	0	0
	0	1	0	0	0	0
	0	1	0	0	1	0
D_4	0	0	0	1	1	1

X

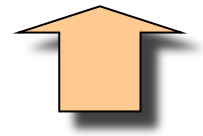
Developers Modified Files (Transpose)

	D_4			D_1
F_1	1	0	0	0
	0	1	1	0
	0	0	0	0
	1	0	0	1
	0	0	1	1
F_4	0	0	0	1

=

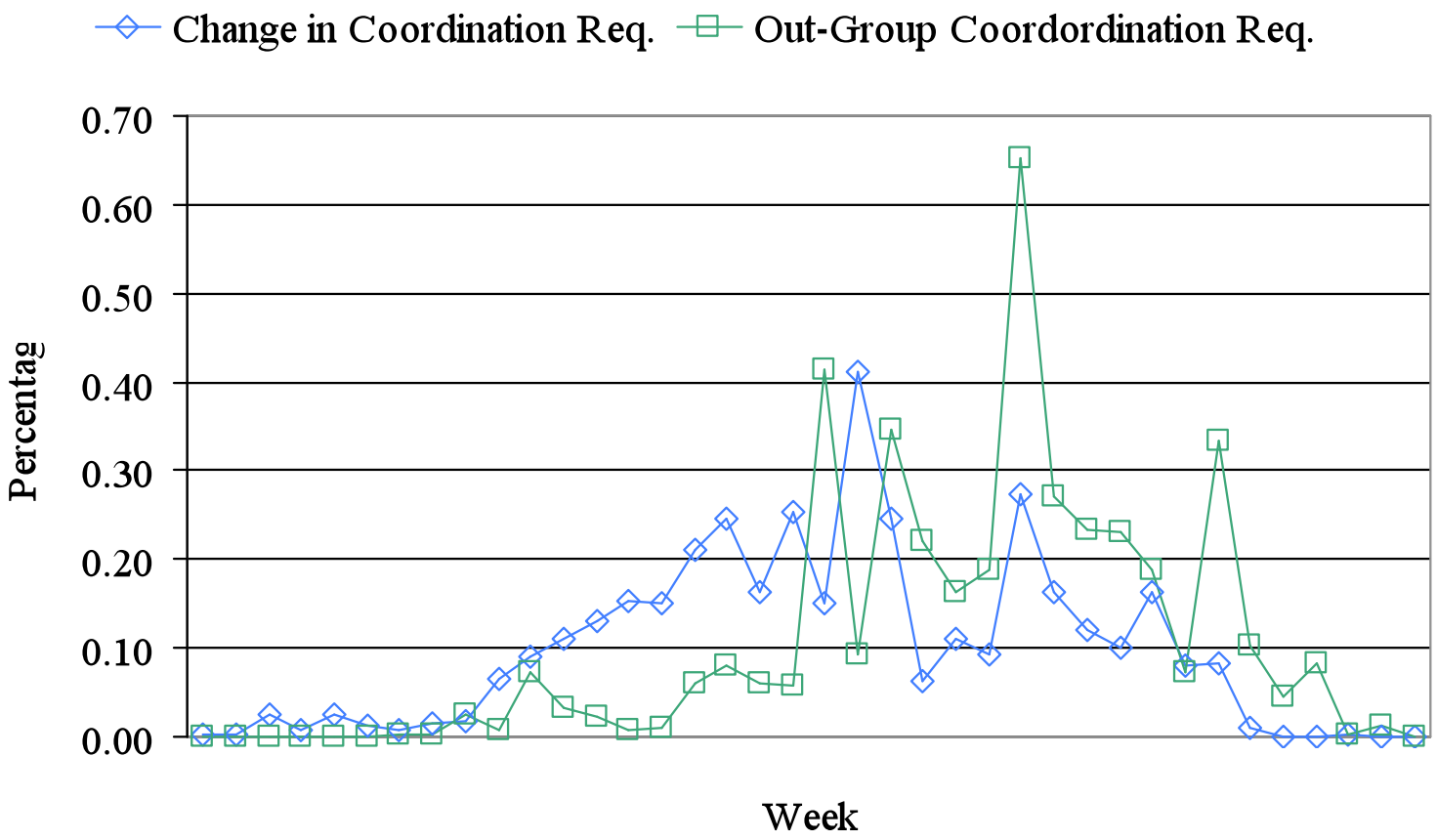
Coordination Requirements

	D_4			D_1
D_1	1	0	0	1
	0	1	0	0
	0	1	1	0
D_4	0	0	0	1



Carnegie Mellon
School of Computer Science

Volatility in Coordination Requirements



Measuring Congruence

Coordination Requirements (C_R)

	W_4			W_1
W_1	1	0	0	1
	0	1	0	0
	0	1	1	0
W_4	0	0	0	1

Actual Coordination (C_A)

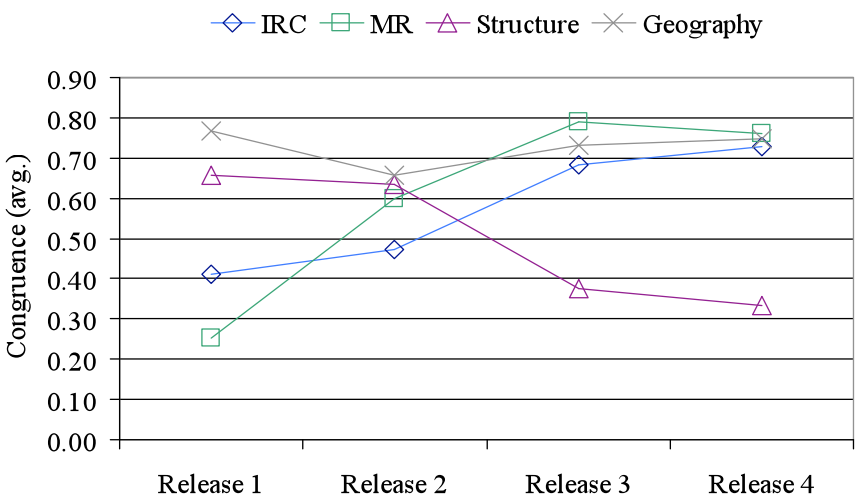
	W_4			W_1
W_1	1	0	0	1
	0	1	0	1
	0	0	1	0
W_4	0	0	0	1

- Team structure
- Geographic location
- Use of chat
- On-line discussion in MR system

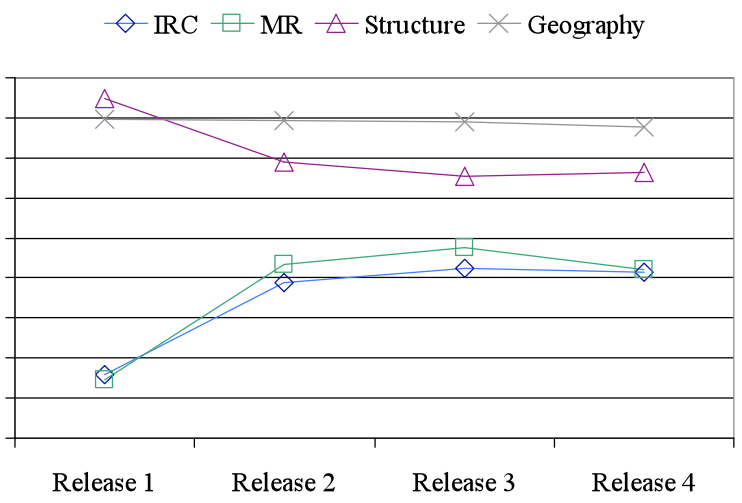
Congruence and Development Speed

- Unit of analysis: Modification Request (MR) (N=1983)
- Constructed regression model
  - Congruence measures as predictors
  - Control variables
  - Resolution time for MR as dependent variable
- Time to complete a work item is reduced by each of the types of congruence
  - Team structure congruence
  - Geographic location congruence
  - Chat congruence
  - On-line discussion congruence

Evolution of Congruence



Top 18 developers



Other 96 developers

Overview

- Coordination requirements and congruence
- In search of a theory: Distributed Constraint Satisfaction
- Initial test of theory: partial confirmation
- Implications
  - For tools
  - For coordination research

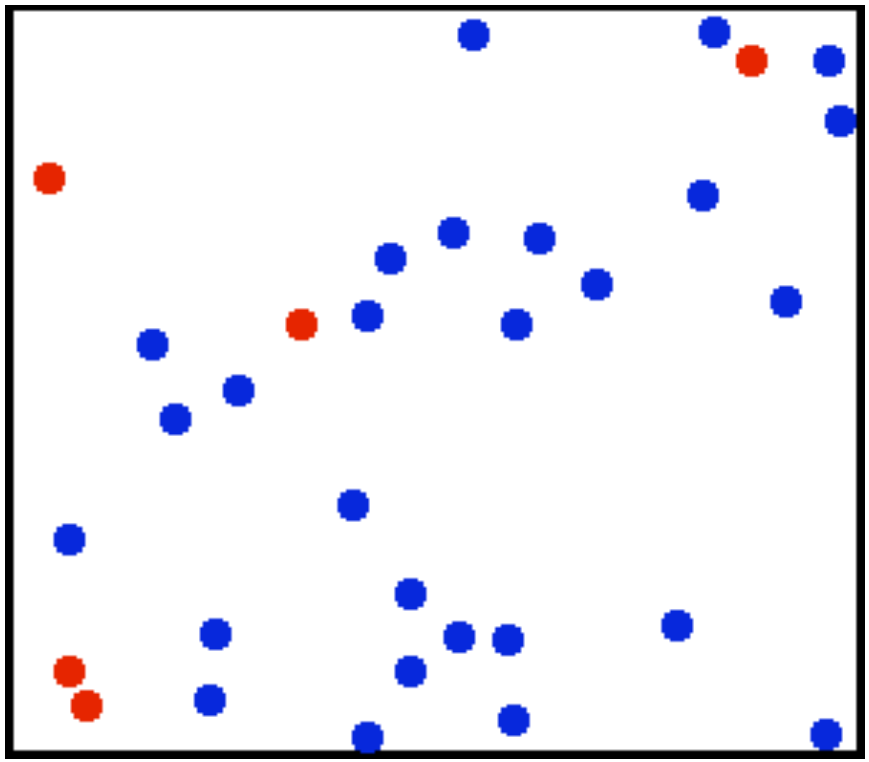
Theoretical Views of Coordination

- Coordination theory (Malone & Crowston)
  - Match coordination problems to mechanisms
  - E.g., resource conflict and scheduling
- Distributed Cognition (Hutchins, Hollan)
  - Computational process distributed over artifacts and people
- Organizational behavior
  - Stylized dependency types, e.g., sequential, pooled
  - Coordination regimens that address each type

Need a Different Approach

- Coordination requirements are generated by rapidly shifting tasks at sub-workflow, micro level
- Not clear that any existing theories apply
- Even if they apply, they do not generate predictions based on micro tasks
- Want to predict observable, macro behavior produced by micro coordination phenomena
- How?

Predicting Macro from Micro:
Kinetic Theory of Gases



Animated gif: Wikipedia

- A gas consists of molecules in constant random motion
- Gas molecules influence each other only by collision
- All collisions between gas molecules are perfectly elastic
- The volume actually occupied by the molecules of a gas is negligibly small

Tom Stretton
http://www2.ucdsb.on.ca/tiss/stretton/CHEM1/gasesx.html

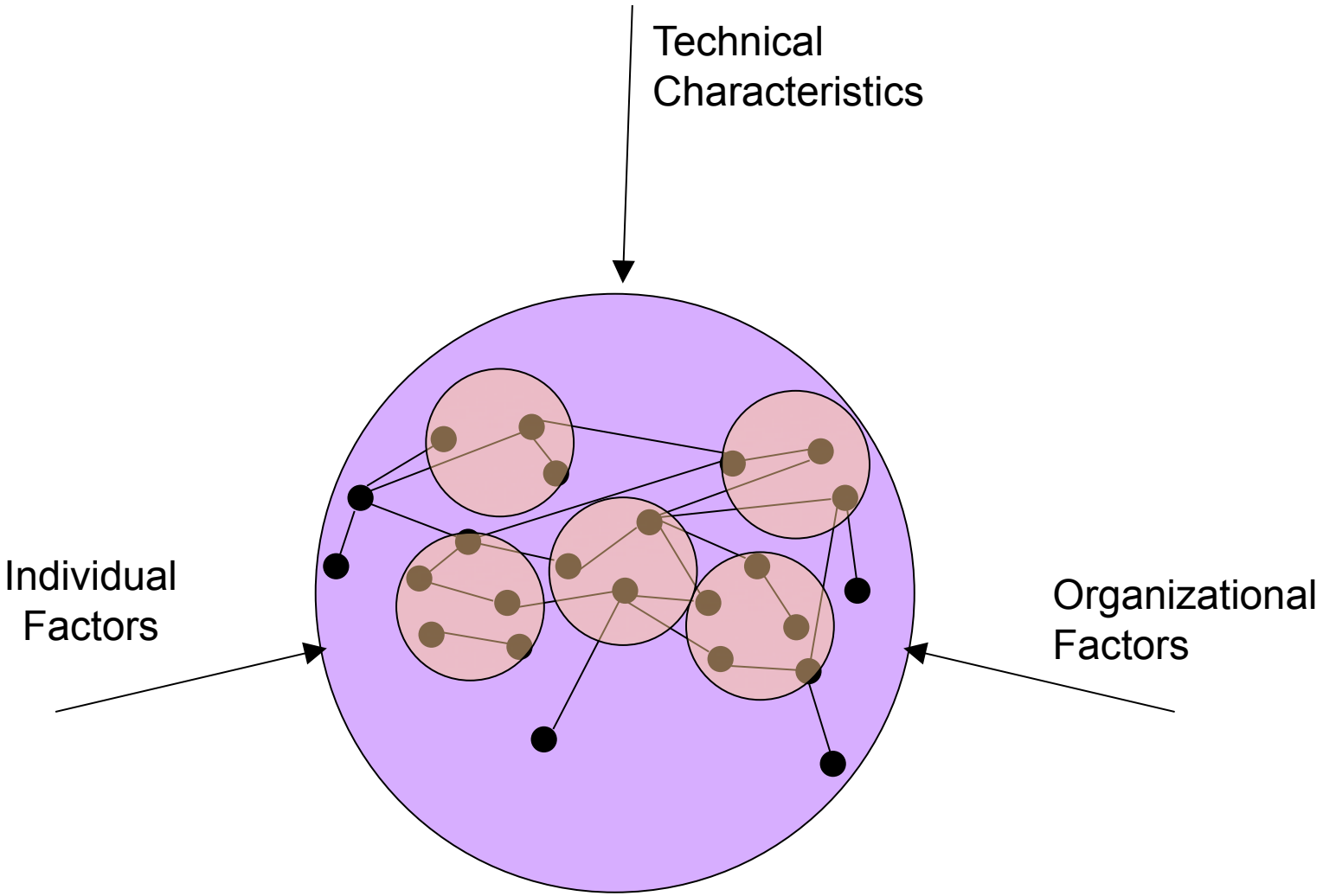
Technical Coordination Modeled as CSP

- Constraint satisfaction problem
  - a project is a large set of mutually-constraining decisions, which are represented as
  - n variables $x_1, x_2, \ldots, x_n$ whose
  - values are taken from finite, discrete domains $D_1, D_2, \ldots, D_n$
  - constraints $p_k(x_{k1}, x_{k2}, \ldots, x_{kn})$ are predicates defined on
  - the Cartesian product $D_{k1} \times D_{K2} \times \ldots \times D_{kj}$.
- Solving CSP is equivalent to finding an assignment for all variables that satisfy all constraints

Formulation of CSP and DCSP taken from Yokoo and Ishida, Search Algorithms for Agents, in G. Weiss (Ed.) *Multiagent Systems*, Cambridge, MA: MIT Press, 1999.

Distributed Constraint Satisfaction

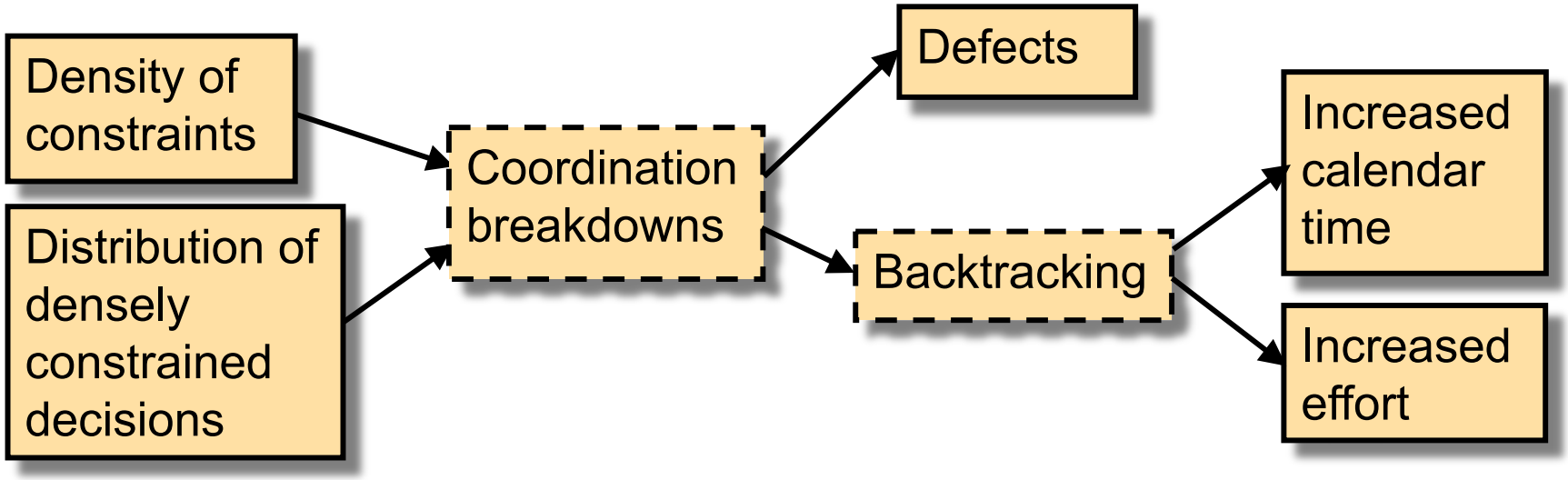
- Each variable x_j belongs to one agent i
- Represented by relation $belongs(x_j,i)$
- Agents only know about a subset of the constraints
- Represent this relation as $known(P_l, k)$, meaning agent k knows about constraint P_l
- Agent behavior determines global algorithm
- For humans, global behavior emerges


Technical
Characteristics
Individual
Factors
Organizational
Factors
Decisions
Constraints
People

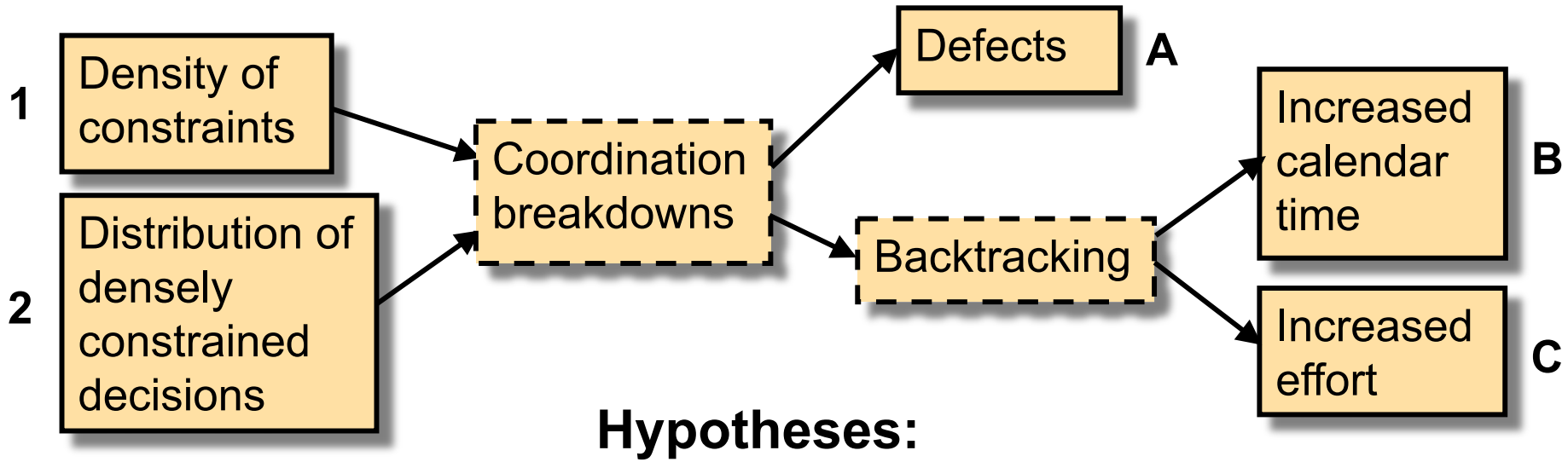
Overview

- Coordination requirements and congruence
- In search of a theory: Distributed Constraint Satisfaction
- Initial test of theory: partial confirmation
- Implications
  - For tools
  - For coordination research

Micro Causes, Macro Effects



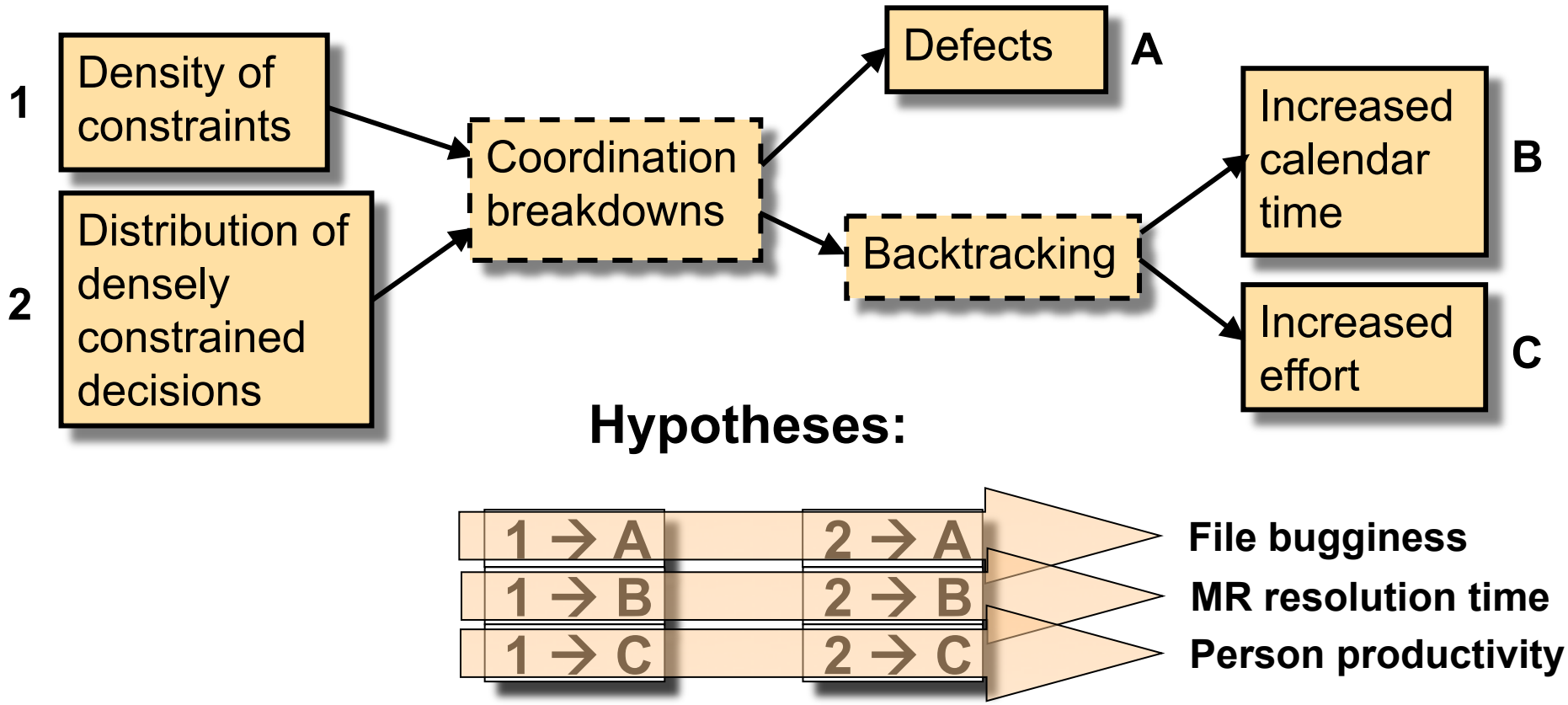
Hypotheses



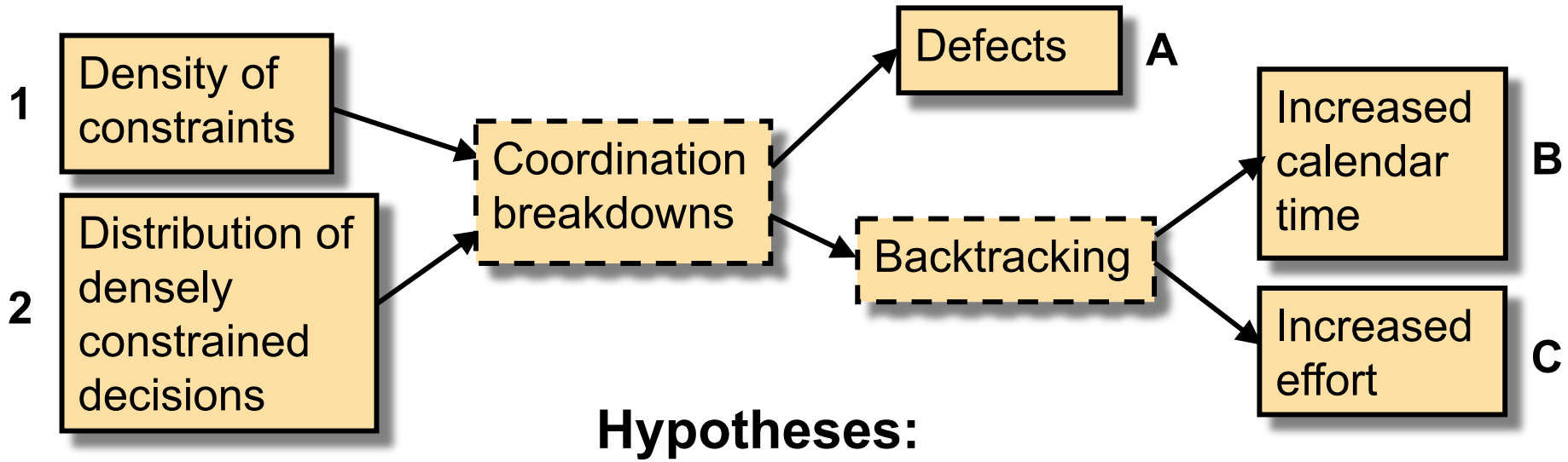
Hypotheses:

1 → A	2 → A
1 → B	2 → B
1 → C	2 → C

Regression Models



Results



Hypotheses:

1 → A	2 → A
1 → B	2 → B
1 → C	2 → C

Overview

- Coordination requirements and congruence
- In search of a theory: Distributed Constraint Satisfaction
- Initial test of theory: partial confirmation
- Implications
  - For tools
  - For coordination research

Implications: Tools

- Just providing self-selected buddy lists is likely not sufficient to support coordination for most people
  - Coordination requirements are too volatile
  - Many people may not know who to select
- Our tools could be doing much more for us in creating and using project data
  - All-encompassing project graph, with all changes to everything, automatic links for pasting, maybe viewing
  - Graph-browsing capabilities to take any node as a start, search across people, artifacts, content history
- Software tools are leading the way here

Implications: Coordination Research

- Coordination requirements are
  - Volatile in short term, evolve over long term
- We need theories that
  - Capture the micro nature of coordination
  - Predict macro scale behavior
- Role of constraint visibility/discoverability
  - May explain difference in “bugginess” effects between call and data dependencies
- There is a complex, time-varying relationship between product modularization and task modularization