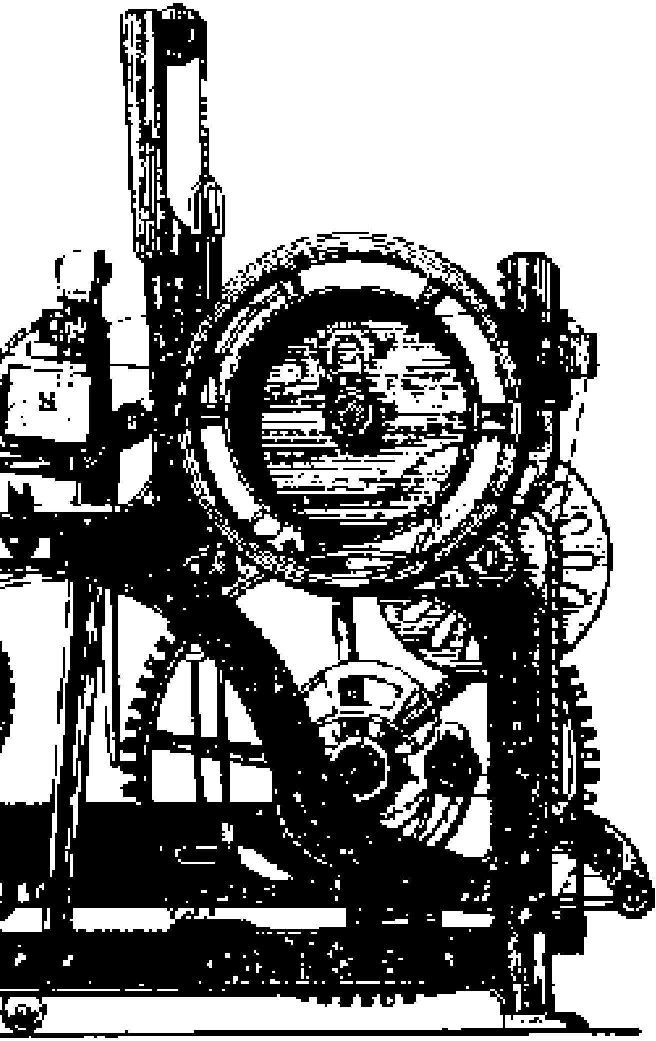
Coordination of Free/Libre Open Source Software Development



Kevin Crowston, Kangning Wei, Qing Li, U. Yeliz Eseryel, and James Howison

Syracuse University
School of Information Studies
crowston@syr.edu
http://floss.syr.edu/

Partially supported by NSF Grants 03–41475, 04–14468 and 05–27457

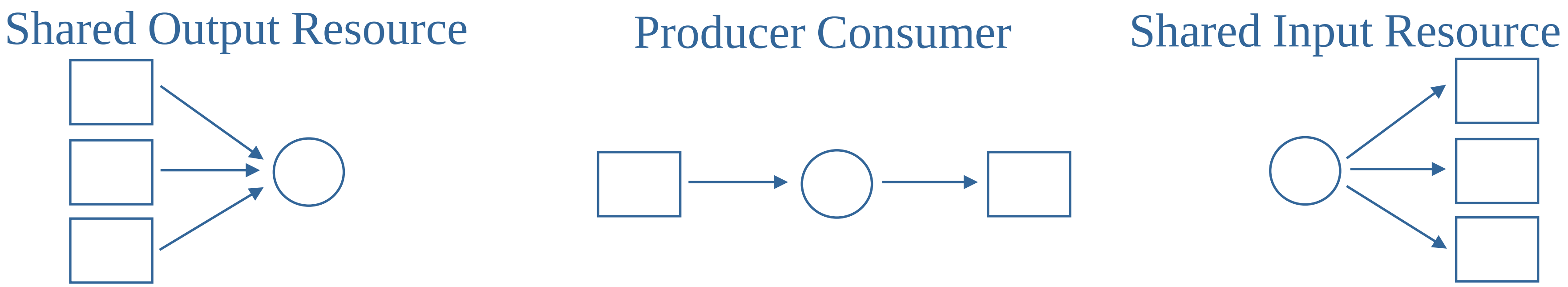
What can we learn from FLOSS?

- Free/libre open source software development has been surprisingly successful

- What can we learn about how to coordinate large numbers of distributed developers?

Coordination theory

- Coordination defined as managing dependencies



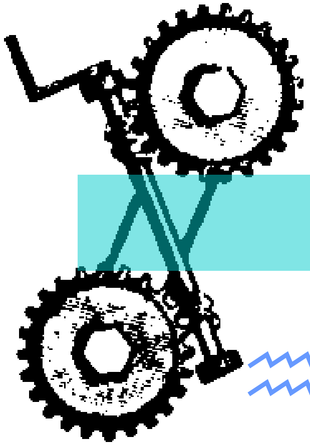
○ Task

□ Resource

→ Task using or creating a resource

Coordination mechanisms

- Dependencies constrain how activities can be performed, requiring additional work (coordination mechanisms)
  - E.g., two programmers working on the same file must avoid overwriting changes
- Maybe a choice of mechanisms to manage a dependency
  - E.g., code ownership, source code control checkout, optimistic concurrency



Example dependencies and coordination mechanisms

- Two activities use the same resource (share dependency) (e.g., same file)
- Two activities create the same resource (common output dependency) (e.g., same patch)
  - Detect and manage overlap
- One activity creates a resource used by the other (flow dependency) (e.g., bug reporting)
  - Manage usability of resource

Specific research questions

- Does FLOSS development use different coordination mechanisms than proprietary development?
- Could those mechanisms be used for proprietary development?

Method

- Inductive coding of FLOSS developer email interactions
- Initial coding scheme based on approach in Crowston & Osborn 2003
- Had to be modified to analyze email rather than observational and interview data
  - Focused on identifying coordination mechanisms and problems rather than dependencies

Data and sample

	EGroupWare	GAIM	Compiere
Programming language	PHP	C	Java
License	GPL	GPL	MPL
Developer count	42	12	44
Data source	egroupware-development	Gaim-devel	Development Chat Forum
Time period	October and November 2004	August and September 2004	January 2001 to November 2002
Messages	665	710	315
Posters (developers)	151 (20)	85 (11)	71 (6)

Results

- Coding system was iteratively developed by 3 coders over several months
- Final system includes 24 categories organized into 3 high-level codes
  - Task-task coordination mechanisms
  - Task-resource coordination mechanisms
  - Resource-resource coordination mechanisms
- Task assignment seemed most interesting
  - Double coded to assess reliability (inter-rater agreement > 0.80)

Findings—Shared output dependencies

- Similar coordination mechanisms for shared output dependency found in both FLOSS and proprietary development teams
  - E.g., marking bug reports as duplicates

    “I suspect that this is the common “cygwin/bin in your path” issue. If you do have cygwin’s bin directory in your path, it causes these symptoms as GAIM tries to load cygwin’s tcl84.dll for the tcl plugin loader. You can read more about it at http://GAIM.sf.net/win32 and in various threads on the forum.”

FLOSS coordination mechanisms

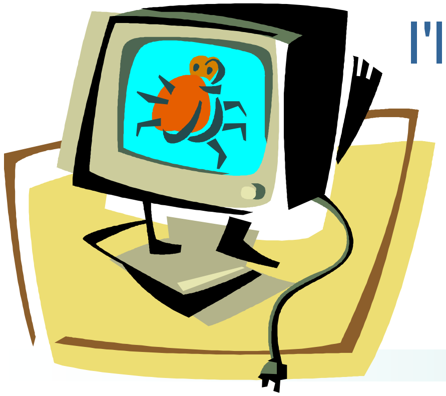
- Avoiding duplicate work also extended to other kinds of development tasks

  “Patches to update the svn tree can be submitted to me, please update this thread with any work you are starting so we can avoid duplication.”

Flow (usability) dependencies

- Common problem in FLOSS and proprietary development is that bug reports are often not usable by programmers
  - Require additional filtering or questions

    “Does Outlook send some sort of error message? What do you mean by crash--does it stop responding, or close? What version of Windows are you running? I'll try to figure this out asap.”



FLOSS coordination mechanisms

- In FLOSS, users can comment on requirements (not just developers) to ensure usability of overall system

  “I have read through this entire thread and seen a lot of suggestions for how to implement this, and as an extensive _user_ of Gaim, let me say that the best one that I have heard so far is the idea of tabbing the buddy list interface.”

Task-programmer dependencies

- Major difference in approach to task assignment
  - Proprietary software team used elaborate system to route bug reports and new work to code owner
  - Other units assigned work to next available programmer

FLOSS coordination mechanisms

- In FLOSS teams, most common form of task assignment was self-assignment

  “Thanx for the wonderful help ... Maybe an idea to make a ‘hello world’ package for on the website ... this kind of standard app .. would be very Helpful for starting eGW developers (like me ;-)) If you’d like, I’ll make the package.”

  “If yes mail me the code or send it to our patch manager on sf.net (http://sf.net/projects /egroupware click on patches) and assign it to me (John).”

Other forms of task assignment

- Ask someone to do something
- Ask an unspecified person

“Can someone please do a brief test, replacing config.php with newconfig.php? If it works for a few people without causing problems, it will help us in the long run.”

“If yes mail me the code or send it to our patch manager on sf.net”

Frequency of task assignment mechanisms

Task Assignment Mechanisms	**Frequency**		
	EGW (%)	**Gaim (%)**	**Compiere (%)**
Self assignment	37(52.9)	60 (59.4)	16 (57.1)
Ask a certain person	15(21.4)	18 (17.8)	9 (32.1)
Ask an unspecified person	12(17.1)	22 (21.8)	1 (3.6)
Ask an outsider (not in the project development team)	0	1 (1.0)	0
Suggest consulting with others	6 (8.6)	0	2 (7.2)
Total of Task Assignment Messages	70	101	28

- Users rarely assign work to someone else, but do volunteer to work on something

Drawback of self-assignment

- People choosing to work on task might not be good at it
  - FLOSS teams need to filter patches to decide if they’re acceptable
- Volunteers might not be reliable
  - FLOSS teams need to work around unreliable contributors
- Several people might choose to work on the same part of the code at the same time
  - FLOSS teams need source code control systems

Implications for conventional software development

- Limitations of self-assignment would seem to make it undesirable for non-FLOSS
- However, “gentle persuasion” is used in other settings, e.g., CERN
- Some companies allow members some discretionary time, e.g., 3M, Google Labs

Conclusion

- Some commonality in coordination mechanisms due to common task
- Less explicit assignment of work
  - No evidence of hierarchy
  - Broader participation in work of the team
  - Assignment similar to “market”
- Many developers are volunteers, motivated by interest rather than pay

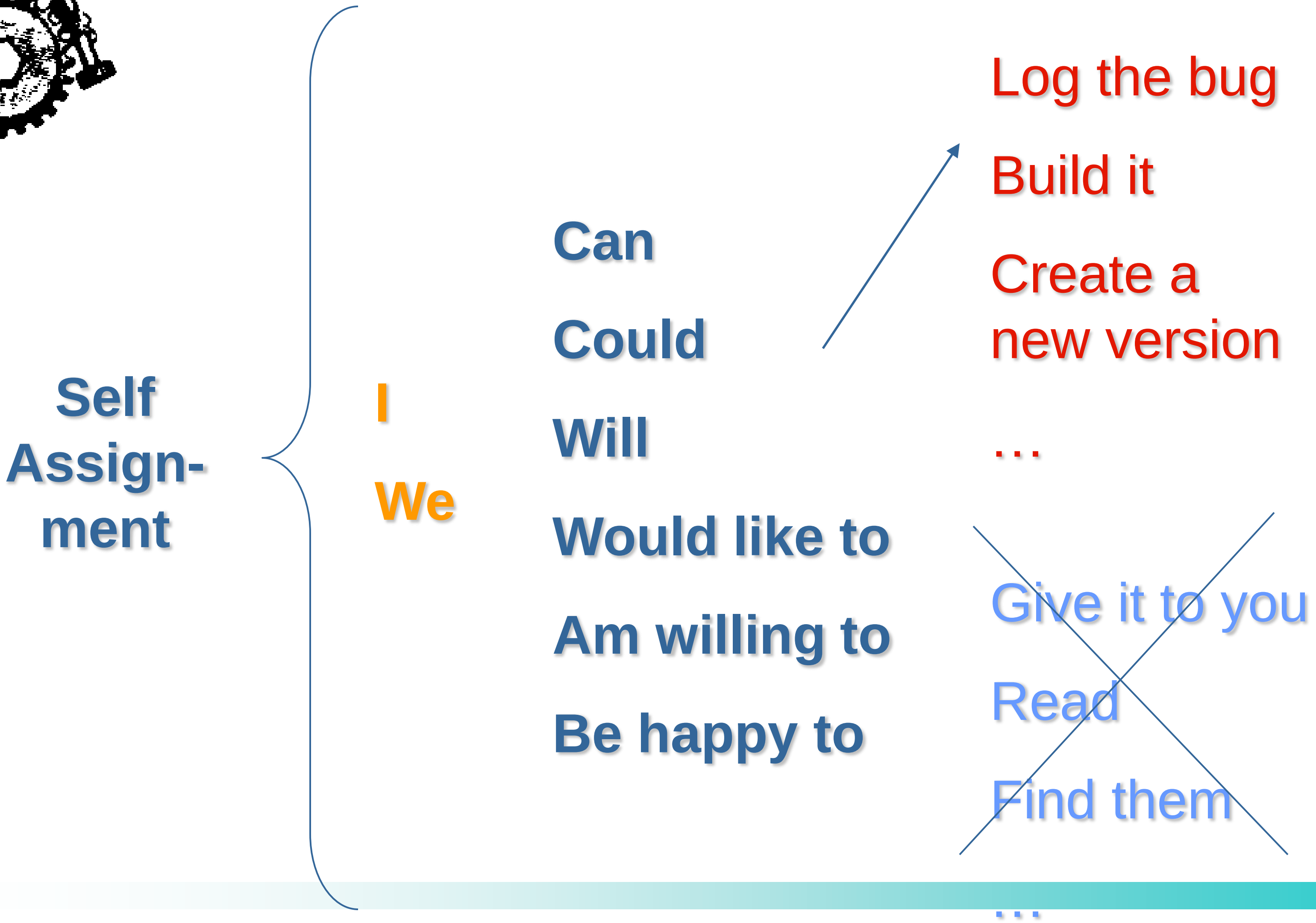
Future research

- Examine link between coordination and other team phenomena
  - E.g., does open nature of FLOSS task assignment help build collective mind
- Study more teams
  - Examine how coordination (esp. task assignment) differs in different kinds of teams (e.g., larger or smaller; those with and without corporate sponsorship)

NLP to study task assignment

- We are developing natural language processing approaches to automatically code task assignment (with Keisuke Inoue)
  - Focus on sentence structures, the frequently used verbs, etc
  - Focused on “Self Assignment” & “Ask to a specific person” and “Assignment to Another”

Example sentence pattern



Self Assignment

I

We

Can

Could

Will

Would like to

Am willing to

Be happy to

Log the bug

Build it

Create a new version

…

~~Give it to you~~

~~Read~~

~~Find them~~

~~…~~

Example sentence pattern

I

We

Can

Could

Will

Would like to

Am willing to

Be happy to

~~Not~~

Log the bug

Build it

Create a new version

…

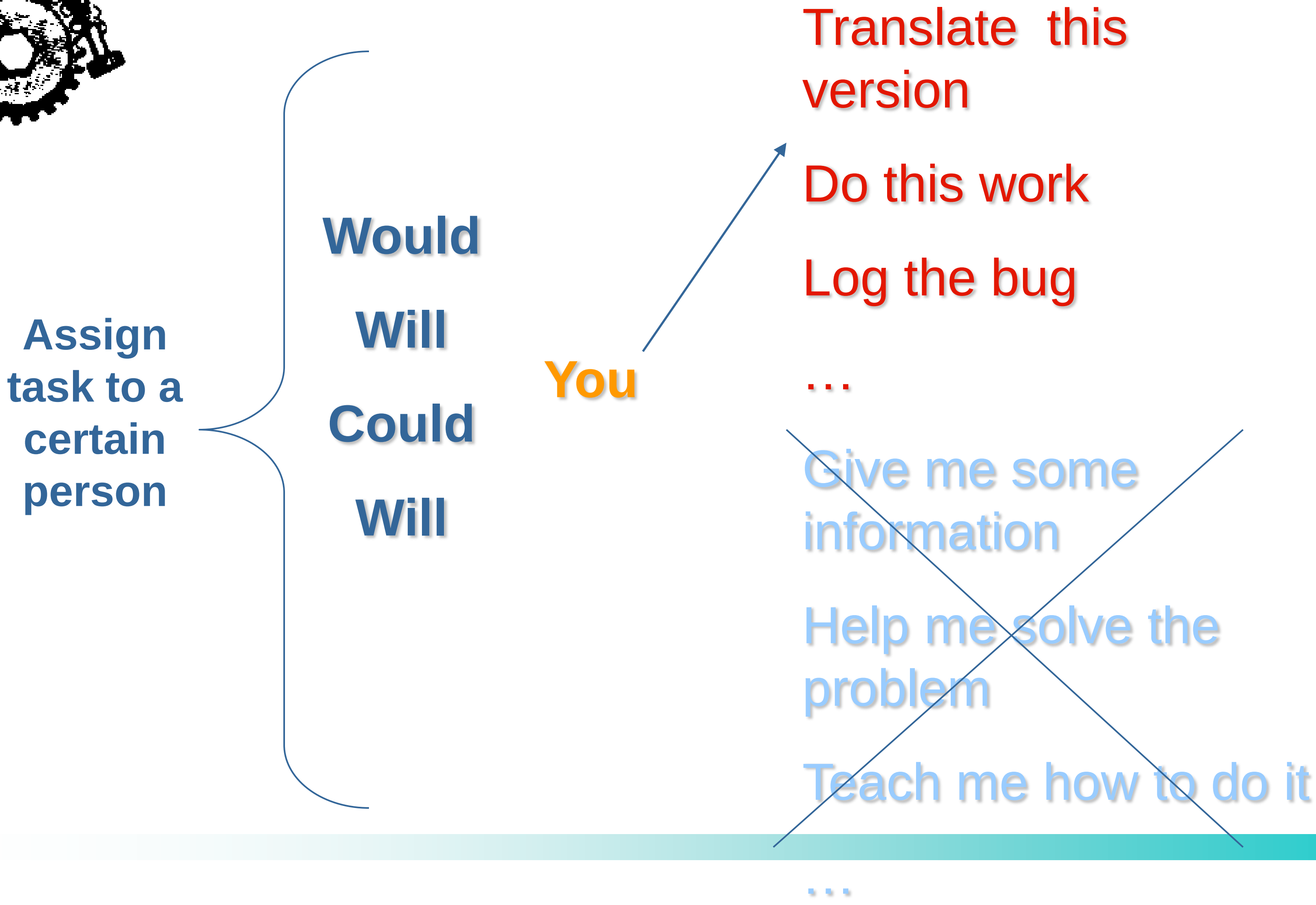
~~Give it to you~~

~~Read~~

~~Find them~~

…

Example sentence pattern



Assign task to a certain person

Would

Will

Could

Will

You

Translate this version

Do this work

Log the bug

…

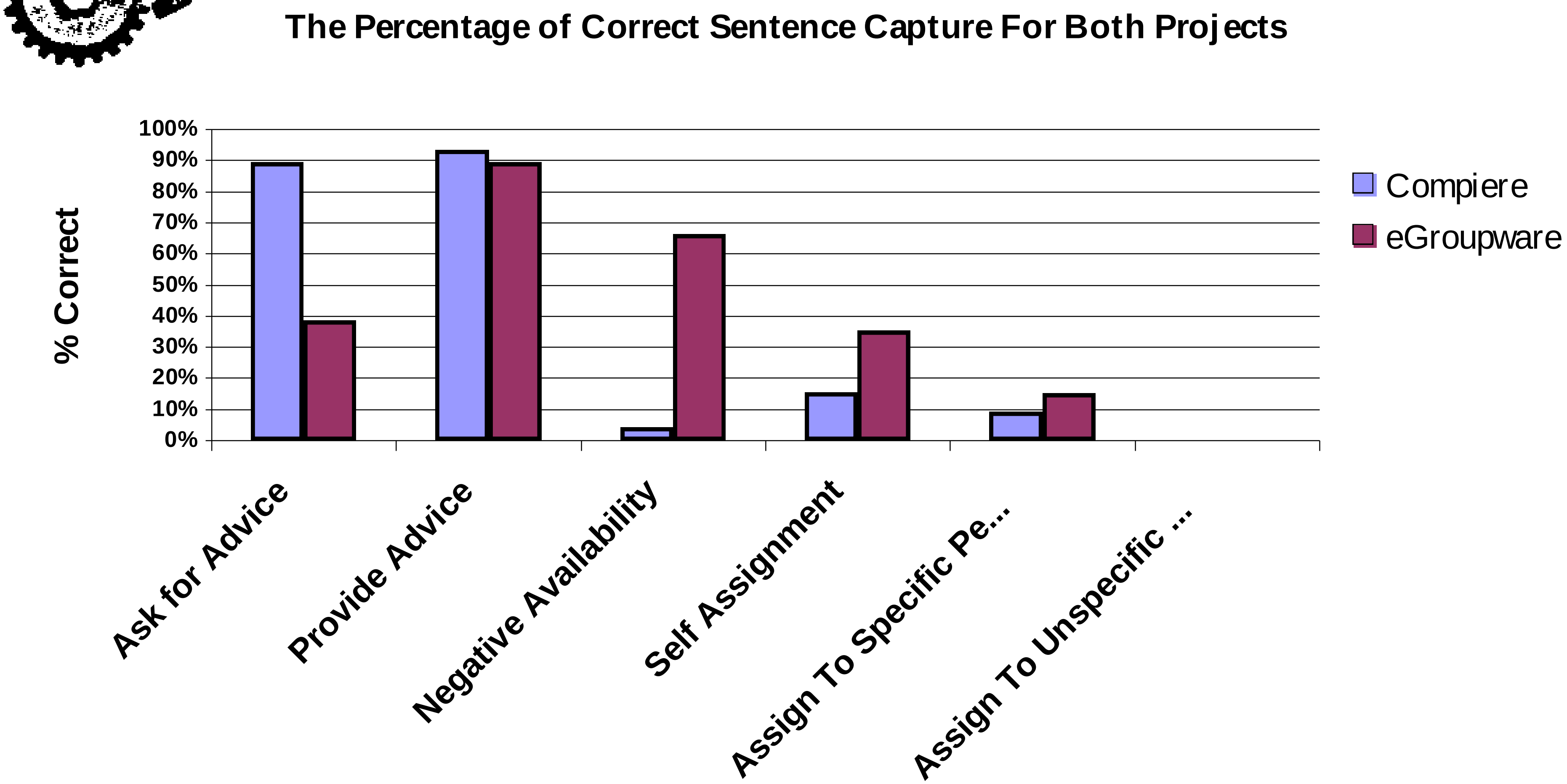
~~Give me some information~~

~~Help me solve the problem~~

~~Teach me how to do it~~

…

Preliminary results


The Percentage of Correct Sentence Capture For Both Projects
% Correct
Ask for Advice
Provide Advice
Negative Availability
Self Assignment
Assign To Specific Pe...
Assign To Unspecific ...
Compiere
eGroupware