COORDINATION AND DISCUSSION IN A VIRTUAL COMMUNITY OF PRODUCTION: A SEMANTIC NETWORK ANALYSIS

TRACK 2

DOES IT MATTER? The organizational impact of information systems

Vitaliano Barberio

Department of Management
University of Bologna
Italy
vitaliano.barberio3@unibo.it

Antonio Mastrogiorgio

Department of Management
University of Bologna
Italy

Alessandro Lomi

Department of Management
University of Lugano (Switzerland)
and University of Bologna (Italy)

EXTENDED ABSTRACT

In this paper we investigate, in a virtual community which aims to produce software, how email-mediated discussions enable and constrain a distributed system of coordination. Our focus is on communication, as a process of structures' enactment, which provides a coherent domain for distributed decision making. According to our point of view, discussion is the blood of the coordination process and the source of both technological and symbolic structures. By means of discussion, management practices are produced, reproduced and selected resulting in institutions; email based discussion itself is a kind of social institution in distributed social systems which like other institutions evolves by means of enactment. In our opinion, discussion is the fil rouge; it represents the fine-grained domain in which the socio-cultural artifact comes into being, arising from the practices of different communication genres (Yates and Orlikowski, 2002; Im et al., 2005). In this paper we present our theoretical perspective of distributed systems of governance claiming that email communication could be thought as a process of reciprocal influence between the material and the cultural domain (Giddens, 1984 Bourdieu, 1990) of production.

Open Source Software (OSS) is a perfect domain to explore our general research question. In fact in OSS: (i) a community of user-developers, distributed around the world, realizes high quality products; (ii) in absence of central planners; and (iii) coordinating production processes by means of email discussions. The emerging literature on governance of OSS projects highlights that the problem of communication has been treated keeping separated the material/technical interaction domain from the socio/symbolic one.

We used semantic network analytical tools in order to empirically explore how discussion enables and constraints distributed systems of coordination in one of the most successful OSS projects ever. In more details we analyzed communication flows, at a micro-interaction level, using email data from the Apache Project, in a time period of two weeks, to cast light on the process of structuration of both material and cultural domains. The first one is traced on data by as "voting action"; while the second is traced as "symbols linked to that voting behavior".

THEORY DEVELOPMENT

1. Open Source Software Coordination and Communication

In 1999 Eric Raymond, a recognized ethnographer of hackers' culture, proposed the metaphor of the 'Cathedral' and the 'Bazaar'[1] describing the OSS development as a distributed-production system, involving a large number of developers and characterized by: (a) the absence of a centralized decision-making unit defining ex-ante the direction of development of the software code; (b) parallel design and debugging; (c) the integration of users into the production of software code; (d) self-selection of programmers for the tasks that best match their abilities.

How coordination works in such environment is still a challenging research question. In particular, most of the literature we reviewed considered communication as an information processing network, or alternatively as a social evaluation network. First we briefly summarize these two positions in literature then we introduce our own theoretical position about it.

1.2. Communication as an Information Processing Network

Organizations (March and Simon, 1958) are supposed to be adequate solutions for complex information processing which provides an integration among interdependent tasks to be coordinated. The most widely accepted interpretation of this information processing view of organizations have been developed by Contingency Theorists (Lawrence and Lorsch, 1967; Galbraith, 1973; Daft and Lengel, 1986). A common argument of all these works is that matching increasing task uncertainty with less formal modes of coordination leads to better performance. A strength of contingency theory is its recognition of the complexity of interdependencies in organizational work. However, when dealing with information-processing it assumes that the environment is

---

[1]   The 'Cathedral' is a metaphoric representation of commercial development, while the 'Bazaar' represents the Open Source way of software development.

predictable enough to characterize existing interdependencies and that predefined mechanisms can be designed for various contingencies (Organizational Design).

For virtual communities, where non formal authority neither central planners are responsible for the so called 'organizational design', organizational structures could be seen as emanating from product architecture (Sanchez and Mahoney, 1996). Following this reasoning, because software has a more modular architecture than other products, virtual communities producing software will tend to have more distributed structures of governance than other productive organizations (von Hippel and von Krogh, 2003; Baldwin and Clark, 2006).

Communication provides the selective principle according to which product and tasks structures co-evolve. Modular architectures have been claimed as a strategy for decoupling interdependencies (Simon, 1962-1996) among tasks. For which tasks that could not be separated some coordination is needed and communication can provide that coordination (von Hippel, 1990). Von Hippel (2007) proposed that open source communities are thinkable as flat networks of user nodes linked by information exchange. According to this view, the content of communication is the mutual assistance that programmers provide each other. From a slightly different point of view Kuk (2006) proposed that programmers use communication as means of epistemic search for the knowledge that they need in order to solve their technical problems. In doing so they try to interact with those other programmers who control more "valuable" knowledge, but also accept a general rule of reciprocity.

1.2.Communication as a Social Evaluation Network

A second point of view on governance structures as coordination and communication networks could be thought as more concerned with the cultural domain of organizations. Community members, by means of repeated interaction, institutionalize norms and values creating what sociologist call an organizational field (Powell and Di Maggio, 1983) that is an ensemble of actors

who share a collectively constructed reality (Berger and Luckmann, 1966). This kind of social structure enhances coordination constraining the possibility of action and the access to resources reducing the uncertainty in decisional processes.

In a Study on the Linux Debian community, O'Mahoney and Ferraro (2007) proposed that individual performance, as the likelihood for candidates of being appointed on community-management formal positions, was affected by the congruence over time between the individual behavior and the socially legitimated kind of merit. Communication as repeated interaction among members provides a means of social evaluation of others and reproduce trust resulting in status positions (O'Mahoney and Ferraro, 2007) and social structures.

In fact, Virtual communities of production have less fluid boundaries than other virtual social networks like for example Facebook and MySpace and repeated interaction among programmers highlight a positional specialization where different roles[2] have different control on development activities. Mockus, Fielding and Herbsleb (2002) study on both the "Apache web server" and the "Mozilla web browser", provided evidence for the existence of teams of 10 and 15 people who controlled the development of the majority of the source code. Some authors (Grewal et al., 2006) showed as the embeddedness (Granovetter, 1985, Uzzi, 1996), as centrality in communication networks, of both projects and individual developers in such social structures increases legitimation and then positively affect access to resources and performance.

In this sense, Neo-Institutional Sociology (Powell and DiMaggio, 1991) goes beyond the separation between action and culture which is latent in the "contingent view". However when dealing with communication, the underlined social structure is often empirically operationalized as a stable entity constraining action possibilities and shaping performance.

---

[2] Both the worlds 'role' and 'position' here is used in the spirit of structural sociology (White et al., 1976). A 'role' is the pattern of relations held by an actor, the 'position' is the connectivity pattern of a role in a system of roles.

2. Avoiding the Separation Among Technological and Institutional Domains

2.1.Organizational Action

Our theoretical perspective of Distributed governance departs from both of the two perspectives presented above when considering the role of communication structures. We propose here that communication is an evolving set of coordination practices. Due to the relative newness of the argument of communication practices in OSS environments we borrowed some concept for our theoretical construction from previous consolidated research on coordination theory (Thompson, 1967) and structuration theory (Giddens, 1984). We also referred very much to practice-based research in the commercial field (Im et al., 2005).

Indeed Thompson's (1967) seminal contribution already avoided this dichotomy between an objective domain of action and a subjective domain of collective sense-making considering the organizational environment as a changing ensemble of tasks to be coordinated over the time. Both technological and institutional uncertainty contribute to explain environmental complexity for tasks to be coordinated. Structures in virtual communities of production, like for example OSS projects, emerge and change as bounded rational agents (Simon, 1957) attempting to take under control the two dimensions of environmental complexity.

Distributed governance in fast growing virtual community is, in our opinion, a general concept which underlines a vision of decentralized/informal decisional processes. Our idea is that each project in a different measure borrows, from both distributed software development experience and the OSS social movement, reproducing work practices. Over the time by means of use of those practices, each project specifies its own governance system.

2.1.Communication Genres and Uncertainty

When considering communication as a fundamental means of coordination for distributed teams of developers, a central issue rises about its role as a set of practices linking both the material domain of action and the symbolic domain of culture. Drawing on Giddens' (1984) structuralist perspective, some author proposed an approach based on Communication Genres (emails, meetings, expense forms, reports, etc.) as a social structure constituted through individuals' ongoing communicative practices (Yates and Orlikowski, 1992; Orlikowski and Yates, 2002). Quoting one of the most recent contributions to this approach (Im et al. 2005):

> *These genres are socially recognized types of communicative actions that are habitually enacted by organizational members over time to realize particular social purposes in recurrent situations (Yates and Orlikowski, 1992). Through such enactment, genres become institutionalized templates that shape members' communicative actions. Such ongoing genre use, in turn, reinforces those genres as distinctive and useful organizing structures for the community ....*

A key argument putted forward by "practice theorists" such as Giddens and Bourdieu, is that neither the material world (the world of action) nor the cultural world (the world of symbols) can exist or be coherently structured independently (Mohr and Duquenne, 1997:309). The duality of culture and practice imply that practices become institutionalized over time by means of use. The ongoing interaction between individuals and institutions could be viewed as a structuration process (Giddens, 1984). Structuration concerns the production, reproduction and transformation of social institutions, which are enacted by the use of social rules. These rules shape the action taken by individuals in organizations; at the same time, by regularly drawing on the rules, individuals reaffirm or modify the social institutions in an ongoing, recursive interaction.

2.1. Communication Genres and Uncertainty

Faraj and Xiao (2006) suggested that in complex knowledge and fast changing

environments the "lens of practice" are more suitable to understand coordination than traditional contingent approaches. Practices as suggested by Bourdieu (1990) have at their principle not a set of conscious, constant rules but practical schemes, opaque to their possessors varying according to the logic of situation. As enacted sets of practices (structures), c*ommunication genres* shape beliefs and actions and, in doing so, they enable and constrain how organizational members engage in communication (Im et al., 2005).

We argue that distributed-governance in virtual communities is a changing entity over time and place. We believe that rather than "contingent", it is "coherent" to a project domain and situation (Mische and White, 1998). Distributed governance systems in OSS projects emerge, according to our point of view, from the use reproduction and transformation of communication practices (*genres*) as the main enacted structure of coordination in an complex (Simon, 1962; Thompson, 1967) environment. In such context the concept of 'trajectories' (Strauss, 1993) as sequences of actions toward a goal, could well emphasize the interplay between contingencies and interactions among actors. Trajectories are also deal with deviations of the course of action from the desired objective. In those 'problematic' scenarios, decisional processes are more dealing with the situation rather than with formal organizational arrangements (Mische and White, 1998).

Our research question is: *how communication genres are flexible to the situation but at the same time provide a coherent domain for action in distributed systems of governance leading discussions to its objectives?* In order to explore this general question we propose in the following of this paper an empirical analysis of email communication extracted from the Apache Project case study.

METHOD AND RESEARCH STRATEGY

3. Methodological Note

Virtual communities are challenging contexts for traditional research methods, then, in this short note, we are going to describe some points to define our

perspective on it. Because of the distributed nature of development, virtual communities of production use of email communication as the main means of coordination.

> *"Mailing lists are the life blood of Apache communities. hey are the primary mode of discourse and constitute a public and historic record of the project. Other forms of communication (P2P, F2F, personal emails and so on) are secondary." ... "The reason is that communications on other than the public mail aliases exclude parts of the community. Even publicly advertised IRC chats can be exclusionary due to time zone constraints or conflicting time commitments by community members who might want to participate[3]".*

We look at communication in OSS projects as a process of interaction which enacts the social structure provided by genres. Because in this environment the most of developers never meet face-to-face, we consider communication in public lists as the only available reality for development practices synthesizing historic traceability, the scope of the process and thematic coherence. One way of understanding discursive genres is to examine the socially recognized or sanctioned expectations around key aspects of communication: purpose, content, participants, form, time, and location (Yates and Orlikowski, 2002).

4. Setting: The Apache Project

The Apache project started in February 1995 when Rob McCool stopped developing his httpd-server program at NCSA[4] and then a small group of users, the so called Apache Group (AG), began a combined effort to coordinate existing fixes to the existing code. After several months of adding features and small fixes, the AG replaced the old server code base in July 1995 with a new architecture designed by Robert Thau.

As the core developers were distributed around the world and all of them were working at the project on a totally voluntary base, both leadership and coordination mechanism were distributed as well to take in account the limited

---

[3] Quotation from Apache community building guidelines.
[4] National Center for Supercomputing Applications (NCSA).

amount of time that each programmer could devote to the project. As Roy Fielding (1999), one of the founding members, pointed out:

"Unlike most open-source projects, Apache has not been organized around a single person or primary contributor"…"There was no Apache CEO, president, or manager to turn to for making decisions. Instead, we needed to determine group consensus, without using synchronous communication, and in a way that would interfere as little as possible with the project progress. What we devised was a system of voting via email that was based on minimal quorum consensus. Each independent developer could vote on any issue facing the project by sending mail to the mailing list with a "+1" (yes) or "-1" (no) vote".

According to the Netcraft survey[5] in few months Apache (blue line in chart 1) became the most used server software in the world and it still is today. Microsoft which is Apache's main competitor (red line in chart 1) also became "involved" in Apache's with a platinum sponsorship in 2008. The amount of work to be coordinated in order to maintain the software over the firs four years of development grew along with the increasing number of users. Then the Apache Group made an important step toward a more formal system of governance.

In 1999 the Apache Software Foundation (ASF) was created to provide: (i) hardware, communication and business infrastructures; (ii) a legal entity for code donations assuring that those resources will be used in the public interest; (iii) legal assistance and legitimation to new projects admitted under its identity umbrella; (iv) protection to the Apache brand from being abused by other organizations. The Apache "software code" from this point will belong to the foundation which aims to maintain it public:

The Apache Software Foundation creates and maintains open source software products for the public benefit utilizing a collaborative, meritocratic approach to software development. Our products are developed by a diverse community of volunteers, a large number of whom use our software products in the course of their own daily lives. Our development discussions are held on public mailing lists. Everyone is invited to join the discussion so long as the usual courtesies of email netiquette are observed. (from the Apache

---

5   Netcraft survey: http://news.netcraft.com/archives/web_server_survey.html.
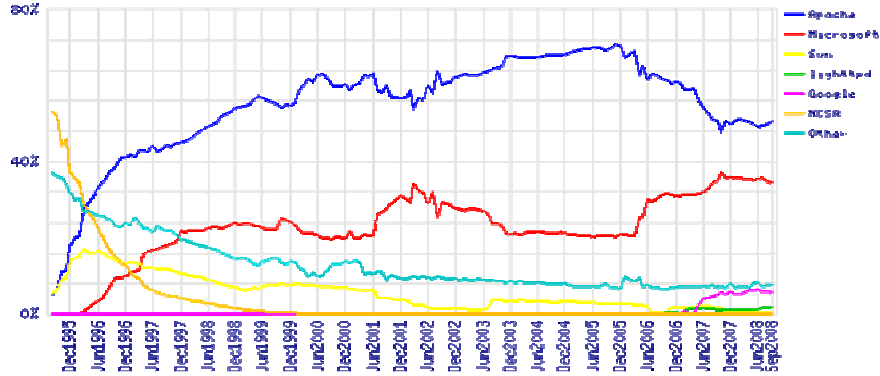
guidelines[6]).

Figure 1. Servers market share across all domains. Number of domains (in percent) adopting the Apache server over time is marked by the blu-line. The main competitor is Microsoft whose market share is marked by the red line.

The Apache meritocratic governance, also called "Apache Way" became over the time an institution for the OSS movement as whole. Literally the govern of merit means that who writes the Apache's (software) code also hold the power in institutional collective decision making concerning the overall direction of development. The Apache Way is a challenging governance system for researchers studying coordination practices. In a very simplified picture, decisions are taken in two steps: (i) generating the consensus/dissensions around a proposal; (ii) vote the emergent/structuring proposal when no consensus is achieved by means of "simple conversation".

5. Data Collection

We gathered data for our study from an infrastructural mailing list belonging to

---

[6] Apache 'code of conduct' is readable online here:
http://wiki.apache.org/incubator/CodeOfConduct

the ASF where topics regarding community building are discussed[7]. The 'community' mailing list was created after a period (1999-2002) of institutional re-organization. The discussion is composed by 155 single e-mails. The temporal extension of the discussion (2 threads) is: 22 October 2002 – 6 November 2002. All the selected emails had the tag [vote]. Following Im et al. (2005) this means that all the selected emails belonged to the same communication genre and then those are supposed to obey to the same institutionalized rules of interaction. In order to make the distributed decisional process working, each Apache voting session should not go ahead for more than 72 hours. Because we looked for a flexible use of genres as "violation" of codes according to situations, we selected these two threads as concerning the same decision expecting that something "did not work properly" in the first session.

6. Analysis

Our analysis is articulated in two main steps: (a) text pre-processing; and (b) semantic network analysis. We parsed the email text to extract single "concepts" and used those concepts to build network representations of the decisional process to be further analyzed (Diesner et al., 2005).

6..1   Pre-Processing

As we are going to explain, text pre-processing is a fundamental requisite for semantic network analysis. We have followed three steps that we call (i) redundant information, (ii) frequency and (iii) thesaurus.

   Redundant information. We have performed our analysis on a flow of 155 email; we have deleted redundant text arising from communications "in replay to" and "forwarded". The presence of automatic copy in those email could be a source of biases. The deletion of redundant information is a fundamental step

---

[7]   The list is: community@apache.org and its public archives are here: http://mail-archives.apache.org/mod_mbox/www-community/

to preserve the text it has been intentionally communicated by individuals.

Frequency. A text is characterized by a number of words that we call concepts; each concept is characterized by a frequency. The distribution of the words in a text follow a Power Law distribution; high-frequency words (trivial concepts) are those represented by commonly used significants such as 'the', 'a', 'have' etc. Low frequency words (idiosyncratic concepts) are those which are not relevant in the domain of discourse. A fundamental step of processing text is the deletion of both tails of the distribution; in other words we have deleted concept with very high (trivial concepts) or very low frequency (idiosyncratic concepts).

Thesaurus. We have reduced the grain of the text to a coarser one by bringing back similar concepts to a single concept; the loss of information represents a gain in terms of synthesis.

6..2   Semantic Network Analysis

The outcome of the pre-processing step is a set of semantic networks (one for each email). These semantic networks are made by nodes and ties, where nodes are concepts and ties are the relationships between such concepts which occur in the discourse. The weight of a tie between two nodes/concepts has been set as the co-occurrence frequency of those concepts in the same sentence. The outcome of the pre-processing step was a set of 155 semantic networks, one for each analyzed email, according to this principle. Here we introduce our semantic analytical techniques.

General Statistics and Symbols. After computing some synthetic statistics for all individual networks, we analyzed those networks looking for concepts with high ranking in both betweenness centrality and degree centrality (Wasserman and Faust, 1994). When also the 'consensus' were high for those nodes the corresponding concept was labeled as a *symbol* (Carley and Kaufer, 1993).

Consolidated Semantic Networks. In order to make easier the interpretation of results addressing our first research question, we consolidated

all those single-mail networks in to a synthetic one. In the resulting *Consolidated Semantic Networks* the weight of ties is given by the occurrence of that tie across multiple messages.

RESULTS

7. Semantic Networks

7..1 General Statistics

General statistics computed across all 155 semantic networks (table 1. and table 2.) show that, on average, emails in Thread_1 had a major number of concepts 37.3068 than Tread_2 (20.4478). Thread_1 is also characterized by an higher concepts' standard deviation (21.8891) when compared to Thread_2 (7.87263). Semantic networks in thread_2 are more densely connected (0.0602428) than networks from Thread_1 (0.045165). At the same time the average diameter is lower in email belonging to Tread_2 (19.9254) than in email belonging to Thread_ (135.7386).

Taken together these two results tell us that emails in thread_1 are, on average, more cohesive in terms of linked concepts than emails in Thread_2. Finally, semantic networks from thread_1 displayed a higher clustering coefficient than those from thread_2. This means that emails of the first set have more concepts' sub-aggregations relatively independent among them.

Table1. Statistics Across all Semantic Networks from Thread_1. Total number of semantic networks = 88

Measure	Min	Mean	Max	Std.dev
Number of concepts	5	37.3068	113	21.8891
Number of isolated concepts	0	0	0	0
Number of links	4	46.3295	162	32.1418
Density	0.0128003	0.045165	0.2	0.0310637
Diameter	5	35.7386	113	21.9689
Clustering Coefficient	0	0.0251622	0.0839026	0.022417

Table 2. Statistics Across all Semantic Networks from Thread_2. Total number of semantic networks = 67

Measure	Min	Mean	Max	Std.dev
Number of concepts	5	20.4478	54	7.87263
Number of isolated concepts	0	0.970149	4	0.869869
Number of links	2	21.8657	76	10.4242
Density	0.0260244	0.0602428	0.166667	0.0208344
Diameter	5	19.9254	54	8.17853
Clustering Coefficient	0	0.0167306	0.166667	0.0304111

7..2 Symbols

Confronting most ranked symbols (high degree, high betweenness and high consensus), we are able to draw some additional qualitative results. In particular even if both Thread_1 and Thread_2 are of the same communication genre [vote] and the same subject 'openness', we observe that concept ranking is slightly different (see table 3. and table 4.).

Table 3. Symbols (high degree, high betweenness, high consensus) in thread_2. There are 74 concepts in this class.

Rank	Concept	Consensus	Degree Centrality	Betweenness Centrality
1	apache	0.0127932	0.086351	0.109545
2	committers	0.0149254	0.0877437	0.0998514
3	vote	0.00852878	0.0835655	0.107672
4	view	0.0140116	0.0738162	0.0930322
5	archive	0.0134024	0.0793872	0.0818668
6	org	0.0130978	0.0584958	0.0887219
7	community	0.0103564	0.0598886	0.0589193
8	sam	0.0124886	0.0529248	0.0597323
9	more	0.00761499	0.051532	0.0473578
10	do	0.00761499	0.0557103	0.0424875

Table 4. Symbols (high degree, high betweenness, high consensus) in thread_2. There are 50 concepts in this class.

Rank	Concept	Consensus	Degree Centrality	Betweenness Centrality
1	vote	0.0080292	0.0597484	0.240226
2	vote1	0.0416058	0.0880503	0.128177
3	need	0.00291971	0.0220126	0.204971
4	committers	0.029927	0.0613208	0.119199
5	vote2	0.0416058	0.0833333	0.0739063
6	-1	0.0321168	0.0204403	0.128142
7	let's	0.0423358	0.0157233	0.118811
8	community	0.00437956	0.0361635	0.129095
9	roy	0.00218978	0.0283019	0.126322
10	no	0.0306569	0.0157233	0.100321

7..3   Consolidated Semantic Networks (CSN)

CSN (see Figure 2. and Figure 3.) helped us to deeper understand the earlier findings coming from individual emails networks' analysis. In order to better 'see' in to the networks we removed links with weight less than 1 and then removed all the isolated nodes. Looking at thread_1 (Figure 2.) the most evident result is that the concept 'view' ranked as 4th in Table 3. is now the most central in terms of weighted degree (more incident high weight lines).

Departing from that node we can find at least two paths corresponding to different proposals about the openness of the community@ mailing list, for example: (i) view → close → except → committers → members → invitees; (ii) view → open → completely → anyone → can → subscribe → post → read. Another interesting path for understanding distributed governance is the one in the bottom right side of Figure 2.: local → governance/governing → bodies → incapable → dealing → trivial → issues → affect. This path expresses the major threat coming with the potential scenario where everybody is allowed to read and write on the mailing list.

Figure 2. Consolidated Semantic Network from email thread_1. Both links

with weigh less than 5.1 and isolate nodes have been recursively removed from the original network in order to offer a clearer representation.

So what is missing in this (Thread_1) representation? As a [vote] thread we expected a very structured communication (few densely connected concepts) and most important we expected to see as very frequent/central concepts those expressing the voting action [+1], [0] and [-1].

All these just mentioned features are actually present in the consolidated semantic network representing Thread_2. In this representation (see Figure 3.) the central sub-group of nodes is clearly expressing an email-mediated voting behavior. Concepts like 'vote1' and 'vote2' are kind of formalized proposals to be voted; '1', '0' and '-1' are the voting actions. Departing from those concepts/nodes we can find again, even if in a more stylized representation, the elements of different proposals: vote1 → 1 → yes → let's → open → everyone; or vote2 → committers → keep → private.
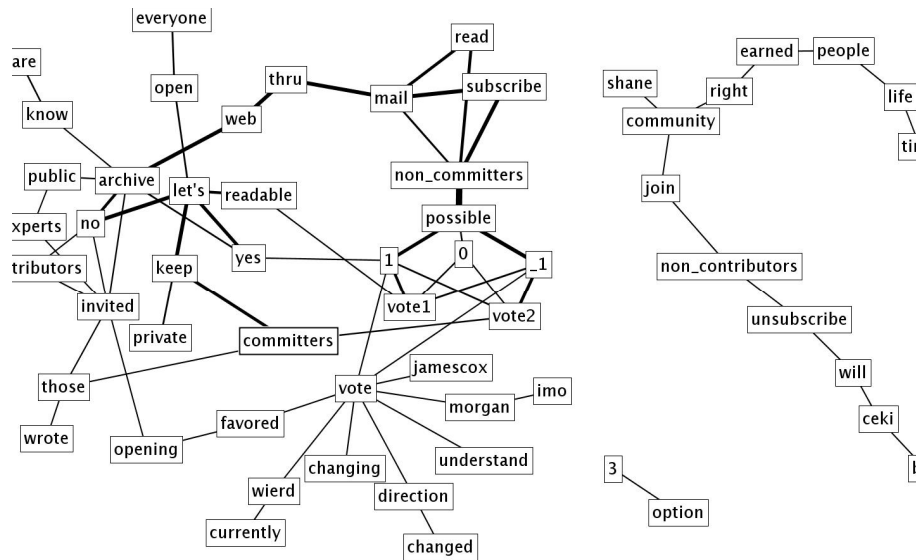
Figure 3. Consolidated Semantic Network from email thread_2. Both links with weigh less than 1.1 and isolate nodes have been recursively removed from the original network in order to offer a clearer representation.

The general governance issue for the analyzed decision was: how open the community mailing should be. In major detail the decision concerned the 'who' should be allowed to do "what". The options for the "who" issue were "committers" and "non-committers", while the options for the "what" issue to be decided were "write" and "reed". Different configurations of this elements have been formalized in a proposal with alternative "scenarios" to be voted by community members in Thread_1. As a second voting session have been required for the same decision we expected that some kind of re-alignment of action (voting) toward a collective decision (how-open) should have been taken. So we checked for an explanation in the first email of the Thread_2 whose body text is reported below (Figure 4.). We see here that "voting" is thinkable as a process leaded by at least two flexible practice: (ii) dialogic consensus formation; and (ii) voting as it.

The mail in Figure 4. confirms that in Thread_1 the conversational practice

used for consensus generating took a very complicated path to be interpreted as a collective decision and in doing so it obscured the 'voting as it'. This mail (Figure 4.) also confirms the flexibility of the "vote" genre according to the situation stating that: first, "*Note: there is no need to indicate the reason for your votes, either for negative ones.*" This happens because the individual points of view already emerged in the precedent session.; second, "*Also, please, don't vote 0.5 or other numbers, let's keep it simple for the final count.*"

```
Please, allow me to restart the voting in order to make it easier to reach some
consensus since it's hard to interpret the results of the previous one.
                              - o -
There are two different concerns for openness:
  - open to read
  - open to write
Let's try to keep them separate.
NOTE: there is no need to indicate the reasons for your votes, either
for negative ones.
Also, please, don't vote 0.5 or other numbers, let's keep it simple for
the final count.
                              - o -
VOTE 1:  would you like to make it possible for non-committers to read
this mail list through a web archive?
  [ ] +1 yes, let's make it readable
  [ ]  0 don't know/don't care
  [ ] -1 no, let's keep it private
                              - o -
VOTE 2:  would you like to make it possible for non-committers to fully
subscribe to this mail list?
  [ ] +1 yes, let's open it to everyone
  [ ]  0 don't know/don't care
  [ ] -1 no, let's keep it for committers only
                              - o -
Please, place your vote even if you already voted in the previous poll. We'll reset
the clock and give 78 hours for the vote. I volunteer to count the results and post
them here.Thanks in advance and sorry for the double poll noise.
```

Figure 4. Body text of the first (in chronological order) email in Thread_2.

DISCUSSION AND FURTHER RESEARCH

In complex knowledge and fast changing environments the 'lens of practice' are more suitable to understand coordination than traditional contingent approaches (Faraj and Xiao, 2006). Practices as suggested by Bourdieu (1990) have at their principle not a set of conscious, constant rules but practical

schemes, opaque to their possessors and varying according to the logic of situation.

Indeed the email voting practice is very much institutionalized in the Apache community. The Apache Group have been using it from the very beginning of its development processes in order to establish both which patches and which new features should have been applied to the existing software code. Even if email voting is institutionalized in a communication genre (Im et al, 2005), in this work, we have found a flexible use of such a genre as consensus generating and voting occurring in emails with the same tag - [vote] -.

We also found that communication genres, like voting, come in handy, to redirect discussions toward an objective (a decision) synthesizing the elements of the decisional process. It seems to us that the distributed governance in Apache is an ongoing synthesis between conversation and situation over time.

Some author proposed that *conversation* is a discussion form in which the 'story' does not tend to a precise final, while *situation* (Mische and White, 1998) is more about the possibility of an unexpected or problematic final (which in our case is the absence of a clear policy for community management). The switching dynamic, from conversation to situation and vice versa, could be a second topic for further research.

Our contribution to the research on distributed coordination practices is a representation of organizational action which brings together the domain of action with the domain of culture. Mohr et al. (2004) proposed the use of "content analysis" and "Galois lattices" aiming to show how institutional action is linked to symbolic categories of recipients for that action. Our research strategy is similar to Mohr's one in its general intent, however we used different analytical tools. In particular the use of semantic analytical tools (Carley and Kaufer, 1993) is still a relatively unexplored research strategy for practice-based representations of decisional processes.

REFERENCES

Baldwin C.Y., Clark. K.B. (2006), "The Architecture of Participation: Does Code Architecture Mitigate Free Riding in the Open Source Development Model?" *Management Science,* 52 (7), pp.1116–1127.

Berger B.L., Luckmann T. (1966), *The Social Construction of Reality*, Doubleday, New York.

Bourdieu P. (1990), *The Logic of Practice*, Stanford University Press, Stanford University.

Carley K.M., Kaufer. D.S. (1993), "Semantic Connectivity: An Approach for Analyzing Symbols in Semantic Networks", *Communication Theory*, 3(3), pp.183-213.

Daft R.L., Lengel R.H. (1986), "Organizational information requirements, media richness and structural design", *Management Science*, 32(5), pp.554 – 571.

Diesner J., Frantz T.L, Carley M.K. (2005), Communication Networks from the Enron Email Corpus "It's Always About the People. Enron is no Different" *Computational & Mathematical Organization Theory*,*Volume* 11 (3), pp. 201-228.

Di Maggio P.J., Powell W.W. (1983), "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields", *American Sociological Review*, 48, pp.147-160.

Faraj S., Xiao Y. (2006), "Coordination in Fast-Response Organizations", *Management Science*, 52(8), pp.1155–1169.

Fielding R.T. (1999), "Shared Leadership in the Apache Project. Commun", *ACM* 42, pp. 42–43.

Galbraith J. (1973), *Designing Complex Organizations*, Addison Wesley, Reading MA.

Giddens A. (1984). *The Constitution of Society. Outline of the Theory of Structuration*, Polity, Cambridge.

Granovetter M. (1985), "Economic Action and Social Structure: The Problem of Embeddedness", *American Sociological Review*, 15(3) pp. 478–499

Grewal R. G., Lilien L., Mallapragada G.. (2006), "Location, Location, Location: How Network Embeddedness Affects Project Success in Open Source Systems", *Management Science* 5 2(7), pp. 1043–1056.

Im H.G., Orlikowsky W, Yates J. (2005), "Temporal Coordination through Communication: Using Genres in a Virtual Start-up Organization", *Information Technology & People*, 18(2), pp. 89-119.

Kuk G.. (2006), "Strategic Interaction and Knowledge Sharing in the KDE Developer Mailing List.", *Management Science*. 52, pp.1031-1042.

Lawrence P.R., Lorsch J.W. (1967), *Organization and Environment*, Harvard University Press, Cambridge.

March J.G., Simon H.A. (1958), *Organizations*. Wiley, New York.

Meyer J.W., Rowan B. (1977), "Institutionalized Organizations: Formal Structure as Myth and Ceremony", *American Journal of Sociology*, 2, pp. 340-363.

Mische A., White H. (1998), "Between Conversation and Situation: Public Switching Dynamics across Network Domains", *Social Research,* 65(3), pp. 695-72

Mockus A., Fielding R.T., Herbsleb J. (2002), "Two Case Studies of Open Source Software Development: Apache and Mozilla", *ACM Transactions on Software Engineering and Methodology*, 11(3), pp. 309-346.

Mohr J.W., Duquenne V. (1997), "The Duality of Culture and Practice: Poverty Relief in New York City 1888-1917", *Theory and Society*. Vol. 26/2-3, pp. 305-356.

Mohr J.W., Bourgeois M., Duquenne V. (2004), *The Logic of Opportunity: a Formal Analysis of the University of California's Outreach and Diversity Discourse*, Research & Paper Series: University of California, Berkeley http://ishi.lib.berkeley.edu/cshe/

O'Mahony S. (2007), "The Governance of Open Source Initiatives: What Does It Mean to Be

Community Managed?", *Journal of Management and Governance*, 11, pp. 139–150.

O'Mahony S., Ferraro F. (2007), "The Emergence of Governance in an Open Source Community", *Academy of Management Journal,* 50 (5), pp. 1079-1106.

Powell W., Di Maggio P. (1991), *The New Institutionalism in Organizational Analysis*, University of Chicago Press, Chicago.

Raymond E. S. (1999), *The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary,* O'Reilly & Associates, Sebastopol CA.

Sanchez R., Mahoney J.T. (1996), "Modularity, Flexibility and Knowledge Management in Product and Organization Design", *Strategic Management Journal*, 17, pp. 63-76.

Simon H.A. (1957), "A Behavioral Model of Rational Choice, in Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting", Wiley, New York..

Simon H.A. (1962), "The Architecture of Complexity. Proceedings of the American Philosophical Society" 106, pp. 467-482.

Simon H. A. (1996), *The Sciences of the Artificial*, MIT Press, Cambridge MA.

Strauss A. L. (1993), *Continual Permutations of Action*, Aldyne de Gruyter, New York.

Thompson J. (1967), *Organizations in Action*, McGraw-Hill, New York.

Uzzi B. (1996), "The Sources and Consequences of Embeddedness for the Economic Performance of Organizations: The Network Effect", *American Sociological Review*, 61 pp. 674–698.

Von Hippel E. (1990), "Task Partitioning: An Innovation Process Variable", *Research Policy*, vol. 19 pp. 407-18.

Von Hippel E., von Krogh G.. (2003), "The Private-Collective Innovation Model in Open Source Software Development", *Organization Science*. 14(2) pp. 209–223.

Von Hippel E. (2007), "Horizontal Innovation Networks by and for Users", *Industrial and Corporate Change,* 16:2.

Wasserman S., Faust K. (1994), "Social Network Analysis: Methods and Applications".

Yates J., Orlikowski W., (1992), "Genres of Organizational Communication: A Structurational Approach to Studying Communication and Media", *Academy of Management Review*, 17(2), pp. 299-326.

Yates J., Orlikowski W. (2002), "Genre Systems: Structuring Interaction through Communicative Norms", *Journal of Business Communication*, 39(1), pp. 13-35.