Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

PhD in Management

Ciclo XXXI

Settore Concorsuale: 13/B3
Settore Scientifico Disciplinare: SECS-P/10

The complexity of exploration and exploitation in organizations

Presentata da Cottini Ciro

Coordinatore dottorato	Supervisore
prof.ssa Orsingher Chiara	prof. Morandin Gabriele
	Co-supervisore
	prof. Soda Giuseppe
	prof. Fioretti Guido
	dott.ssa Rossello Simona

Esame finale anno 2020

To my family

We ran the race
The race was won
By running slowly

I. Anderson

Tu volera dans un toile d'araignée
Mais tu gagnera, tu gagnera
De nouveau ta liberté

Nathalie

از گهواره تاگور دانش بجوی

Persian proberb

Contents

1 Introduction **9**

1.1 The aim of the research 11

2 Exploration and exploitation **17**

2.1 March's model 17

2.2 The evolution of the exploration and exploitation 22

2.3 March's model evolution 30

3 The research question **41**

4 Agent Based Model **49**

4.1 Agent based models in the research 50

4.2 The anatomy of an agent based model 55

4.3 Emergence 58

4.4 KISS vs KIDS 61

4.5 Agent based model criticism 62

5 The proposed approach **67**

5.1 General structure 69

5.2 Agent's topology 73

5.3 Knowledge representation 75

5.4 Knowledge exchange 76

5.5 Meetings 77

5.6 Self-search 79

5.7 Physical layout 81

5.8 Friendship 88

5.9 Knowledge distortion 91

5.10 Turmoil 91

5.11 External interaction and unlearning 92

6 Model description **95**
6.1 Objective . . . 95
6.1.1 Purpose . . . 95
6.1.2 Entities, state variables, and scales . . . 96
6.1.3 Process overview and scheduling . . . 97
6.2 Design . . . 99
6.2.1 Emergence . . . 99
6.2.2 Adaptation . . . 99
6.2.3 Sensing . . . 100
6.2.4 Stochasticity . . . 100
6.2.5 Observation . . . 101
6.3 Details . . . 101
6.3.1 Initialization . . . 101
6.3.2 Input data . . . 102
6.3.3 Knowledge array . . . 105
6.3.4 External reality adherence . . . 106
6.3.5 Knowledge exchange . . . 106
6.3.6 Knowledge distortion . . . 106
6.3.7 External reality change . . . 107
6.3.8 Unlearning and external interaction . . . 107
6.3.9 Agent stack and stack update . . . 108
6.3.10 General stack . . . 109
6.3.11 Channel assignment . . . 110
6.3.12 Meetings . . . 110
6.3.13 Self-search . . . 112
6.3.14 Physical layout . . . 113
6.3.15 Friendship . . . 118

7 Simulations programme and output **119**

8 Duality **129**

9 Pathfinders **137**
9.1 Network size and learning rates . . . 148
9.2 Network analysis . . . 151

10 Cloisters **161**
10.1 Wider propinquity . . . 174
10.2 Network analysis . . . 181
10.3 Stack level and energy . . . 186
10.4 Office footprint . . . 193

10.5 End points convergence 196

11 Gatherings 201
11.1 Impact of physical layout 212
11.2 Network analysis 213
11.3 Hierarchy 228

12 Friends 233

13 Complexity and emergence 241
13.1 Unlearning 243
13.2 Unstability and divergence 244

14 Summary and contribution 249
14.1 Environment and March's effect 249
14.2 Emergence in exploration exploitation dilemma 252
14.3 The importance of networks 253
14.4 Contribution 255

15 Limitations and possible expansions 259

16 Appendix A: regression diagnostics 263

Chapter 1

Introduction

Many decades passed since Plato defined knowledge a *"justified true belief"*. Indeed the environment has changed, the concept has been thoroughly discussed, sectioned and recomposed many times to find a satisfactory explanation for one of the most complex concepts mankind created. Despite the incredible profuse effort, knowledge is still a live struggling topic for both scholars and practitioners. But we live in a "knowledge society" and the always increasing importance of knowledge demands new insight. Undoubtedly, this new society calls for innovation and for this reason, knowledge is of paramount importance for organizations; without knowledge hardly any company could survive long.

Many different theories have been brought to the attention trying to cover different aspects and mechanisms of knowledge in its multiple facets. Knowledge has been studied from different perspectives: creation, conversion, transfer, management, reuse, protection, sharing. Within these different topics, one captured the attention of academics for its prominence and impact on organization: the exploration-exploitation problem.

Exploration is the quest of the new whereas exploitation is taking advantage from the already possessed knowledge. Exploration is considered synonymous with new ideas, new knowledge, new technologies, often assimilated to experimentation, test, and change. Exploitation is rather considered synonymous with refinement, execution, and efficiency. Both tendencies are vital for any given organization, which is called to identify a trade-off between the two in order to reach a sustainable long-term prosperity.

Hence the presence of this tension create a dilemma. A trade-off dilemma for companies while they compete for scarce resources, making an implicit or explicit choice between exploration and exploitation: explicitly through calculated decisions about investments and strategies, implicitly through procedures, practices and organization forms (March 1991). Exploration and exploitation activities differ in the locus of action, in the degree of certainty and time of possible results. Exploitation activities are more certain in the results, nearer in time and, since the organization learns from experience, more appealing than exploration activities. For these reasons, adaptive processes are more likely to produce rapid and near in time effects, de facto slowing the adoption of explorative mindset. As March pointed out, *"reason inhibits foolishness; learning and imitation inhibit experimentation"* (March 1991, 73).

March (March 1991) proposed a simulation model to discuss the equilibrium in the organizational learning arena. He started from the appreciation that although knowledge is created and diffused in a variety of ways within an organization, a specific process called mutual learning can be isolated. The mutual learning process encompasses a concurrent learning process by both individuals and organizations: while individuals learn from and influence the organization, the organization learns and influences the individuals. March demonstrated the impact that individuals and their propensity toward exploration or exploitation have on the overall organizational knowledge.

In the last decades the exploration and exploitation topic has been debated extensively and many assumptions or results March proposed have been challenged, expanded, enhanced by following scholars. Gupta et al. (Gupta, Smith, and Shalley 2006) in a dedicated forum argued about fundamental aspects of the balance as orthogonality versus continuity, ambidexterity versus punctuated equilibrium, duality versus specialization; Axlerod (Axelrod 1997) and Miller et al. (Miller, Zhao, and Calantone 2006) argued about the spatial relationships among individuals, Rodan (Rodan 2005) wrote about organizational policies and mutual learning, Fang et al. (Fang, Lee, and Schilling 2010) about the impact of isolated and semi-isolated group, Kunz (Kunz 2011) focused on inter and intra group relationship, Mitomi et al. (Mitomi and Takahashi 2015) argued about low levels of learning rate and Aven et Zhang (Aven and Zhang 2016) considered the social distance in the scenario.

The balance of exploration and exploitation has been approached with at least two different intents: (1) defining the best balance for an organization, (2) defining which are the levers to modulate in order to obtain a desired level of balance (Gupta, Smith, and Shalley 2006). The first approach is not only extremely hard to achieve, but also the identification of the right mix of exploration and exploitation could be of limited utility for a manager and his/her company. The environment where contemporary organizations operate is continuously changing, compromising the validity of the optimum over time and the organization has sometimes to introduce projects or activities that are necessary, but not coherent with the optimum balance already found. Eventually, some projects are necessary also if the balance has to be compromised. The second approach aims to understand which organizational dimensions a manager has to leverage in order to obtain the desired ratio between exploration and exploitation. Therefore, the remit shifts from finding the optimum mix given the (extremely variable) boundary conditions to the ability to gain a desired mix level modulating it in time.

In this light it is important to understand the characteristics and implications that the exploration and exploitation process has in an organization, especially in science-driven companies. Taking for example a R&D organization, the understanding of the intimate interactions mechanisms could be of paramount importance in facing the challenges of today markets, where not only new knowledge is required to develop new products but also the ability to incorporate it and to exploit is also pivotal. The number of projects made in cooperation with external entities such as academia or CRO/CMO[1] is increasing, stressing the need to let new knowledge flow into the organization and to make it available in a short time, making organizational junction points and loosely coupled subgroups important. Moreover, new paradigms such as open innovation challenge the traditional way an organization learns, projecting the organization within a community where interpersonal relationships are essential.

Indeed, the results proposed have a big impact but there is still the need to better investigate the mechanisms for different reasons and a wider approach is needed.

1.1 The aim of the research

The present work aims to challenges the output that literature obtained in the exploration and exploitation problem through the mutual learning mechanism. The challenge starts from consideration that the topic could be represented by a tripartite diagram (reported in figure 3.1): there is an exchange mechanism (the mutual learning one), there is a context where individuals live and there is the individual's connection topology. Literature so far avoided to treat the problem in its entire nature, scholars either focused on particular aspects of the exchange mechanism either focused on specific feature of the context.

Hence, the challenges to literature spans in three directions. First, the context has been underestimated in its complexity and its power to shape the

[1] Contract Research Organisation and Contract Manufacturing Organisation

output of individuals' interaction (at least when exploration and exploitation was treated with modelling techniques). March's output could be less general than expected when a bigger part of the context is considered in the model.

The scope of the initial March's study (March 1991) needs to be broadened, uncovering the relationship among different channels such as formal connection, friendship, individuals' proximity, intersection between departments and projects team, revealing the potential impact they could have in the organizations exploration exploitation balance. Modern organizations are complex and complex organizations have received little attention in exploration and exploitation modelling literature. Relevant literature covered only partial aspects of an organization behaviour and no attempts have been made to merge the most important ones. It is widely accepted, for example, that governance and hierarchy play an important role in shaping the choice an individual could make. At the same time, it is widely recognized worthwhile to include informal networks (for example, friendship network) in the discussion uncovering some essential dynamics. However, literature has also demonstrated that office layout could influence the behaviour of individuals (Oldham and Brass 1979, Oldham and Rotchford 1983, Boutellier et al. 2008, Sailer et al. 2009). But, works on multiple layers of relationships such as formal, informal and spatial are still missing as the impact of hierarchy policies on mutual learning process is lacking.

The underlying *fil rouge* starts from the belief that the exploration and exploitation problem could not run out only at the individual level trying to generalize the results summing the output. The problem has to be considered at higher level: not only do individuals take part and give contribution to the problem but also the organization, the layout and the governance play an important role.

Second, the connections network have a not negligible effect in the interactions scenario and forcing the system to a particular topology could be wrong. The connections networks has been only recently considered and treated through pre-ordered topologies without rooting the characterization of individuals on real data. It has been demonstrated that network topology has a big impact on exploration and exploitation balance (Fang, Lee, and Schilling 2010, Aven and Zhang 2016). As consequence, connections networks should be let free to evolve and researcher should capture insight on exploration and exploitation also from their characteristics.

Third, in the era of complex systems, the wish is to explore this territory from a different perspective. Modern organization shares many aspects with complex systems and emergent properties of such systems could only be discovered considering all the relevant features at the same time.

Despite the enormous and valuable amount of results provided by the existing literature, the problem has been studied trying to isolate some features and debating them. But, if the idea that we are living in a complex world could be acceptable, then it should also be acceptable to expect some phenomena emerging from the behaviour of the organization that could not be foreseen in advance just looking at the structure. There is the need to reveal potential emergent phenomena, to understand and to discuss about the role of individuals, organization, layout, and governance in the mutual learning problem.

The present research tries to advance in the discussion rooting the methodology in the computational social science field.

Accordingly, the research walks through the path of simulation models traced by March (March 1991) and subsequent scholars (Axelrod 1997, Rodan 2005, Gupta, Smith, and Shalley 2006, Miller, Zhao, and Calantone 2006, Fang, Lee, and Schilling 2010, Kunz 2011, Mitomi and Takahashi 2015, Aven and Zhang 2016) landing in the agent based modelling territory.

This technique is recognized as extremely powerful when the model encompasses interactions among agents and it deals with different levels of communications (Secchi 2015, Edmonds and Meyer 2015, Gilbert and Troitzsch 2005). It is also recognized as indicated to deal with complex systems and emergent phenomena (Secchi and Neumann 2016, Edmonds and Meyer 2015, Troitzsch 2009).

Indeed, *"artificial society modelling allows us to grow social structure in silico demonstrating that certain sets of microspecifications are sufficient to generate the macrophenomena of interest"* (Epstein 2006, xi).

As Epstein well described (Epstein 2006, 5), the agent-based computational model is a fundamental tile in the generative social science since it is suitable to answer to the generativitist's question *"how could the decentralized local interactions of heterogeneous autonomous agents generate the given regularity?"*

Agent based models are suitable to answer the previous question mostly when the system to be studied owns some specific characteristics. The first characteristic is heterogeneity that is, the individuals are not all the same, they could differ in some traits like sex, religion, country belonging and so on. The second is the autonomy, when the system does not have a sort of top-down or central control over the behaviours of the individuals. Hence *"no central controllers or other higher authorities are posited ab initio"* (Epstein 2006, 6). This leads to the co-evolution of micro and macro structures. Explicit space and local interactions are two other attributes since the individuals *live* in a space which influences and is influenced by them. This influence is part of the interaction among the individuals during the evolution. Last, bound rationality is a feature of the system where individuals could only have partial information and visibility and the act accordingly.

The research aims also to become a useful clue for managers called to lead the whole complex organizations in a turbulent environment with the remit to obtain positive results from it and without the chance to switch off undesired layers or interactions to simplify the management.

Therefore, as a drop in the ocean of extant literature on the matter, the present work is devoted to knowledge. Knowledge as a fundamental asset in an organization, knowledge that flows within the organization, knowledge in its myriad of facets including the exploration of the new and the exploitation of the known. In particular, this work is focused on the exploration exploitation problem in a complex system, through the glasses of evolutionary networks theory and the agent based theory.

The present document is constituted by four different parts. The first one is dedicated to literature review and research question. Chapter 2 gives a contextualization of exploration and exploitation problem, starting from the description of March's model and related output (March 1991). Then the chapter splits into two parts. The first dedicated to the evolution of the exploration and exploitation problem discussing about important characteristics and achievements reached in the following years. The second part focuses on the discussion of subsequent elaboration of March's model, remaining in the modelling literature's furrow.

Chapter 3 is dedicated to the research question and its formalization.

The second part of the document is dedicated to the presentation of the model used for the research. Chapter 4 discusses about the modelling techniques used to answer the research question. A brief history of the technique starts the chapter, then its role in the research is developed. The anatomy of an agent based model (henceforth ABM) is treated and then its link with emergence is explained. The closure of the chapter is dedicated to the design approach and to a brief discussion about the technique's limits.

Chapter 5 introduces the model used in the simulation. This is not yet a formal description of the model but it is believed to be important as well. Since the presented model is extremely articulated, a general discussion about the assumptions made and decision taken (also in light of extant literature) is important to understand the subsequent chapters and the results.

Chapter 6 gives the formal description of the model following the ODD[2] procedure.

The third part of the document is dedicated to the analysis of the re-

[2]Objective, design, details (Grimm et al. 2006)

sults. Chapter 7 opens this part with a high level glance of the simulations programme, its structure and the remit of various simulation runs. The simulations table is thoroughly explained as the main output and charts used in the following chapters.

Chapter 8 includes the demonstration that the mutual learning problem is more composite than just the interaction of the individuals. Chapter 9 discusses the simulation of autonomous search, chapter 10 about the layout impact on the dynamics and chapter 11 discusses meetings. Chapter 12 includes the informal relationships analysis. Chapter 13 starts scratching the surface of emergent phenomena obtained in comprehensive simulations.

The last part of the document is dedicated to a summary of the results obtained and the contribution that this work could bring to literature and managerial world (Chapter 14) and to the discussion of the limits that this work inevitably has and the potential future expansion (Chapter 15).

Chapter 2

Exploration and exploitation

2.1 March's model

The balance between the exploration of new ideas, knowledge, possibilities and the exploitation of the already possessed ones is a central concern for any company. At the time March issued his influential paper (March 1991), this topic was well known among scholars since decades (Schumpeter 1934; Holland 1975). In "the tenacious past," Kuran (Kuran 1988) proposed a review of a number of theories which tries to explain why companies or societies do not always adapt to a changing environment, involving the so-called personal or collective conservatism. In rational models of choice the compromise between exploration and exploitation was studied in terms of rational search assuming it is possible to choose among different alternatives (Hey 1982). Within limited rationality theories the debate assumed that the search is in some way inhibited by the fact that the preferred alternative is above or under the target. Organizational learning studies argued on the difference between refinement or invention of new technologies (Levinthal and March 1981) and evolutionary theories viewed the balance in terms of process of variation and selection (Hannan and Freeman 1987).

March (March 1991) depicted exploration using words like search, innovation, risk taking, experimentation, play, flexibility and exploitation with terms like refinement, choice, production, execution and efficiency. Of course, these two alternatives tend towards opposite directions putting the organization into a serious dilemma. If the organization is more prone to exploration excluding exploitation, it could suffer the cost of bringing the new inside without gaining properly from its benefit. Too much innovation could be pursued but living in a state of underdevelopment, without the creation of competences. Vice versa, an organization exhibiting too much exploitation could tend to a suboptimal stable level. In addition, since both exploration and exploitation compete for scarce resources, the balance of the two is extremely important to guarantee the existence of the organization. As result, an organization makes a choice between exploration and exploitation either explicitly and implicitly: in an explicit way through strategy to be deployed or aware calculated decision and in an implicit way through rules, practices and routines. A subtle problem is that exploration is in a less advantaged situation; indeed, the return provided from exploration is always less certain, more remote and organizationally distant than the return provided from exploitation. This because, as March stressed, what is good in the long run is not always good in the short one, and what is beneficial for one historical moment or part of organization could not be beneficial for another historical time or another part of the organization. On top of that, future is by definition uncertain. On the contrary, since organizations learn from previous experience, the feedback coming from exploitation is certain, quick and clear and these advantages sum up over time bringing the organizations to prefer exploitation. The consequence is that the tendency to prefer exploitation forgetting exploration is potentially self-destructive on the long term. However, it is not an easy task to tune the balance also because organization changes over time and the variability of the environment is not always controllable.

In order to have an insight about this dilemma, March proposed a model in the organizational learning scenario. Knowledge is diffused among individuals in a variety of ways (instructions, documentation, examples, etc.) and what happens is that at the same time, through socialization, individuals learn from and influence the organization and the organization learns from and influences the individuals in a process called mutual learning. The balance between exploration and exploitation highlights the conflict between short and long-term strategy and between personal and collective knowledge.

But how could mutual learning be associated with exploration and exploitation problem? As Miller et al. highlighted, *"March modelled exploitation as rapid learning from a code that quickly changes to reflect best practices in an organization. Exploitation produces rapid conformity to codified beliefs and practices throughout the organization. March modelled exploration as slow learning from the organizational code, resulting in greater diversity of beliefs over a longer period of time"* (Miller, Zhao, and Calantone 2006, 709). March himself stated that *"slower learner allows for greater exploration of possible alternatives and greater balances in the development of specialized competences. Slow learning on the part of individuals maintains diversity* *longer, thereby providing the exploration that allows the knowledge found in the organizational code to improve"* (March 1991, 76). Thus, March mapped the exploration exploitation problem into the fast and slow learner in the mutual learning scenario.

The proposed model considers a certain number of individuals, which over the time socialize within the organization, mimicking the mutual learning phenomenon. An external reality with different dimensions and immutable has been considered as the source of the new and the organization has been modelled through the so called organization code. The organizational code comprises *"languages, beliefs, and practices"* (March 1991, 74). Organizational code concept should not be reduced to a repository of standard procedures or best practices; it is much more richer, it could be considered encompassing also the unaware knowledge an organization owns.

More in detail, the model envisages three distinct entities: external reality, a certain number of individuals and the organizational code. The organizational environment (external reality) is composed by m dimensions, it is independent from the beliefs individuals have about it and it is, at least in the first run, immutable (March also considered variation of external reality in turbulence experiment).

At the beginning of the simulation, external reality is populated in each of the m dimensions by equal probable values 1 or -1. A stochastic approach is used to assign the value to each dimension. Individuals and organization code have for each dimension a belief of what could be the external value and this belief could change over time. Again, at the beginning there is a stochastic assignment of values -1, 0 and 1 in each dimension for all individuals and organizational code. In this case there is a further value, 0, which represents the case "no belief" about a particular dimension, while 1 and -1 represent commitment and anti-commitment to a particular external reality value.

The model encompasses two mechanisms: "learning from the code" and "learning by the code". The learning from the organizational code part is visible when individual beliefs change or adapt to organizational beliefs in each dimension. If the value is 0 the belief is not affected (because there is no belief) but if the value is -1 or 1 and it is different from the value the organization code has on the same dimension, the value could change adopting the organizational code one. The probability to change is p_1 and it is called "learning rate", showing the effectiveness of socialization (learning from the code). The opposite mechanism, learning by the code, is modelled first creating a group of individuals which beliefs about external reality correspond better than the organizational code one. In this case the organizational code could learn from this group of individuals with probability p_2.

Letting the model run, the studied output was the proportion of reality

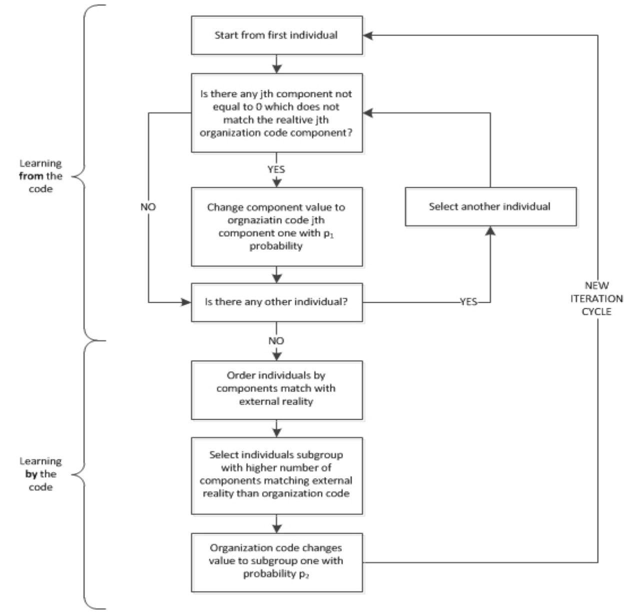


Figure 2.1: March's model workflow

that is correctly represented in the organizational code and the proportion of reality that is correctly represented by the individuals. A visual explanation of the flow is shown in figure 2.1.

March analyzed the output focusing on three main topics: the impact of learning rates, individual turnover and external turbulence. Higher learning rates bring to the equilibrium faster, and slower socialization leads to greater knowledge at the equilibrium. In detail, when socialization is slow, organization has higher knowledge at the equilibrium if its learning rate is higher; vice versa, when socialization is fast, higher knowledge is reached with slower learning by the code (figure 2.2). The knowledge peak is reached when the code learns very rapidly from individuals which socialization is low. This is related to the fact that the organization code could only learn from individual deviating from it and therefore having different knowledge to be learnt. As long as an individual keeps diversity from the code, she is a potential source of knowledge. Therefore a diversity kept longer provides the exploration able to improve the organization. An interesting point is that fast learning individuals tend to have a positive first-order effect on their personal knowledge but a negative one on organizational knowledge on the long term.

March suggested that *"there might be some advantage to having a mix of fast and slow learners in an organization"* (March 1991, 76), giving the start to a debate about possible ambidexterity of the organization. Turnover

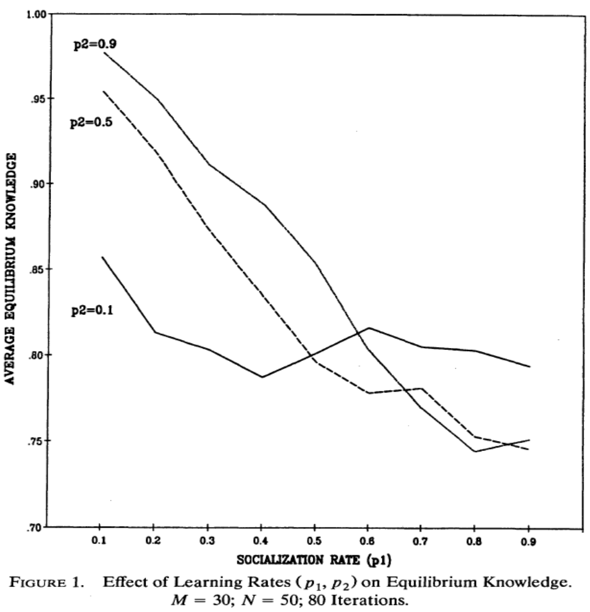


Figure 2.2: March's first experiment outcome

could be an option to introduce variability into the organization when the code could not learn from the individuals (since no more deviating individuals are present). However, the effect of turnover is complicated and it depends on the combination between learning rate and turnover rate since, for example, slow learning together with high turnover could not guarantee the right exploitation. External turbulence highlights the fact that external environment could change and adaptation is crucial for the organization. Therefore, mutual learning has a dramatic effect on the long run considering that once the equilibrium is reached, it will be kept forever. Hence, the convergence between individuals and organization is in general beneficial for both parties as long as the code has the chance to learn from individuals before they converge toward the code itself.

The output presented by March had an enormous resonance in literature and many scholars have been investigating the problem since the issue of the paper. For the rest of the proposed work two aspects still need to be further discussed. First, the exploration and exploitation dilemma evolution with particular attention to the possible balance and to the mutual learning mechanism. Second, the evolution of March's model since in the following years it has been further developed.

2.2 The evolution of the exploration and exploitation

After March (March 1991) seminal paper, the idea of exploration and exploitation has been lengthily discussed in literature. The discussion happened at all levels of the problem but three discussions are particularly relevant for the study. First, the definition of exploration and exploitation seems to be in some way elusive and a discussion of the definition could be useful. Second, there is the need to clarify the type of knowledge associated to exploration and the type associated to exploitation. Third, the idea of balance needs some comments.

Notwithstanding a fertile discussion in literature, a common agreement on the definition of exploration and exploitation is not yet reached. Gupta et al. (Gupta, Smith, and Shalley 2006) in their paper rose the provocative question "what do exploration and exploitation really mean?" witnessing this lack of agreement. Although many authors, among others Benner and Tushman (Benner and Tushman 2002), Katila and Ahuja (Katila and Ahuja 2002), and Lavie and Rosenkopf (Lavie and Rosenkopf 2006), tried to provide a theoretical background and empirical evidence to March's model output, a full consensus is still to reach.

Indeed the most influential papers on exploration and exploitation root in different theories and the overwhelming following literature spans over many fields of research. Clearly these two aspects contribute to a still not well defined situation. Moreover, the argument is so wide and with myriads of implications that there is still the need to argue and discuss about it. In the last five years two interesting reviews of works about the concept were edited. These papers are interesting because tried to shed light on the exploration and exploitation panorama, one more oriented to find the fundamental works and the other one more oriented to find the legacy of the exploration and exploitation idea in the literature.

In their innovative review, Almahendra and Ambos (Almahendra and Ambos 2015) agree with Gupta et al. (Gupta, Smith, and Shalley 2006) and stress that researchers created a set of definitions and re-conceptualizations of the idea that brings to some inconsistency and ambiguity in the understanding and interpretation of the exploration exploitation paradigm. Moreover, they also highlighted the lack of analysis of extant literature, the absence of agreement on which are the most important articles and the fact that the intellectual structure of exploration–exploitation have not yet been adequately discussed. Almahendra and Ambos try to bridge the gap with a meticulous review of 145 papers from the literature of the last 20 years. The remit is

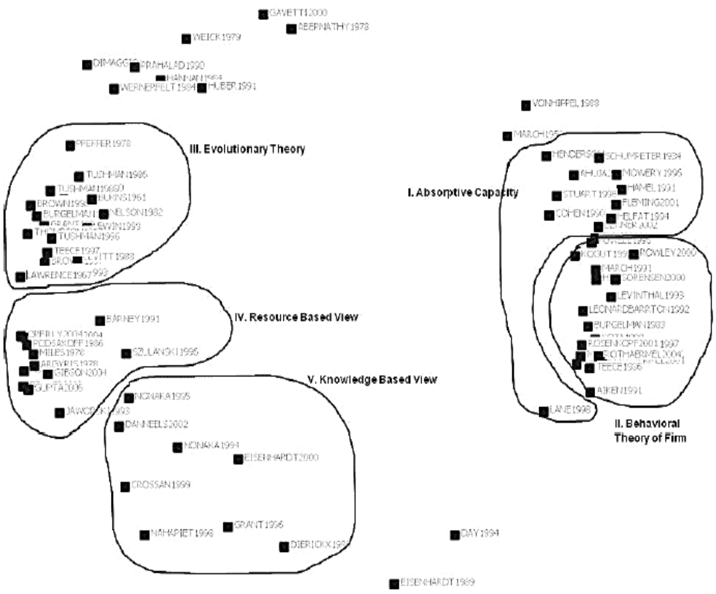


Figure 2.3: Almahendra and Ambos clustering

twofold: find the root theories upon which the discussion is based and to discover emerging trends in the discussion. The analysis revealed that at least five main groups used the concept of exploration and exploitation within which papers could be considered homogeneous in the topic of management research (Resource Based View, Absorptive Capacity, Behavioral Theory of Firm, Evolutionary Theory and Knowledge Based View). Figure 2.3 shows the obtained clustering.

Since March (March 1991) argued that exploration and exploitation compete for scarce resources, the problem has been studied from the Resource Based View perspective trying to discover the influence of firm-specific resources (assets, skills, organizational processes, information and know-how) to its learning activities. Resource Based View has become one of the most influential theory in the management arena and it tries to debate on what is essential to maintain a competitive advantage and how to use the resources in the best way. Resources could be thought in a broad meaning, including assets, processes, skills, know-how. In this scenario, scholars argued about the role of knowledge as driving force in firm growth, focusing on the ability to create new knowledge (exploration) and to integrate and combine it (exploitation).

If the problem is studied from the adaptive learning side, the behavioural theory of firm is the main stream of debate (Cyert and March 1963). The theory considers the organization changing over time encoding from previous history, experience and old routines with the aim to create procedures and strategies to gather positive future outcome. March also suggested that a company should learn either from internal and external experiences. Exploration is then defined as creation of variety of experiences and exploitation is the creation of reliability in the experiences (Holmqvist 2003).

Cohen and Levinthal (Cohen and Levinthal 1990) defined as absorptive capacity the ability to assimilate and utilize new external knowledge form the environment. Lane et al. (Lane, Koka, and Pathak 2006) elaborated the definition arguing that the usage of external knowledge is obtained through three sequential processes: (1) exploratory learning, recognizing and understanding the potential value of external knowledge, (2) transformative learning, assimilating new knowledge, (3) exploitative learning, re-creating new knowledge. In this field, the ability of a company to gain knowledge from external sources became an important topic after Cohen and Levinthal published their work (Cohen and Levinthal 1990). Their finding suggests that a company is sensitive to its environment and part of the strategy is to allocate the absorptive capacity and resources. Moreover when a company needs knowledge not related to its ongoing activities, it must create new adsorptive capacity.

Nelson and Winter (Nelson and Winter 1982) created the foundations for an evolutionary theory. Borrowing from biology, they argued that routine should be considered as the cornerstone in order to understand how an organization changes. Then, many scholars attempted to understand how routines are involved in the balance between exploration and exploitation (Sidhu, Commandeur, and Volberda 2007).

But exploration and exploitation dilemma overflows the initial application and the concept is applied in several different fields. In 2018, Wilden et al. (Wilden et al. 2018) performed a very interesting study on the concept of exploration and exploitation. They started with the idea to understand the legacy of March paper in a broader literature, with the aim to *"analyze the impact it has had on scholarly thinking using a comprehensive and structured review of the diverse research inspired by its publication ... without limiting our review to specific research domains or journals"* (Wilden et al. 2018, 353). With a sophisticated statistical analysis they found five distinct clusters of inspirations (figure 2.4).

First, studies focusing on organizational learning which is, in turn, split in two sub clusters. The first is more focused in the co-evolutionary adaptation, the second is more marketing oriented. Second cluster is made by studies on international learning and collaboration. Third and greatest cluster is dedicated to dynamic capabilities and absorptive capacity. Fourth is dedicated to

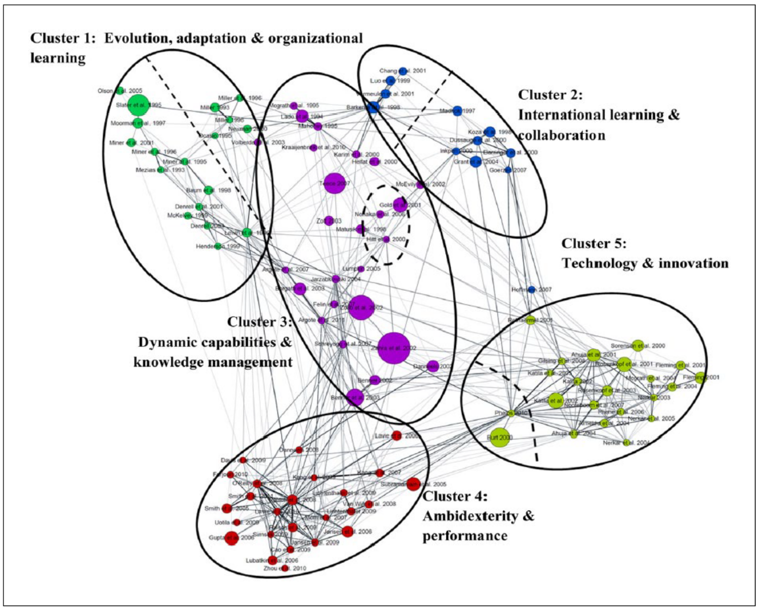


Figure 2.4: Wilden et al. clustering

organizational exploration exploitation with attention to ambidexterity and performance. Last cluster is dedicated to technology and innovation.

What emerges from the analysis is that papers related to cluster 4 and 1 (learning and marketing, alliances and acquisition) are older than papers in the other clusters. Recent works have been published dealing with organizational structure, ambidexterity, and performance (belonging to cluster 3) and innovation and technologies (cluster 5). But authors also argued that recent papers left the original investigation of exploration and exploitation embracing more a theme of innovation. Moreover, as they wrote, *"March's idea has been ≪stategyfied≫"* (Wilden et al. 2018, 361), thanks also to the emergence of the behavioural theories.

Another interesting result is the progressive shift in the study towards the organization level while March emphasised the importance of the individual level learning.

This recent review confirms that the topic is still vivid and calls for future research either in better defining the concepts either in discovering the unexplored properties: *"our analysis reveals prospects for extending the notions of exploration and exploitation to new domains, but we caution that such domains should be clearly delineated. We conclude with a call for further research on the antecedents of exploration and exploitation and for studying their underexplored dimensions"* (Wilden et al. 2018, 352).

Despite Wilden et al. call, researchers built on the original idea of exploration and exploitation and some paper brought a significant contribute to the discussion. A prolific discussion has been made on what knowledge could be associated to exploration and exploitation problem. Whereas exploration is mainly associated to learning and innovation, exploitation gathers much less consensus on its definition. Closely looking at the problem, there are two levels of questions about it: is exploration/exploitation problem linked to existence/absence of knowledge? Does exploration/exploitation problem deal with different kind of knowledge? If yes, which one? Empirical studies seem to have answered in a positive way to the first question: the exploration and exploitation problem is not linked to presence/absence of knowledge. Baum, Li and Usher (Baum, Li, and Usher 2000), Benner and Tushman (Benner and Tushman 2002), He and Wong (He and Wong 2004) positively associated exploration and exploitation with knowledge but of different kind. As Baum et al. suggested, *"exploitation refers to learning gained via local search, experiential refinement, and selection and reuse of existing routines. Exploration refers to learning gained through process of concerted variation, planned experimentation and play"* (Baum, Li, and Usher 2000, 768). As cited, He and Wong argued exploitation as *"technological innovation activities aimed at improving existing product-market domains"* (He and Wong 2004, 484) whereas exploration is a *"technological innovation aimed at entering new product-maker domains"*. Benner and Tushman involved trajectories in their explanation arguing that *"exploitative innovations involve improvements in existing components and build on the existing technological trajectory whereas exploratory innovation involves a shift to a different trajectory"* (Benner and Tushman 2002, 679).

Shifting slightly the focus, Rosenkopf and Nerkar (Rosenkopf and Nerkar 2001) associated exploration with all the activities related to learning and innovation and reserved exploitation to the usage of past knowledge. On the same side is Vermeulen and Barkema's (Vermeulen and Barkema 2001) paper which defines exploration as *"search for new knowledge"* and exploitation as the *"ongoing use of firm's knowledge base"*. Gupta et al. reasoned that all activities include at least some learning and *"it is more logical to differentiate between exploitation and exploration by focusing on the type or amount of learning rather than on the presence or absence of learning"* (Gupta, Smith, and Shalley 2006, 694).

Exploitation is about making best use of what we already know. This could be achieved avoiding past mistakes others have made, gaining the end faster and at less cost. Exploitation of current knowledge includes best-practices transfer and vicarious learning from those who seem to have more knowledge than we do. But, exploitation of knowledge brings to a convergence in beliefs: excessive exploitation can result in premature consensus (Levitt and March 1988). Exploration mitigates this because it reintroduces variation into the system, which is essential to any evolutionary process. However, Katila and Ahuja (Katila and Ahuja 2002) argued that exploitation is important not only as refinement or efficiency gaining but also for creating new knowledge. Indeed, although exploration is important in knowledge creation, providing new solutions, exploitation combines existing solutions generating new combinations.

What could be said about the balance among exploration and exploitation? Chanda and Ray stress that *"extant research has not provided a satisfactory answer to the question as to what mix of exploration and exploitation activities is ideal for organizations"* (Chanda and Ray 2015, 248), identifying three schools with three different ideas. The first posits to allocate the minimum possible resources and to devote the rest to the other activities (Levinthal and March 1993), the second school suggests to keep the organization in equilibrium among the two activities (He and Wong 2004) and the last one recommends to intervene when the organization starts decreasing the performance and to leverage on the activity which seems to give better results.

In a forum dedicated to exploitation and exploration, Gupta et al. (Gupta, Smith, and Shalley 2006) not only they posed the aforementioned questions about the meaning of exploration and exploitation, but also focused on three other aspects of the trade-off. The first aspect refers to orthogonality versus continuity that is exploration and exploitation could be seen as two sides of a continuum or being orthogonal. True that part of the answer depends on the definition of exploitation and exploitation but the theory behind the organizational learning path depends on which vision is considered. By consequence, also the balance between the two depends upon the framework chosen. March (March 1991) asserted that exploitation and exploration are incompatible and presented arguments supporting this thesis. First, both compete for scarce resources and if resources are allocated in exploitative activities, they could not be allocated to explorative activities, and vice versa. Secondly, these activities are self-reinforcing and they tend to repeat themselves: exploration calls for more exploration and exploitation calls for more exploitation. Finally, the mindset required for these activities is very different hence combining the two activities is at least problematic if not impossible. March stressed the idea saying that *"exploiting interesting ideas often thrives on commitment more than thoughtfulness, narrowness more than breadth, cohesiveness more than openness"* (March 1996, 280). But, according to Gupta et al. (Gupta, Smith, and Shalley 2006) even if March logic is not questionable, the assumptions could. In doing that, it is possible to rethink about orthogonality. There are several studies that treat exploitation and exploration as simultaneously present in an organization and then, orthogonal. Moreover, what emerges is that the orthogonality could depend on the level of analysis. Indeed, in theory, an organization could be split into different parts each of one could be devoted entirely to exploitation or exploration. In this case, the two activities could be performed in parallel. If the level is lowered to individual, it is shared opinion that no one could excel in both routines. In any case some conclusions could be drawn: (1) the degree of orthogonality is related to the scarcity of resources, (2) in a single domain (at individual level, for example) exploitation and exploration are mutually exclusive, (3) across different and loosely coupled domains exploitation and exploration could be considered orthogonal. Although scholars agree on the need to have a balance between exploitation and exploration, it is not so clear the way to achieve that balance. The first proposed is ambidexterity (Weick 1976, Levinthal 1997, Burgelman 1991, Benner and Tushman 2003, Christensen 2013) while the second one is punctuated equilibrium (Tushman and Romanelli 1985, Levinthal and March 1993, Vermeulen and Barkema 2001, Burgelman 2002, Siggelkow and Levinthal 2003). Ambidexterity is the synchronous use of both exploitation and exploration through loosely coupled subunits in an organization. Punctuated equilibrium is linked to temporal differentiation of the usage of exploitation and exploration. It is clear that ambidexterity and punctuated equilibrium are two mechanisms completely different. Ambidexterity is exhibited by organizations highly differentiated with weakly integrated subunits. Hence exploratory units are small and decentralized with loose cultures and processes while exploitation units tend to be larger and more centralized, exhibiting tight cultures and processes. Moreover, exploratory units tend to proceed through try and error processes, while exploitation units succeed reducing variability, increasing efficiency and control. The latter are the preferred locus for project management efforts (Benner and Tushman 2003). The distinction between ambidexterity and punctuated equilibrium is intertwined with orthogonality and continuity. If the analysis is bounded in a single domain (again, at individual level) exploitation and exploration could be achieved considering them as a continuum, hence ambidexterity is not feasible. This means that punctuated equilibrium is the only option viable. Conversely, if the analysis permits orthogonality of exploitation and exploration then ambidexterity could be appropriate as path. Extending the vision, it could be argued that in certain conditions the balance between exploration and exploitation could be achieved at a broader level than organizations. Hence, organizations could specialize in exploration or exploitation. He and Wong (He and Wong 2004) demonstrated the positive relationship between ambidexterity and sale performances since explorative activities are necessary to bring a new product on the market and exploitative activities needed to improve existing product on the market could also improve firm performances. A further dichotomy to discuss is the relationship between duality and specialization. Gupta et al. (Gupta, Smith, and Shalley 2006) also argued about the necessity that an organization exhibits both exploration and exploitation or if, under certain conditions, the balance could be considered avoidable. From daily experience, it could be noted that organizations live in a broader social system and hence the balance could be achieved at the level of this social system rather than at the organization level. In this scenario, an organization could specialize in exploration activities (or exploitation) and managing the trade-off at higher level. This is, in some way, the generalization of ambidexterity at higher level. The authors suggest that this configuration is likely to succeed if the following conditions subsist: (1) the two organizations with different specializations have to mutually control complementary resources, (2) the exploration specialization has a dynamic environment, the exploitation specialization a stable one, (3) the level of co-specialization is low.

Tushman and O'Reilly (Tushman and O'Reilly III 1996) discussed about ambidexterity in firms. They differentiated the exploiting units which remit is the incremental innovation from the exploring units dedicated to radical innovation. Starting from the semiconductor industry, they argued that only two out of 10 leaders of vacuum technologies in the 50's were present in the 70's. The causes could be disparate: the company decided not to invest in new technologies or it invested in the wrong ones. Or, the cause could be cultural. What emerges is that companies always experiment periods of incremental changes and period of disruptive innovation and these discontinuities could be ascribed to technologies, competitors, political and market conditions. Hence, authors argued that the companies face a dilemma, on the short run they have to be aligned to strategy, structure and culture but in the long run to be successful they necessarily have to destroy this alignment. Hence, ambidextrous companies are needed to solve the dilemma.

Rosenkopf and Nerkar (Rosenkopf and Nerkar 2001) discussed about spatial impact on exploration and exploitation taking as framework of their study the optical disk industry. They suggested that path-dependant mechanism that involves search in different dimensions is crucial for the learning of a company. Taking into account four types of different explorations (local, internal boundary-spanning, external boundary-spanning and radical) coming from the differentiation between technological and organizational boundaries, they found that drawing knowledge from different technological areas needs competences and skill dramatically different from drawing knowledge from different firms.

Katila and Ahuja (Katila and Ahuja 2002) gave contribution to the topic at least in two ways. They fostered the concepts of research scope and depth together with local and distant search and, they argued that the propensity of a company to cite different patents (exploration) or to cite similar patents (exploitation) could give insight on its future product development. Then, they challenged the idea that exploration and exploitation is always struggling with scarce resources since, with public resources like patents, exploration and exploitation could be made at the same time. Second, they also challenge the idea that seeking knowledge is always an explorative activity. Indeed, R&D is struggled between the need to further adapt the current knowledge (local search exploitation) and experiment new knowledge (distant search exploration).

Benner and Tushman (Benner and Tushman 2003) adopted the contingency theory to discuss about the balance of exploration and exploitation mechanism. They argued that there is not necessarily a more efficient organizational structure. Rather, a company needs to dynamically balance exploration and exploitation and this requires a complex management system and extremely competent team.

Gibosn and Birkinshaw (Gibson and Birkinshaw 2004) tried to overstep the idea of structural ambidexterity introducing the idea of contextual ambidexterity. There is no more the need to structurally separate the exploration and exploitation activities, rather individuals must develop the ability to exhibit both behaviours. Firms need to create the dual capacity helping people defining an organization system and adequate procedures.

2.3 March's model evolution

March's seminal paper inspired a huge number of following studies[1] gaining an important role in literature. The model has been improved and discussed from a theoretical perspective, highlighting different aspects of the mutual learning process.

These studies could be classified broadly in two main tracks. In the first track, exploration and exploitation balance has been studied empirically, through real organizations data; whereas the second track relies on formal models (mainly computational models).

Since the presented research will be conducted with the aid of a simulation tool, the literature review is focused on main papers which elaborated the initial March's model from a computational perspective.

[1] According to Chanda et al. (Chanda, Ray, and Mckelvey 2018) more than 20,000 citations could be found in Google Scholar (on March 2018).

The basic assumption in the original model is that mutual learning between individual and organization is of great importance in developing knowledge in different environments. What emerges from the key features of the model is that March considered that the belief of one individual could not directly influence the belief of other individuals. There must be a mediation through the organizational code, superior belief actually migrates toward the organizational code. Then, organizational code conforms to these beliefs and then other individuals could adapt to that new belief. In this view, March's idea was to estimate the output distribution of knowledge and knowledge exchange within ecologies through stochastic variability (March 1991).

Assuming that mutual learning could happen only through the code, March ignored the contribution of individual interpersonal networks. Nevertheless, the importance of interpersonal network contribution has been widely recognized by different authors (Allen 1977, Miller, Zhao, and Calantone 2006). Moreover social learning theory envisages individual's knowledge dependence on interaction with other people in the social context (Schwandt 2008).

Most of subsequent studies argued on the idea of interaction among individuals. This at least at three levels: spatial, organizational and cultural. In the first case, individuals could interact on spatial dimensions, there is a pre-ordered framework of possible interactions within which the interaction could take place (e.g. the villages in Axelrod 1997). In the second case, agents could interact on organizational level, they could belong to groups that interact or not, or they could interact with manager or not (e.g. in Rodan 2005). Last, some authors study the problem of knowledge transfer and the impact of different background and individual characteristics. Topics like forgetting and distortion have been also considered (e.g. in Blaschke and Schoeneborn 2006 and Aven and Zhang 2016).

Axelrod (Axelrod 1997), in his paper about the dissemination of culture, discusses the impact of distance in exchange. In his paper, Axelrod talks about *culture* defining it as a set of attributes that are subject to social influence and uses a different approach respect to March's one since agents could interact among themselves. Again, there is not a central authority and agents followed simple rule to exchange culture. In a square grid, a sample of 100 individuals are plotted, arrayed on a ten by ten grid, where each individual interacts with four different neighbours (north, east, south, and west) except for the individuals along the edges and corner who have three and two neighbours respectively. The exchange rule is very simple: the more two agents share the same culture, the more they have chance to share other tracts ot culture. Once the interaction takes place, a treat which is different becomes equal. Of course, if two agents do not share any treat, they could not interact at all. At the end, the simulation shows only the stable regions, which could not interact among themselves. Axelrod investigates the dependency of their number on different characteristics like number of culture features, traits per feature, size of the territory and influence of interaction. Some final outputs are interesting: a local convergence could lead to global polarization, the individual interplay could shape the entire influence process and even simple mechanism could lead to counter intuitive results. In this case, the number of stable regions decreases with the increase of cultural features and large territories. Axelrod moves a step forward in considering the spatial distribution of agents and relative interaction. He considered the square grid as villages and he studied the connection with neighbours. But, the focus is on close distances: by no way distant villages could interact directly. True that most of the sharing is made with nearby neighbours, but also sporadic contacts with distant individuals could have effect on the social dynamics.

In 2005, Rodan (Rodan 2005) starts from March's paper and tries to answer the question how organizational policies impact on mutual learning, deep diving in the variety of individual- and organizational- level processes that affect exploration and exploitation of beliefs. The author speculates on two mechanisms of variance-inducing. The first one is focused on the propensity of individuals to experiment and the following influence of two forms of restraint on experimenting, studied either at individual and organizational level. The second one is focused on turnover of members in an organization. From Rodan (Rodan 2005) point of view, experimentation typically involves making choices when outcomes are unpredictable. Indeed, choices even if well intentioned and considered could be not distinguishable from a random selection which, in some way, could be assimilated as the representation of stochastic alteration in individuals' beliefs. However, considering the collective learning as the exploitation of the knowledge owned by the most knowledgeable members of the organization, beliefs would converge on those. It is clear that, since the most knowledgeable individuals do not have anybody's knowledge to exploit, if there is no random alteration of their beliefs the final knowledge level could not be higher than the most knowledgeable individual at the beginning of the simulation. By consequence, if the system needs to learn more, there should be a way for individuals to jump over these most knowledgeable individuals, and this way is experimentation. Since experimentation encompasses risk-taking actions, it allows other individuals to remove most knowledgeable ones from their positions, maybe by chance or trial and error, and to gain a more accurate set of beliefs about the environment. Exploration allows an organization to overcome the learning limits imposed by the most knowledgeable individuals at the inception while exploitation increases the efficiency (but making the organization more vulnerable to environmental changes). Nevertheless, continuous experimentation might be as bad as absence of experimentation, because it does not use prior knowledge. Moreover, individuals are not all equal in the willingness to experiment, someone is comfortable in taking risk and to decide with little information, some other is more cautious. Rodan (Rodan 2005), suggests that the individual predisposition to experiment and the constrains imposed on their exploration activities modulate the rate at which variation is inserted into the organization. Excessive, unconstrained exploration could have negative effect on learning. The author, then, argues that *"there is an inverted U-shaped relationship between experimentation and learning, and an optimal level of experimentation that provides sufficient variation without discarding existing knowledge"* (Rodan 2005, 411). A merit of Rodan's model is the introduction of experimentation and the attempt to represent the policy making and dissemination mechanism. But, the model still presents limitation on the interplay between agents. Agents could be part of the policy-making elite (if their score is greater than the threshold) and once there, they create the organizational code. Then other individuals acquire knowledge from the organizational code. This means that the interaction is still indirect.

Blaschke and Schoeneborn (Blaschke and Schoeneborn 2006) introduce the forgetting ability for the agents. They start from the consideration that March's model represents well established concepts as knowledge and learning while memory is completely disregarded. Even following scholars, progressing on organizational learning theory, did not consider this issue. In this view, authors introduce the idea of organizational memory. And forgetting is a primary function of memory and, as Luhman states, selective forgetting let the system to be sensitive to new solicitations (Luhmann 1997). Authors extend the model creating an unified view of learning, knowledge and memory. Forgetting is represented by the change in beliefs from 1 or -1 to 0 and it is modulated by two new parameters (p_5 and p_6[2]). Forgetting affects each of the different beliefs independently, this means that the code could forget also when individuals remember and vice versa. The output is in some way interesting: including code and agents forgetting does not necessarily increase the equilibrium knowledge. But according to Luhmann's theory, they found that forgetting insert a sort of instability in the system, as sensitiveness to new situations. Authors suggest to keep March's mechanism of mutual learning as foundation for future expansion of exploration and exploitation, perhaps including new characteristics as knowledge latency and reality enactment. Again, Blaschke and Shöneborn extend the initial idea but they do not chal-

[2] p_5 for the individuals and p_6 for the code.

lenge the inability of agents to interplay. They extended the initial model, maintaining unchanged the interaction dynamics.

Still in 2006, Miller et al.'s study (Miller, Zhao, and Calantone 2006) uses March's model to create two new dimensions of interpersonal learning and location. In recognizing that much of the learning that goes through an organization occurs directly from person to person, they considered that, allowing interpersonal learning, the location of individuals and their resulting networks become essential considerations. What happens is that people tend to learn from those near to them, suffering from spatial myopia (Levinthal and March 1993). Moreover, individuals tend to create dense social networks with proximate colleagues. Allen (Allen 1977) already suggested that building layout has an influence on interpersonal communication in R&D. To encompass interpersonal relationships, Miller et al. add to March's model two assumptions: 1) the individuals are situated within a grid in which each has four neighbours. This differs from Axelrod model since the grid is without edge in order individuals to have all the same number of neighbours, 2) individuals learn through engaging in local and distant search. Local one tries to find the best performer among neighbours then updating each belief to that of the superior neighbour with probability p_3[3]. Distant search involves randomly drawing four individuals from population and choosing the best performer among them. The searcher adopts the superior belief with probability p_4. Miller et al. also consider that knowledge is never fully explicit and they recognize that it could be partially converted to codified knowledge. In order to create a more realistic model they also consider that codification efforts tend to be episodic and not continual since organizations pay little attention and devote scarce resources in codify knowledge. Then, two more assumptions are considered: 1) a proportion of the beliefs is tacit. Organization only moves explicit knowledge and it does not convey the tacit one. 2) Learning from code is episodic. Every t period, the explicit elements of the organizational code are updated and individuals learn from this updated code within the same period. This incorporates the assumption that individuals learn from code immediately after it is updated. If the code is not updated in a given period, individuals pay no attention and focus on learning from other individuals. They also find that *"the small-world effect of learning through distant search becomes redundant if the organizational code facilitates knowledge transfer among distant individuals"* (Miller, Zhao, and Calantone 2006, 716). A strength of the presented model is the ability to examine the code-related learning rates (p_1 and p_2) in correspondence with the local and

[3]the terms p_3 and p_4 are so called in the models to keep continuity with March models which has p_1 and p_2. Probability p_3 and p_4 are added in Miller et al. study.

distant search ones (p_3 and p_4). When distant search is possible, the system could exhibit superior knowledge. Moreover, *"learning from the code and learning through distant search are substitutes"* (Miller, Zhao, and Calantone 2006, 716). Although Miller et al. considered the spatial dimension in a more sophisticated way than previous authors, agents still are obliged to remain in a fixed grid and to interact among themselves on a rigid way. Distant search has been introduced but with a sporadic random interaction rather with a codified and systematic way. But absolutely interesting is the finding that distant search acts ad learning through the code.

In 2007, Kane and Alavi (Kane and Alavi 2007) investigate the impact of technologies on exploration exploitation problem. The central idea is to understand how *IT-enabled learning mechanisms* impact on the exploration and exploitation problem and, more broadly, on organizational learning. To incorporate IT mechanisms, Kane and Alavi made three changes to March's model. First they allowed individuals to learn from one another arguing that *"learning is operationalized within March's simulation as a function of the individual and not of the environment or of the organization, introducing learning between individuals is consistent with the theoretical basis of learning found in the model"* (Kane and Alavi 2007, 800). Hence an individual who seeks for knowledge assembles a group of others (called Φ-group) perceived as generally more knowledgeable. Moreover, individuals are organized in groups or teams of equal dimension. Second, the model encompasses also different IT channels for knowledge exchange. First, knowledge repositories and portals with a three levels mechanism: contribution to repositories, synthesize of knowledge and dissemination. Second, team rooms used either for knowledge repository and communication. Here, individuals learn from the Φ-group made by others belonging to the same project team. Third, communities of practice where individuals assemble a network to learn from. The first observation authors make is that the different IT-channels have diverse effects on the exploration and exploitation process: knowledge portals and teams rooms tend to have more exploitative nature bringing to a rapid plateau, communities of practice have more an explorative behaviour meeting the plateau later. Once the three mechanisms play together, team rooms seems to lead the blending effect. A second interesting effect is that communities of practice seems to be influenced by learning rate but knowledge portal and team rooms mechanism are not so influenced.

Two years later, Kim and Rhee (Kim and Rhee 2009) intervene in the discussion presenting an extension of March's model including a better conceptualization of environment dynamism defining its amplitude and frequency and better clarifying the notion of internal variety in an organization. Authors believe that literature showed the importance of external environment on exploration and exploitation problem but few scholars actually focused on environmental contexts and its structure. In the proposed model, they allow the external reality to be changed and modulated by two parameters: the portion to change and the frequency of the change. Furthermore, they support the idea that a overarching organizational process need to be added to the exploration and exploitation activities of a company. Hence, adopting March's model (March 1991) and following Miller et al. extension (Miller, Zhao, and Calantone 2006), Kim and Rhee consider three organizational practices: vertical socialization, horizontal socialization and turnover. The aim is to understand how these practices influence the organization internal variety which, in turns, influences the ability to adapt to external environment. Vertical socialization is modelled by considering the interaction with the organizational code, as in March's model. Regarding horizontal socialization, authors follow Miller et al. (Miller, Zhao, and Calantone 2006) idea of local search, abandoning the distant search arguing its little effect in combination with vertical socialization. Turnover is modelled by replacing old members of the organization with new members with new knowledge (randomly assigned). The inner part of the model is then represented by the vertical socialization (learning from the code) and horizontal socialization (local search). Authors suggest that a slow vertical and intense horizontal socialization brings to the highest level of code knowledge, whereas moderate vertical and horizontal socialization leads to the lowest one. Hence, *"the level of internal variety, along with the mechanisms by which each practice influences internal variety, affect adaptations of organizational knowledge. Managing internal variety through a combination of strong complementary practices, rather than anchoring on moderate levels of those practices, can achieve the balance between exploration and exploitation"* (Kim and Rhee 2009, 11). External turmoil has indeed a strong effect on the ability of the organization to adapt. It has found that higher internal variety has different effects, depending on the frequency and amplitude in changes of the external environment. Kim and Rhee's output is extremely important, stressing the profound interconnection that the external environment has with the organization. But the model is borrowed from Miller et al.'s one with its limitations already discussed. Moreover, the environment is not oriented considering turbulence.

Fang et al. (Fang, Lee, and Schilling 2010) argue in 2010 about the exploration and exploitation balance in presence of isolated or semi-isolated subgroups in an organization. They base the work elaborating on Wright (Wright 1932, Wright 1964) idea that evolving populations actually grow divided into small groups with limited chance to interact. This limitation allows the subgroups to keep the genetic diversity and to explore different solutions in the space of possibilities. The parallel is then made with the decentralization of learning process to subunits in the organization and relative barrier to knowledge diffusion. Clusters are responsible of the existence of richer and greater amount of information because fragmentation dampers the consolidation of a paradigm and organizational convergence. Including interpersonal learning, authors consider different topologies or organization relationships: the nearly-isolated, semi-isolated and random networks. The model diverges from March's one in that the organization is seen as a complex system where all individuals could interact among themselves. Authors trust in the recent advances in graph theory and understanding of connection networks so they presented different topologies for the individuals connections starting from Watts connected caveman model (Watts 1999) and modulating the isolation with a parameter β. The output suggests that an improvement in exploration and exploitation balance could be achieved splitting the organization into small semi-autonomous units with a small fraction of cross-groups links. Semi-isolated groups help to preserve the heterogeneity of the organization and to explore more and wider. The presence of a small number of cross group relations allows the knowledge to diffuse into the organization. Fang et al. make a step forward in the extension of spatial dimension and they start studying the exploration and exploitation problem introducing also elements of graph theory. But, the network used is pre-ordered and the system is not free to evolve in terms of connections.

In 2011, Kunz (Kunz 2011) develops the paper starting from extant literature which studies the influence of different organizational levels on exploration and exploitation. In her paper, Kunz's interest is on how within-groups and between-groups exploration and exploitation activities differ and how task complexity and breadth influence the organizational learning process. With a sophisticated model in which genetic algorithms are used to simulate knowledge complexity and breadth, Kunz reveals that the introduction of between-group processes changes the performance and evolution an organization could achieve. There is a superior behaviour in organization where between-groups is present compared to organization where this level is absent. All the rest being equal, exploration and exploitation equilibrium depends on task complexity.

Schilling and Fang (Schilling and Fang 2014) extend March's model in two ways: they allow direct learning among individuals, removing the organizational code, and they build the interpersonal network with a degree of *"hubbyness"* based on Xulvi et al. (Xulvi-Brunet and Sokolov 2002) and Barabasi et al. (Barabási and Albert 1999) works. Retracing in some way Fang et al. (Fang, Lee, and Schilling 2010), the structural parameter affects the speed of the organizational learning. When the parameters is set to 0, the resulting network is a scale free one and the organization learn very rapidly. Interesting to note that scale free organization does not have the best performance in terms of knowledge as when it is *"hubby"*. The maximum is reached when the organization exhibits a moderate *"hubby"* situation. Authors suggest that the cause is linked to the presence of few extremely well connected actors in the network in the scale free and *"hubby"* configuration. Being so well connected, they consolidate and propagate superior knowledge very rapidly. Hence, hubs create shortcut in the organization and they have beneficial effects until they increase to much. In this case they *"may quickly out compete other ideas in the organization, extinguishing diversity too quickly"* (Schilling and Fang 2014, 25). Again, distinct hubs behaviours as forgetting, lying and playing favorites seems to have positive long term effect, maintaining the diversity within the organization.

The following year, Mitomi and Takahashi (Mitomi and Takahashi 2015) elaborate the initial model further detailing two aspects. The first step is to better define the behaviour of the model when organization code has to learn from superior group. March did not specify how the model behaves when the superior group has no individuals or there is no majority opinion, or what happens when the majority opinion of one component of the superior group has the same value as the component of organizational code. Moreover, they slightly change the learning probability to account the number of individuals in the superior group. A similar reasoning is applied for the learning from the organizational code side, detailing the behaviour in the grey areas. Their focus is also on a missing part of the range of values of learning rate, since March actually studied the values from 0.1 to 0.9. The authors focus on the range from 0 up to 0.1 discovering different conclusions respect to March's one. They challenge the result that slower socialization (lower p_1) leads to greater knowledge at the equilibrium and also they suggest that the initial model is not fully supportive the fact that slow learning on the part of individuals maintains the diversity longer, bringing organizational knowledge to improve. This paper could be seen as a refinement of March's initial model rather than an extension. Even if it clarify some aspect of the original model it still preserves all the limitations.

In the same year, Chanda and Ray (Chanda and Ray 2015) span the entire state space of exploration and exploitation following the orthogonal conception introduced by March (March 1991). Their basic idea is not to extend somehow the initial model or to add new functionalities rather to find an answer to a couple of puzzling questions: what explains successful companies in balancing exploration and exploitation and which is the appropriate managerial response to the environmental dynamism. The output shows that several combinations lead to high knowledge confirming that strategies including more exploration could have the same results as strategies including more exploitation. This paves the road toward the idea that the manager could have an intentional strategy toward exploration or exploitation.

Aven et Zhang (Aven and Zhang 2016) challenge March assumption that knowledge is transferred perfectly and without distortion between members. This assumption does not hold in organizations where knowledge sender and receiver are separated by social distance. Since the greater the social distance the higher the chance of transformation, authors define knowledge tranformation the intentional or unintentional modification of the original solution. Of course, this modification is likely not to be an improvement. The simulation uses a predetermined social network and it considers three different ways of transmission: (1) individuals could search the solution only using their contacts in the social network (2) individuals could search across the entire organization using public repository and (3) a mixture of the two. The results suggest that once the search is made through the public repository the homogeneity is higher and the convergence is faster. Clusters preserve knowledge heterogeneity and slow the convergence. Considering the distortion dimension the result is the opposite, with public repository the distortion is less.

Again in 2016, Miller and Martignoni (Miller and Martignoni 2016) come back on exploration and exploitation problem and its relationship with forgetting, stressing that previous literature ignored forgetting as natural part of the human learning experience. In this view, they distinguish the term belief which could be true or false from knowledge which is only correct and they modelled the agents with the inability to remember everything. Practically, it is the conversion of a belief from one value to one of the two possible others[4]. Differently from Blaschke and Schoeneborn (Blaschke and Schoeneborn 2006), Miller and Martignoni allow interplay among the agents which are randomly assigned to a connection network at the beginning of the simulation. The model randomizes the order of learning and forgetting for all the individuals. Without the presence of forgetting the output of simulations confirms March and subsequent scholars indication: learning rate and long term knowledge are inversely related. When forgetting is switched on but the learning rate is low, forgetting eliminates beliefs, resulting in a loss of diversity. When the learning rate is high, beliefs rapidly converge. In the intermediate values of learning rate, forgetting maintains belief diversity by causing agents to learn about different aspects of the environment. Forgetting maintains diversity into the organization because it reverts agents to ignorance about a specific belief.

[4]Possible values are 1, -1 and 0.

Chanda in 2017 reasons about the contribution of manager since as the author wrote *"the difficulty in definitively linking outcomes of managerial action to organizational outcomes has been a festering issue in organizational research"* (Chanda 2017, 61). In his study, Chanda introduces the idea of complexity as the extent of interdependence among elements of a managerial decision and a new variable called the probability of organizational success. According to Chanda, there is a link between the organization knowledge and the probability of success: the called complexity link. Then he compares the output also with his previous work (Chanda and Ray 2015) which strongly suggests the potential presence of managerial intentionality toward exploration or exploitation. Chanda final position is that higher complexity reduces the scope of application of managerial intentionality, restricting the optimal exploration and exploitation configuration. But, organization could regain some managerial intentionality by trading-off some performance in stable environments. If the environment is dynamic, the managerial intentionality could be regained by lowering the aspiration level.

Still Chanda in 2018 (Chanda, Ray, and Mckelvey 2018) introduces the collective human capital (CHC) and introduces also different populations of capital (*marchian*, *supra-marchian* and *sub-marchian*) exploring the continuum of exploration and exploitation. He relaxes one of March's assumption: agents could start with a predefined set of beliefs overriding the random assignation made in the original model. Hence, marchian population is created with random assignation and it exhibits moderate CHC. *sub-marchian* population is created having a CHC lower than *marchian* one and *supra-marchian* population has higher CHC than *marchian* population. Substantially, the model has the same features of March's one. March's output suggests that is preferable to be prone to exploration than exploitation but, from Chanda study, it seems true only for organizations with moderate or high CHC. Conversely, with low level of CHC, organizations should be more devoted to exploitation than exploration. This result holds with the assumption that the organization has limited access to heterogeneous knowledge and it operates in a moderate environmental turbulence.

March's initial model has been challenged, extended and improved since its first appearance in the literature. Scholars clarified many important aspects of exploration and exploitation problem but there is still space for further development.

Many concepts have not yet found place in scholars works and much of the complexity organizations experiments everyday is not yet understood.

Chapter 3

The research question

As previously mentioned, there are different ways of thinking about exploration and exploitation dilemma. One assumes that the balance between exploration and exploitation needs to be found and possibly the right mix needs to be achieved. One of the main challenge about this view is to find a meaning of right mix between exploration and exploration.

Although reachable, the answer to this question could be in some circumstances lacking of practical implications from a managerial point of view. A new development project suddenly added to an R&D pipeline portfolio could compromise the balance achieved. At that point, regardless the reasons why the project has been added, there is no added value in arguing whether or not it fits in the right mix neither which could be the new optimum given the project. An historical interesting example appeared on Aviation Week and Space Technology on October 15th 1979 dealing with Lockheed L-1011 Tristar airplane project. The Tristar plane was a revolutionary aircraft from a technological point of view with the introduction of many advances and features but the exploration rate on this project was too high comparing with the time to exploit the innovation within the project. The point here is that the disequilibrium between exploration and exploitation was not due to managers' inexperience or incapacity (most of them were highly skilled and professional) but due to the necessity to be on the market soon in the historical moment of air transport expansion. Therefore, the project was an economic disaster but a strategic success. It is evident that in this situation the best exploration and exploitation mix is of relative importance. This was not, of course, an isolated case and modern organizations often face stresses that compromise the balance. Nevertheless, the rate of change in the industry world makes this effort bigger and bigger.

We prefer a second way to consider the exploration and exploitation dilemma: it assumes that the organization could be simulated and parametrized. In this view, it is possible to assess which is the ratio between exploration and exploitation changing the value of the parameters. For example, it could be interesting to understand how the balance changes according to the increase of projects number or which could be the parameters to leverage in order to obtain a desired balance. In this case, once an external challenge comes along and once the managers have chosen the appropriate balance level, it is possible to understand which characteristics of the organization should be leveraged in order to cope with the desired level.

Again, considering a modern organization, it is extremely difficult to know in advance, which is the output of exploration and exploitation ratio only looking to the structure. An illuminating example could be a typical R&D organization. Possibly, the organization is a functional area made up by different departments with the aim to develop new products. For example, organization scope could span from new candidates identification up to market launch. In order to cope with complexity related to the broad number of activities and the concurrent number of projects, the organization is managed as matrix organization. For each new product, a project team is put in place with a delegate of each technical department and managed by a project manager. The remit of the project team is to drive the product to market according to milestone and budget, assuring a coherent and plausible strategy. Technical departments have the goal to deal with potential different topics, trying to insert innovative solutions to face development issues. In light of this vision, the exploration and exploitation problem could be viewed such as the balance between project teams dealing with exploitation of consolidated knowledge and department endeavouring with exploration of new knowledge. Often, there is no clear cut between project teams and department activities since the team member could be a department delegate and since it is the department itself that executes the activities but it is rather evident that project team tends to prefer exploitation in order to fit activities with project constrains (time, budget). The department tends toward exploration to solve technical issues during development updating the state of the art of internal knowledge to face increasing technical complexity.

Albeit the huge work done and the extremely positive results obtained, in a complex scenario the exploration and exploitation problem it is far from being completely understood in it all facets and contexts.

The problem could be seen as a tripartite scenario: there is an exchange mechanisms, there is a context and there is a connections network of agents. All these three parts mutually influence themselves as represented in figure 3.1 and they should be studied together. The context gives the environment

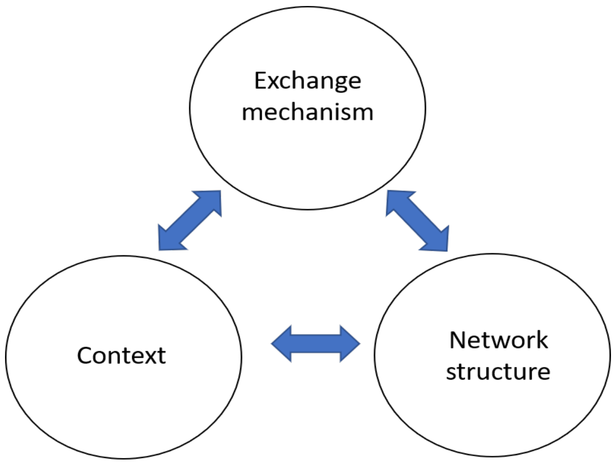


Figure 3.1: The proposed exploration and exploitation problem structure

for the exploration and exploitation problem, it gives the boundaries and the rules and it influences the topology of the social networks: different contexts lead to different networks. At the same time, context influences the exchange mechanisms output, the same mechanism could behave differently if the context in which is used changes. An open source software communities or a rigid and structured organization could lead to different results even sharing the same exchange mechanism. At the same time, the context could dominate the underneath social structure. Rules, policies and constraints could foster or prevent some links to happen. Of course, the exchange mechanism itself has an impact on the context mostly during output analysis. Different mechanisms in the same context could, in principle, result in different interpretations. Also the network structure is potentially impacted by the mechanism as it could impede the formation of certain link or it helps maintaining it according to its inner rules.

More in details, the exchange mechanism is the way individuals share knowledge, in March's paper the selected one is the mutual learning. This mechanism has been widely used in literature and well studied and commented. Its way of work is clear and also its behaviour is well known. In the proposed work we adopt the mutual learning mechanisms as the mechanism individuals have to share knowledge. This because the aim is to study the effect of complex contexts and network structure in an exploration and exploitation framework. Hence, the idea is to use a well known mechanism to be able to compare the output with literature and to suggest possible future research.

Second, there is a context in which the mechanism is absorbed. The context is the social framework where the mechanism acts. The context brings the rules, the way individuals meet or interact. Previous literature never attempted to consider more than one interaction layer at the same time in the models. With the exception of Kane and Alavi (Kane and Alavi 2007), authors have modelled only one communication channel. Again, Kane and Alavi, studied the impact of IT related channels without considering others. Considering the scenario presented before, there are different elements concurrently affecting exploration and exploitation balance form a context point of view: (1) hierarchy and project governance, shaping the backbone of the organization (2) informal relationships present in an organizational network, (3) spatial constrains, because a big organization with highly specialized knowledge necessarily occupies ample physical layout. Historically, the formal interactions were firstly studied by scholars but it is not only a problem of hierarchy, organization and governance, some other aspects need to be considered to effectively address the underlying mechanisms: networks among individuals play a pivotal role. Formal, informal and spatial distance networks could contribute in shaping the organization behaviour. Aven and Zhang pointed out that *"investigating the effect of social network distance on organizational learning is particularly challenging, because real-world organizations rarely present opportunities to disentangle the influence of social distance from a host of other confounding variables such as organization size, network structure, density, and member characteristics"* (Aven and Zhang 2016, 1105).

Informal networks play an important part in the picture even if, as Soda and Zaheer said, *"research has only recently begun to explore the organizational consequences of the simultaneous existence and interplay between formal and informal elements of an organization"* (Soda and Zaheer 2012, 3). But, they need to be considered together because, as authors suggest, *"informal and formal organizational elements need to be examined together in terms of their mutual interplay – not only because taking a holistic approach provides with a more complete understanding of organizational functioning, but also because the two sets of elements act in tandem and their interrelationship has implications for performances of individual organizational actors"* (Soda and Zaheer 2012, 751). Spatial configuration is the third important point to consider because *"it has been demonstrated empirically that space matters"* (Sailer and McCulloh 2012, 47). Spatial configuration could create different level of propinquity[1], which, in turn, gives rise to clusters. Local clusters could heavily impact the interplay among actors and then the exploration

[1] In the presented work the term propinquity is used instead of proximity since the physical closeness is the object of the study.

and exploitation balance. Moreover, closely looking to the real composition of the organization, matrix elements could be sought. These elements highlight the importance of junction points balances as discussed before. Hence, the impact of the organization elements is supposed to be important and by consequence not avoidable in the study. Hence:

Proposition 1. *Different communication channels lead to different results in exploration and exploitation scenario.*

Proposition 2. *The superimposition of different communication channels influences the exploration and exploitation ratio.*

Proposition 3. *Not in all the configurations the presence of slow learners leads to superior knowledge on the long term (henceforth March's effect).*

Encompassing all these characteristics of an organization, it is evident that exploration exploitation dynamic shows some treats of a complex system: (1) individual behaviour is non linear and can be characterized by threshold, if-then rules, or non linear coupling, (2) individual behaviour exhibits memory, path-dependence, and hysteresis, non-markovian behaviour or temporal correlations, including learning and adaptation, (3) interactions are heterogeneous and can generate network effects. Convinced that emergent phenomena could rise from this scenario, the proposed approach aims to attack the exploration and exploitation problem to a different perspective. Rather than focusing on particular feature of the problem, a broader perspective is adopted. Hence:

Proposition 4. *The superimposition of different communication channels creates a complex scenario an hence emergent phenomena are expected.*

Third, the connections network changes in two dimensions: time and context. The connections network has its own natural evolution over time, according to the need it has to cope with the assigned task. Not all the connections are always useful for all the time period considered. At the beginning of knowledge exchange, for example, more connections are required to establish a good sharing whereas at the end the connections could take different shape, when there is less knowledge to share. Moreover, changing the shape while performing the task, the network could assume different topologies during the evolution. This not only in terms of more or less connections but in terms of distribution of links. What is still missing in the extant literature about exploration and exploitation is a focus on the connections evolution in terms of link and unlink. Authors who dealt with the spatial problem in the exploration and exploitation problem always started with a defined structure of relations and they never modified it during the simulations. The ability to unlink is a feature of the presented approach, modelled with the agent's characteristic of keeping trace of all her previous interactions and maintaining only the useful connections (i.e. those fruitful in terms of knowledge sharing).

The other dimension is the context. The social network assumes different shapes depending to the context in which it is studied. Literature developed this topic in three different ways also linked to the maturity degree of social network analysis. In the first period, scholars did not pay attention to the social structure, also March in his paper did not allowed interpersonal relationships. Some of the subsequent works (Rodan 2005, Blaschke and Schoeneborn 2006, Mitomi and Takahashi 2015, Chanda and Ray 2015, Chanda 2017, Chanda, Ray, and Mckelvey 2018) adopted the same model, de facto ignoring the interpersonal relationships. The following evolution is represented by the extension of the model to simple fixed interconnections or almost fixed connections (Axelrod 1997, Miller, Zhao, and Calantone 2006, Kane and Alavi 2007, Kim and Rhee 2009, Kunz 2011, Aven and Zhang 2016). In this view agents are free to connect on a grid with immediate neighbours or to converse with distant individuals randomly chosen. Third evolution is the adoption of network topologies borrowed from well know social network analysis models (Fang, Lee, and Schilling 2010, Schilling and Fang 2014, Miller and Martignoni 2016). Random, small world or caveman networks create the background for the model. Although extremely interesting and important output came out, a topology is pre-ordered at the beginning of the simulation. Moreover, when scholars compare networks, the topology is selected *ex ante*. This means that for a specific context and mechanism, literature compares the performance of small world networks compared to scale free ones, assuming that a particular topology would have been chosen by the context itself to execute the task. In other words, there is the need to understand whether or not a particular context brings to a specific topology. This is done starting from the absence of a connection network and let the system create its own one. Those two aspects together call for a network free to evolute, free to born, rise, and possibly to die. Hence:

Proposition 5. *Different exploration and exploitation contexts bring to different network topologies, given the same exchange mechanism.*

Proposition 6. *In a context where the unlink feature is present, network topologies evolve differently over time according to different contexts.*

As consequence, the idea to study the mechanisms affecting exploration and exploitation balance in a complex scenario poses a methodological issue. Historically, March and Simon (March and Simon 1958) suggested the paradigm that the problem of understanding an organization could be decomposed into sub-problems treatable with pieces of research. Once all the pieces have been studied, a grand theory could be created summing up all the results. Now, due to the intimate and intricate interconnections of the different aspects of the problem this approach might not be the best one. Hence, instead of paradigm-driven research, problem-driven research could be more suitable. As Davis and Marquis wrote, *"problem-driven work is distinguished by its orientation toward explaining events in the world—starting with the question ≪why is it that . . . ?≫Paradigm-driven work, in contrast, begins with hypotheses deduced from theory intended to be general"* (Davis and Marquis 2005, 334). Simulation technique could help to cope with problem-driven research because not only does it allow to understand the hidden mechanisms of a complex system but also it *"could help the scholar to explore research questions highlighting the implications of the hypothesis they make"* (Fioretti 2013, 228). To develop the proposed study, a simulative approach is adopted. The simulative way is preferred not only for its characteristics in line with the type of problem to solve but also since it is not easy to have access to individual level data on knowledge transfer and field experiment are extremely challenging (Levine and Prietula 2012). Among the plethora of possible different simulations techniques, we select the Agent Based Model (ABM) which main benefits are discussed in the next chapter.

Chapter 4

Agent Based Model

The idea to use simulations in social sciences is not new. Troitzsch (Edmonds and Meyer 2015) poses the foundation more or less at the birth of computer era since eminent scientist as John von Neumann was also a pioneer in social science and also Herbert Simon, a founder of social science, was an early adopter of computers as aid in his research. Although such scientists were involved in simulations, the idea of computer simulation in social science had a very difficult beginning (Troitzsch 1997). After few initial experiments, the idea of simulation became more vivid only after some years.

Forrester in mid-1950s developed the approach called *system dynamics* trying to solve complex problems with a set of mathematical equations (mainly differential equations). The idea behind this approach is that the system could be described by aggregated variables even if it is made by different parts or individuals.

As Gilbert and Troitzsch pointed out (Gilbert and Troitzsch 2005), these first models were too focused on prediction while social scientist is more interested in explantation and understanding. Always in the 1950s Orcutt developed the *microsimulation* approach where the system is modelled by individuals owning different states and the evolution is studied through transition probabilities estimated from statistics. Then, aggregated statistics could be derived from the output. This approach collected a good success especially in some parts of the worlds and it influenced also the policy makers. One of the distinctive aspect of Orcutt's approach is the absence of explanation, it simply tries to forecast. Moreover, the individuals could not interact and there is no mention of their intentions or motivations.

First simulations in 1960s tried to predict the output of relevant problems. Worth to mention is the tentative to predict the future of companies

(Forrester 1961) and the prediction of referendum campaigns, especially the fluoridation referendum (Abelson and Bernstein 1963). The latter could be classified as a forerunner of modern agent based models since it simulated 500 individuals exposed to information coming from different communication channels and with the ability to interact among themselves.

In the 1970s and 1980s great interest was given to new approaches hoping to simulate abstract processes. Cellular automata and the famous game of life (Gardener 1970) gave new pulse to the discussion. One of the most cited work based on cellular automata it the segregation study performed by Schelling (Schelling 1971) which provided a provocative output: segregation could not be avoided, it happens also within tolerant people. Another important contribution was given by Axlerod (Axelrod 1984) with the Tit-For-Tat model based on the prisoner's dilemma.

These simple models became the background for always more complex studies and scenarios where individuals are pro-active, goal-directed and autonomous. In the so-called generative social science (Epstein and Axtell 1996) the researcher could grow the experiment and understand the functioning mechanism of a system. It is in the 1990s that simulation gained a big improvement when some techniques like non linear dynamics and artificial intelligence were borrowed from other disciplines as biology, physics and mathematics. This led to the development of multi-agent models.

A graphical representation of the evolution of simulations in social sciences is shown in figure 4.1. The grey area represents models based on mathematical equations, the white one is dedicated to objects, event or agent based models.

4.1 Agent based models in the research

Modelling means creating a *model* of a target which is simpler to study than the target. Hopefully, the conclusions collected from the study of the model could be applied also to the target (Gilbert and Troitzsch 2005). Also, modelling is the imitation of part of a real system in a certain moment of time (Davidsson and Verhagen 2013). The model is always simpler since any possible model could not consider all the features, facets and nuances of the reality, especially in social sciences. In social sciences, the target is often a dynamic system that changes over time, interacts with the environment and has an impact on the environment too.

But, as Prietula et al. argued (Prietula, Carley, and Gasser 1998), what is lost at the level of detail is regained by the researcher as control on the environment. She could models some aspects that are difficult to observe

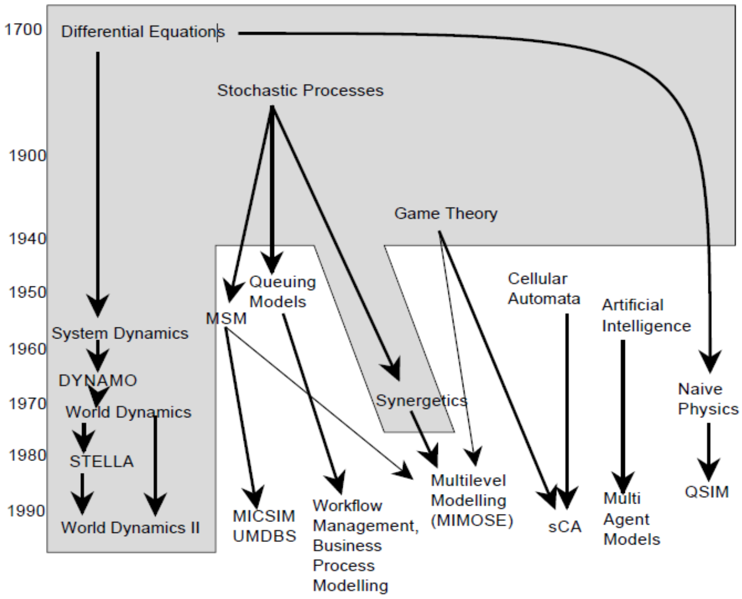


Figure 4.1: The development of contemporary approaches to simulation in the social sciences (Edmonds and Meyer 2015)

in the reality as the system dynamics and all the non-linear effects. Hence, she can vary the parameters and study the output in what has often been called *"virtual experiments"*. Coen resonates writing that *"simulation allows researchers to learn about the processes and mechanisms of dynamic systems unavailable from correlational models"* (Coen 2009, 2).

There are different systems to model and Davidsson and Verhagen identify four macro groups[1]: human-centered systems, natural systems, socio-technical systems and artificial societies (Davidsson and Verhagen 2013). In the first group it is possible to count models dealing with human societies which are groups of individuals with goals and activities, models dealing with organizations which consider structures of peoples with an assigned tasks or activities and models dealing with economic systems with a particular focus on economic activities as trading or servicing a particular market. Natural systems could be, in turn, split in animal societies where the interest is, for example, on ant colonies or bird stocks and ecological systems where also the interaction with environment, plants and humans are considered. The socio-technical groups encompasses hybrid realities where natural individuals interact with artefacts as in public transport modelling. Last, in artificial

[1]excluding physical, chemical, mathematical, biological models

societies the main topic is as set of software or hardware entities.

Generally speaking there are several ways to build a model which serves diverse purposes. A model could be created to manage a system and to support the decisions to take in order to maintain it. A model could be useful also in the design or engineering phase where with simulation it is possible to choose among different scenarios and to understand which of them is the most promising or robust solution. Models are adapt also for theory or hypothesis verification, comparing two (or more) of them. Education is another field in which model could have a role, giving to the user a better insight of the system mechanism. Strictly related to this, is the use of modelling for training. In this case, models could act as a counterpart of the learner and give a solid support in the apprenticeship. Another flourish field of action for modelling is entertainment where users manage different systems as government, cities, social relationships as so on.

Finally, models are extremely useful to understand and to gain further knowledge about specific domains. Those are called explorative studies and the peculiar characteristic is the absence of a theory or model to verify: the intent is to study the phenomenon. Modelling is then supportive in later theory creation or verification. Gilbert and Troitzsch (Gilbert and Troitzsch 2005) argue that there is an increasing interest in this approach supporting the discovery and theories formalization. They also point out that there is an advantage using simulation over mathematics approach. First, simulations require computer program and, hence, programming languages. Programming languages are more expressive and less abstract than the mathematical way. This of course enlarges the access also to non-mathematicians. Second, programmes are naturally adapt to deal with parallel processes and not rigid defined sequences. Mathematical model has some difficulties in managing these structures. Modularity is another advantage: programmes could be more easily improved or equipped with new functionalities than a set of equations. Last, the modelling approach allows to naturally incorporate heterogeneity. Agents could, in principle, have all different characteristics.

The simulation is basically a run of the model with the researcher witnessing what is going on.

There are mainly two different approaches in modelling: statistical modelling and simulation modelling (Gilbert and Troitzsch 2005) or equation-based and individual-based modelling (Davidsson and Verhagen 2013). In the statistical approach, the relationship among target and model is quite well understood: a set of equations (with all relevant parameters) is created, data are collected and the model is created estimating the equations on data. Then, the analysis is made by prediction in which the model gives a forecast based on collected data and, secondly, the analysis of the parameters

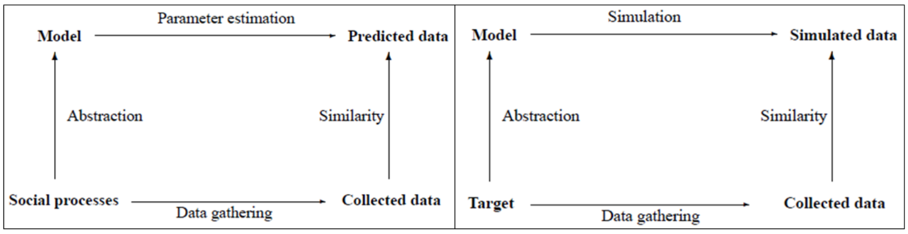


Figure 4.2: The logic of statistical modelling (left) as a method and the logic of modelling (right) as a method (Gilbert and Troitzsch 2005).

and their impact is done (figure 4.2, left side). Modelling through simulation shares most part of the approach (figure 4.2, right side) even if the core is different. In this case the estimation is made on a computer program and not on statistical equations. This difference has a profound implication: statistical models typically explain correlations among measured variables, simulation models are more related to processes dealing with objects that could not be represented in statistical or mathematical way.

According to Secchi and Neumann (Secchi and Neumann 2016), agent based models (ABM) are extremely important in the organizational behaviour field of research. Given its extremely broad meaning, organizational behaviour encompasses elements coming from management, applied psychology, organizational sociology and economics. Of course the presence of diverse elements creates the need for the scholar to cope with two distinct and opposite tendencies: in one direction there is the need to incorporate new insights from different disciplines to extent the research areas and in the other direction there is the need to keep the identity of the organizational behaviour discipline. According to the author, this tension among specialization and cross contamination could be reconciled also with the help of agent based models: *"ABM [is] a game changer and [is] something that could serve as a mediation to facilitate cross-disciplinary research without compromising specialization"* (Secchi and Neumann 2016, 3).

Agent based simulation helps for abstraction and abstraction could be interpreted by different disciplines. ABM also fosters the keeping of discipline-dependant elements with the ability to provide diversification. But, ABM has other features that make it a good a candidate as standard tool in organizational behaviour. First, the simulation implies an insight in the social dynamics. Just the mere act of running the simulation allows the researcher to have a privileged seat in looking at social changes. Second, ABM could consider diversity: there is no limit in the potential customization of agents and environment. Potentially all agents could be different. Moreover, also

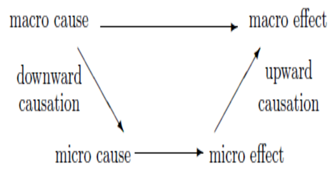


Figure 4.3: The downward and upward causation (Coleman 1990).

aggregates of agents could be simulated and studied. Last, agents are socially embedded that is they do not imitate other agents but they could be influenced by them.

Agent based models are extremely adapt to social science also for their flexibility to design social structures. As Troitzsch noted (Troitzsch 2009), social simulation needs to cope with systems which exhibit a degree of autonomy and capability to deliberate. Those ingredients, he noted, are absolutely necessary for any theory and model, given that it is difficult to accept humans interpreted as mere deterministic or stochastic automata. Likewise, dealing only with aggregate variables or high level modelling, only the macro level of the society is represented. This concept brings into the discussion the known Coleman boat (Coleman 1990). In his book, Coleman defined the link between micro and macro levels as figure 4.3 shows.

The central idea of the diagram is that human actions are determined by the social environment and at the same time they influence the social environment itself. This results not only in a change of the society (that is the macro level) but also it results in a change of the individual actions (that is the micro level). Of course, the diagram operates a sort of simplification, since the reality does not necessarily have only two levels and the interaction among individuals and society could happen through intermediates.

As Troitzsch highlighted, this upward and downward causation diagram well depicts the two social levels (micro and macro) needed in many social theories ad models (Troitzsch 2009).

There is a vast literature regarding ABM in managerial science (Wall 2016), but some arguments are more adequate to be treated with ABM techniques. Davis et al. (Davis and Marquis 2005) argued that ABM and in general simulation is effective for developing a theory when the research question involves a tension or trade-off. Bonabeau (Bonabeau 2002) suggested the use of ABM when potential emergent phenomena could be present. There are several situations where these phenomena could rise, as previously described.

Macal and North (Macal and North 2006) suggested several situations which could benefit by ABM, for example, when there is a natural representation as agents or decisions and behaviour could be defined discretely, when it is important that agents adapt, change, learn and engage in dynamic strategies, when spatial component are important, when past does not predict the future.

Fioretti (Fioretti 2013) explored further the potential field of application of the technique arguing that ABMs are suitable when the researchers could not ignore the structure of the interactions between social actors (structure matters), when the interest is in the overall behaviour coming from the interactions among actors (bottom-up approach) and when the research question is focused on out-of-equilibrium dynamics where the potential presence of more than one equilibrium is not important to the question itself (out-of-equilibrium interest). Put it simple *"a more precise statement would be that if relations are intricate and the structure matters, if an actor-to-structure, bottom-up perspective is sought and equilibrium is not a concern, then ABMs are likely to be the only tool available beyond qualitative research"* (Fioretti 2013, 233). According to Miller (Miller 2015), the creation of a theory involves the representation of phenomena. A true representation is linked to the real features of the phenomena being represented. This view is consistent with the work of the modeller since *"scientists use models to represent aspects of the world for various purposes"* (Giere 2004, 747).

4.2 The anatomy of an agent based model

A formal agreement on the definition of agent based models (ABM) has not yet reached among scholars and practitioners (Secchi 2015). According to Gilbert (Gilbert 2008) ABM is a computational method that enables a researcher to create, analyse and experiment with models composed by agents that interact within an environment. ABM is then used to derive findings from the system's behaviour ("macro level") and from the agent's behaviuor ("micro level") (Bonabeau 2002, Epstein 2006). ABM reproduces agents' interactions and it is not a tool for handling data but it could help the scholar to explore research questions highlighting the implications of the hypothesis they make (Fioretti 2013). The main idea behind is to create a computational model of the system under investigation and then let the system evolve and observe the behaviuor of the agents and the possible emerging properties of the system. Regarding this kind of approach, Axelrod (Axelrod 1997) defined it as a third way of doing science.

An agent based model is basically made by three elements: agents, environment and rules (Gilbert 2008).

The agent is the peculiar characteristic of a ABM and it is through it that the research is conducted (Secchi 2015). The agent could impersonate different entities such as organization, individuals, aggregates of individual, depending on the remit of the study. Every agent is autonomous and she has some defined characteristics. Virtually there is no limit for the owned characteristics which could be few in simple models or much more in complex systems. The characteristics shape the agents and they could be assigned independently for each agent. Heterogeneity is then another ABM feature which allows the researcher to tailor the reality representation. The second fundamental characteristic of the agent is the ability to interact. Agents could interact following pre-ordered rules or could interact as output of emerging phenomena during the simulation. Again, the modelled agent could see the environment around her and interact with it without having the perception of what happens at higher levels. This is an important point since is generally perceived as unrealistic that an individual could manage at the same level all the interactions and could own such extensive knowledge. If the agents represents a human, than it is possible to model her state or behaviour, that is the decisions and action taken. State could be made up by physical and mental state (Davidsson and Verhagen 2013). Physical state is much more straightforward to model than the metal state and typically is represented by a vector containing information as sex, age, health status and other aspects. The mental state needs the sentiments, intentions and desires to be taken into account. Different solutions have been proposed and two of the most known are the BDI and BOID models. The behaviour of an agent could be modelled as deterministic or stochastic and it could be influenced by the state of the agent itself (when the agent relies basically on her own state to behave), by the state of the environment (when the surrounding environment is able to shape agent's behaviour), by the state of other individuals (when also others agents could influence her decisions) and by social states (if the agent in the model could reason at social level). These features could be static or dynamic: in the first case the decision rules do not change over the simulation whereas in the second case they could change as consequence of learning or adaptation.

The environment is the space where agents *"live"* and interact. According to Gilbert and Terna (Gilbert and Terna 2000), the environment is a multidimensional area where agents are located. Of course, the environment is defined before the simulation and before posing the agents inside. The environment could be the representation of a physical space when the location is important or could be abstract when, for example, the interactions among companies are studied. According to Davidsson and Verhagen (Davidsson and Verhagen 2013), there are three important aspects to consider when dealing with the environment. First, spatial explicitness should be carefully modelled. In models where agents are firms and the their relations is the object of the study, there is no need to explicitly define the environment and the resulting code could be easier. But, in some other models, the location plays an extremely important role as in the cellular automata model (Gardener 1970). The location could be specified either as relative position or absolute position. Cellular automata model is an example of the first approach since it is important how the cells are displaced but it is not important to know where physically they are. In the second case, when absolute location is important, it could be possible to incorporate information as the geolocalization (Schüle, Herrler, and Klügl 2004). In their work, Schüle et al. propose the coupling of geographical information with the agent representation to capture the richness of spatiality which is extremely important in certain environment. A couple of traffic simulations are provided using information from SeSAm (Shell for Simulated Agent System) and Arcview systems. The second important aspect to consider is time. Time explicitness could be avoided if the problem does not need its management or it could be expressed by a sequence of steps. The last point to consider is the presence the exogenous events. These events happens regardless the status of agents and their influences. The presence of these events create a stochastic environment rather than a deterministic one.

The rules complete the list of the pieces of an ABM model. Rules are defined by the social scientist and allow agents to behave. The rules could be behavioural, interactional o time-dependant. Behavioural rules define what an agent could do whereas the interactional rules define how agent interact with other agents or the environment. The time-dependant rules are used to change agent or environment characteristics during the simulation. The rules are what makes the agent *autonomous*, *pro-active*, *reactive* and *social.* The agent could have a task to perform or a goal to reach but she could also react to external trigger or event and she interacts with other objects.

Figure 4.4 represents the agent based model structure as described in the work of Gilbert and Terna (Gilbert and Terna 2000).

Agent based models are extremely capable to create a controlled (in silico) experiment where the social scientist could observe the interaction of the agents. This interaction permits to focus on the micro-macro link that is the possibility to understand how general macro properties of the system are generated by actions at micro level. Moreover, as noted by Squazzoni et al. (Squazzoni, Jager, and Edmonds 2014), the emergent properties influence in turn the behaviour of the agents. Hence ABM is the mean for sophisticated representation anyway keeping in mind that they are *"a simplification – smaller, less detailed, less complex, or all of these together – of some other structure or system"* (Gilbert and Terna 2000, 2).

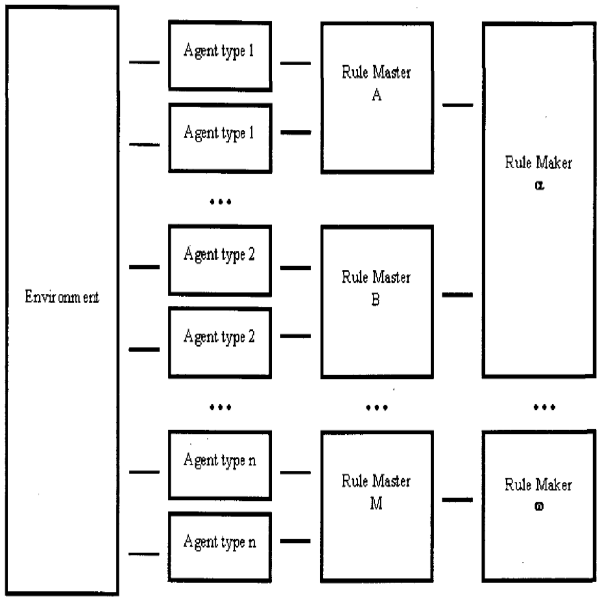


Figure 4.4: The Environment-Rules-Agent framework (Gilbert and Terna 2000).

4.3 Emergence

One of the reason why ABM is interesting is the emergence of phenomena (or complexity). These phenomena are not hard coded in the program rather they come up through the interaction of agents with their characteristics and with the environment. An interesting aspect of emergence is the impossibility to predict *ex ante* its appearance and behaviour. Edmonds and Meyer define complexity as *"the phenomena of interest result from the interaction of social actors in an essential way and are not reducible to considering single actors or a representative actor and a representative environment. It is this complexity that (typically) makes analytic approaches infeasible and natural language approaches inadequate for relating intricate cause and effect"* (Edmonds and Meyer 2015, 4). Therefore, a property of the system *emerges* from the local level and there is no chance to derive it from the functional relationships of individual agents (Epstein 2006; Epstein and Axtell 1996).

Complexity is an ubiquitous concept, shared by many disciplines as meteorology, biology, mathematics, chemistry and physics. For this reason Manson (Manson 2001) argued that the term complexity theories is better depicting the concept than theory only. In this light, complexity has different nuances, according to the discipline-perspective used. In natural sciences, practitioners adopt the idea of complexity arguing that disequilibrium is often necessary for the systems to grow. The idea of complexity first started in meteorology thanks to the work of Lorenz (Lorenz 1993) and then blossomed in a myriad of declinations. In 2002, Stacey et al. (Stacey, Griffin, and Shaw 2002) tried to summarize the complexity of definitions and theories arguing that three are the key ones: chaos theory, dissipative structure theory and theory of complex and adaptive systems. The first theory directly descends from Lorenz studies and it considers systems as constantly in transformation in an irreversible way. Chaotic system are not linear and they exhibit complex patterns that could not be predicted form the causes. This implies that linear causality and Newtonian mechanics are refused. The second theory relies on the idea of energy dissipation where semi-stable configurations passes through different states. These configurations may reach instability points called bifurcations where they self-organize in new configurations that cannot be predicted. The third one consider a number of agents which behave according to rules. Stacey (Stacey 2007) argued that the main difference among the first two theories and the last one resides on the fact that the formers try to create mathematical models at macro levels whereas the latter uses the interaction at micro level.

Regardless the nature of the theory considered there are three common concepts to take into account. First, chaos. As Burnes wrote, *"from the complexity perspective, chaos describes a complex, unpredictable, and orderly disorder in which patterns of behaviour unfold in irregular but similar forms"* (Burnes 2005, 79). Stacey found three types of order-disorder configuration in complex system: the stable equilibrium where the system at the end dies, the explosive instability where system goes out of control and the bounded instability where system transforms itself to survive (Stacey 2007).

Second, the edge of chaos. Under the condition of bounded instability, the system is always on the edge between chaos and order. Third, the presence of order-generating rules, which let the system exhibits order from chaos.

In this scenario, many researchers attempted to show that the complexity could be explained with simple rules trough computer simulations. The parallel with organization has already been done considering the organizations as dynamic, complex and non linear systems, therefore, complexity is a valid concept also in social sciences research (Wheatley 1994, Tetenbaum 1998, Styhre 2002, Stacey 2007).

Many authors discuss about complexity in social science, admitting similarities with natural phenomena but also underlining some peculiarity of the former. Troitzsch (Troitzsch 2009) argued that social systems not only are subject to existing forces as the natural systems, they also react to them, they are proactive and they have goals which sometime could be also conflicting. Then, a mathematical reductionism, as proposed by mathematician and physicist, could not capture all the interactions at micro level which could be several and diverse. Hence, the bottom up approach better describes the interactions and the behaviours at individual level and it is preferable in simulating social systems. Of course this approach gained some critics since it has been noted that arguing the macro level behaviour looking at the micro level could be misleading as it is trying to predict the behaviour of a heat flow in the gas just looking at the single molecules. But the appreciation that individuals have a conscious behaviour leads to justify the bottom up approach.

Moreover, in social system, the superimposition principle does not work, forces do not sum in a linear way and not always the strongest force prevails on the other.

Again, authors recognize complexity in social science as multifacet: Troitzsch, for example, attributes to complexity four different dimensions: domain, time, approach and composition. Multi agent models could be used to represent nested levels of the society (recalling also Coleman boat) with their usual features: autonomy, reactivity, proactivity and social embeddedness. In this way different domains could be modelled. At the same time, encompassing such different levels and spanning different domains, also different time scales are naturally considered. Different systems or nested system could have different time representations. Again, the flexibility of ABM gives the researcher the chance to model all these differences. The third attribute of social complexity is linked to a variety of possible approaches either from an implementation point of view and from a discipline point of view. Finally, agents could models also different systems: an agent could part of a family, a school class, a military organization and, extremely important, at the same time.

Dealing with social phenomena, there is a last important point to highlight. Individual could recognize the emergent phenomena and behave accordingly. Individuals could change their characteristics and actions based on this recognition. This emergence is called second-order emergence to differentiate it from the so called first-order emergence that is the phenomena created by the interactions of the agents at micro level.

There are a couple of implications on emergence form data: first, ABM eliminates some typical limits of equation based model. There is no more the need to find the solution of the equations rather system evolution and dynamics are realigned (Secchi 2015). Second ABM becomes a natural way to model complex structures.

4.4 KISS vs KIDS

How deep in the details does a modeller have to go in building the model? This question always accompanies the researcher. There is a trade off to reach between a detailed model and a more simplified one. From one side a detailed model encompasses more features, it is more adherent with the system under observation but it less manageable in terms of implementation, debugging and improvement. On the other hand a lighter model is easier to keep under control but it lacks in representativeness.

For many reasons, the approach to simplify as much as possible gained lot of favours. This approach has it own acronym: KISS - keep it simple stupid. Many reasons were brought to support this approach as the less complex a model is the easier is to maintain it and, again, a simple model could be investigated effectively when something goes wrong.

This approach is also founded in our human being that is naturally led to think that everything could be reported to something simple even if in appearance is complex.

But this approach found some opponents. Edmonds, for example states that simplicity is not a necessarily a truth indicator (Edmonds 2007). The central idea of the critics is that the model should be complex in the right way to serve the purpose it is built for. Even if some emergent phenomena exit from simple models, this is not an indication that all models must be simple and that emergent phenomena could be necessarily reduced to simple models. The behaviour of many systems could not be expressed by simple models or expressed by aggregated values which output could deviate from the target system.

Agent based models enter in this discussion and they could play a pivotal role. Edmonds and Moss observed that *"adopting a multi-agent model represents a move toward descriptive accuracy"* (Edmonds and Moss 2004, 132). Indeed the target system is represented in a ABM by agents which exchange messages. This reduces the gap between what is observed and what is modelled.

Hence, the model should not be simplified *ex ante* but *ex post*; only once the researcher has enough information about the environment she is studying she knows which features could be left out. This new approach is called KIDS - Keep it descriptive stupid. Hence, the model should start with the simplest description of the system as available data and resource allow, permitting following adjustment and improvement based on the gained understanding. Edmonds and Moss exemplified the two approaches in figure 4.5.

This discussion is also important to separate the intended theory from the implemented model. In KIDS perspective, the theory should be as de-

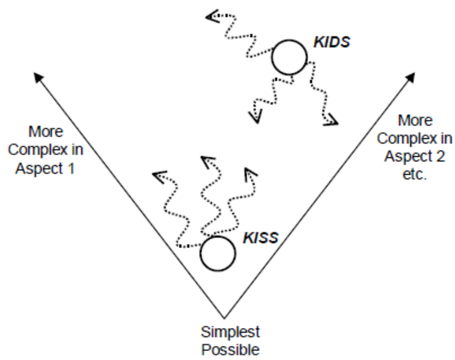


Figure 4.5: KISS and KIDS approach (Edmonds and Moss 2004).

scriptive as possible and then the implemented model should be its direct representation. By consequence, any model feasibility consideration should not in any case propagate back to theory and simplify it.

Corinne Coen (Coen 2009) felt the urgency to consider this topic as an sensitive one since although both approaches give important output for the discipline, it is very hard to integrate them. Saying that the question has been discussing since the origin of occidental philosophy, Coen supports Burton and Obel position (Burton and Obel 1995). In their view, the simplistic versus realistic question should not be answered *"per se"* rather it should be answered in light of the research question. As Secchi wrote (Secchi 2015) is the *why* that should drive the right level of complexity in designing the model.

4.5 Agent based model criticism

Despite important characteristics and a wide range of possible applications, ABM receives some critics and scholars tend to be reluctant to adopt the technique for their studies.

A first cause could be ascribable to the still unclear role of modelling in social science. The main question still open is whether models are more useful to build theories or to validate some findings. Moreover, scholars still tend to consider modelling as an exotic technique, sometime outside the box of traditional tools (Secchi 2015). This technique suffers of another cons: the researcher should master some programming language and have some IT capabilities to create a proper model (Fioretti 2013).

There are also reasons more linked to the inner nature of a model: it is an artificial representation of the reality and any possible relation with the reality inevitably passes through the interpretation of the researcher (Gilbert and Troitzsch 2005). It is in principle possible to create models which mitigate this problem but the effort could become too high in light of the benefit.

Another reason lies in the fact that agent based models are used to simulate complex structures or part of the reality. As mentioned above, simplification intertwines with modelling so, necessarily, models could only represent isolated or limited characteristics of the real world, not the complete picture (Grüne-Yanoff and Weirich 2010; Mäki 2011). It goes without saying that the explanation provided is partial, giving potentially accurate but incomplete accounts (Weirich 2011).

Moreover, working with ABM encompasses at minimum three activities: find the constituent mechanisms, dealing with emergence and postulate simplified assumptions. The first activity is linked to the explanatory strategy the modeller wish to use. Bunge (Bunge 2004) states the methodological rule "no mechanism, no explanation", echoed by Epstein's notion of generative explanation in computational modelling: *"If you didn't grow it, you didn't explain its emergence"* (Epstein 2006, xii). The modeller codifies her mechanism in a computer program enhanced with comments describing the underlying schemas. In ABM framework together with the environment, the agents' capacities are de facto the mechanisms. Hence, scholars use these models to explore their output rather the real ones (Grüne-Yanoff and Weirich 2010, 25).

By consequence, emergent properties depend on the constitutive objects and relative link which make the whole picture but there is also the necessity to understand the kind of disconnection between parts and the whole and the causal power the whole has. Cunningham (Cunningham 2001) made a distinction between ontological-emergent and epistemologically-emergent property. An ontologically-emergent property is more than its components. This idea of emergence precludes naturalistic scientific explanation. Hence, ABM could not deal with ontological emergence since the emergent phenomenon is not reducible to the properties of the lower level objects.

Vice versa, the epistemologically-emergent property has another issue, the components are not sufficient for its theoretical explanation. This because what is epistemologically-emergent is heavily based on the possessed knowledge and the properties could change over time. Moreover, these properties can be (a) diversely realizable and/or (b) interactively complex. Phenomena that could be realized in different manners are difficult to theorize regardless the method used. Interactively complex emergent outputs are difficult to infer (Humphreys 2008) even if not inexplicable in principle (Bedau 2013; Epstein 2006). Agent-based modelling is best suited for this form of epistemological emergence. ABM offers a way to avoid part of the computational burden that impedes the theorization of interactively complex phenomena, because ABM, at least in social sciences, presents the emergence as the collective output of interpersonal interactions (Epstein and Axtell 1996). Agents do not remain unchanged after the interactions, and causality does not only come from individuals to collective output. Agents' properties are both causal and caused.

Often ABM deals with lateral causation (agent-to-agent) that makes the agents change. But ABMs could also encompass the presence of contexts or artefacts that exist regardless the agents and could shape the behaviour of the agents over time. In making simplifying assumptions, the modeller has to choose a position in the range from concrete to isolated. A model is called concrete when it reproduces the details of the social process and is called isolated when such details are reduced with the aim to focus on a particular causal mechanism (Windrum, Fagiolo, and Moneta 2007). Carley (Carley 2009) discussing on computational modelling highlighted the problem of model veridicality in terms of both agent properties and context details. Other than linking veridicality to external validity and policy relevance, she also stressed some problematic aspects. Veridical models tend to require more code than the others do and it is less likely that this could be done with an off-the-shelf software. This requires more programming and computational resources, presenting obstacles to verification and use by others. Again, Miller (Miller 2015) argues that a central point of modeller's work is to answer the question about what should be true in the model in order to produce the observed dynamics. Indeed, the phenomenon-driven research fosters the backward problem solution (searching the underpinning mechanisms) than the forward problem solution (exploring the implication of known mechanisms). As Boero and Squazzoni (Boero and Squazzoni 2005) wrote, the modeller draws upon theory, prior models, and empirical data in acts of creative bricolage. And the creation of a model is a sort of interactive process starting from an initial knowledge about the phenomenon, the relevant theory and the modelling tool and it builds upon progressive understanding. The modeller acts as an interpreter throughout this process.

ABM proved to be very useful in many fields of research but it is not exempt from criticism due to some limitations or potential ones. Miller (Miller 2015) states that ABM is particularly suited for developing theories of interactively complex epistemologically-emergent phenomena. However, he argued that ABM methodology suffers the under-determination, the epistemic opacity of models, the cost of verifying models and correcting errors, and the restriction of agents to rule-based and pseudo-random behaviours. Underdetermination is linked to the fact that models show sufficiency and not necessity. Although this is true with models it is also true in all forms of theorizing. Then a model needs to be verified empirically and this work is never complete and it is always subject to falsification and revision. Opacity is related to what Humphreys noted, that is, *"in many computer simulations, the dynamic relationship between the initial and final states of the core simulation is epistemically opaque because most steps in the process are not open to direct inspection and verification"* (Humphreys 2004). Hence, a model is opaque to the degree of complexity and impossibility to decompose emergent properties.

Last, the restriction of agent behaviour is linked to the theory of agency as the ability to deliberate resulting in non-deterministic answer. But, agency, if thought as the intentional capacity to act based on free will, could not, at least in principle, be codified in a fixed algorithm.

Chapter 5

The proposed approach

The intention of the study is to broader the boundaries of exploration exploitation problem, considering the complexity and emergent phenomena that could come out by including different layers of knowledge exchange into a system. The literature so far has focused on particular aspects of the problem, in this model we embrace the opposite perspective introducing autonomous search, physical propinquity, informal relationship and governance. Those channels are present and active in an industrial company especially a knowledge-driven one.

A pharmaceutical industry R&D has been taken as reference to build the model. Typically, the R&D of a pharmaceutical industry is made up by different departments with their own different remits. Again, it is possible to define three main activities in the R&D: preclinical, development and clinical. In the first one new targets, new molecules and new products are found. In the second these new molecules are developed and a formulation and a process are associated. There might be the chance to develop a synthetic route for the active ingredients and a device, if required for the dosage form. The third sector is involved in any clinical operation that is to prove safety and efficacy of the products being developed. Of the three mentioned areas, the second has been selected as a prototype for the model. There are different reasons for this choice: first, in a pharmaceutical medium size company ($\sim$ 5000 employees) the R&D sector is structured with a well defined governance. This means that people has precise tasks to perform and there is a clear organizational structure. This is considered helpful in defining the model. Second, the department considered to build the model has nearly 130 employees, with different units absolving different parts of the development chain. This brings to an interest scenario. Completely different activities are carried out by the departments and all together contribute to the final goal that is the product development. The process chemistry unit develops the best possible synthetic path for the active ingredients to be used, the formulation unit creates a product around it, finding the suitable ingredients, quantities and dosage form. The process development unit has the goal to scale the process up starting from small amount of product manufactured by hand and targeting a commercial production with automatic production lines and quantities involved of different order of magnitude. The clinical trial supply units is responsible to package the product for the clinical study according to the phase of the study and its design. Moreover, a product is developed only if there are analytical methods to check if the specification are met. Hence, an analytical development unit is dedicated to find the right measurement technique and the set up of the right parameters. Finally, a device unit is dedicated to the design and development of devices needed to provide the product to patients. Within the department, products are developed by project teams with delegates by every unit. This structure is considered interesting for the exploration exploitation problem since units have the need to explore new ideas to be able to answer to always new challenges, instead projects have the tendency to rely on exploitation to deliver on time the projects related milestones. Third, also the development phase of a product brings into the scenario an exploration exploitation problem: typically in phase I a project is more oriented toward the exploration since the formulation, the process, the analytical methods and devices are in the building stage. When a project reaches the phase III, is more prone to the exploitation side since the time to market pressure increases and since any change is extremely difficult to manage due also to external constraints. For this reason it relies on the knowledge developed in the previous 5-10 years. Fourth, a department with the indicated dimension and with laboratories and production plants has necessarily to deal with physical space. The employees could not stay all together in the same office and they are split in different buildings and floors. Since the impact of physical layout is an objective of the study, this aspect was considered beneficial. Fifth, in such a structure the non formal connections a present and they could be not trivial. The friendship connections could be found but the dimension of the structure could bring to small cluster of informal relationships. Sixth, this department has been chosen since the actual information were available: the governance rules, the physical layout, the informal networks are known.

The idea is to incorporate these aspects in an agent based model. The proposed approach tries to model five different mechanisms: exchange through self search of the agents, exchange through organization meeting, exchange through project meetings, exchange through the friends networks and through the propinquity network, that is through the neighbours in the offices. These channels could be split into two groups: formal and choral exchange, and informal and individual exchange. In the first groups are enumerated the meetings with their double flavour: department and project. Department meetings are internal discussion within subgroups of the entire organization. Their aim is either managerial, to discuss internal governance (burning issues, planning, etc.) and scientific, to discuss about product development related to a specific discipline. Project meetings are instead cross-department activities and their remit is to discuss about milestones, strategy or development route.

The second group is focused on actions that single employee could take during the working day to sort out part of the duty. Typically there are three main channels to exchange knowledge: looking around that is to ask to neighbours, to ask to friends or to ask to employee considered experts or belonging the needed information. The list is considered to be quite exhaustive of the everyday life in this kind of organization.

The rest of the chapter is dedicated to a conversational description of the approach, a detailed and structured description of the model is postponed to chapter 6.

5.1 General structure

The core idea of the model is to consider the exploration exploitation problem in an environment with multiple way of interactions. In every model cycle, an agent could have an interaction through one of the informal channels and could attend meetings. Just as reference, a model cycle could be thought as a working day. During a working day, all agents could take actions and exchange knowledge.

The entire simulation lasts after N_{run} cycles in which the entire organization is involved in a series of actions. This suggests the second type of iteration. Since every agent takes at least one action and there could be multiple meetings, the order of these actions must be managed[1].

Therefore, in the model there are two kinds of iterations: the time iteration and the agents' iteration. The first iteration simulates the time passing, for the considered organization it could be compared to a working day. The second iteration is related to the action agents perform during each day, this because it is plausible that every agent starts at least one knowledge exchange per day. Hence, the second iteration let all the agents take actions. This does not necessarily mean that one agent is involved only in one knowledge exchange per day, because other than the action voluntarily started, an

[1]Typically in a agent based model the actions are randomised in every cycle to simulate reality.

agent could be the target of someone else knowledge research.

The number of agents could change between different runs of the model but never within the same run. Thus, once the number of agents is fixed, it will not change until the end of the simulation. This, as Chanda commented, to signify resource constraint (Chanda 2017).

As mentioned, every cycle starts with the definition of the actions sequence shuffling agents and meetings and creating every time (cycle) a different list. The number of agents (N_{ag}), the number of department meetings (N_m^d) and the number of project meetings (N_m^p) are parameters of the model, hence they could be changed and their effect considered.

Once the list is available, every employee (agent) is assigned to an action. The assignment is made on a stochastic base according to the professional job role the agent has. March and all the following scholars never attempted to model heterogeneity in the action profile, they all considered the agent equals in their characteristics (March 1991, Rodan 2005, Miller, Zhao, and Calantone 2006, Gupta, Smith, and Shalley 2006, Fang, Lee, and Schilling 2010, Kunz 2011, Mitomi and Takahashi 2015, Aven and Zhang 2016,Mueller, Bogner, and Buchmann 2017). This model also takes into account different roles within the organization with different jobs. A laboratory technician is supposed to do her daily work between the laboratory and the office hence the chance she could attend a project meeting is very low. On the other hand, a project team member is more likely to spend the day in different meetings than to stay in office all the day. Again, administrative worker spends most of the time in the office and neighbours will be the preferential channel to exchange knowledge and information.

The cycle continues actualizing the actions but, to be able to study the evolutionary nature of the system there is the need to record what happens during the simulation. Therefore there is the need of a completely new mechanism. The underlying idea is to equip the agents with a tool able to trace the interpersonal relationships and let the agents “move freely” into the model performing tasks.

Hence, a stack (S) is give to every agent. This stack has the structure presented in figure 5.1.

The stack has the number of columns equals to the number of knowledge categories[2] (defined by the parameter N_{ER}^{cat}) plus one dedicated to store the identity of the agents. The cells store the number of positive exchanges the ith agent had on a particular topic with jth agent . Looking at the figure 5.1 and assuming it is the stack of agent 1 she gathered knowledge 5 times from agent 33 on topic category 2. Agent 2 passed knowledge to agent 1 12 times

[2]Knowledge structure is discussed in next sections

Agent	Knowledge cat. 1	Knowledge cat. 2	...	Knowledge cat. M
33		5	...	
2	1		...	12
77	2		...	

Figure 5.1: Agent stack representation

on knowledge category M and 1 time on knowledge category 1. The number of rows (that it the number of remembered colleagues) could be set at the beginning of the simulation.

Any time a positive exchange takes place, the counter within the specific cell increments, vice versa, any time the exchange does not take place the counter decreases (of course never going below 0). This mechanism mimics the memory an employee has when searching for information, that is, an agent could come back to the colleagues who gave her positive feedback to her request the previous times. As it will explained in further details in section 5.6, any time an agent needs to start an autonomous search for a particular topic, she will start looking at her stack in order to find her best answering colleague.

It is important to highlight that the stack represents the agent perspective and it depends on the history of agent's relations. The stack and its mechanism act as a reinforcement learning exhibiting two central features: the tendency to repeat successful task and to avoid unsuccessful ones and to assess the performance against a desired level (Puranam and Swamy 2016). The first feature is linked to the increase or degrease of the goodness of the interaction; as long as the interaction is positive the score is increased and the agent become a preferred interlocutor for a colleague on a specific topic. For the second feature, the agent when consulting her stack seeks for the best scoring colleague for interaction.

The implementation of a reinforcement learning is also based on previous literature showing that it is a good way to model the behaviours of individuals with limited knowledge of the environment (Puranam and Swamy 2016, Camerer 2011, Erev and Roth 1998).

The model has also a general stack $\hat{S}$ which reports the summary of all the agents' relationships. It works as the collective memory of all the exchanges. It could behave as a knowledge index that facilitate the person-to-person interaction (Levine and Prietula 2012). Within the model, the general stack plays an important role during the autonomous search giving an agent the expert list to consult when the personal stack is not enough. Again, this part will be better explained in section 5.6.

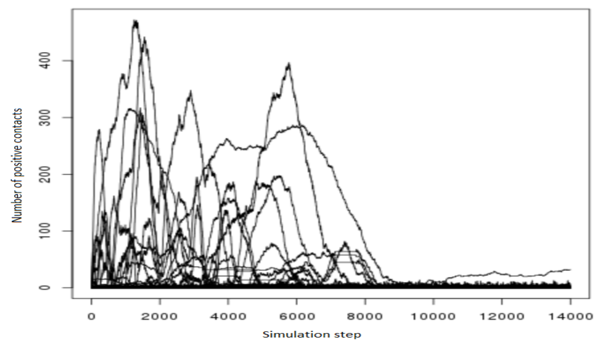


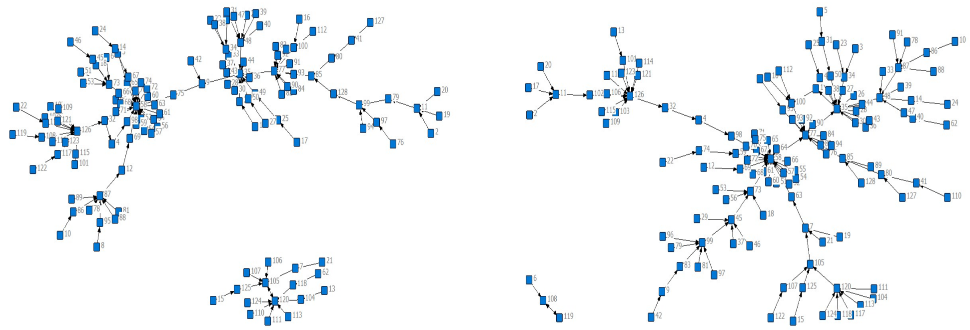
Figure 5.2: Agent stack evolution during simulation

Indeed the agents' stack is a dynamic object and the counters will always change during the simulation evolution. Picture 5.2 shows a typical agent's stack evolution during the simulation, where abscissa represents simulation run step and the ordinate represents the number of positive contacts and each line is associated to a different colleague. It is possible to observe its intrinsic dynamic structure.

Agent's stack is like agent's memory of what happened in the past. This memory effect could be modulated in order to study how the system is dependant on the ability agents have to remember. Defining *ex ante* the length of the stack it is possible to limit the number of colleagues an agent could remember. The parameter S_i^L defines it. If S_i^L is equal to $N_{ag} - 1$ (where N_{ag} is the number of agents) then the stack has the length to host all the other agents in the simulation[3]. In this case there is always space to remember an agent and relative performance. If S_i^L is less than $N_{ag} - 1$, the stack could not host all the other agents and only the best performer could be recorded. When the stack is full and a new agent needs to be inserted, the model seeks for the contacts who performed worst and randomly selects one of them and remove it form the stack. The new agent could take her place.

A final remark is worth about agents' stacks. Through the recording of the evolution of all the stacks it is possible to build the network of connections and to monitor it during the evolution. Changing the values in the stack, the network changes accordingly and the change gives an useful insight. The evolutionary nature of the model and the stack mechanism lead to a dynamic network. Picture 5.3 shows the connections network at two particular evolution time steps (after 1000 and 2000 cycles): the change is quite evident.

[3]the stack is limited to $N_{ag} - 1$ since the Nth agent is the agent owning the stack.



(a) network after 1000 run (b) network after 2000 run

Figure 5.3: Example of network evolution

5.2 Agent's topology

March's model states that all the agents interact with the organizational code. The topology of the model is then a star with organizational code in the centre and all the individuals connected to it. An important aspect of the model is the immutability of the connections schema. The connections are the same throughout all the simulations.

However, more recent studies on knowledge exchange stress the importance of direct interaction among agents. Mueller et al., for example, in their work on the effect of structural disparities on knowledge diffusion state: *"the network is therefore the list of all pairwise relationship between agents"*, and *"an exchange therefore takes place if two agents are directly linked and if both agents can receive new knowledge"* (Mueller, Bogner, and Buchmann 2017, 617).

Authors dealing with spatial influence in exploration exploitation problem faced the topic in two different ways[4]. First, authors like Axelrod (Axelrod 1997) and Miller et al. (Miller, Zhao, and Calantone 2006) defined an interconnection between the agents but with rigid characteristics. Axelrod included the geographical distribution of agents, *"a simple 100 sites, arrayed on a 10 by 10 grid"* (Axelrod 1997, 208). Each individual could only interact with its immediate four neighbours[5] and the edge sites only with three adjacent sites. Miller et al. used Axelrod approach but considering the grid without edges, so every site has four neighbours. Second, authors

[4]Actually there is a third approach proposed by authors presenting NK models. This approach is not considered in the list due to the divergence of modelling technique respect to the proposed one.

[5]North, south, east and west.

like Mueller et al. (Mueller, Bogner, and Buchmann 2017) argued about the importance of the connection network structure imposing a defined topology *a priori*. In their paper, for example, authors compared four different topologies: Erdös-Renyi, Barabasi-Albert, Watts-Strogatz and evolutionary net.

But, in last years, literature experienced a further change in the approach. Knowledge diffusion has became an important topic in the literature touching a twofold objective: understanding how knowledge diffuses into social networks and how social networks evolve accordingly. Gross and Blausius (Gross and Blasius 2008, 259) argued that *"the majority of recent studies revolve around two key questions corresponding to two distinct lines of research: what are the values of important topological properties of a network that is evolving in time? And, how does the functioning of the network depend on these properties?"*

These questions have been touched in literature in studies about dynamics of networks and in studies about network of dynamics. Recently scholars realised that the two veins could be merged in a unique and prolific field of research involving the so called adaptive co-evolutive networks. Although the application of this concept is endless, some general insight could be already distilled and, as Gross and Blausius wrote (Gross and Blasius 2008, 260), *"that certain dynamical phenomena repeatedly appear in adaptive networks: the formation of complex topologies; robust dynamical self-organization; spontaneous emergence of different classes of nodes from an initially inhomogeneous population; and complex mutual dynamics in state and topology"*.

Picture 5.4, taken from Gross and Blausius (Gross and Blasius 2008), is the graphical representation of the idea behind the adaptive co-evolutive networks.

Many papers exist studying this phenomenon from different perspectives and new impulse was given after the seminal concept of small-world networks (Watts and Strogatz Watts and Strogatz 1998) and scale free networks (Barabasi and Albert Barabási and Albert 1999). In another recent work Luo et al. (Luo et al. 2015) studied the evolution of directed connected agents.

The model deviates dramatically from the previous studies, because we appreciate the importance of direct links between agents and also the importance of dynamicity of the relations networks itself. But believing in the generative social science principle there is no reason to impose a particular structure to the network of relationships. Any imposition could behave as bias in the understanding of the environment effects on mutual learning mechanism. As consequence, the model does not start with a predefined network, it grows with the evolution of the model.

The only predefined structures are the offices layout and the friendship

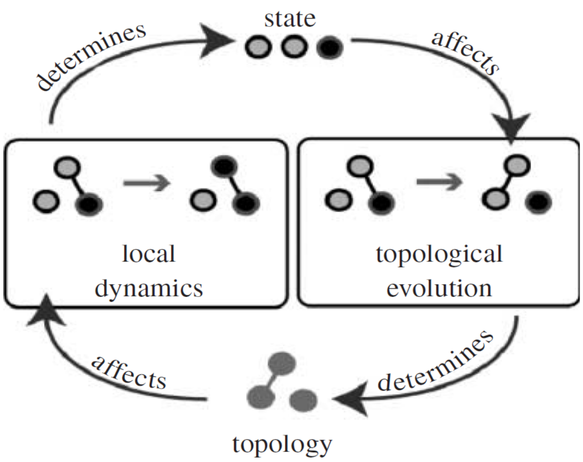


Figure 5.4: Adaptive coevolutive networks

relations for reasons which will be clarified in the rest of the chapter.

5.3 Knowledge representation

Many scholars studied the process of knowledge transfer in social network scenario in the last years (Cowan and Jonard 2004, Cowan, Jonard, and Zimmermann 2007, Morone and Taylor 2004). This because the transfer mechanism is extremely important in the study of interaction. Agent based models are a proper tool to simulate any problem related to knowledge transfer within social network (Guechtouli 2014). Different models of knowledge have been presented, with different characteristics but Cowan et al. (Cowan and Jonard 2004) state that modelling knowledge as a scalar is not suitable if the process of diffusion should be considered and modelling knowledge as a stockpile was criticised by other scholars (Morone and Taylor 2004).

Guechtouli (Guechtouli 2014) tried to assess the best representation of knowledge to be used in agent based models. He focused on three different mechanisms having knowledge modelled as a binary vector, as a single stockpile where knowledge could be accumulated and a multiple stockpile for a more sophisticated representation. Results show that the transfer of knowledge is facilitated when it is represented by a binary vector. The present model is then based on knowledge modelled essentially as a binary vector, also following the approach March adopted in his work (March 1991). As in March, the knowledge enters into two representations: the external reality and the beliefs every agent has about external reality. In his paper, the external reality is represented by a vector of n dimensions where each dimension could be either 1 or -1 with equal probability. One of the limitation of March's model is that external reality is not differentiated according to different skills or competences present in a real organization, hence, all the individuals play using the same external reality. The present model offers a different view, the external reality is made by N_{ER}^{cat} different categories mimicking the different competences of the real world (as in figure 6.2). All N_{ER}^{cat} categories are composed by N_{ER}^{dim} dimensions. The motivation of having knowledge categories relies on the observation that, in a structured organization, different departments focus on different competences. Moreover, if the organization is efficient, these competencies are only partially overlapping: indeed the competence needed for defining a synthetic route is extremely different from the competence needed to develop an analytical methods or to scale a production process up. In the model there are as many knowledge categories as the number of departments in accordance to what already explained.

5.4 Knowledge exchange

Knowledge exchange takes place in the same way March proposed: *"if the code is 0 on a particular dimension, individual belief is not affected in each period in which the code differs on any particular dimension from the belief of an individual, individual belief changes to that of the code with probability, p_1"* (March 1991, 74). March with p_1 intended the learning rate of the agents. The same approach has been used also from scholars in following works, as for example in Miller et al.: *"in any given period, individuals learn from the code with probability p_1. This probability of learning from the organizational code reflects the strength of socialization into organizational norms. Individuals' beliefs are unaffected by elements of the code with 0 values"* (Miller, Zhao, and Calantone 2006, 710). Fang et al. used the same approach underling that *"similar to March's original work, individuals can observe the pay-off of their overall belief set but they cannot directly observe how each element of the belief set contributes to this"* (Fang, Lee, and Schilling 2010, 630). In a potential exchange two agents are involved: the receiving and the giver one. The adherence of both is calculated and if the adherence of the giver is greater than the adherence of the receiver the exchange could take place. Hence, according to the learning rate of the agent (represented by parameter L in the model), the knowledge could be exchanged.

5.5 Meetings

The meetings exchange paths trace what happen in the everyday life in a medium-large organization. Indeed a modern R&D organization manages the development organizing activities in projects and splitting topics that are of projects competence and topics that are of organization competence in a sort of cross-related two dimensions. As mentioned, departments are the locus of knowledge creation where the practical activities take place. To organize such activities, internal meeting are needed. Vice versa, projects are the locus of exploitation of knowledge where to meet milestones and striving with time and budget.

In the studied organization there are a number of units, each of them is focused on a particular facet of the entire knowledge. Meanwhile, every project team is responsible to meet the product development milestones assuring the delivery as agreed. This intersection of goals creates the dichotomy in the meetings. There are project meetings in which the team discusses about the project issues and strategy and there are department meetings in which the internal organization and activities are discussed.

The meeting mechanism is based on three actions: definition of the topic to discuss, creation of the list of attendees and meeting execution. As in the real life, every meeting has an objective and topic to discuss and this must be agreed in advance. Considering the discussion about external reality and knowledge, the number of topics is equal to the number of departments. This to cover all the potential topics to be discussed within the organization. The selection happens in two ways: a department meeting discusses about its own piece of knowledge whereas the project meeting could discuss all the topics. This is fairly correct considering that department meeting involves employees of the department and the discussion is on internal stuff. A development project extends to a wider number of topics since to delivery a new product it is necessarily needed the contribution of all the competences. Hence the topic is randomly selected among the possible ones and once defined, different attendees are selected. The selection is made starting from the group of agents which for the current run step have been assigned to the meeting, then only those who fit with the type of meeting are taken. For department meeting all the agents not belonging to the department whose topic is going to be discussed are discarded. For project meeting all the agents are retained.

The following step is to define the number of attendees: indeed not all the meetings have the same number of attendees. Different meetings have different number of attendees: a face to face requires only two attendees, a formal project milestones review requires a substantial number of attendees. By consequence, the number of attendees is randomly selected in the range from two to the number of available attendees[6] (i.e. the number of agents retained from previous step). Once the topic and the number of attendees are available, the meeting could take place. The meeting has two steps: the definition of the most influencing agent and the update of the other agents' beliefs. The first task is performed as a hierarchical research: first the agent(s) with the highest rank in the organization are selected[7]. If the selection returns only one agent, this is the most influencing agent of the meeting. Otherwise a selection is needed among them and the selection is made on attendees popularity. Again, if this further selection returns only one agent, this is the most influencing attendees. If even this second selection returns more than one agents then the agents have the same position in the organizational rank and the same popularity. In this case there is not a most influencing agent and a mean belief is calculated[8]. The mechanism reflects what usually happens in a real organization: the higher position in the organization rank has the power to determine the output of a meeting consciously or unconsciously. The big manager monopolizes the attention and could take a decision based on his feeling and judgement. The ability to model this phenomenon permits to explore the influence of the organization (through managers) on the knowledge exchange flow. This is an important difference from March's model (March 1991). If big managers are not present, popular people could lead the show during a meeting and influence the final decision. Usually these people are those considered the most knowledgeable or the most charismatic and all the other agents tend to follow their thoughts or opinions. Missing also popular people, the discussion during the meeting is made with all attendees at the same level and hence there is no dominant position. It could be considered fair enough to close the meeting with a sort of common agreement, represented in the model by the mean of the beliefs. The model derives individuals' ranking and popularity in two different ways. Ranking is an attribute of the agent, it is something that could be set at the beginning of the simulation. The higher is the ranking number of an individual, the higher is the position in the hierarchy. Popularity is a dynamic attribute of the individuals and it is extracted from the general stack. The individuals with the higher score for a particular topic are the most popular. This because the general stack stores all the times an individual shares knowledge and, to share knowledge, she must own superior knowledge.

[6] A modified mechanism is also considered to study the physical layout impact. Details are in section 6.3.12.

[7] Of course the selection is made within the list of attendees, hence the highest rank is intended among the attendees to the meeting.

[8] The estimation is made in accordance with March (March 1991) summing by component all the agents' beliefs.

Hence, she could be considered an expert by the community and then her popularity is high.

Once the meeting output belief is decided, the following step is to check whether or not other attendees update their beliefs adopting the meeting output one. The principle of belief update is the same March used in his model (March 1991). There is a knowledge donor and a knowledge receiver and the latter actually changes her belief in the donor one with probability p[9] (learning rate).

It is important to remark that only the beliefs related to the topic discussed during the meeting are update, not all the beliefs. Once all the attendees have (or have not) update their beliefs, the following step is to update their stack.

The last step is to update the general stack with the output of the meeting: the most influencing attendee(s) is (are) recorded. Then the meeting is considered closed.

5.6 Self-search

March's initial model did not give the chance to agents to communicate freely. They could only exchange knowledge with the organizational code and vice versa. As already mentioned, this feature of the model has been criticised by following scholars. For example Miller et al. wrote: *"we recognize that face-to-face interaction can be critical to knowledge transfer. Interpersonal learning is a decentralized process that takes place without the mediation of an organizational code"* (Miller, Zhao, and Calantone 2006, 711). One of the main limit of the absence of interplay among agents is the fact that the knowledge exchange does not depend on personal acquaintances. Again, they suggested that *"much of the learning that goes on within organizations occurs directly from person to person and is not limited to exchanges mediated by organizational codes."* (Miller, Zhao, and Calantone 2006, 711).

Of the same idea are also Fang et al. when they write that *"organization is seen as a complex system wherein individuals directly interact with one another."* (Fang, Lee, and Schilling 2010, 629).

Mueller et al. stated that *"informal cooperation is the rule rather than the exception. The same holds true when it comes to more complex systems"* (Mueller, Bogner, and Buchmann 2017, 617).

[9] March actually did not model the direct knowledge exchange among two agents. Agents only could interact with the *organizational code* and viceversa. This aspect of March model has been criticized in literature (Miller, Zhao, and Calantone 2006, Gupta, Smith, and Shalley 2006).

The proposed model follows this tendency recognizing the extreme importance of the network of contacts that could grow when the system evolves over time, and the pivotal role of informal contacts among individuals in a real world organization. Hence it tries to overcome this limit introducing the self-search mode. Self search is more complex in the mechanism if compared with the other features in the model, for at least two reasons: first it does not start from pre-constituted network of information and second the search is more sophisticated in the rules. As already explained in section 5.1, every agent is equipped with a table where she traces the history of all the relations with other agents and how they ended up. Every time an agent enters in contact with other agents, two are the possible outcomes: the knowledge is exchanged, the knowledge is not exchanged. The agent and general stacks are then update accordingly. The flow starts with the need of knowledge by an agent and following decision to seek it through the relationship with other agents. This research is not linked necessarily with the department neither the project the agent belongs. In the model, an agent is free to contact any other agent in the organization.

Miller et al. (Miller, Zhao, and Calantone 2006) introduced the concept of local and distant search: the agent starts looking around herself and, if the knowledge is not greater than the one she possess, she tries to exchange with other agents in the grid selected by chance. The proposed model follows this vision considering both local and distant search. Self-search is considered the distant search since there is no limitation on the contact to others.

Agent starts the research identifying the topic for which they need knowledge. This because external reality is actually split into several different topics. Once the topic is defined[10], the agent looks at her stack to find whether or not some experts are presents. Expert means someone who already had positive contact with her and therefore could be a good candidate for a further exchange (figure 6.7 shows the entire flow).

The model allows the one-way exchange that is the knowledge flows from the more knowledgeable agent toward the less one even if there is no payback (that is the reciprocal flow from the asking to the giver agent).

If the expert is present, the adherences of the two agents are compared[11] and if the contact's one is greater then the knowledge exchange could take place. If not, the agent tries to find the expert in the contact's contacts simulating a possible suggestion coming from the colleague. Again, if the expert is present in the contact's list, the adherences of the two are compared

[10]Technically, the assignation is made randomly during the simulation. In every iteration, if an agent acts in the self-search channel, a random generation is done to select the topic for the agent.

[11]the comparison is made according to March model (refer to section 5.4).

and if the colleague's one is greater, the exchange could take place. If also in the contact's list there are no experts, the agent could take a look in the general stack list. This last step simulates the collective perception of experts. If the contact in the general list has superior knowledge, the exchange could take place.

5.7 Physical layout

Gullahorn argued that *"is not safe to ignore the sheer fact of contiguity as factor in interaction"* (Gullahorn 1952, 123). In observing the work in an office for a couple of months he discovered that *"friendship has been shown to be of some importance in determining the frequency of interaction, but distance appears to be the most important factor"* (Gullahorn 1952, 131). The impact of layout on social interactions and behavior has been recognized as fundamental since long time. Oldham and Brass stated that *"numerous studies have demonstrated that architecture and physical layout can substantially influence variables such as patterns of communication and social interaction"* (Oldham and Brass 1979, 267). Oldham and Rotchford (Oldham and Rotchford 1983) stressed the importance of office layout as important factor modulating reactions and behaviour of individuals. They considered characteristics as openness, density, darkness but they also considered architectural accessibility as the extent to which an employee's individual workspace, such as desk, is accessible by others. This characteristic was measured as the number of walls and partitions surrounding the workspace and the result was that individuals behaved differently if they have *visual access to others.* Hatch discussed physical barriers writing that *"physical settings provide contexts for behaviour. They are thought to have influence through their ability to support the range of activities that becomes associated with them and to constrain other forms of activity"* (Hatch 1987, 387). Moreover, she argued that physical structure could impact on interaction at least on three aspects: i) the choice of interaction partner, ii) type of interaction, iii) the amount of interaction. In more recent years, Boutellier et al. (Boutellier et al. 2008) studied the influence of geometrical layout on knowledge creation, passing through communication. The idea that individuals, groups, departments but also rooms, floors and building could act as island of knowledge brings to the necessity to modulate physical layout in order to modulate communication. In addition, as author said, in a science-driven business it is pivotal such communication in order to gain the desired knowledge creation and exchange. In 2012 Sailer and McCulloh (Sailer and McCulloh 2012) echoed with the statement *space matters*, arguing that propinquity brings to clusters that could locally exhibit reciprocity and transitivity. In 2019, in a new empirical study, Lee (Lee 2019) demonstrated that the spatial proximity has a huge effect on individual exploration. Studying a South Korean e-commerce company relocation, he focused on the effect of changing the neighbours and the relative impact on the level of engagement in the exploration activities in the individuals. The conclusion of the study highlights a very important point: *"this study suggests that organizations should be strategic when deciding how to rearrange the spatial proximity between individuals when the organization's goal is to increase exploration. In particular, this study presents evidence that not all individuals may benefit equally from being collocated with previously separated peers. At least two factors seem to matter: an individual's prior organizational experience and ties with previously separated peers"* (Lee 2019, 19).

But, traditional methodologies did not take into account network structures, investigating interaction frequencies. From the concise literature review about impact of space layout it could be possible to distil two important messages: i) in an exploration exploitation study it could be counter productive to ignore physical layout as conveyor of informal relationship ad occasions of mutual learning, ii) a new tool is probably needed in order to take into account networks. This second point is even more true in the proposed research work: complex formal and informal information flows to be modelled by ABM are spread via networks. Having a similar approach also for physical topology would ease the work and keep homogeneity of methodology. Space syntax (Hillier 1996) could be thought as a research program investigating the relations between human societies and space with particular attention on inhabited space in all its forms: buildings, settlements, cities, landscapes. The starting point of space syntax is recognizing spaces as key and necessary resource of human societies, needed to organize the societies themselves. For this reason, the space is configured and the act made is constraining the continuous space into a set of connected discrete units. The advantage of this operation is the ability to apply different labels to the different parts; these parts could, in turn, be assigned to groups, people or activities. Hence different rules of behaviour or convention could be created and associated to the parts of the space and these parts could be recognized bringing symbols or cultural charge. Space syntax aims to create paths of description for these spaces (buildings, settlements, complexes) so that the underlying covered social construct could be highlighted. The importance of this step relies in the possibility to move forward: secondary theories or practical explanation could be created unfolding the effect of such spatial configuration on social variables. As Sadek et al. (Sadek and Shepley 2016) argue, the built environment has to be considered to accurately predict its influence on

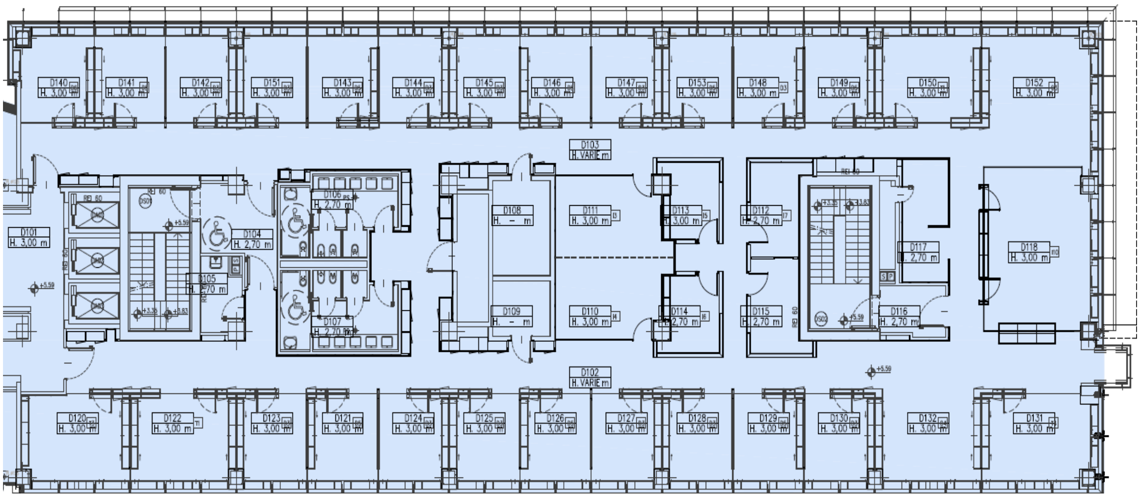


Figure 5.5: Typical offices layout

people's mental-health, behavioural interaction, and engagement with the environment. Consequently, researchers have adopted Space Syntax, along with other tools and methods, to achieve this aim. The concept of Space Syntax, which was originally introduced by Hillier (Hillier 1996), concentrates on revealing the underlying social logic of spaces by developing strategies to describe their configurations and their effects on various social and cultural attributes (Bafna 2003). It comes with a different quantitative descriptions of environments configurations with particular attention on buildings and urban contexts, focusing on interconnection. Montello (Montello 2007) argued that the rigorous description allows for potential explanations of a variety of physical and psychological responses such as user movement, experiences, and cognitive knowledge of place. Space syntax recognizes graph theory as a powerful tool so the main methodological problem has been the problem of reducing any configured space to an appropriate graph. Indeed, the issue is related on how to convert a continuous entity into a discrete one. The most used approach is the use of convex or axial maps at least dealing with buildings. The efficacy of the convex map lies in its ability to capture the sociologically relevant relationships embedded within a plan. For the purpose of the study, these maps are the relevant tool to translate physical offices layout (figure 5.5) into a network. Once the network is obtained, it is possible to find agent neighbours based on real topological data.

The construction of the network starts from the study of the layout of the building where the physical offices are. As Hillier wrote, *"encountering, congregating, avoiding, interacting, dwelling, conferring are not attributes of*

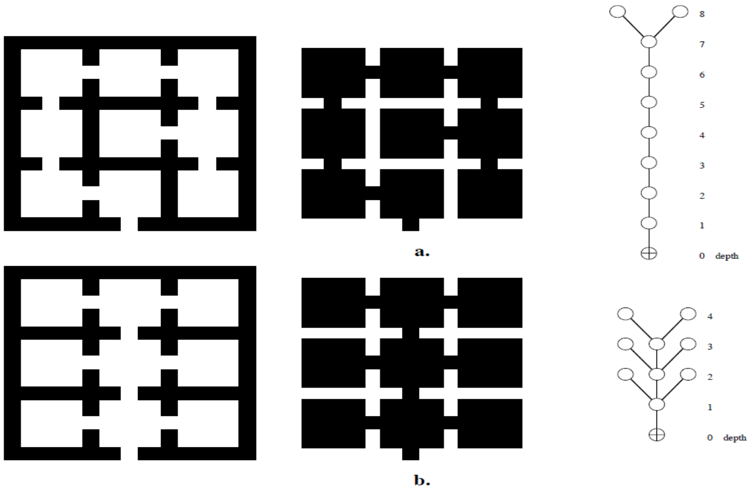


Figure 5.6: Notional courtyard buildings (from Hillier 1996)

individuals, but patterns, or configurations, formed by groups or collections of people. [...] We should therefore in principle expect that the relation between people and space, if there is one, will be found at the level of the configuration of space rather than the individual space." (Hillier 1996, 20). In a well understandable picture in his book (figure 5.6), Hillier (Hillier 1996) showed how the physical layout could have an impact on human activities.

The first column represents the normal way buildings are drawn. They are the physical elements of the buildings. The column in the middle represents the spatial element, the complementary of the physical ones.

Looking at the two examples, the building elements are the same, the external perimeter, the numbers of cells and the physical displacement of these cells. What changes is the location of doors. It seems a minor change but, actually, the difference could change the way individuals perceive and use the space; the configurations are not equal.

The third column gives the layout representation (called j-graphs) of the two configurations where the first is nearly a straight line with only a fork at the end and the second is shorter and with more branches.

Again, Hillier commented on that picture that *"the pattern of permeability (for example, the doors) would make relatively little difference to the building structurally or climatically, that is, to the bodily aspect of buildings, especially if we assume similar patterns of external fenestration, and insert windows wherever the other had entrances onto the courtyard. But it would make a dramatic difference to how the layout would work as, say, a domestic interior."* (Hillier 1996, 22).

Hence, different corridors, floors and building create a disruption in the linkage of agents. The question is how to represent the concept of neighbourhood in an agent based model where the underlying structure relies on networks. Although space syntax is a comprehensive and wide theory, for the purposes of the model, only the basic concepts are used. The idea is to transform a j-graph into a network that is the most important object in the model.

The building where the author works has been selected as the reference to build the structure. Hence, starting from the layout of the buildings (partially shown in figure 5.5), the corresponding j-graphs are derived. Each node in the j-graph represents a physical space where an individual has the office. Some spaces are individual, others are shared and others are also open-spaces. In each space there are the individuals' seats. Then, the network could be derived replacing the spaces with individuals. The organisation taken as reference is spread in all the company site and therefore the network is not connected because of stairs, changing of corridors or even buildings. The first assumption made is that it is not possible to consider as neighbours two individuals separated by a stair, or two individuals seating in different corridors, floor or building (this also from Space Syntax theory).

There is a second aspect to further discuss and figure 5.5 could be taken as reference. If the upper corridor is considered, there is a long displacement of offices. All these offices are joined in a j-graph but from a neighbourhood perspective it is hard to say the the individual in the left most office is a neighbour of the individual of the right most office.

For this reason the derived network has been further elaborated. To quantify the level of neighbourhood, the index of proximity $d_{i,j}$ is introduced[12]. This index is created starting from the the first agent a replicated for all the remaining ones. In figure 5.9 section A a possible combination of offices is considered. Now, in section B the candidate individual for which to calculate $d_{i,j}$ is highlighted with the green label. Individuals in the right office take the value +1 since they are the nearest neighbours. Moving right, there is a big room with two open spaces. All the individuals in the first one take the values +2 and the other in the the second one the values +3.

There are two more important assumptions, as shown in section C. If the candidate individual seats in a shared office, her room-mate takes the value +1 since he is the nearest neighbour. Second, the index is calculated in every direction the layout spreads. In this case the corridors moves on in the left direction hence, also the neighbour on the left takes the value +1. Section D shows a more complex example where the candidate individual is in an open

[12]where $d_{i,j}$ is the geodesic distance among ith and jth agents

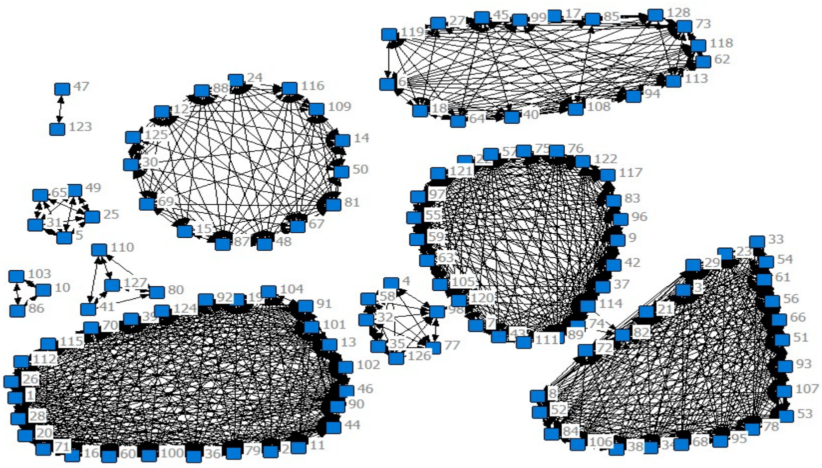


Figure 5.7: Propinquity network

space. Their nearest open space room-mates take the value +1 and then the other are calculated as explained.

The final assumption made in the model is that $d_{i,j}$ could not exceed 6. Beyond this value, two individuals are not considered neighbours.

When the process is completed, the neighbourhood network is obtained. Figure 5.7 shows the output of the process.

The network is actually the adjacency matrix where the connection between agent i-th and j-th is maintained only if $d_{i,j} \leq 6$. The matrix is therefore a weighted matrix where in the entrance $\{i, j\}$ there is the value of $d_{i,j}$. A detail is shown in figure 5.8.

The network is the basis upon which the physical layout workflow starts from. Differently from self search exchange mode, the physical layout does not allow a free search of the contact. This because seats in the offices are fixed and the neighbours all pretty much always the same. Hence when a person seeks knowledge from the surrounding environment she could have access to the other persons seating nearby. In the proposed model, when an agent is assigned to physical layout path the first step to perform is to identify which are the neighbours. The neighbours are identified based on two parameters defined at the beginning of the simulation. These two parameters are G_{min} and G_{max}, which represent respectively the minimum and the maximum distance between agents in terms of offices. This means, for example, that if G_{max} is 3, the available neighbours will be searched within the distance of 3 offices. These two parameters allows a precise governance of the influence sphere of the offices in order to better understand the impact

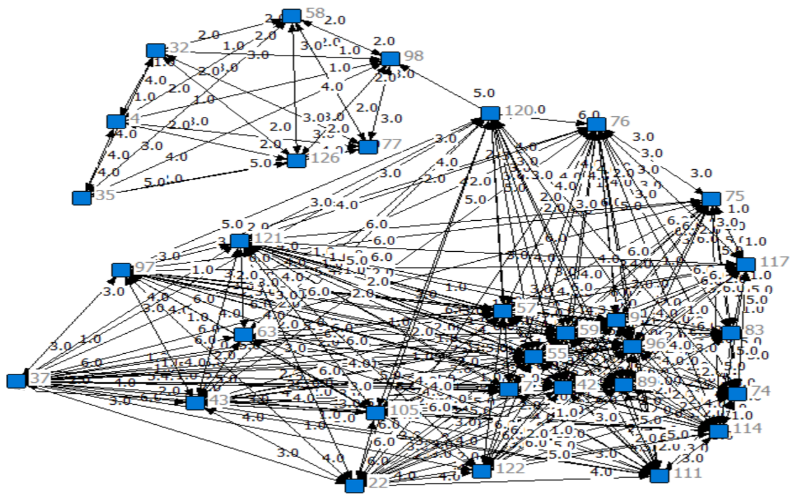


Figure 5.8: Detail of weighted neighbourhood adjacency matrix

on knowledge exchange.

Clearly, once the available neighbours are found, one of them is sampled and exposed to the potential exchange. As usual, the knowledge of the contact is checked and the adherence is calculated. At this point a fork is present in the model, and the superiority in knowledge of the contact could be considered or override. This is done through a parameter (F^{ksp}) that could be defined at the beginning of the simulation and allows the system to behave accordingly. If the superiority is kept, the exchange is only possible if the contact has superior knowledge respect to the agent leading the search. Vice versa, if the superiority is override, the exchange could take place regardless the level of adherence of the two agents. This is allowed in the model since knowledge transmission through neighbourhood is often not subject to judgement by the individual and possibly also distorted or false information are moved (as it happens for rumours). This because the idea to ask the neighbours a piece of information instead of asking to the expert (who could seat in a distant office) is rooted in the easiness of the task. This means that they will be more prone to accept the answer without checking in depth its goodness.

Once checked whether or not the exchange is permitted, the exchange could take place according to the same rules already explained considering the learning rate of the individuals. If the superiority is required but the counterpart has not the requirement for the exchange, then the flow does not perform the task.

The last step for all the branches is to update the agents' stacks and the general stack according to the path followed.

5.8 Friendship

Friendship or, more in general, informal relationship is a central topic in literature with a long tradition of studies. Reagans and McEvily in 2003 wrote that *"informal interpersonal networks are thought to play a critical role in the knowledge transfer process. Our understanding of how informal networks affect knowledge transfer, however, remains unclear because the effect of networks on knowledge transfer has yet to be examined directly"* (Reagans and McEvily 2003, 240).

Phelps echoed in 2012 *"knowledge networks research has explored a variety of characteristics of formal and informal relationships that influence knowledge outcomes"* (Phelps, Heidl, and Wadhwa 2012, 1120).

Burns and Stalkers in 1961 already divided the formal from the informal structure (Burns and Stalker 1961) but, although informal channel is not new, it captures the attention of many scholar for its extreme importance and elusive behaviour.

Many are the evidences that informal channels contribute to knowledge diffusion. Reagans himself with Zuckermann (Reagans and Zuckerman 2001) found that interactions among different scientists not belonging to the same networks help the knowledge diffusion and the productivity since *"the optimistic view is founded on the hypothesis that teams that are characterized by high network heterogeneity, whereby relationships on the team cut across salient demographic boundaries, enjoy an enhanced learning capability"* (Reagans and Zuckerman 2001, 393).

Bell et al. (Bell and Zaheer 2007), Mäkelä et al. (Mäkelä and Brewster 2009) as well Caligiuri et al. (Caligiuri 2014) reported that friendship ties increase the sharing of knowledge and the information flow.

Informal relationships have been the focus of many works each of one trying to understand the contribution in knowledge transfer from different perspectives. Lawson et at. (Lawson et al. 2009) argued about the mechanisms underlying the formal and informal socialization relationships among supplier and customers in the product development value chain.

Allen et al., starting from the evidence that informal networks are present in any organization, discussed their role in a R&D environment. The results are pretty clear: *"through a better understanding of the informal organization of R&D staff, they can more successfully capture and exploit new ideas"* (Allen, James, and Gamlen 2007, 179). Link et al. (Link, Siegel, and Bozeman 2007) argued about the role of informal technology transfer in the uni-

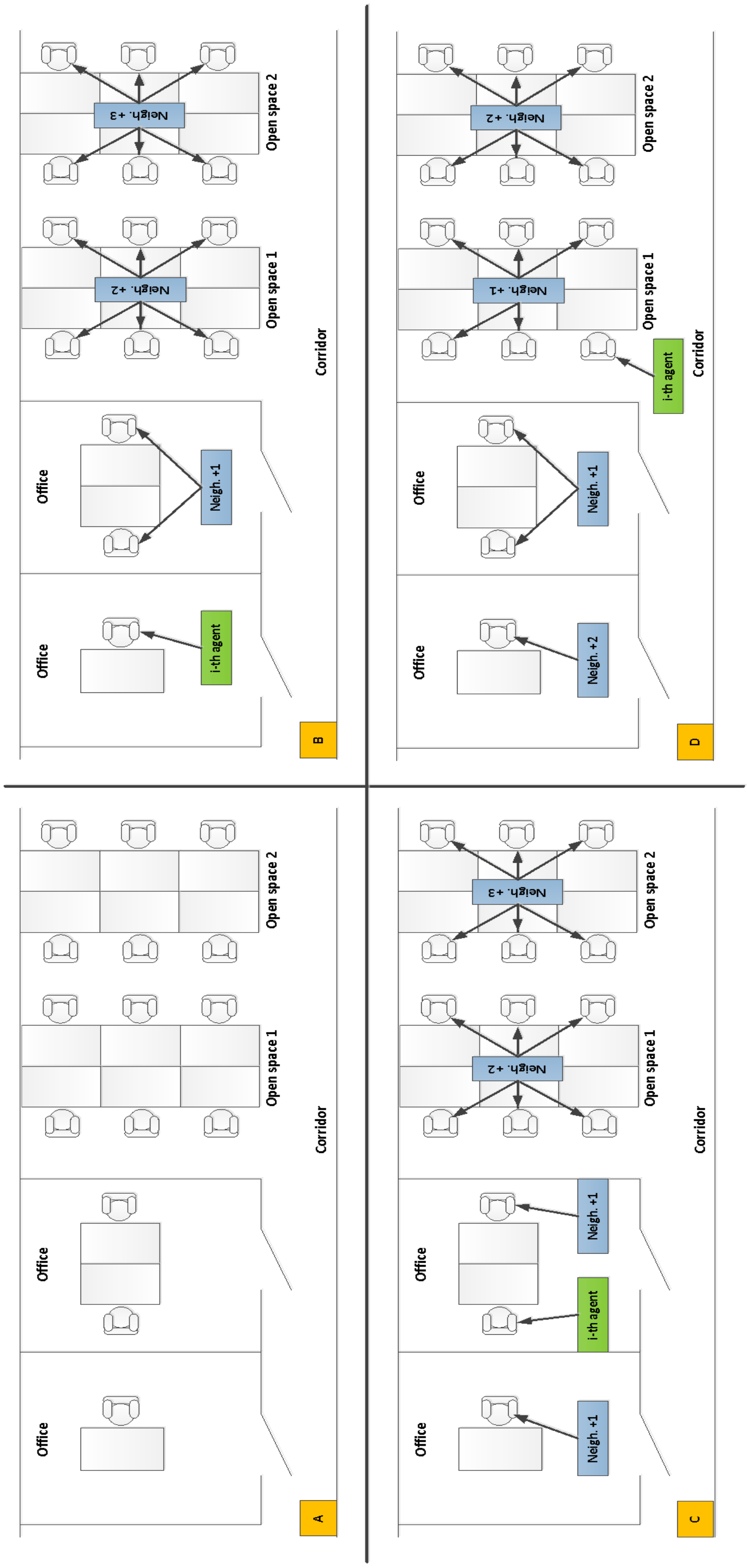


Figure 5.9: Office distance rule

versity and Dahl et al. (Dahl and Pedersen 2004) discussed about the role of informal relationships in the industrial clusters.

Some scholars attempted also to model friendship creation and persistence through agent based model. Singer et al. found that *"friendship is fundamentally different from the behaviour of other social networks in that they are single scale networks and show a small world effect"* (Singer, Singer, and Herrmann 2009, 1).

From this brief introduction is clear that the informal channel has an important role in the organization, reason for which the present model takes it into account. Picture 6.9 shows the model workflow for the friendship[13] exchange channel.

In the presented work, friendship relationships are not considered mutable. Hence, relationships are given for granted and they could not change over the simulation cycles. This mainly because the focus is not to understand how friendship links actually are created, maintained or removed[14], rather the impact of an already existing informal network on the general dynamics is the focus. For this reason, friendship channel starts with a predefined adjacency network. To build the network, the individuals of the reference organization have been asked to state who are their friends. These links have been anonymised to preserve the personal data and only the adjacency matrix is used in the model. The present model is run with organizations having different numbers of agents, hence to cope with that a portion or a composition of the adjacency matrix has been used to fit with the number of agents in the simulation.

If an agent is assigned to friendship channel for a particular cycle all her friends are retrieved and one is selected for the exchange (if someone exists). For sake of simplicity the assignment is not considered symmetric and only the perspective of the seeking agent is managed. This means that if the i-th agent is assigned to friendship and the j-th agent is selected (because is her friend) the contact will happen regardless the fact that the j-th agent is herself assigned to friendship for the same cycle.

Once the friend is selected, the agent could directly go to the knowledge exchange part without assessing the reciprocal knowledge level. This is somehow a difference respect to the other sharing channels and the main reason for that is that an individual could trust a friend. The relation among friends is often intimate and the level of trust is so high to exclude the challenge in what is shared. Just because the exchange is made though a friend the per-

[13]In the present document friendship and informal channel are considered interchangeable terms.

[14]For a review on this refer to Singer et al. (Singer, Singer, and Herrmann 2009).

sonal individual stack is no altered. A friend is a friend and it remains as it is: hence regardless the sharing of knowledge the link is kept unchanged. Only the general stack is update because if something is shared, it increase the common perception of knowledge owned by the giver.

5.9 Knowledge distortion

"A key assumption of March's model and subsequent research is that the solutions are transferred perfectly and without distortion between members". Aven and Zhang (Aven and Zhang 2016, 1104) started in this way their reasoning about distortion of knowledge which has the apex in the sentence *"we challenge the assumption of perfect knowledge transfer. In particular, we contend that, as the knowledge sender and receiver are separated by greater social distance, the degree to which the knowledge will be transformed increases as well"* (Aven and Zhang 2016, 1104). Distortion is considered an important topic in the model since the organization deals with different aspect of knowledge. As already explained, knowledge is composed by N_{ER}^{cat} categories each of them representing different topics an organization has to deal with. Naturally, agents in the organization could not be expert in every topics and by consequence the problem of knowledge distortion increases its importance. This is not the solely reason to incorporate distortion in the model since knowledge distortion is actually present even if agents expertise is overlapping. Distortion is modelled by a random change of knowledge bits during the transfer. Hence when the piece of knowledge is going to be transferred from an agent to another, it could be changed due to distortion[15]. The model gives the ability to tune the amount of knowledge to distort, acting as distortion effect modulator.

5.10 Turmoil

Just after the impact of learning rate, March studied the effect of external turbulence on the organizational equilibrium, changing the external reality as he wrote: *"suppose that the value of any given dimension of reality shifts (from 1 to -1 or -1 to 1) in a given time period with probability p_4"* (March 1991, 79). What he found is the dramatic impact of external turbulence on the organization performance especially on the mid long term. And, again, he stressed that *"once a knowledge equilibrium is achieved, it is sustained*

[15] Of course the new random value could not be equal to the changing one.

indefinitely" (March 1991, 79). Form the simulation perspective this phenomenon is linked to the fact that once the agents share the same belief about the external reality, there is no more predominance of any individual knowledge and, by consequence, any knowledge transfer is switched off.

The present model proposes the same feature with some refinements. First, the external reality change could be activated or not depending on the simulation parametrization. Second is could be set the precise moment in which let the change happen. Third, the change in external reality could be one shot in the simulation o periodic. As in March model, the external reality is inverted in all its dimensions.

5.11 External interaction and unlearning

Two last interesting characteristics are considered in the model: interaction with the external world and unlearning effect.

March's model considered the organization and the agents as belonging to a closed system. Agents, besides not having inter-relationships, do not have any relations with individuals outside the organization. This is virtually impossible in the real world in which agents actually meet other individuals outside their organization and they share knowledge and they change their beliefs about the external reality. These changes are then carried into the organization when they come back to meet their colleagues. Bocanet et al. in this sense wrote than *"in March's model and in most of the works proposing agent-based models to deal with internal dynamics, the system is a closed one, rigid to any interaction and omniscient. But organizations in an open system have to deal with their business environment in order to survive."* (Bocanet and Ponsiglione 2012, 28).

Chanda et al. (Chanda and Ray 2015) give a slightly different reading of the exploration-exploitation problem in March model. As opposed to Miller et al. (Miller, Zhao, and Calantone 2006), they suggested that *"the learning rate of organizational members maps to exploitation. This follows from the fact that organizational members learn locally, from the organizational code. The organizational code, in turn, gets its wisdom from the elites of the organization. The rate of personnel turnover maps to exploration. Distant search is involved in turnover because the knowledge embodied in a replacement person is not constructed by local search"* (Chanda and Ray 2015, 255). For the authors is the interaction with external environment that drives the exploration.

Sachdeva in her review (Sachdeva 2013), pointed out that a missing piece in the original March's model is the unlearning effect, considered a crucial weakness also by Hedberg (Hedberg 1981). This effect found place in Blaschke and Schoeneborn work (Blaschke and Schoeneborn 2006) dealing with organizational memory. They stressed the fact that fading memory is actually a beneficial effect since it frees space for new solicitations.

These two effects are presented together since could be represented with the same mechanism. Unlearning could be modelled by a change in the belief due to the forgetting. Hence, the picture of external reality could slightly fade over time and it could assume different values. A change in value represents the fact that an agent has a distorted remembrance of external reality.

External interaction with other individuals should follow the same rules in changing the beliefs but this would imply to model also the external contact with their beliefs and the contacts of the contacts and so on. This is clearly an huge computational effort that could be avoided by changing some agents beliefs mimicking the results of knowledge exchanges with others.

It is clear that the result of both the activities is a change in the belief of the agents. Therefore, the model permits to select how many individuals per cycle could undergo this change and it could be selected how many dimensions to change. Once selected the number of dimensions, each of one is replaced with one of the other two values, sampled with the same probability[16].

[16]This means that if the value of the dimensions is 1 it could be replaced by 0 or -1, if it is 0 by 1 or -1 and if it is -1 by 0 or 1. The selection among the two alternative is made by sampling one value from the uniform distribution.

Chapter 6

Model description

Different ways exists to describe a model in social sciences. In the last decades one protocol became extremely relevant in the agent based model literature with the aim to provide a formal description of the model and to propose a sort of check list to cover all relevant and important aspects. Originally proposed by a group of ecologysts (Grimm et al. 2006), the ODD (Objective, Design, Details) protocol has experienced a crescent consent within scholars (Edmonds and Meyer 2015) becoming *de facto* a sort of standard. The main benefit of this approach is to facilitate the communication among the modellers, to ease models replication and comparison and to foster the iter-disciplines dialogue. The protocol is composed by three blocks. The first one (Objective) is dedicated to the purpose of the model without entering into its details. The block should have a comprehensive overview of the model, its main remit and the rationale behind its conception. This block is made by three sections: purpose, entities sates variables and scales and, finally, process overview. The second block (Design) provides an overview of the model structure with information about its characteristics. The protocol offers a sort of list of aspects to cover, as emergence, adaptation and so on. The third part (Details) covers all the needed details useful to implement the model. It is made by three parts: initialization, input and submodels.

6.1 Objective

6.1.1 Purpose

The purpose of the model is to study how the adherence to an external reality evolves in a system made by individuals who could interact among each other in different ways. In particular five different channels of communication are considered here: autonomous search, exchange by propinquity, exchange through organization governance, exchange through project governance and exchange through informal network. The core idea of the model is to show three effects. First, the adherence trend proposed by March and discussed by following literature is heavily impacted by different channels especially when active at the same time, then the balance of the exploration exploitation could be in principle reached but the equilibrium is unstable and volatile. Second, there are emergent phenomena that could not be described focusing only on specific aspects of the exploration exploitation problem. There is the need to consider it in a more comprehensive way. Third, the network of connections is dynamic and tends to shape itself according to environment it is exposed to. Moreover, not only the shape of network is affected by the channels of communication but also by the temporal evolution of the system. This means that the network topology changes over time. Even if the agent's task is the same, the connection characteristics are not always the same. The model is based on real data to give heterogeneity to the agents: data are acquired or derived from the organization that is used as a reference. Hence, office layout, friendships connections, organizational structure and job roles used as parameters in the simulations are provided using the real ones.

6.1.2 Entities, state variables, and scales

The first entity considered in the models is the external reality. External reality is given at the beginning of the simulation and it does not interact with other entities of the model. It is taken as reference by agents during the simulation. External reality is characterized by a sequence of two possible values (1, -1) and this sequence is conceptually split into parts of equal length called topics. If needed, it is possible to activate the change of external reality during the simulation. The change does not happen through interaction but is imposed by the simulation at the desired time.

Another entity is the individual which mimics a member of the organization. The agent acts into the organization with the goal to seek for knowledge trying to reach the knowledge owned by the external reality. The agent is endowed by a representation of the external reality (also called belief). This representation has the same length of the external reality and it could take three values (1, -1 and 0). If a number in the individual sequence is equal to the relative number in the external reality sequence it means that individual is aligned with external reality. If not, she is not aligned. If the number is 0, she does not have an opinion about the specific part of the external reality. The organization has many agents and the goal of the agent is to interact with other colleagues exchanging knowledge. The agent will seek someone with better representation of external reality in order to exchange it and to better her situation (that is the adherence of her representation to the external reality).

To interact in this model, agents are given with further characteristics. First, Individual has memory of all the interactions with other individuals and uses this memory to seek newer knowledge during the interaction. Second, the agent has a own propensity to accept others knowledge during the interaction. This propensity is called learning rate. Third, agent has a position in the organizational hierarchy, she belongs to a specific unit in the organization and she has a own job role in it. These information are used to simulate the organizational governance, that is department and project meetings. Individuals have also friends and offices.

On the time basis, one step of the model ideally represent a working day in which an individual could seek and acquire knowledge about a specific part of the external reality. The evolution ends once the number of predefined cycles has been reached.

6.1.3 Process overview and scheduling

The model has a sequence of steps which are repeated during the simulation. For every time step, different events could take place[1]. The model has two nested levels of loops: the outer one cycles on the number of run steps (mimicking the working day) and the inner one cycles on the activities list to perform during the run step. The sequence of the operations is the following: first, the eternal reality is changed based on the decided strategy. Second, some individuals are sampled and let interact with the external world. Third the sequence of agents interactions and meetings for the run step is created. The list is created sampling randomly from the agents list and meeting list. This list is the schedule for the inner iteration. Hence, looping on the activities list, meetings, autonomous search, propinquity search and informal search are carried out. Last operation is the snapshot of individuals' rank, adherence, popularity and the snapshot of the network of connections. All data are saved in table format and processed at the end of the simulation[2]. Figure 6.1 show the structure of the model.

[1]Not necessarily all the events are present in every simulation. For example if in a simulation the external reality change is switched off, the step is not performed.

[2]This means that there are no real time output during the simulation. All pictures and analysis are done after the simulation ends. This choice was made to reduce the computational time and to perform simulation and analysis in parallel.

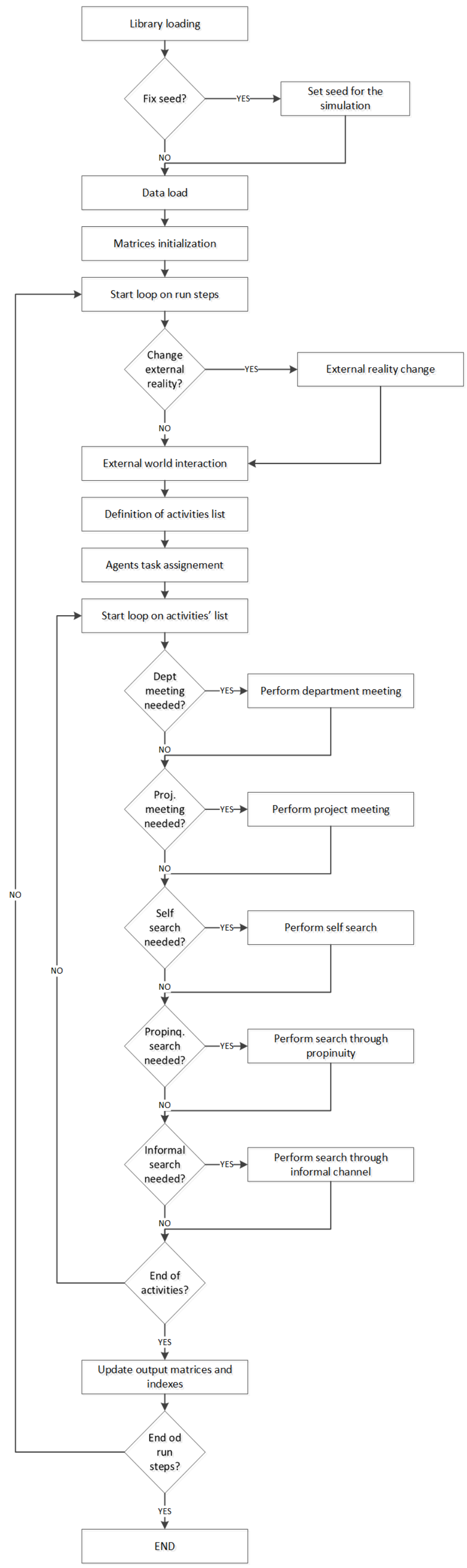


Figure 6.1: Model structure

6.2 Design

6.2.1 Emergence

Two simulation outputs are saved to study the desired phenomena. The first output is at agent's level and the second one at connections' network level. Agent's adherence to external reality is the main output of the simulation, with particular attention to its evolution over time. Collecting all the adherences, the mean adherence of the entire organization emerges during the evolution as consequence of interaction among individuals and as consequence of interaction among individuals and external world.

The connections networks also emerge and evolve as consequence of these interactions. Network evolution over time is considered as the second main output. Network adjacency matrix is then saved at regular intervals and the evolution of the network and its characteristics is examined.

6.2.2 Adaptation

Individual has different rules to follow during the interaction.

When interacting in the autonomous search, the individual could choose the counterpart based on the results of previous relationships. For the selected topic, the individual must search in her stack who is the best responding colleague and select her. Once selected, the interaction takes place if the adherence of the counterpart is greater than the owned one. If not, the individual could seek in the counterpart's stack to find a potential other interlocutor. If the individual has not other colleagues in the stack, she could consult the popularity stack[3] and select from it. Again, the exchange could take place only if the counterpart adherence is grater than the owned one.

If the individual is involved in the search through propinquity, the choice of the counterpart could only be made among the colleagues that are seated near the individual. The possible choices are defined by the physical layout network, provided at the beginning of the simulation. This network does not change over time. The exchange could happen either seeking better knowledge or simply exchanging the owned one.

If the individual is involved in the informal channel, the interaction could happen only with a person selected among her friend's list. This list is fixed and does not change during the simulation or between simulations. In in formal interaction there is no seek for better knowledge.

Individual involved in a department meeting could attend only if they are in the attendance list and they belong to the same organizational unit. For a

[3] That is the general stack $\hat{S}$.

project meeting only the first constraint is valid. During the meeting, agents exchange knowledge with the attendees with higher organizational rank. If more than one has the same rank, the most popular is chosen as reference. If, again, more than one has the same popularity, an average of their beliefs is calculated and used as reference.

6.2.3 Sensing

The agent has three variables to consult during the interaction: one is internal (stack), one is from the environment (general stack) and the last one is a result of an evaluation of the interlocutor. Before the interaction, the agent could consult her stack where is traced the output of previous interactions and decide who could be a potential counterpart for the next interaction based on that. Moreover, if the owned stack is not enough to decide, agent could have access to the general stack, that is the ranking of popularity. Accessing this stack, the agent could find a proper interlocutor. During the interaction, any agent knows the output about the comparison of her representation of external reality and the representation of the other individual she is interacting with. Hence she knows whether or not the counterpart owns a better representation of external reality. Finally, agents know the organizational rank of the colleagues.

6.2.4 Stochasticity

The model presents stochasticity at different levels. During the initialization phase external reality, agents' beliefs, and general stack encompass a stochastic step. Random selection among values 1 and -1 is made to create external reality, a random selection among 1, -1, 0 is made to create the belief for all the present agents. General stack is randomly populated to let the evolution start. If all agents' stacks and general stack were empty, no interaction could take place since an agent has no memory or previous interactions (not yet happened) neither there is a common list of experts to consult.

Closed the initialization phase, stochastic steps are present during the execution of the run step.

First, for the interaction with external world, a number of agent is selected random from the list of all the agents.

Second, the creation of the activity list is made by reordering randomly all the activities to be managed during the run step.

Third, every agent is assigned to a communication channel for the run step and this assignation is made choosing randomly. Here the probability to be assigned to a specific channel is not equal but is weighed by the job role[4]. Details of this mechanism are presented in section 6.3.11.

Fourth, the knowledge topic is randomly selected before any interaction and any meeting.

Fifth, all the attendees to a meeting are randomly selected according to the rules specified in sections 6.3.12.

In self search, propinquity and informal interaction, the selection of the interlocutor is made on a random choice among the candidates if there are more than one potential interlocutor. The same approach for general stack consulting, if more than one candidate is present, a random choice is made.

At lower level, stochastic processes are present during knowledge distortion and knowledge exchange.

6.2.5 Observation

Average adherence value at the asymptote and the time to reach it are used as main output and figures showing the trend of these values, changing simulation parameters, are used to derive the results. Network indexes as density, Q factor, QAP index and energy are used to discuss the topology evolution.

6.3 Details

The part pertaining detail starts with two standard parts, initialization and input data, where it is described how the model starts the simulations. The following part of the section gives a detailed explanation of all the important parts the model is made of (henceforth submodel).

6.3.1 Initialization

Simulation starts with the initialization of the relevant parameters. The first groups of parameters is made by the flags used to shape the scenario to be simulated. All five channels of communication, interaction with external world, change of external reality (and type of change) and suppression of knowledge superiority in propinquity channel could be selected or deselected independently and a dedicated flag is present. Changing the values of the flags allows to create different scenarios where the agents could interact.

[4]For example, a lab technician has higher chance to attend to a department meeting than a project meeting.

There is a flag to manage the random engine, the model could run with a fixed seed number of an unfixed one. A dedicated flag is available as a dedicated parameter to impose the seed number.

The second group of parameters is used to modulate the value or intensity of specific actions. The shape of external reality (and then of individual representation) could be set modulating the number of topic present in the reality array and relative length. The number of agents in the simulation is another parameter. The number is selected starting from the dimension of the selected organization (128) and choosing a number which could be easily doubled or halved more than once for sake of simplicity. Hence the first value is 128, 256, 64 and 32 are then derived. No values greater than 256 are considered in standard simulation since it is difficult to have departments with this dimension. Considering that March simulated 50 agents, 64 and 32 are considered suitable[5]. Strictly related to number of agents is learning rate array.

The number of simulation steps is provided. The number is great enough to let the system evolve properly and it is based on preliminary run test. This because the system evolves with different speed changing the environment setup. This number drives the outer loop of the model (refer to 6.1.3).

To manage governance channels (department and project meetings) the numbers of meetings are provided. Their values enter in the creation of the activities list during the simulation. One value is for department meetings and the other for project meetings.

Other modulating parameters are the portion of individuals to involve in the external world interaction and the portion of knowledge array to involve.

Finally, a couple of parameters are needed to modulate the influence of physical layout during propinquity channel simulation. The two parameters define how distant an individual could search for knowledge through office layout.

Table 6.1 reports the list of the parameters. The first column has the symbol of the parameter, the second column has the relative description and the third has the equivalent column name as reported in table 7.1.

6.3.2 Input data

The simulation requires different input data to run. The input data provided to the model are the real data derived from the characteristics of the organization taken as reference for the model.

[5]During the simulation campaigns, also 16 and 2560 agents were simulated but for specific needs as explained in next chapters.

Table 6.1: Model parameters and data

Symbol	Description	Table 7.1 ref.
F^{sel}	Activation or deactivation of self search channel	SC
F^{dpt}	Activation or deactivation of department meeting channel	DMC
F^{prj}	Activation or deactivation of project meeting channel	PMC
F^{pro}	Activation or deactivation of propinquity channel	PC
F^{fri}	Activation or deactivation of informal channel	IC
F^{ewi}	Activation or deactivation of interaction with external world	AEI
F^{rnd}	Activation or deactivation of fixed random seed	SRS
F^{erc}	Activation or deactivation of external reality change during simulation	ERC
F^{ape}	Type of external reality change: periodic or aperiodic	ERCT
F^{ksp}	Activation or deactivation of knowledge superiority in propinquity channel	PS
F^{kd}	Activation or deactivation of knowledge distortion during exchange	KD
N_{ag}	Number of agents in the simulation run	NAG
N_{ER}^{cat}	Number of external reality knowledge categories	Ncat
N_{ER}^{dim}	Dimension of each knowledge category	NDcat
N_{run}	Number of simulation steps	not applicable
N_m^d	Number of department meetings per step	Ndm
N_m^p	Number of project meetings per step	Npm
L	Agents learning rate array	LR (.xls file)
S_{rnd}	the seed number, if F^{rnd} is set to TRUE	SN
T_{dep}, T_{prj}, T_{aut}	Agent job role definition	.xls file
N_{dd}	Portion of knowledge to involve in the distortion during exchange	NDD
G_{min}	Minimum distance between agents in propinquity channel	Pmin
G_{max}	Maximum distance between agents in propinquity channel	Pmax
τ_{er}	Frequency for external reality change	Tec
N_{ed}^{ag}	Number of agents per step to involve in the interaction with external world	NAGE
τ_{ewi}	Frequency for interaction with external world	Tei
N_{edd}	Portion of knowledge to involve in the interaction with the external world	NDE
N_{kd}	Activation or deactivation of knowledge distortion during exchange	KD
D_T	Agents' department belonging	.xls file
R^{ag}	Agents' organizational rank	.xls file
d	Geodesic distances among agents	.xls file
N^{fri}	Agents' friendship connections	.xls file
S^L	Agents stack dimension	SL
k	Stack update constant	SK

First, two networks of connections are provided: the physical layout one and the informal one. Propinquity network is a geodesic matrix, where the *"office"* distance ($d_{i,j}$) among each pair of agents is reported. This network is essential when an agent has to select the neighbours to interact with. Since the matrix exhibits the distance, it is possible to modulate the effect of office interaction forcing the agent to search only among neighbours within a specific distance. The creation of the propinquity model followed the space syntax approach starting from the real office layout of the reference organization.

Friendship network is the second provided to the model. This network is created elaborating the results of a dedicated survey where all the members of the organization were asked to answer to the following request:

"Please select the colleagues you consider good friends. If they are colleagues in your competence area they should be persons you spend time with or share social informal activities. If they are in geographical distant areas, they should be persons you have regular link through social media or you would like to spend time with if was possible"

All members were coded to guarantee the privacy and the results led to the friendship adjacency matrix.

Propinquity and friendship matrices do not change over time hence a portion or a multiple of them are used when the number of agent in the simulation is not equal to 128. If the number is less (64, 32, 16) a portion of the initial matrices is provided, if the number is greater (256) the matrices are doubled[6] .

The second part of input data shapes the agents. Department belonging is the first information provided. This information is necessary to create the department meeting attendees and to apply the knowledge distortion. The organization has 7 different departments and the number of knowledge category is the same: in the simulation it is possible to exchange knowledge about 7 different topics. Departments have been coded from 1 to 7 and the agents received a value accordingly. In the simulations agents are coded with numbers as in friendship matrix. It is important to stress that the codes given are not the same for privacy protection so agent number 23 in the the department list is no the same as in friendship network. This discrepancy has been judged as not important for the simulations considering that friendship could link anyone in the organization.

[6]The simulations with 2560 agents do not need these matrices since only self search is considered.

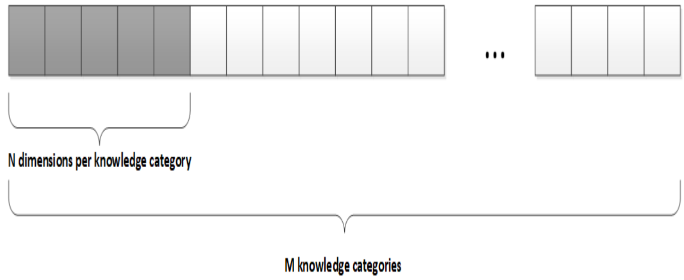


Figure 6.2: Knowledge representation

Job role is the second data provided to the model. Starting from the real organization 11 different roles were derived such as line managers, researchers, lab technicians and assistants. Those roles are coded from 1 to 11 and for each of them the time dedicated to self search, department meetings and project meetings are given to the model. For each role these three values sum up to 100 and they are used to create the activities list.

Apart from real data, stack level and learning rate for all the agents are provided. Stack level $\left(S^L\right)$ is used to limit the number of interlocutor an agent could remember and then it limits the number of interactions to remember. Learning rate L is the array of all the learning rate needed for knowledge exchange.

6.3.3 Knowledge array

The external reality and the individual representation are modelled with an array of values. The admissible values for the external reality array are 1 and -1, for the individual array the values are 1, -1 and 0. Since the model takes into account also the differentiation of topics, the array is actually made by juxtaposing N_{ER}^{cat} arrays where N_{ER}^{cat} is the number of topics.

Also, the length of the arrays could be changed with the parameter N_{ER}^{dim}. This means that if the number of topic is 5 and the relative length is 10 the overall array is long 50. Figure 6.2 shows knowledge representation. The external reality and individual arrays are created at the beginning of the simulation and each position is populated sampling from the admissible values with equal probability.

During the knowledge exchange, only the array portion of selected topic is involved. Of course changing the topic, the relative portion of the array changes.

6.3.4 External reality adherence

Knowledge adherence is calculated as the number of entries in the individual array that are equal to the relative entries in the external reality array divided by the length of the array. The adherence H of the ith agent is then calculated with the following formula:

$$H_i = \frac{\#(A_i[j] = E_j)}{N_{ER}^{cat} \cdot N_{ER}^{dim}} \tag{6.1}$$

where j runs over the interval $\left[1 \,..\, N_{ER}^{cat} \cdot N_{ER}^{dim}\right]$.

During the simulation the comparison is made on a specific topic therefore the formula used is the following:

$$H_{i,T} = \frac{\#(A_i[j] = E_j)}{N_{ER}^{dim}} \tag{6.2}$$

where $T \in [1 \,..\, N_{ER}^{cat}]$, $j \in [1 + N_{ER}^{cat} \cdot (T-1) \,..\, N_{ER}^{cat} \cdot T]$. The adherence at a specific time step s is defined as $H_{i,T}^{s}$.

6.3.5 Knowledge exchange

Knowledge exchange takes place only if the adherence of the donor is higher than the adherence of the seeker for the specific topic T ($H_j^T > H_i^T$). Once the exchange starts among agents i (receiver) and j (donor), for every entry of the reality representation array A owned by the receiver, the following formula is used:

$$A_i[n] = \begin{cases} A_i[n] & if \quad A_j[n] = 0 \\ A_j[n] & if \quad A_j[n] \neq 0, A_i[n] \neq A_j[n], \lambda < L_i \end{cases} \tag{6.3}$$

where $T \in [1 \,..\, N_{ER}^{cat}]$, $i \in [1 + N_{ER}^{cat} \cdot (T-1) \,..\, N_{ER}^{cat} \cdot T]$ and $\lambda \in U(0,1)$.

6.3.6 Knowledge distortion

Knowledge distortion happens during the knowledge exchange. The distortion is applied on the counterpart representation[7] before the knowledge exchange starts if the two agents do not belong to the same department and if the distortion is active. A predefined number of entries in the array is then subject to distortion $\left(N_{dd} \leq N_{ER}^{dim}\right)$. To find which entries in the array are

[7]The counterpart belief is copied, distorted and used for exchange. The copy is discarded at the end of the exchange. The original counterpart belief is not changed.

selected for distortion N_{dd} samples from $U(1, N_{ER}^{dim})$ are sampled forming the subset D. Hence for every $i \in D$ the following formula is used:

$$A[i] = \begin{cases} 0 & if \quad A[i] = 1, 0 < \delta \leq \frac{1}{2} \\ 1 & if \quad A[i] = 1, \frac{1}{2} < \delta \leq 1 \\ -1 & if \quad A[i] = 0, 0 < \delta \leq \frac{1}{2} \\ 1 & if \quad A[i] = 0, \frac{1}{2} < \delta \leq 1 \\ -1 & if \quad A[i] = -1, 0 < \delta \leq \frac{1}{2} \\ 0 & if \quad A[i] = -1, \frac{1}{2} < \delta \leq 1 \end{cases} \tag{6.4}$$

where $\delta \in U(0, 1)$. Posing $N_{dd} = 0$ means to suppress the knowledge distortion effect (This parameter is contemplated in the columns KD and NDD in table 7.1).

6.3.7 External reality change

Turmoil si simulated changing the external reality array. A simple way to creates turmoil was chosen: swapping every entry of the external reality E array. This means that every entry equal to 1 becomes -1 and vice versa. The formula used is:

$$\hat{E_T} = -\hat{E} \tag{6.5}$$

6.3.8 Unlearning and external interaction

Unlearning and external interaction is made with the same mechanism as already discussed. An agent subset K of dimension N_{ed}^{ag} ($N_{ed}^{ag} \leq N_{ag}$) is randomly chosen the interval $[1 \,..\, N_{ag}]$. After this operation, the entries of representation that undergo to interaction or unlearning need to be defined. The subset I is created sampling N_{edd} numbers in the interval $[1 + N_{ER}^{cat} \cdot (T - 1) \,..\, N_{ER}^{cat} \cdot T]$, where T is the selected topic. Then the following formula is used:

$$A_k[i] = \begin{cases} 0 & if \quad A_k[i] = 1, 0 < \delta \leq \frac{1}{2} \\ 1 & if \quad A_k[i] = 1, \frac{1}{2} < \delta \leq 1 \\ -1 & if \quad A_k[i] = 0, 0 < \delta \leq \frac{1}{2} \\ 1 & if \quad A_k[i] = 0, \frac{1}{2} < \delta \leq 1 \\ -1 & if \quad A_k[i] = -1, 0 < \delta \leq \frac{1}{2} \\ 0 & if \quad A_k[i] = -1, \frac{1}{2} < \delta \leq 1 \end{cases} \tag{6.6}$$

where $\delta \in U(0, 1)$, $k \in K$ and $i \in I$. These parameters are specified in columns NAGE and NDE in table 7.1.

6.3.9 Agent stack and stack update

Every individual is associated with a stack S to keep trace of the interactions she has during the simulation. The form of the stack is the following:

$$S_i = \left[\begin{array}{c|c} \hat{A} & \mathbf{R}^i \end{array}\right] = \left[\begin{array}{c|ccc} A_1 & R^i_{1,1} & \cdots & R^i_{1,N^{cat}_{ER}} \\ \vdots & \vdots & & \vdots \\ A_N & R^i_{N_{ag},1} & \cdots & R^i_{N_{ag},N^{cat}_{ER}} \end{array}\right] \tag{6.7}$$

where $\hat{A}$ is the vector containing the remembered agents (from 1 to S^L) and $\mathbf{R}^i$ is the relative matrix containing the cumulative results of the interactions between agents i and j on different topics. Hence S_i has dimensions $S^L_i \cdot (N^{cat}_{ER}+1)$ where $\left(S^L_i < N_{ag}\right)$. The general notation for the agent's stack entry is $S^s_{i,j,T}$ where i is the agent owner of the stack, j is the counterpart, T is the topic and s a moment during the execution of the model. Hence $S_{32,45,4}$ represents the cumulative results of the interactions among agent 32 and 45 on topic 3[8]. By construction $R^i_{j,k} \geq 0$ for $i \in \left[1 \,..\, S^L\right]$, $j \in \left[1 \,..\, S^L\right]$ and $k \in [1 \,..\, N^{cat}_{ER}]$.

After every interaction, the agent stack needs to be updated. At the step $s+1$ the stack entry is changed in the following way:

$$S^{s+1}_{i,j,T+1} = \begin{cases} S^s_{i,j,T+1} + k & if \quad H^s_{j,T} > H^s_{i,T} \\ S^s_{i,j,T+1} - k & if \quad H^s_{j,T} \leq H^s_{i,T} \\ 0 & if \quad H^s_{j,T} \leq H^s_{i,T}, S^s_{i,j,T+1} = 0 \end{cases} \tag{6.8}$$

where s and $s+1$ represent the moment before and after the exchange and $k > 0$ is a model parameter. Being k positive, the entry in the matrix increases when an agent interacts with a more knowledgeable agent and it decreases when the adherence of the counterpart is lower than the owned one. Once an entry reaches the value 0, any subsequent unfruitful interaction will not decrease the value any longer, since all the entries in the matrix need to be not negative. Of course, if a positive interaction takes place, the entry becomes again greater than 0 and hence further decreases could be again possible. In the initialization phase, all the agent's stacks are set to 0. For the simulations presented, k is set equal to 1. Figure 6.3 summaries the entire work flow regarding the agent stack update.

In the flow is also detailed the mechanism adopted to free space in the stack when $S^L_i < N_{ag}$. The worst performant colleagues are found creating the list L_{elim} defined by:

[8] This because the first column of the stack is dedicated to agent's label.

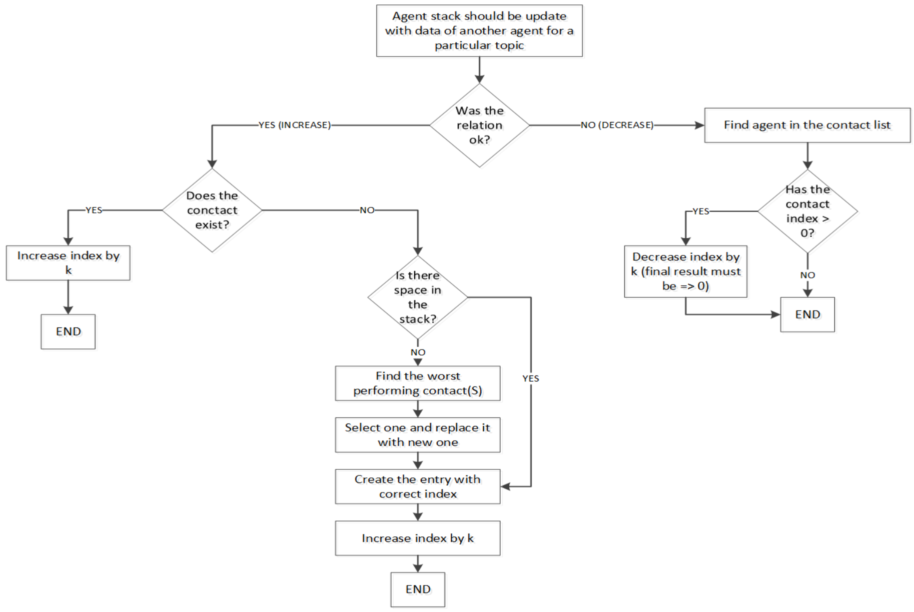


Figure 6.3: Agent stack update workflow

$$L_{elim} = \left(j | S_{j,T+1} = \min_{1 \le j \le S^L} (S_{j,T+1}), 1 \le j \le S^L \right) \tag{6.9}$$

if $\#L_{elim} \geq 1$ then an agent is randomly selected and discarded. The empty space could be replaced with the new agent protagonist of a positive exchange.

6.3.10 General stack

The model is equipped with a general stack $\hat{S}$ which traces the popularity of the agents. The stack is a matrix with N_{ag} rows and N_{ER}^{cat} columns. In the initialization phase all the entries are set to 0 and a defined percentage $\hat{S}_1$ is set to 1 to let the interactions start. $\hat{S}_1$ is set to 0.06% in the presented simulations.

The update of the stack is made with the following criterion:

$$\hat{S}_{j,T}^{s+1} = \begin{cases} \hat{S}_{j,T}^{s} + k & if \quad H_{j,T}^{s} > H_{i,T}^{s} \\ \hat{S}_{j,T}^{s} - k & if \quad H_{j,T}^{s} \le H_{i,T}^{s} \\ 0 & if \quad H_{j}^{s} \le H_{i,T}^{s}, \hat{S}_{j,T}^{s} = 0 \end{cases} \tag{6.10}$$

where s and $s+1$ represent the moment before and after the exchange and $k > 0$ is a model parameter. For the simulations presented, k is set equal to 1.

6.3.11 Channel assignment

In the initialization phase, the workload for all the individuals is created. Every individual is provided with three workload parameters: T_{dep}, T_{prj} and T_{aut} which represent the probability an agent could spend time in department meeting, project meeting and autonomous activities (self search, search though propinquity and search through informal channel).

From the three parameters, for every run step s, the following set are created:

$$\begin{array}{l} De_s = (1|\#De = T_{dep}) \\ Pr_s = (2|\#Pr = T_{prj}) \\ Au_s = (3|\#Au = T_{aut}) \end{array} \tag{6.11}$$

where $\#De_s + \#Pr_s + \#Au_s = 100$. After that the set $W_s = De_s \cup Pr_s \cup Au_s$ is created.

At the beginning of every run step, a random sample from W_s is withdrawn. If the number sampled is 1, the agent will be assigned to department meeting, if the number is 2 to project meeting and if the number is 3 to autonomous search. When an agents is assigned to autonomous search, then a further random selection is made choosing which channel will be used (self search, search though propinquity and search through informal channel). This operation is repeated for each agent.

6.3.12 Meetings

Meeting submodel is composed by four sections: topic selection, attendees list definition, meeting belief definition and knowledge exchange. The submodel is the same for department and project meeting except in one step in the attendees list definition. By assumption, in a department meeting all the attendees must belong to the same department. This constraint is not necessary for project meeting.

In the first section, the topic is selected among the interval $[1 \,..\, N_{ER}^{cat}]$. The second section starts from the list of all potential attendees to find the final list of attendees for the particular meeting. This list (L_{pot}) comes from the channel assignment from which all the agents assigned to the desired type of meeting are selected (groups De_s and Pr_s).

If the meeting is of type department, the list of attendees is intersected with the list of department agents ($L_{pot} \cap D_T$). The list D_T is provided as input data at the beginning of the simulation.

Hence the candidate list L_{can} is equal to L_{pot} for project meeting and equal to $L_{pot} \cap D_T$ for department meetings. The final list of attendees L_{fin} is then derived randomly sampling A_{meet} individuals from L_{can} where the number A_{meet} is randomly chosen in the interval $[2 \, .. \, L_{L_{can}}]$ where $L_{L_{can}}$ is the length of L_{can} array.

With the final list set, the submodel enters in the belief definition stage. Here the goal is to define the belief that exits from the meeting and all the attendees have to interact with. The first selection criterion is the hierarchical rank: the attendees with the higher rank are selected. If only one agent is selected, her belief will become the meeting output belief. If more than one agents are present, the submodel looks for popularity in the general stack and it selects the attendees with the highest score. If only one agent remains, her belief will become the meeting output belief, otherwise an average belief of all the selected agents will be created and used as the output belief. The average belief is calculated with the following formula:

$$B_j = \operatorname{sgn}\left(\sum_{k \in L_{fin}} A_{k,j}\right) \tag{6.12}$$

where $j \in [1 + N_{ER}^{cat} \cdot (T-1) \, .. \, N_{ER}^{cat} \cdot T]$.

The last section is the knowledge exchange. All the attendees enter in the knowledge exchange with the meeting output belief and their stack is then updated. As final step also the general stack is updated.

Simulations subgroups DN, DO and MB (refer to chapter 7 for subgroups explanation) have actually a different mechanism for L_{fin} definition, for reasons explained in chapter 11. This new mechanism is called *"restricted"* whereas the former is called *"unrestricted"*. The model, as a first step, selects the meeting room and chooses the attendees from the panel according to the maximum capacity of the room. The number of attendees for a meeting could be from a minimum of 3 to a maximum equal to the capacity of the meeting room in which the meeting is held. This means that the length of L_{fin} and A_{meet} is equal to the maximum attendees capacity of the selected room.

The list of available room is derived from the studied layout and it is reported in table 6.2. The table shows for all the different capacities how many rooms are available.

Table 6.2: Number of meeting rooms and relative capacity

Room capacity	Number of available rooms
3	5
5	3
6	9
8	11
10	1
12	1
14	4
20	1

Considering all the meeting rooms together, it is possible to have the number of attendees per meeting with relative probability as figure 6.4 shows. For example the 17.5% of times, a meeting will be held with 3 attendees. The chart shows a more realistic scenario where most of the meetings (61%) are hold with a limited number of attendees (≤ 6).

At this point, A_{meet} individuals from L_{can} are randomly sampled to obtain L_{fin}. The list L_{can} is derived in the same manner, as explained before.

Figures 6.5 and 6.6 show the complete flows with unrestricted and restricted attendees' selection mechanism.

6.3.13 Self-search

The search starts with the selection of the individual protagonist of the quest. Following step is the selection of the topic T. This selection is randomly made, choosing a number in the interval $[1 \ldots N_{ER}^{cat}]$.

Once individual and topic are selected, the following step is to check the presence of a counterpart in the agent stack. If the counterpart exists, the agent could interact with her, if the stack does not provide any chance, individual could consult the general stack. This action is made by checking that $\underset{1 \leq j \leq S_i^L}{max}(S_{i,j,T+1}) \neq 0$. If the condition is true, all the indexes corresponding to the maximum value are retrieved creating the subset C.

Hence if $\underset{1 \leq j \leq S_i^L}{max}(S_{i,j,T+1}) > 0$ and $\#C = 1$, the two knowledges are compared and if the counterpart one is greater $\left(H_{j,T}^s > H_{i,T}^s\right)$ the knowledge exchange takes place as described in 6.3.5. If $\#C > 1$, one of individuals is randomly chosen.

If $H_{j,T}^s \leq H_{i,T}^s$ the exchange could not take place between agent i and j, the stack and general stack are in any case updated (in a negative way)

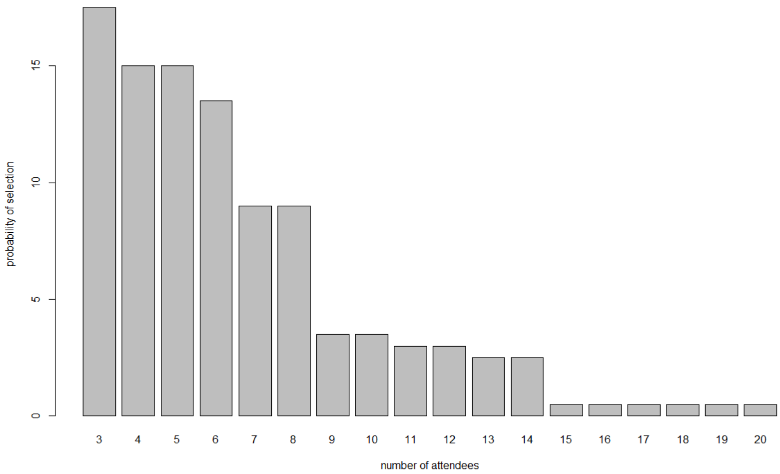


Figure 6.4: Probability of meeting attendees numbers

and ith agent could search in jth stack to find a second agents to meet for knowledge. If $\max_{1 \leq k \leq S_j^L}(S_{j,k,T+1}) \neq 0$, ith agent could select a kth agent and enter in the knowledge exchange.

If $\max_{1 \leq j \leq S_i^L}(S_{i,j,T+1}) = 0$, ith agent does not have other agents in the stack to talk with and then she could search in the general stack. The mechanism then is the same. If $\max_{1 \leq k \leq N_{ag}}(\hat{S}_{k,T}) \neq 0$, ith agent could select a lth agent and enter in the knowledge exchange.

The self search sub-model could also incorporate the knowledge distortion. When ith and jth agents interact but they do not belong to the same department, the knowledge distortion is activated according to 6.3.6.

Figure 6.7 shows the entire flow for self search.

6.3.14 Physical layout

For an agent, propinquity search starts with the determination of available neighbours. This submodel loads the network derived from the offices layout (pic. 5.7) and it defines a subset o potential neighbours according to the following formula:

$$P_i^N = \{j | j \in (1, N_{ag}), G_{min} \leq d_{i,j} \leq G_{max})\} \tag{6.13}$$

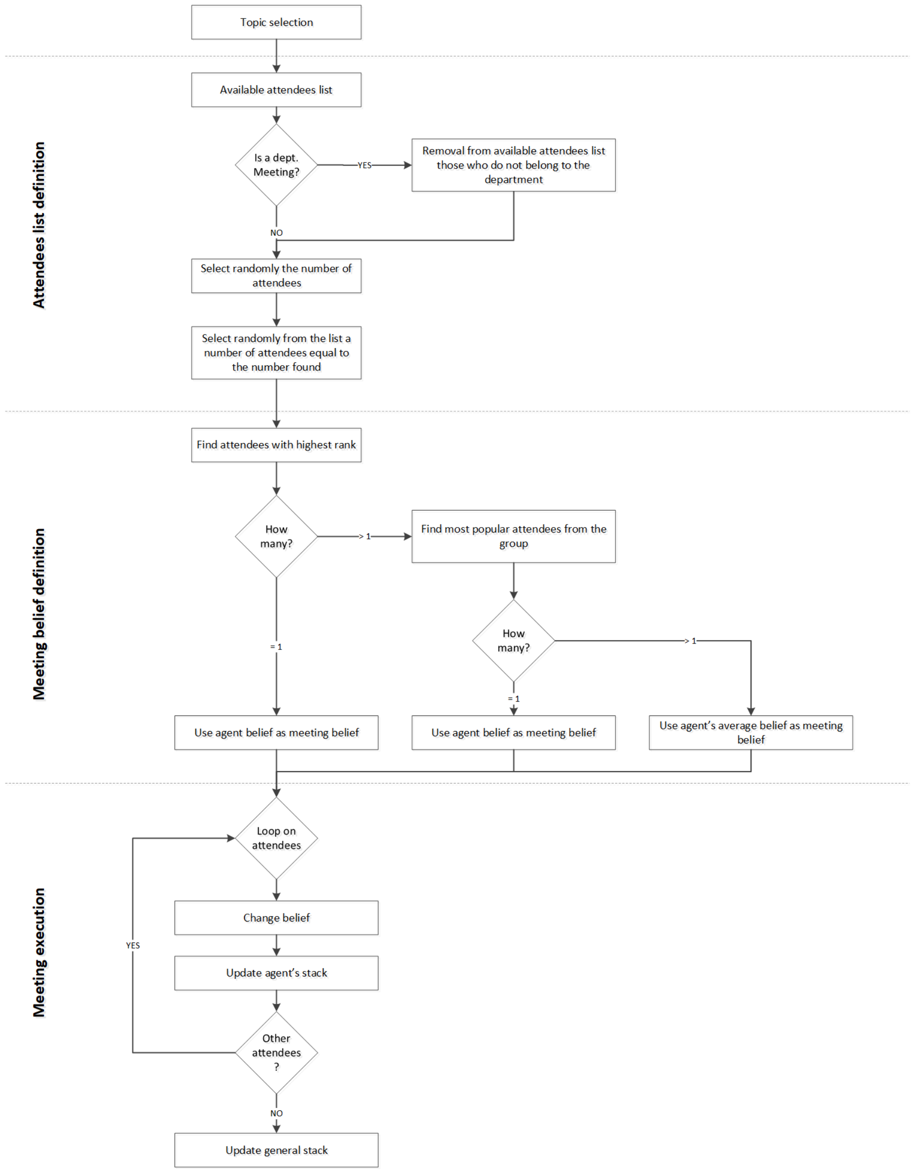


Figure 6.5: Meeting submodel flow with unrestricted mechanism

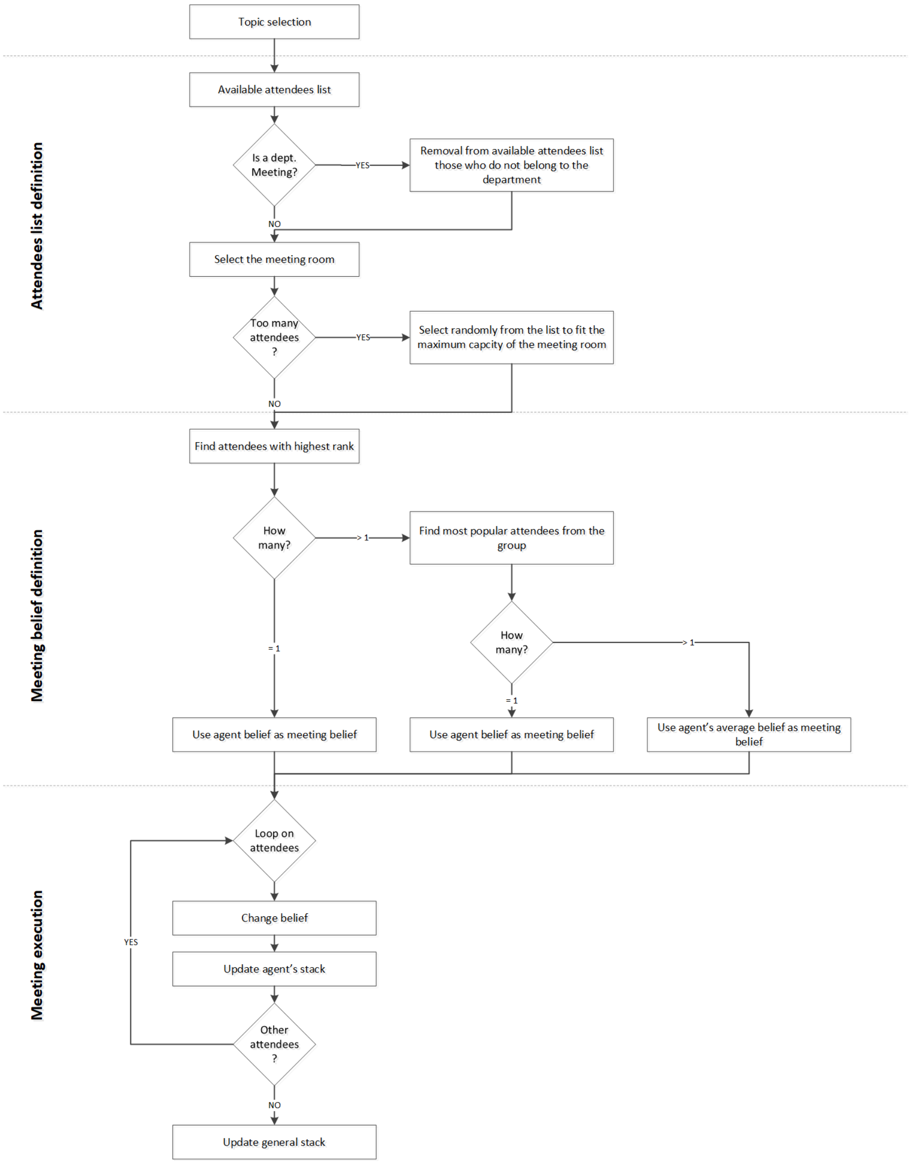


Figure 6.6: Meeting submodel flow with restricted mechanism

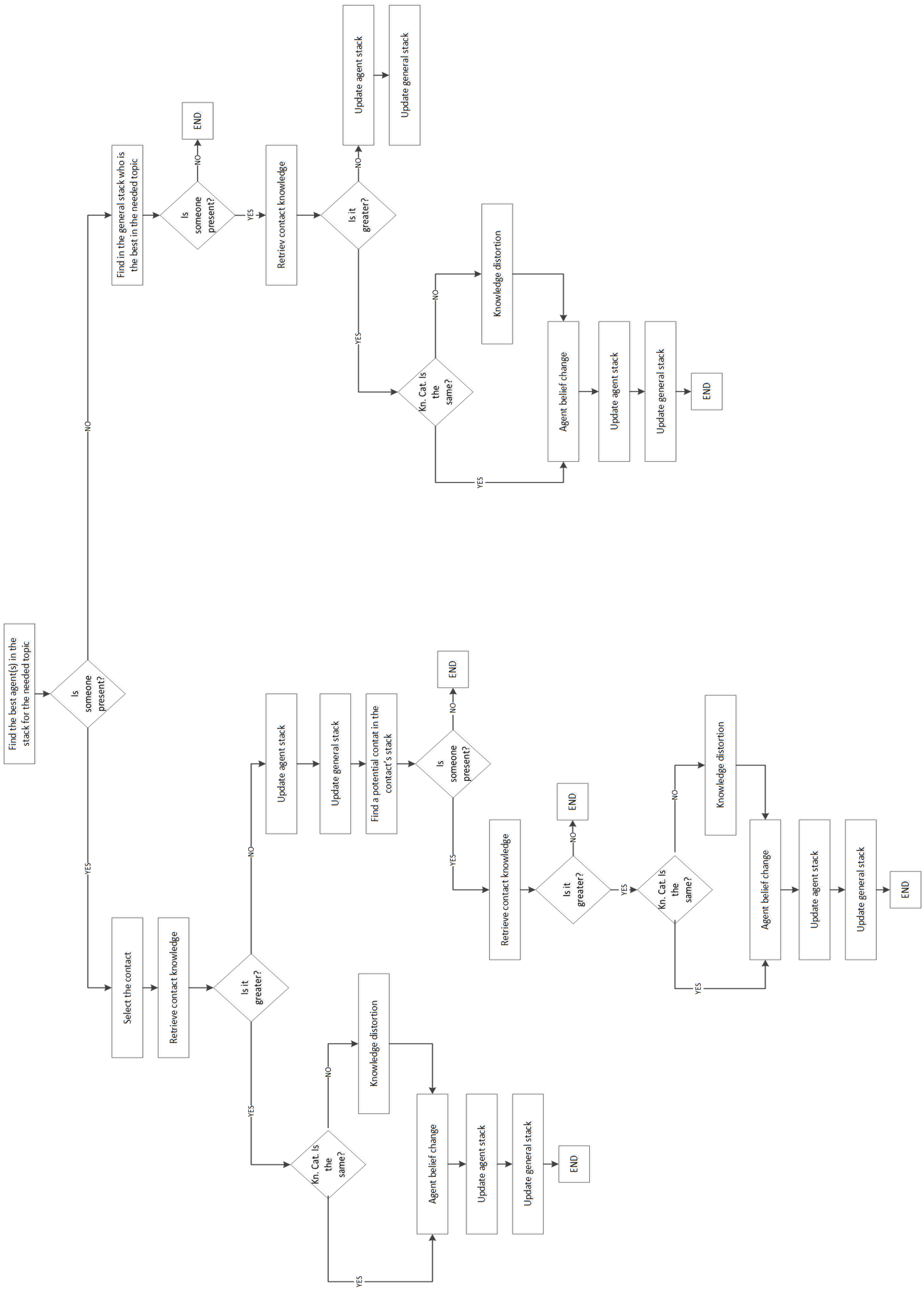


Figure 6.7: Self search submodel flow

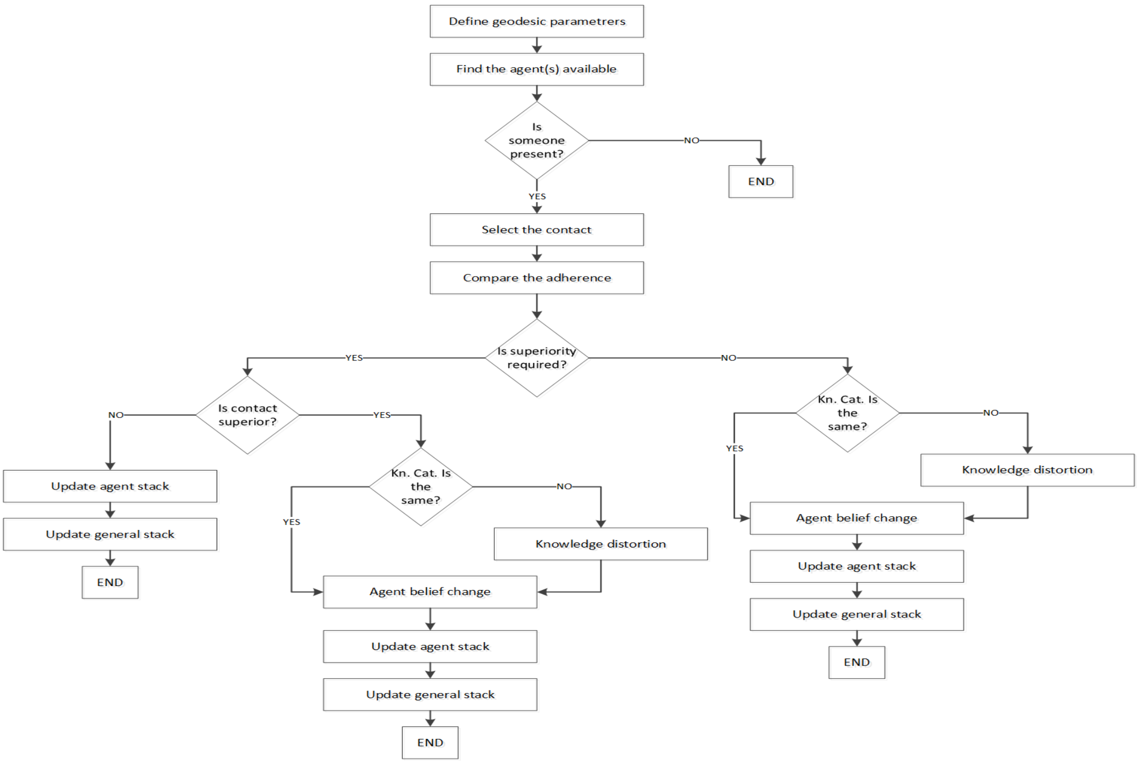


Figure 6.8: Propinquity submodel flow

where $d_{i,j}$ is the relative distance between ith and jth agent. One neighbour is selected sampling randomly from P_i^N, where all the indexes have the same probability.

Once the interlocutor is found, the way is split according to model parameter F^{ksp}: if it is set to 1, the model checks for knowledge superiority, if it is set to 0, knowledge superiority is skipped. In the first case, the two adherences are compared and if the counterpart is superior, the knowledge exchange could happen as in 6.3.5, once the topic is selected. After this operation, the individual and general stack are update. Again, knowledge distortion mechanism is available and selectable. If the alter agent does not have superior knowledge, the two stacks are update in negative way and the submodel ends. If superior knowledge is not required, the exchange could take place in any case and the steps are the same as above.

Picture 6.8 reports the entire flow of the submodel.

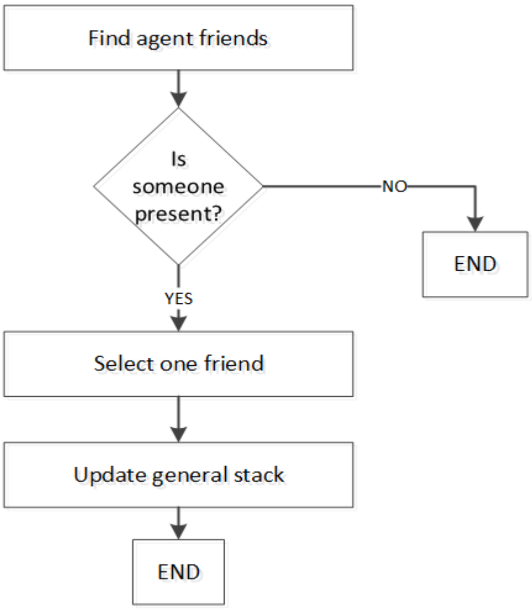


Figure 6.9: Friendship submodel flow

6.3.15 Friendship

Informal channel submodel starts with the definition of the topic and the definition of the friends subset of the individual of interest. The subset is defined as follows:

$$F_i = \left\{ j | j \in (1, N_{ag}), N_{i,j}^{fri} = 1 \right\} \tag{6.14}$$

where N^{fri} is the friendship network loaded at the beginning of the simulation. Among the friends in F_i, one agent is randomly selected. The two agents are friends and the model, by assumption, does not compare the two adherences.

The submodel enters directly in the knowledge exchange phase and only the general stack is then updated.

Picture 6.9 shows the entire flow for the submodel.

Chapter 7

Simulations programme and output

This chapter deals with the simulations programme and it clarifies the output of the simulations studied in the rest of the document.

As explained in the previous chapter, the simulations cover a huge number of parameters which, it turn, span over relative values ranges. The backbone of the experiments is the communication channel. This means that all the simulations are grouped according to the main communication channel under study. In first group (blocks A and B) all the simulations related to self search are considered. In the second (C) the propinquity effect is studied, with and without the self search channel. In the third group (D) meetings are considered together with self search. Propinquity is both considered and excluded. In this group both kinds of governance are considered: department and project meetings. Group E considers friendship. Group G is related to a dedicated analysis of relationship between March's effect and number of agents. Last group (M) considers all the channels together and the maximum heterogeneity.

Table 7.1 shows the entire prospectus of the simulations. The table needs a further explanation.

Column G (group) reports the group of simulations as previously mentioned. The same division is also adopted in the discussion of the results. Groups A, B are treated in chapter 9, groups C and G in chapter 10, group D in chapter 11, group E in chapter 12 and group M in chapter 13.

Column SG (subgroups) reports the subgroup the simulations are split into. Each group has been split into subgroup, covering a particular aspect of the entire study. This division in subgroup is functional to simulations run scheme than to results analysis. Most of the simulations were run in parallel to save time, hence an appropriate schema was needed to easily merge data

Table 7.1: Simulations high level prospectus

G	SG	NS	SC	PC	DMC	PMC	IC	AEI	NAG	LR	KD	NDD	ERC	ERCT	PS	Ncat	Ndcat	Ndm	Npm	Pmin	Pmax	NAGE	NDE	SL	Sme	PR	SRS	SN	Tec	Tei	SK
A	A	108	yes	no	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	yes	4	no	NA	no	7	20	0	0	1	2	12,25,50,100	0	L,M,H	NA	no	yes	9,102,345	NA	1	1
A	B	108	yes	no	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	yes	4	no	NA	no	7	20	0	0	1	2	12,25,50,100	2	L,M,H	NA	no	yes	9,102,345	NA	1	1
A	C	108	yes	no	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	yes	4	no	NA	no	7	20	0	0	1	2	12,25,50,100	10	L,M,H	NA	no	yes	9,102,345	NA	1	1
A	D	108	yes	no	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	yes	4	no	NA	no	7	20	0	0	1	2	12,25,50,100	20	L,M,H	NA	no	yes	9,102,345	NA	1	1
A	G	12	yes	no	no	no	no	yes	2560	0.1,0.5,0.9	yes/no	4	no	NA	no	7	20	0	0	1	2	0,1000	0,2,10,20	L,M,H	NA	no	yes	9,102,345	NA	1	1
A	R	96	yes	no	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	yes/no	4	no	NA	no	7	10,30	0	0	1	2	0,1000	0,2,10,20	L,M,H	NA	no	yes	9,102,345	NA	1	1
A	S	24	yes	no	no	no	no	yes	2560	0.1,0.5,0.9	yes/no	4	no	NA	no	7	10,30	0	0	1	2	0,1000	0,2,10,20	L,M,H	NA	no	yes	9,102,345	NA	1	1
B	A	108	yes	no	no	no	no	yes	32,64,128 256	0.1,0.5,0.9	no	0	no	NA	no	7	20	0	0	1	2	12,25,50,100	0	L,M,H	NA	no	yes	9,102,345	NA	1	1
B	B	108	yes	no	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	no	7	20	0	0	1	2	12,25,50,100	2	L,M,H	NA	no	yes	9,102,345	NA	1	1
B	C	108	yes	no	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	no	7	20	0	0	1	2	12,25,50,100	10	L,M,H	NA	no	yes	9,102,345	NA	1	1
B	D	108	yes	no	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	no	7	20	0	0	1	2	12,25,50,100	20	L,M,H	NA	no	yes	9,102,345	NA	1	1
C	A	216	no	yes	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	0	1	2	12,25,50,100	0	L,M,H	NA	no	yes	9,102,345	NA	1	1
C	B	216	no	yes	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	0	1	2	12,25,50,100	2	L,M,H	NA	no	yes	9,102,345	NA	1	1
C	C	216	no	yes	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	0	1	2	12,25,50,100	10	L,M,H	NA	no	yes	9,102,345	NA	1	1
C	D	216	no	yes	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	0	1	2	12,25,50,100	20	L,M,H	NA	no	yes	9,102,345	NA	1	1
C	E	216	no	yes	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	0	1	6	12,25,50,100	0	L,M,H	NA	no	yes	9,102,345	NA	1	1
C	F	216	no	yes	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	0	1	6	12,25,50,100	2	L,M,H	NA	no	yes	9,102,345	NA	1	1
C	G	216	no	yes	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	0	1	6	12,25,50,100	10	L,M,H	NA	no	yes	9,102,345	NA	1	1
C	H	216	no	yes	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	0	1	6	12,25,50,100	20	L,M,H	NA	no	yes	9,102,345	NA	1	1
C	I	216	yes	yes	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	0	1	2	12,25,50,100	0	L,M,H	NA	no	yes	9,102,345	NA	1	1
C	J	216	yes	yes	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	0	1	2	12,25,50,100	2	L,M,H	NA	no	yes	9,102,345	NA	1	1
C	K	216	yes	yes	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	0	1	2	12,25,50,100	10	L,M,H	NA	no	yes	9,102,345	NA	1	1
C	L	216	yes	yes	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	0	1	2	12,25,50,100	20	L,M,H	NA	no	yes	9,102,345	NA	1	1
C	M	216	yes	yes	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	0	1	6	12,25,50,100	0	L,M,H	NA	no	yes	9,102,345	NA	1	1
C	N	216	yes	yes	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	0	1	6	12,25,50,100	2	L,M,H	NA	no	yes	9,102,345	NA	1	1
C	O	216	yes	yes	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	0	1	6	12,25,50,100	10	L,M,H	NA	no	yes	9,102,345	NA	1	1
C	P	216	yes	yes	no	no	no	yes	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	0	1	6	12,25,50,100	20	L,M,H	NA	no	yes	9,102,345	NA	1	1
C	Q	54	yes	yes	no	no	no	no	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	0	1	2	0	0	8,11,14	NA	no	yes	9,102,345	NA	0	1
C	R	36	yes	yes	no	no	no	no	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	0	1	2	0	0	20,25	NA	no	yes	9,102,345	NA	0	1
D	A	288	yes	yes/no	yes	yes	no	no	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	L,H	L,H	1	2	0	0	H	U	no	yes	9,102,345	NA	0	1
D	B	288	yes	yes/no	yes	no	no	no	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	L,H	0	1	2	0	0	H	U	no	yes	9,102,345	NA	0	1
D	C	288	yes	yes/no	no	yes	no	no	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	L,H	1	2	0	0	H	U	no	yes	9,102,345	NA	0	1
D	J	288	yes	yes/no	yes	yes	no	no	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	L,H	L,H	1	2	0	0	H	R	no	yes	9,102,345	NA	0	1
D	K	288	yes	yes/no	yes	no	no	no	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	L,H	0	1	2	0	0	H	R	no	yes	9,102,345	NA	0	1
D	L	288	yes	yes/no	no	yes	no	no	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	L,H	1	2	0	0	H	R	no	yes	9,102,345	NA	0	1
D	N	288	yes	yes/no	yes	no	no	no	32,64,128, 256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	L,H	0	1	2	0	0	H	R	yes	yes	9,102,345	NA	0	1
D	O	288	yes	yes/no	no	yes	no	no	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	L,H	1	2	0	0	H	R	yes	yes	9,102,345	NA	0	1
E	A	72	yes	no	no	no	yes	no	32,64,128,256	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	0	1	2	0	0	H	NA	no	yes	9,102,345	NA	0	1
G	A	9	yes	no	no	no	no	no	16	0.1,0.5,0.9	no	0	no	NA	yes/no	7	20	0	0	1	2	0	0	H	NA	no	yes	9,102,345	NA	0	1
M	B	24	yes	yes	yes	yes	yes	yes	32,64,128,256	unif	no	0	no	NA	yes	7	20	unif	unif	1	6	12,25,50,100	10	unif	R	yes	yes	9,102,345	NA	1	1

Table 7.2: Propinquity channel's simulations

No self search - Short range	No self search - Long range
Self search - Short range	Self search - Long range

at the end. Hence, in the next chapters, the analysis does not necessarily follow the simulation subgroups scheme.

The simulations programme starts from the analysis of self search alone. This to find which is the relationship among adherence and important parameters as number of agents, knowledge distortion, interaction with external world and stack level. Moreover, the aim is also to define how broad is March's effect in the simpler scenario. Again, since the connection topology is not defined *a priori*, part of the analysis is devoted to understand the basic dynamics.

These first simulations show that the knowledge distortion is not important and hence it has been excluded for the remaining part of the simulations.

The second group of simulations considers the propinquity channel. Propinquity channel is studied alone and together with self search channel. Furthermore, propinquity channel is studied also with short and long influence. The high level simulations path for propinquity is represented in table 7.2.

During these simulations involving self search and propinquity, the three parameters T_{dep}, T_{prj} and T_{aut} are set to (0, 0, 100) so all the agents are involved only in autonomous searches. Stack level is studied only in specific subgroups (CQ and CR) where the impact of its shape is studied. For all the other simulations the stack level takes three levels as reported in table 7.5. Again, agent ranking is not relevant for these simulations so was deactivated (meaning that all agents have the same ranking).

Third part of simulations programme is devoted to governance channel that is department and project meetings. The impact is studied on self search alone or together with propinquity. Moreover the impact of governance is explored with the two channels together or separated (only department meetings or only project meetings) and shaping the number of meetings and the time allocated to them. This last parameter is set modulating T_{dep}, T_{prj} and T_{aut} according to table 7.3. Subgroups DA, DB and DC consider these simulations and the unrestricted selection mechanism.

This schema is also replicated with the restricted selection mechanism in groups DJ, DK and DL. Subgroups DN and DO have the same structure of subgroups DL and DM but with the prestige feature switched on to estimate the effect of ranking position in governance channel.

Following part of simulations is dedicated to informal channel. Supposed a mild effect, informal channel is studied only together with self search. This

Table 7.3: Workload for formal channels simulations

	Low level			High level		
Meeting scenario	T_{dep}	T_{prj}	T_{aut}	T_{dep}	T_{prj}	T_{aut}
Department meeting only	20	0	80	90	0	10
Project meeting only	0	20	80	0	90	10
Both meetings	5	5	90	45	45	10

in light of the influence propinquity and governance have on the results.

The two last groups of simulations are for special purposes: group GA is used to better study March's effect as discussed in chapter 10 and group MB is devoted to the union of all the channels and introduction of whole agents' heterogeneity. This groups is used to start discussing about emergent phenomena.

Now, all the remaining columns of table 7.1 are explained and discussed.

Column NS (number of simulations) shows of how many simulations the subgroup has.

Columns SC (self search channel), PC (propinquity channel), DMC (department meeting channel) PMC (project meeting channe) and IC (informal channel) indicate which communication channel is active or not during the subgroup simulations. A subgroup reporting *"yes/no"* includes some simulations with the channel active and other with the channel switched off.

Column AEI (Activation external interaction) enables the study of external interaction and unlearning effect.

Column NAG (number of agents) shows the number of agents used in the simulations. The numbers of agents used in the simulations are 32, 64, 128, 256. The choice is based on the appreciation that the organization taken as reference to create the model was composed by 128 persons. Starting from this number the other are calculated considering three aspects: first, the numbed should easily doubled or halved for sake of computational simplicity, second, the range should cover the number of agents March used (50), third, at least one point above 128 is necessary to cover the range of the reference organization. There are subgroups with 2560 and 16 agents. The first subgroup is related to simulations intended to discuss about the leaning rate effect with extremely large organization and the number is taken 10 times greater than the greatest value of typical range. The second subgroup is devoted to a simulations intended to demonstrate the break of March's effect when the number of agent is small. Following the same approach, the number is calculated halving 32.

Column LR (Learning rate) reports the values for the agents' learning rates. Three vales have been chosen: 0.1, 0.5 and 0.9. Being a probability of knowledge exchange, the values must be within the range $(0, 1)$. According to extant literature (refer, for example, to figure 2.2), 0.1 is the value for slow learner individuals, 0.9 is for fast learner individuals and 0.5 for individuals neither prone to fast or slow learning. Although some scholars have reported many more points in the possible range of learning rate (Mitomi and Takahashi 2015), three points have been considered enough to create a trend and to understand the underlying behaviour without loading too much effort on the computational side (considering the huge number of parameters combinations studied). Group M has learning rate sampled form the uniform distribution $U(0, 1)$. Learning rates are provided to the model as a file.

Column KD (knowledge distortion) reports whether or not the knowledge distortion is active during the knowledge exchange (refer to 6.3.6). Subgroups with both yes and no, follow the same logic already explained. The knowledge distortion is demonstrated as marginal in chapter 9, hence switched off in the remaining simulations.

Column NDD (number of dimensions distorted) is the amount of knowledge that could be distorted during the exchange. The number is set to 4 since it is the 20% of 20 (the dimensions of each category). Since every exchange takes place for a specific knowledge category, an amount of 20% of knowledge distortion is considered sufficiently high to study its effect.

Column ERC (external reality change) shows if the external reality is changed within the simulations. Following column ERCT (external reality change) shapes the kind of change: it could be periodic or aperiodic.

Column PS (propinquity superiority) reports whether or not the simulations have the propinquity superiority active that is if the exchange happens only after a superior knowledge is found (refer to 6.3.14).

Columns Ncat (number of categories) and NDcat (number of dimensions per category) rule the external reality and individual beliefs representation according to section 6.3.3. All the simulations take 7 knowledge categories. The number is derived from the number of departments the reference organization has. For this reason, also the number of department is 7. Each category is made by a number of dimensions that typically is 20. The number was select to lay near the number used in extant literature and also considering that a knowledge represented by 140 bits is long enough for the remit of the study[1]. Two subgroups presents different values for the dimension (10 and 30) because their simulations were run to study the effect of knowledge complexity on network density (refer to chapter 9).

Column Ndm (number of department meetings) and Npm (number of

[1]The length is calculated multiplying N_{ER}^{cat} and N_{ER}^{cat}, as described in section 6.3.3.

Table 7.4: Number of meetings for different number of agents

	Low	High
32	10	40
64	20	80
128	40	160
256	80	320

project meetings) report the number of dedicated meetings when relative channels are active. The table shows the term L (low) and H (high) since the actual value depends on the number of agents present in the simulations. Table 7.4 shows the real numbers. The lower values is determined considering that only the 30 % of the agents attends a meeting in a (simulation) day[2] and the upper values considering that all the agents attend a meeting and 30% of them actually attend to 2 meetings. Rounding to integer number the values are 10 and 40. The other are calculated doubling the number as the number of agents doubles.

Column Pmin (minimum propinquity range) and Pmax (maximum propinquity range) represent the influence of propinquity channel when active. The numbers are derived as explained in section 5.7 and 6.3.14.

Column NAGE (number of agents involved in interactions with external world) declares the number of agents that are involved in the interaction with external world or unlearning effect at the beginning of every run step. The vales are related to the number of agents, keeping the number of involved agents below 40%. This is considered a satisfactory number appreciating that every run (day) almost half of the agents are involved in the interaction (or are subject to unlearning).

Column NDE (number of dimension involved in the interaction) reports the number of dimensions involved in the exchange with the external world or unlearning effect. The levels 0, 2, 10 , 20 are related to NDcat column values. Considering that typically the extension of a category is 20 the proposed numbers mimic no interaction (0 dimensions), low interaction (10% of the dimensions), medium interaction (50% of the dimensions) and high interaction (100% of the dimensions).

Column SL (stack level) rules the memory effect of agents. Again, since it is related to agents' number, the values are Low (L), medium (M) and high (H), but table 7.5 reports the actual numbers.

The low value is 5 for any number of agents as considered the minimum number of colleagues an individual could remember. The medium value is

[2]Represented in the model by a run step.

Table 7.5: Stack levels for different number of agents

	Low	Medium $\left(\frac{N_{ag}}{2}\right)$	High $(N_{ag} - 1)$
32	5	16	31
64	5	32	63
128	5	64	127
256	5	128	255

set to 50% of the number of present colleagues and high value to the whole number of them. A couple of clarifications: for simulation with 2560 agent the stack level was not changed since for only self search the stack level is not a limiting factor. Two subgroups (CQ and CR) report different values (8, 11, 14, 20 and 25) since the simulations were dedicated to the study of the stack level (refer to section 10.3). The last subgroups only report the High value since the aim is to study the effect of channels without constraints coming from memory effect.

Column SMe (selection mechanism) reports the mechanism used for the creation of the meeting attendees. There are two different mechanisms: U (unrestricted) where the selection is made without considering physical layout constraints and R (restricted) where the room capacity is taken into account.

Column PR (prestige) shows subgroups where the prestige is active. As the previous column, this parameter is only applicable to meeting channel. In groups DN and DO, a number of agents equal to 12.5% of the population is randomly promoted to higher rank (4 in 32 agents simulations, 8 in 64 agents simulations, 16 in 128 agents simulations and 32 in 256 agents simulations). The percentage is selected considering the hierarchy present in the reference organization.

Column SRS (set random seed) defines whether the random seed should be imposed or not. Column SN (seed number) gives the three seeds used in the simulations. This parameter is active if SRS is set to TRUE.

Column Tec (Tau external change) gives the pace of external reality change if ERC is set to TRUE and ERCT is set to *"periodic"*.

Column Tei (Tau external interaction) gives the pace of interaction with external world and unlearning effect.

Column SK (stack k parameter) set the amount of increase and decrease during stacks update (refer to sections 6.3.9 and 6.3.10).

Last row of simulations (group MB) has values as *"unif"*. In this case, the parameters are fed with a file containing the values previously created sampling from uniform distributions (refer to chapter 13 for detail).

As explained in chapter 6, every simulation is made up by several run steps and for every run step the adherence of all the individuals is recorded. Adherence is calculated as the amount of individual's belief that correctly represents the external reality (details are in section 6.3.4). Moreover, every 25 steps, the connections network is recorded for the analysis. Considering that there is always a knowledge exchange, the individual adherence is supposed to increase over time[3]. This because the mutual learning mechanism seeks for superior knowledge. Hence, the recorded output presents for all the individuals a crescent curve of adherence over time.

Picture 7.1 presents on the left a typical simulation output. All the black lines represent the individual adherences to external reality from the beginning of the simulation till the end. In this case the simulation stops after 1000 steps and all the curves reach a stable end point. On the x-axis the run step is reported whereas in the y-axis the value of the adherence is reported. By construction adherence is a value spanning over the range $(0, 1)$. If the value is 0, the individual has a belief that is completely different from the external reality whereas if the adherence is 1 the owned belief coincides with the external reality.

The average curve could be considered (red curve) without a loss of information. By construction, the red curve is made by averaging at every run step the individuals' adherences. March himself wrote *"the proportion of reality that is correctly represented in individual beliefs (on average) can be calculated for any period"* (March 1991, 75). Hence, the analysis is made on the average values of the adherences during the simulations.

The mutual learning mechanism calls for an asymptote reached when there is no superior knowledge among the individuals and then there is no more knowledge to share. Also, Levine and Prietula observed that knowledge accumulates more rapidly in the early stages and accumulates slowly in the terminal phases (Levine and Prietula 2012). This value is what March called the equilibrium point and reported in his paper (figure 2.2). Once the average representation was found, March studied the values of the red curve once reached a stable values, arguing that *"an equilibrium is reached at which all individuals and the code share the same (not necessarily accurate) belief with respect to each dimension. The equilibrium is stable"* (March 1991, 75). In the present work the same value is recorded and used for the analysis. This value is actually made by two information: the value adherence takes and the moment it takes it (called onset point). Being the model regulated by run steps, the moment of the onset is recorded as a run step. In the figure 7.1 the two values are the intersection of the blue line with the x-axis and

[3]This is true at least in the simplest scenarios.

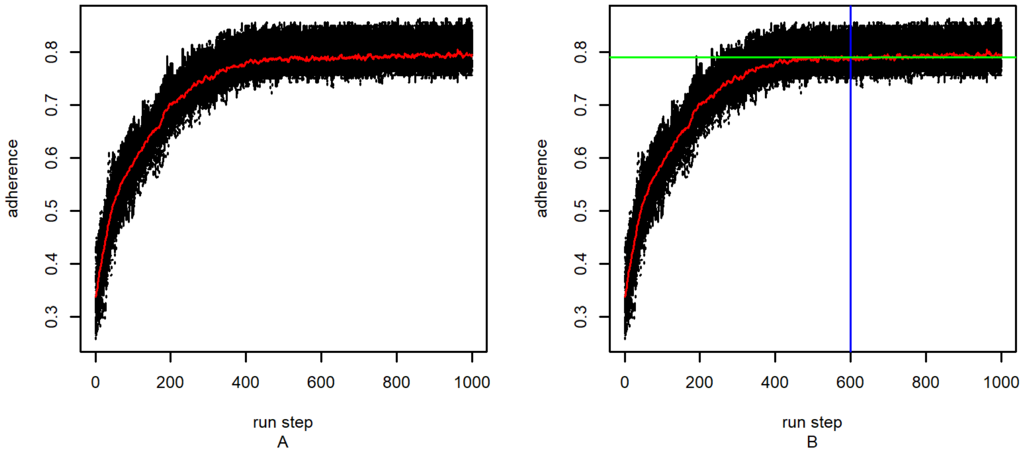


Figure 7.1: Output example

the intersection of the green line with y-axis.

To better show the effect of the context on exploration and exploitation at different values of individual learning rates most of the figure in the document are not shown as in figure 7.1 but as in figure 2.2. This means that the average adherence end point is shown in graphs where the axes denote values of the parameters set for the simulations.

Regarding network analysis, the individual stacks are recorded every 25 run steps and the directed adjacency matrix is derived from them. The presence of a non zero entry in correspondence to a colleague in an individual's stack means that there is a connection and the adjacency matrix will have an not null entry. The obtained adjacency matrix is then used to calculate relevant indexes according to the needed analysis.

Chapter 8

Duality

This chapter is called duality since it starts discussing about the importance of the context in the exploration and exploitation problem. As claimed in propositions 1 and 2 (refer to chapter 3), the impact of the context is presumed great in shaping the amount of knowledge the organization holds at the end of the simulations. This means that, *ceteris paribus*, if only self search is available as communication channel the final average adherence owned by the organization is different from the one owned when propinquity channel is active. Moreover, the superimposition of different channels could bring to different adherence values.

For this first analysis all the end points and onset values of the subgroups reported in table 8.1 are used. Respect to table 7.1, the subgroups and CQ, CR and GA have been removed since these subgroups contain simulations with special parameters values. MB subgroup is not considered for its different behaviour.

Table 8.1: Used simulations subgroups

Group number	Subgroup	Group number	Subgroup	Group number	Subgroup
1	AA	12	CA	23	CL
2	AB	13	CB	24	CM
3	AC	14	CC	25	CN
4	AD	15	CD	26	CO
5	AG	16	CE	27	CP
6	AR	17	CF	28	DJ
7	AS	18	CG	29	DK
8	BA	19	CH	30	DL
9	BB	20	CI	31	DN
10	BC	21	CJ	32	DO
11	BD	22	CK	33	EA

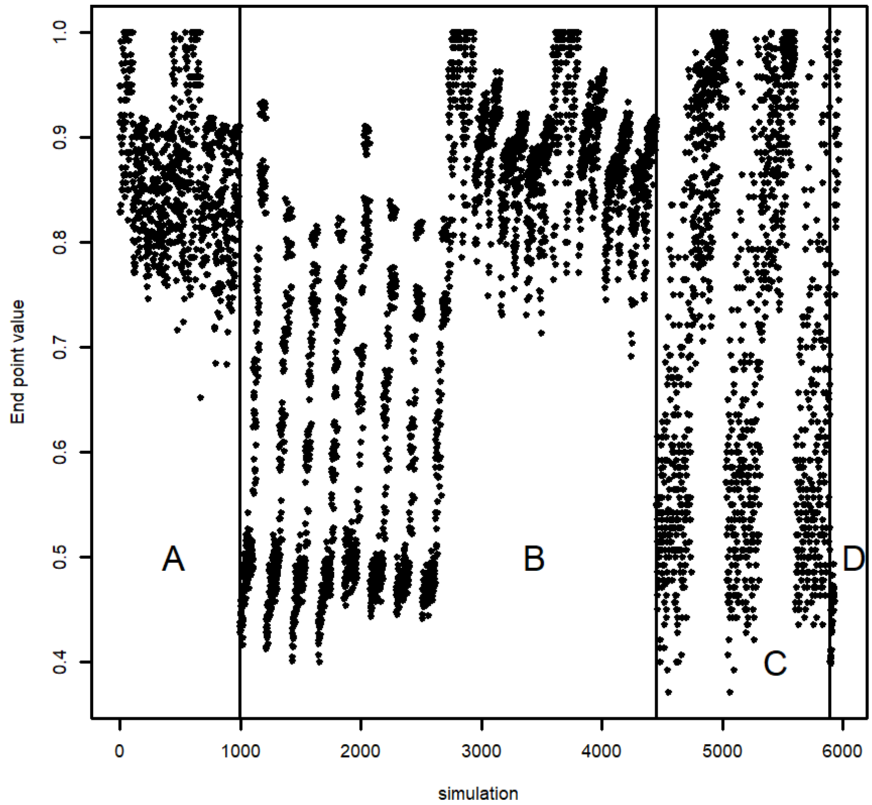


Figure 8.1: End points values

Picture 8.1 gives an overview of the adherence values for 5964 simulations. On the x-axis is reported the simulation as incremental index and on the y-axis the adherence value. The figure is split in four regions demarcating four groups of simulations. The leftmost part (A) shows the output of simulations where only the self search channel of communication is active: individuals could only interact in this way. The central region B, shows the results when propinquity is considered, region C shows the results when also formal channel is added and region D shows results of informal channel inclusion.

The simulations exhibit remarkable different values either within and between the four main groups. This evidence creates the foundation for the whole elaboration that is based on the following considerations. First, the differences within the groups of simulation demonstrate that even if the communication channel is the same, changing the conditions (i.e. the model

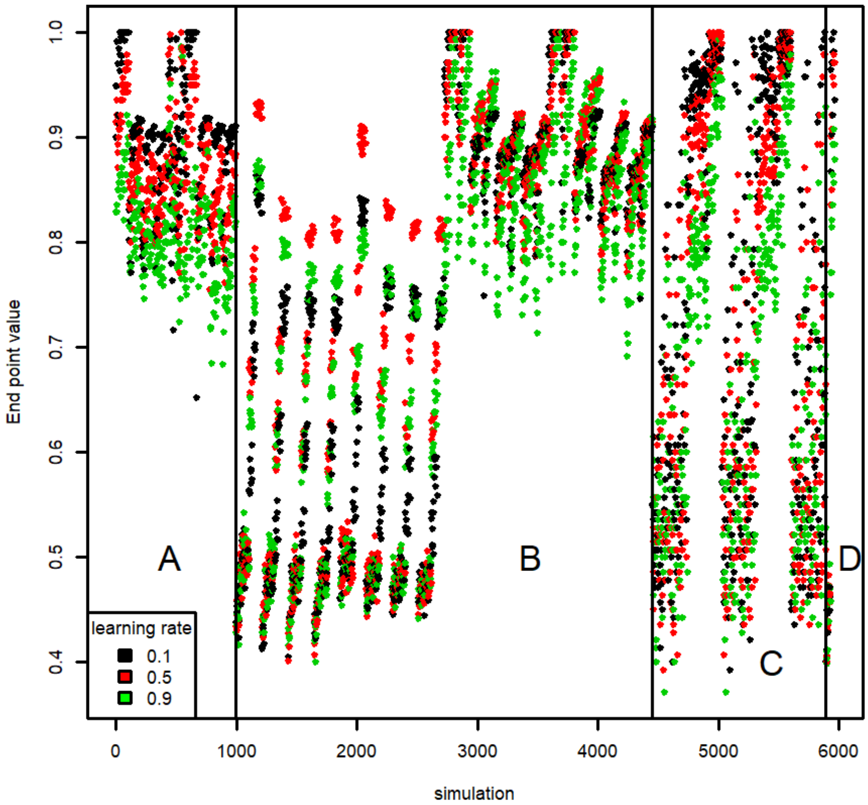


Figure 8.2: End points values coloured by learning rate

parameters values) the output could be extremely different. This effect is particularly evident in the second and third group. Second, the differences between groups demonstrate that, adding or changing the communication channels, the output could be very diverse. Again, this supports the idea that the environment is able to change the behaviour of the system.

Picture 8.2 presents the same output but coloured by learning rate. A black point means that during the simulation all the agents had a learning rate equal to 0.1 (they were all slow learner). If the point is green, all individuals in the simulation were fast learner (learning rate equal to 0.9). The red one considers individuals with a learning rate in the mid point of the possible range (0.5). According to March output, *"slower socialization (lower* p_1[1]*) leads to greater knowledge at equilibrium than does faster socialization"*

[1]this means value equal to 0.1.

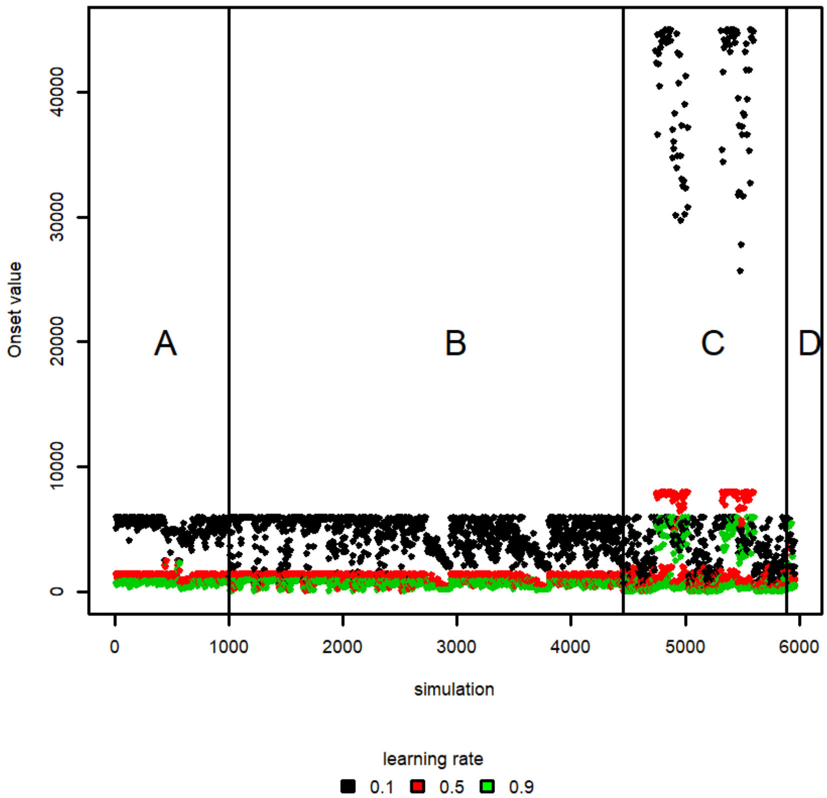


Figure 8.3: Zoom of end points values onset coloured by learning rate

(March 1991, 75). This is almost true in the first group of simulation where self search is the only channel active, black points tend to lie above red points and red points above the green ones.

But looking at the other groups of simulation this consideration does not hold any more. This means that, increasing the complexity of knowledge diffusion, the output is no more linear or predictable and the external environment has a strong effect. Moreover even if the organization is equipped with slow learners, a superior output is not granted.

What seems to remain valid is March's output about the time required by the organization to onset the end point value as figure 8.3 shows. Slow learners have a longer time to reach the equilibrium and consequently black points have higher onset values. Fast learners have shorter time to onset and then green points have lower values.

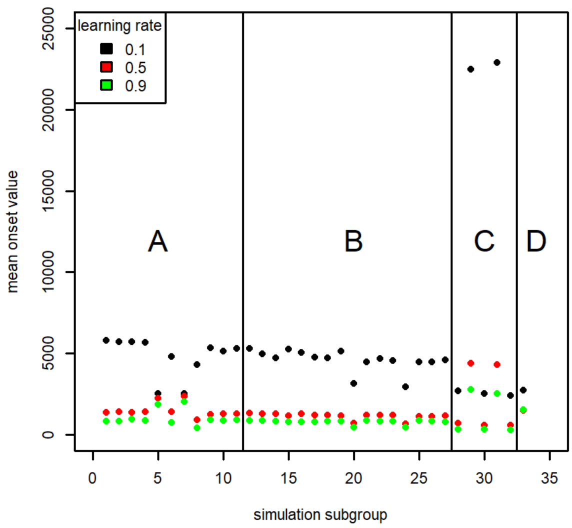


Figure 8.4: average end points values onset for different subgroups

These values could be further analysed. Each group of simulations is actually made by subgroups as reported in table 7.1. The total amount of subgroups is 33 and for each the average values of the onset at different learning rate could be calculated as reported in figure 8.4.

The figure is split into the 4 aforementioned groups. A principal component analysis could better cluster the data in order to have a clearer vision of the situation. Variables are the onset values at the three different learning rates and observations are the mean vales for each subgroup. Picture 8.5 on the left shows the entire score plot considering the first two components, and on the right part shows a zoom of the central group of points.

At least 5 different clusters could be identified: the two central ones (blue and brown circles), the cluster with subgroups 5 and 7 (yellow), the cluster with subgroups 29 and 31 (red) and cluster made by subgroup 33 alone. Blue cluster is considered the main one where most of the simulations landed. The interesting thing is that subgroups 5 and 7 have a largest number of agent respect to all the other[2] (for a reason explained in chapter 9, these simulations were run with 2560 agents). It seems that increasing a lot the agents' number, also the onset valued tend to drift from the main group. This could be correct considering that the organization needs more time to involve all the agents

[2]Typically a group has simulations with number of agents spanning from 32 to 256.

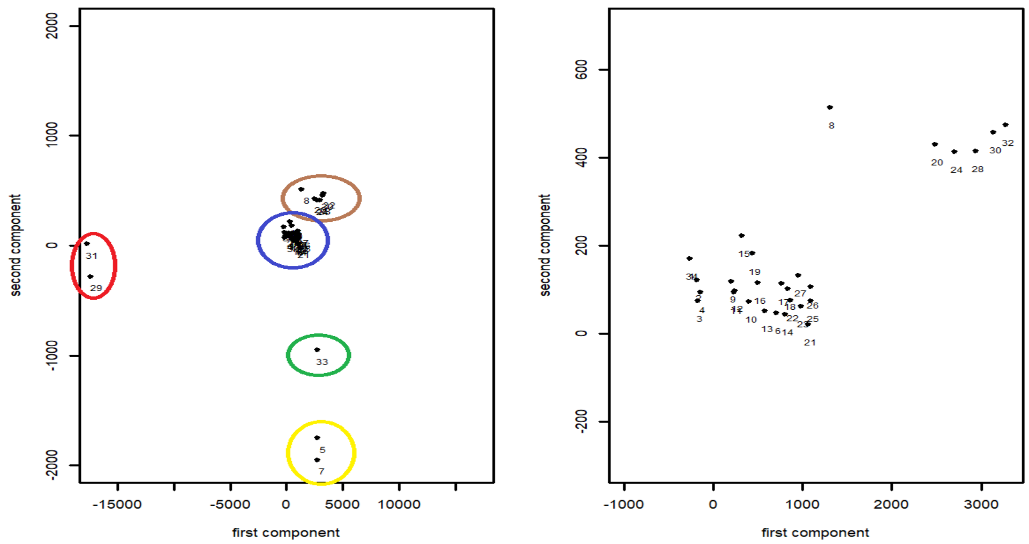


Figure 8.5: PCA on subgroup onset values

toward a consensus. Cluster with subgroups 29 and 31 represents simulations where the department meeting is active. This channel seems to stretch the time needed by the organization to reach a stable point. This is in some way expected with department channel stressing the inter department sharing, delaying the overall sharing. Cluster with subgroup 33 is the simulations dedicated to informal channel. The effect of homogenization is seen in the analysis of subgroup EA (refer to chapter 12).

Changing focus, the brown cluster is made by two type of subgroups: 8, 20, 24 and 28, 30, 32. The first triplet is featured by simulations where there is no interaction with the external world whereas the second triplet is characterized by simulations where project meetings channel is active.

As a final evaluation, figure 8.6 reports in scatterplots the end point value and the end point onset for each single simulation within the simulation groups. A general consideration could be made looking at this figure: the exploration exploitation problem is not merely a problem of learning rate. This is evident from two considerations: fist, the output is extremely different with different channels and second, within the same channel, different values of parameters lead to extremely different results.

Summarising, the environment is of paramount importance in shaping the results; agents and environment create a duality that could not be ignored. The exploration exploitation problem could not be studied solely from the

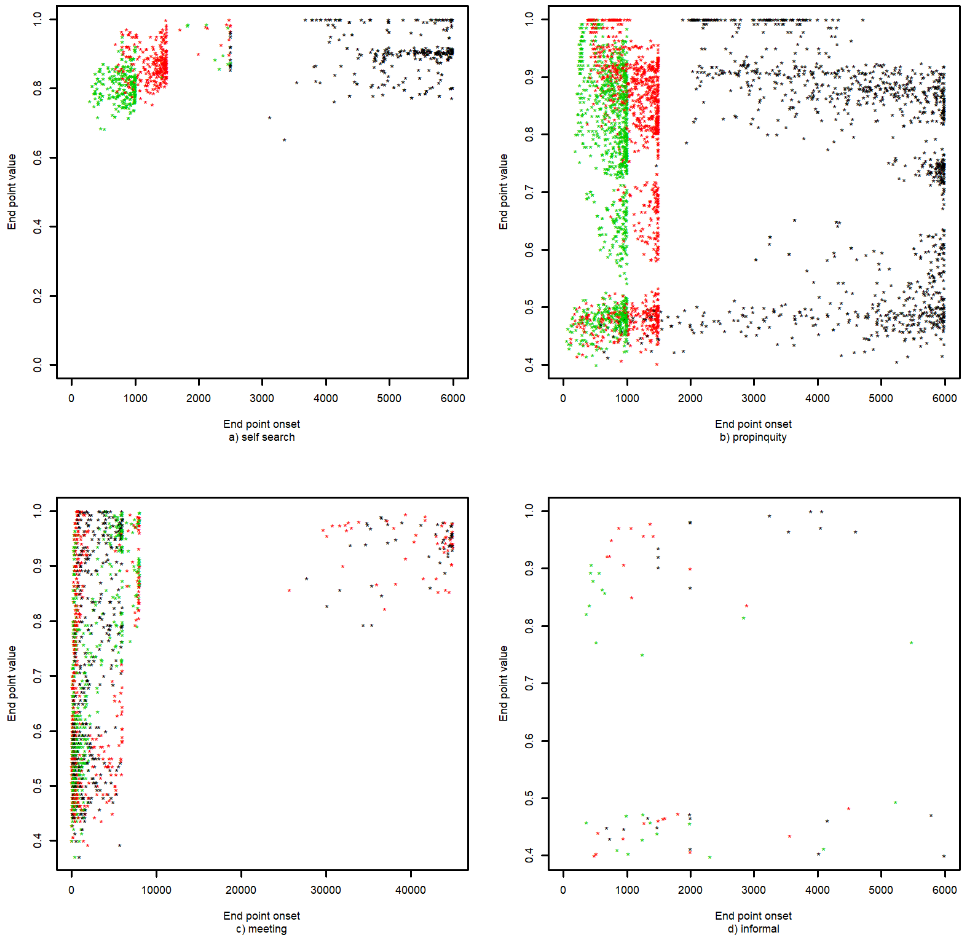


Figure 8.6: End point vs Onset time

agent perspective or focusing on a small portion of the context.

This result is aligned to part of the literature, as for example, Miller et al. supported this perspective writing that *"March's model affirmed learning at both the individual and organizational levels. Hedberg (Hedberg 1981) and Fiol and Lyles (Fiol and Lyles 1985) had argued earlier that organizational learning could not be merely reduced to learning at the individual level"* (Miller, Zhao, and Calantone 2006, 710-711).

With this first important result it is possible to deep dive into self search data analysis.

Chapter 9

Pathfinders

This chapter is dedicated to the analysis of the simulations of blocks A and B. In these blocks the self search channel is studied in detail. This means that in the first group of simulations, meeting, propinquity and friendship channels are switched off and only self search is active.

In these simulations, learning rate, number of agents, knowledge distortion, stack level and interaction with external world are taken into account.

Self search is tested in the two main configurations: closed, where no interaction with external world is considered, and open, where interaction is considered and modulated at three different levels. The three levels mimics low, medium and high external interaction (corresponding to 2, 10 and 20 knowledge dimensions). This four combinations creates the fundamental simulation package.

Every external interaction package is, in turn, simulated at different number of agents.

In extant literature, the number of agent is not a precise point of study[1] considering that in many papers all the agents could only interact with organizational code and the code with a mean of super agents. In the present work the effect of the number of agent is deeper investigated. One of the main reason is that the model does not start from a pre-ordered network configuration hence the number of agents and the structure of the connections became extremely relevant in the analysis. Network dimension is modelled through the number of agents which play in the simulation and four different network sizes are considered: 32, 64, 128 and 256 agents.

[1] In March model the number of agents was fixed at 50 for all the simulations and never changed.

At this point, to avoid possible artefact of the simulations, every subsequent combination is run three times with three different seeds. Hence, for every level of number of agents parameter, the sub-packages are replicated three times.

The following level of simulations grouping is the stack level. As explained in section 5.1, self search is based on agents' ability to find proper colleagues to gather knowledge from. The agent has memory of previous relations with other colleagues and relative level of goodness. The memory effect is modelled through the stack dimensions that is the number of agents an agent could remember. Again, three levels of memory are considered: low, medium and high. The levels are set as following: high level means that an agent could potentially remember the interactions with all the other agents, medium with half of them and low with few of them (5). Then, at high level, the memory effect is huge because no agent is discarded from the stack to be replaced by other agents. In medium and low levels the stack capacity is limited hence agents are inserted and removed according to the performance of the interaction.

The last level of simulation differentiation is the learning rate. Learning rate is studied at three levels: 0.1, 0.5 and 0.9: that is slow learner (0.1), fast learner (0.9) and a value in the middle (0.5).

Figure 9.1 shows the entire structure of the simulations with explicit reference to the five layers: external interaction level, number of agents, random seed, stack dimension, learning rate.

The experimental path of fig. 9.1 is then replicated considering also distortion in exchanging knowledge. In this group of simulations, the two involved individuals could be considered culturally distant from a knowledge perspective if they do not belong to the same department. In this case, as proposed Aven et al. (Aven and Zhang 2016), there could be a misunderstanding while knowledge is passed resulting in a distortion between the sender and the receiver. In this simulation 20% of the transferred knowledge is subject to distortion. The value of knowledge distortion is chosen high to verify whether or not the knowledge distortion has a main effect on the whole simulation.

Counting all the combinations, the experiment package consists of 864 simulations.

The rest of the chapter is logically split into two sections: the first arguing about the end point values and the second discussing about the network's topology.

A remarkable output of March's model is the effect of leaning rate on the knowledge level the system reaches at the equilibrium. A possible explanation is reported later in the same paper: *"slower learning allows for greater*

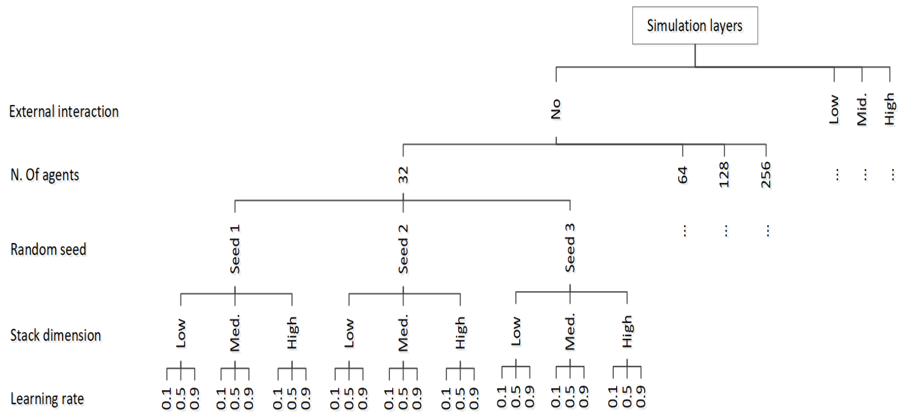


Figure 9.1: Self search experiment structure

exploration of possible alternatives and greater balance in the development of specialized competences" (March 1991, 76).

Figure 9.2 shows all the end points values for the 432 simulations without distortion and figure 9.3 shows all the 432 end points for the simulations with distortion.

This kind of figure is quite complex and worths an explanation. The figure is split into 4 boxes containing simulations from 1 to 27, 28 to 54, 55 to 81 and 82 to 108. The first group contains the simulation made with 32 agents, the second 64, the third 128 and the last one 256. The other main division is by color representing the level of external interaction or unlearning effect. Red lines are related to no external interaction, green lines to low interaction, blue to moderate-medium interaction and purple to high interaction. Within each box related to a certain number of agents there are three subgroups linked to the level of the memory effect (low, medium and high). Inside there are three-points lines: they represent the three end points at different learning rates (0.1, 0.5 and 0.9). Each three-points line is replicated three times considering the three different random seeds. As consequence, the three points line represents the effect of learning rate on that particular set of simulation parameters.

Continuing with the same approach used in the chapter 8, the end point values considered are the average values of the end points of all the agents at the equilibrium. So, the visualization gives a partial information about the output because standard deviation could complement the scenario. In figure 9.4 the end points of all the simulations are plotted against their standard deviation at the equilibrium. This figure shows the output from different perspective. The leftmost part (graph A and C) reports the output of the

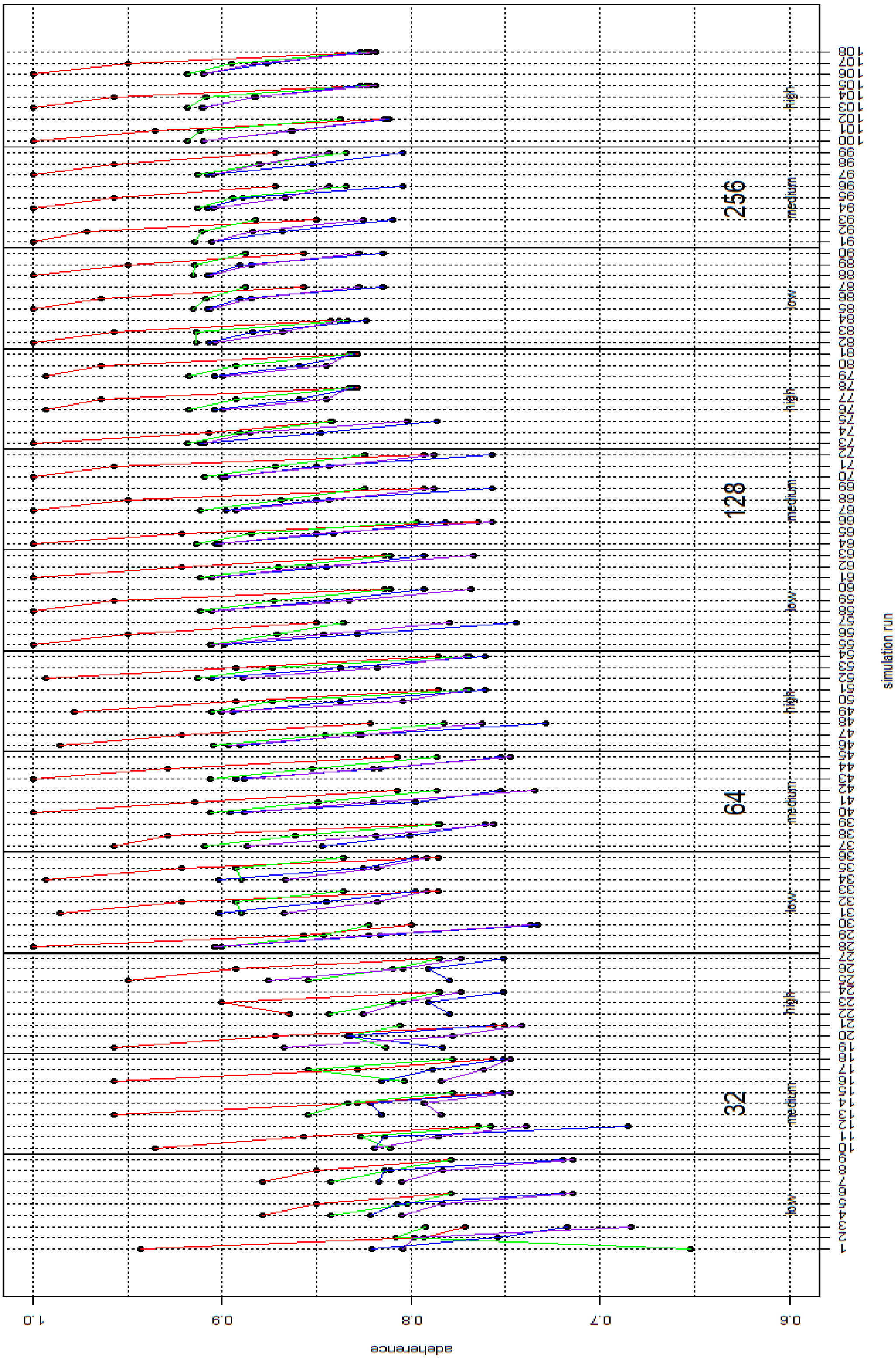


Figure 9.2: Knowledge end points without distortion
(n. dist. dim: red = 0, green = 2, blue = 10, purple = 20)

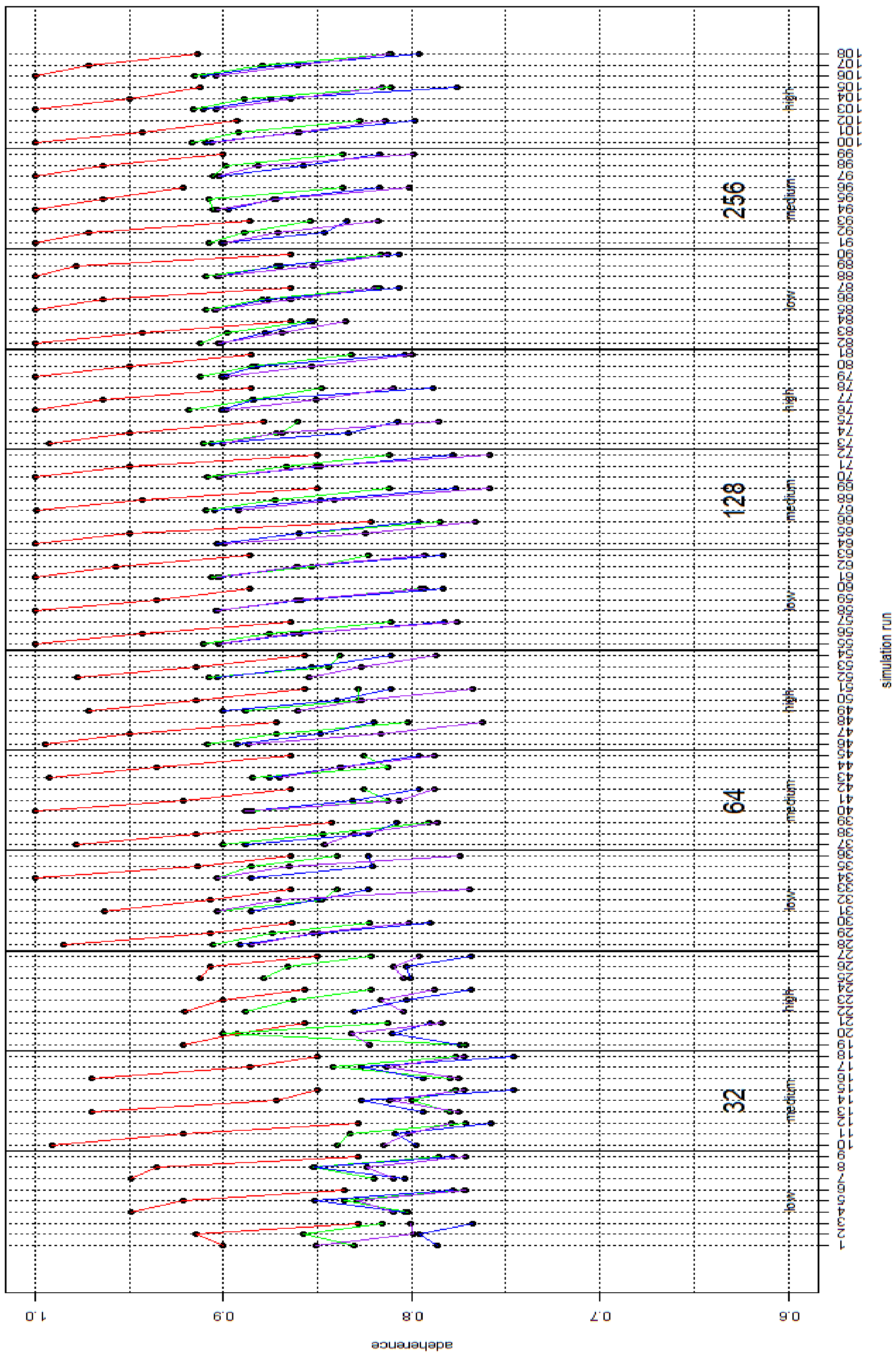


Figure 9.3: Knowledge end points with distortion
(n. dist. dim: red = 0, green = 2, blue = 10, purple = 20)

simulations without the knowledge distortion during knowledge exchange whereas the rightmost pat (graph B and D) shows the output when knowledge distortion is active. The colour scheme used in the upper part is the level of interaction with external world (or unlearning), hence, red is no interaction, green is low interaction, blue is medium and purple is high interaction. The scheme used for the lower part is number of agents: black is for 32 agents, red for 64, green for 128 and blue for 256 agents. This means that, for example, graph C shows the output of the simulations without knowledge distortion coloured by number of agents.

A straightforward result is clearly visible: increasing the external interaction (or unlearning effect), the standard deviation at the equilibrium increases, regardless the presence of knowledge distortion. But, not linearly with the portion of interaction: green points represent low level of interaction, blue medium, and purple high. Medium and high level of interaction tend to have the same variability at the equilibrium. The external interaction (or unlearning effect) seems to reach a saturation point beyond that any increase does not increment more the procured noise.

The lower part of the figure demonstrates that there is not an evident path in the data due to the number of agents. This result seems to come from the environmental conditions rather from the number of individuals present in the organization.

Looking at the charts presented, there are some considerations to pose.

First, March's initial output is generally replicated if the simulation has at least 64 agents[2]. If March's general statement holds, all the three-points lines should be strictly decreasing. In the region of 32-agents simulations, the trend is not always respected (figures 9.2 and 9.3).

Second, on average the adherence level at the end points is higher when external interaction is not involved, when it comes into play, the average level is lower. This would mean that an open world is more noisy and the final output is lower than the one obtainable with a closed world. The same consideration holds with unlearning effect: the higher the cultural distance the lower the knowledge obtained at the equilibrium. But an effect of saturation as seen for standard deviation could also seen here on the level of knowledge. Considering that blue and purple lines in figures 9.2 and 9.3 are overlapped, an increase of 100 % of the impacted dimensions of knowledge categories does not give an equal effect on the decreasing in knowledge adherence. The organization seems able to defend itself against turbulence brought by external interaction or unlearning effect.

Third, the level of external interaction (or unlearning) seems to have a

[2]March simulations had 50 agents.

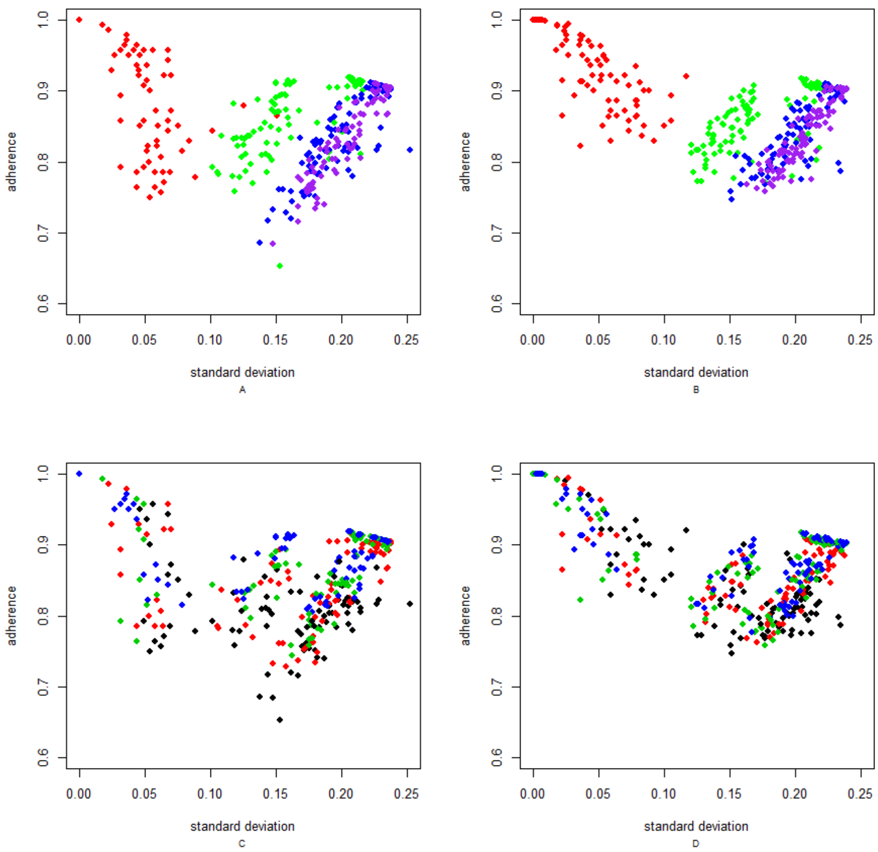


Figure 9.4: Standard deviation vs end point values

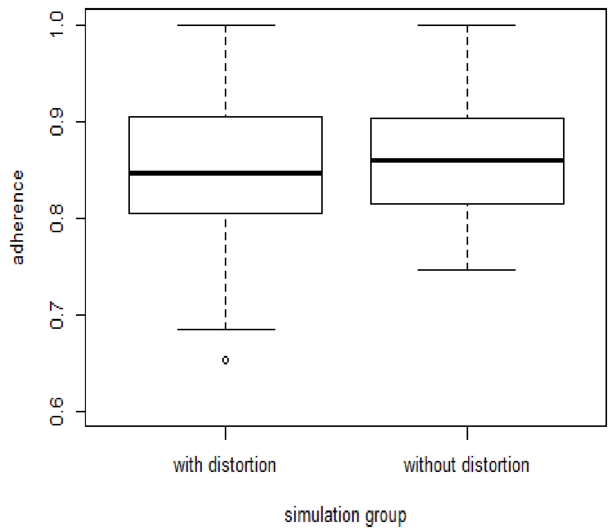


Figure 9.5: Simulation groups comparison

moderate impact on the equilibrium level. The comparison of the output with three levels highlights the good overlapping of the curves.

Fourth, the equilibrium end point seems to depend also on the number of agents present in the network: the higher the number, the higher the end point value.

Fifth, the memory effect seems not to have any impact on the end points values.

Last, all these interesting outputs seems to hold regardless the distortion during the exchange.

Before discussing in detail all the points it would be interesting to have a gross comparison between the simulation with and without distortion. Boxplot (figure 9.5) reports the comparison for the whole end points for the two simulation groups. There are no evident differences, or better, the presence or absence of distortion has no macro effect on the output of the simulation.

This result deserves a bit more discussion. As the simulation architecture is symmetric, it is possible to evaluate the net contribution of knowledge distortion for every combination. This because a simulation is run with and without distortion, *ceteris paribus*. Picture 9.6 reports the difference among each pairs of simulations (with and without distortion). What is evident are the values all close to 0 with a boxplot median almost completely overlapped

Table 9.1: Differences between simulations

Index	External Interaction (or unlearning)	Minimum	Median	Mean	Maximum
End value	No	-0.121	-0.018	-0.027	0.043
	Low	-0.178	0.005	0.000	0.075
	Medium	-0.091	-0.002	-0.009	0.051
	High	-0.117	0.001	-0.004	0.072
	All together	-0.178	0.000	-0.010	0.075
Onset point	No	-2284	-544	-789	-113
	Low	-1907	-17	-158	1007
	Medium	-2148	-106	-240	333
	High	-2082	-37	-159	516
	All together	-2284	-179	-338	1007

to 0.

A quantitative analysis is presented in table 9.1 where the differences are summarised clustering the simulation by the degree of external interaction of the agents.

Again, there is no evidence of knowledge distortion in act and it is somewhat interesting considering that in the present model 20% of the knowledge is distorted every time two agents have a social distance (that is they do no belong to the same department). Not only does the external interaction cover the knowledge distortion but also this is true when the organization is in the isolated condition.

Of course, the effect of knowledge distortion is visible on the efficiency of the organization since the equilibrium is reached with a time approximately 14% lower when the distortion is not present (refer to table 9.1)). Hence, knowledge distortion is considered negligible for the rest of the analysis.

Coming back to the points listed before, to assess all the hypothesis a linear model has been performed for the two groups of simulations, considering the following structure:

$$y = f(\text{learning rate, stack level, n. of agents, ext. interaction level})$$

Learning rate could take three levels (0.1, 0.5 and 0.9) as stack size (low, medium and high). Number of agents and interaction levels have four values, 32, 64, 128, 256 and no interaction, low, medium and high respectively.

Table 9.2 reports the output of the linear regression (model 1 and 2). Model 1 refers to the simulations without knowledge distortion whereas model 2 refers to the simulations with knowledge distortion. Basically all the aforementioned points are verified: all variables are significant but stack level is not. Figures in appendix A report the diagnostic of the two regression models.

Table 9.2: Regression models 1-4 output

	Dependent variable: end point value			
	(1) no dist.	(2) dist.	(3) no dist.	(4) dist.
learning rate	−0.140 (0.005)	−0.108 (0.005)	−0.140 (0.005)	−0.108 (0.004)
	$p = 0.000^{***}$	$p = 0.000^{***}$	$p = 0.000^{***}$	$p = 0.000^{***}$
stack	0.00002 (0.00003)	−0.00000 (0.00003)		
	$p = 0.585$	$p = 0.929$		
agents	0.0003 (0.00003)	0.0002 (0.00003)	0.0003 (0.00002)	0.0002 (0.00001)
	$p = 0.000^{***}$	$p = 0.000^{***}$	$p = 0.000^{***}$	$p = 0.000^{***}$
ext. dist.	−0.003 (0.0002)	−0.003 (0.0002)		
	$p = 0.000^{***}$	$p = 0.000^{***}$		
distortion			−0.061 (0.004)	−0.084 (0.003)
			$p = 0.000^{***}$	$p = 0.000^{***}$
intercept	0.907 (0.004)	0.919 (0.004)	0.930 (0.005)	0.955 (0.004)
	$p = 0.000^{***}$	$p = 0.000^{***}$	$p = 0.000^{***}$	$p = 0.000^{***}$
Adjusted R^2	0.719	0.638	0.763	0.805
Residual Std. Error	0.036 (df = 427)	0.036 (df = 427)	0.033 (df = 428)	0.027 (df = 428)
F Statistic	277.040^{***} (df = 4; 427) (p = 0.000)	191.036^{***} (df = 4; 427) (p = 0.000)	464.074^{***} (df = 3; 428) (p = 0.000)	594.826^{***} (df = 3; 428) (p = 0.000)

Note: *p<0.1; **p<0.05; ***p<0.01

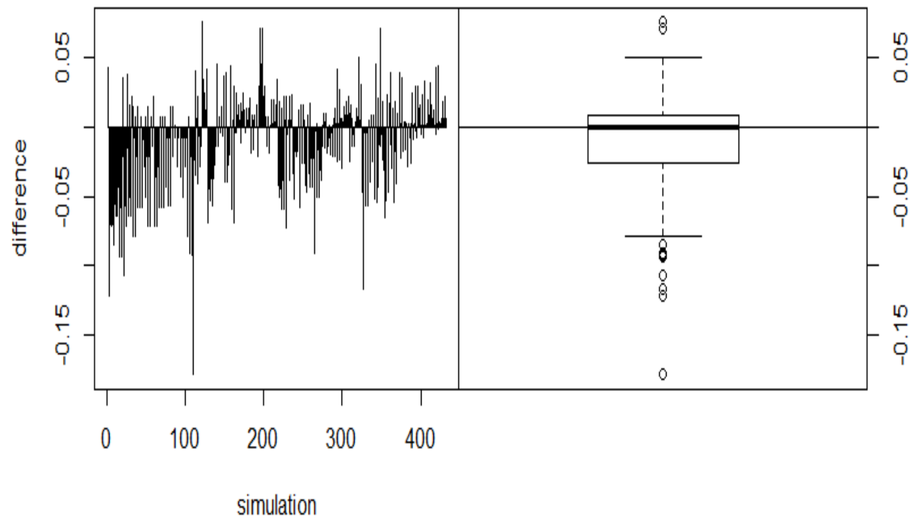


Figure 9.6: Differences between simulations

The model could be further elaborated considering in more detail the contribution of external interaction. Performing a two-way ANOVA ($\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}$) on the end point values with the two variables number of agents and external interaction level it is possible to consider to change the variable external interaction: from a continuous variable to a dichotomous one carrying the presence or absence of interaction (see figure 9.7). The figure reports in the y-axis the adherence and on the x-axis the values clustered by external interaction level and number of agents[3]. From the figure it is possible to appreciate that the behaviour could be split into two groups: the first without external interaction (or unlearning) and the second with external interaction (or unlearning). The external interaction (or unlearning) level seems not able to modulate the final adherence value. The variable is then substituted with a dummy one with value 1 when external interaction is present and 0 when not.

The linear models are then reprocessed, creating models 3 and 4, with the output shown in table 9.2. The the overall model is more significant with an increased adjusted R^2 (from 71.9% to 76.3% and from 63.8% to 80.5%).

[3]The first number reports the level of interaction and the second the number of agents. Hence 20.64 means the cluster of simulations with higher level of interaction and 64 agents.

Diagnostic of models 3 and 4 are reported in figures in appendix A.

From the two models could be evinced that learning rate, number of agents and distortion have an influence on the exploration exploitation balance. But, the role of the number of agents is still not completely explained. This is the focus of the next section.

9.1 Network size and learning rates

As briefly outlined before, the end point values have a dependency on the number of agents present in the network (or in the simulation). To better understand the phenomenon a new chart is presented (figure 9.8) indicating the average end point trends as function of the number of agents.

It is possible to state that the average end points increase with the number of agents with a curve that is not linear. To prove this statement, 4 non-linear models are presented following the preceding division: with and without distortion in the knowledge transfer and then with and without external interaction. The last division is made in the following way: the points with no external interaction create the first model and the three groups with the external interaction (low, medium and high level) are kept together to create the second model. This is justified considering the boxplots in figure 9.7.

Since a non linear dependency is suspected, to built more robust models a fifth point has been calculated simulating the self search with 2560 agents, hence increasing of one order of magnitude the size of the organization. Then, the models are calculated again, according to the following formula:

$$y = \frac{Ax}{B + x} \tag{9.1}$$

where:

$$\text{adherence} = \text{f(n. of agents)}$$

Table 9.3 summaries all the models (models 5-8) and figure 9.9 shows the graphical output. The coloured dotted lines represent the fitted models.

All models are significant and indeed also the dependency from the number of agents is significant. If the trend follows the proposed formula, the average end point sooner or later will reach the maximum value allowed[4]. This result has 3 main implications: i) in bigger organizations the knowledge could be transferred reaching higher values of adherence and, above

[4] That is A, considering that $\lim_{x\to\infty} \frac{Ax}{B+x} = A)$

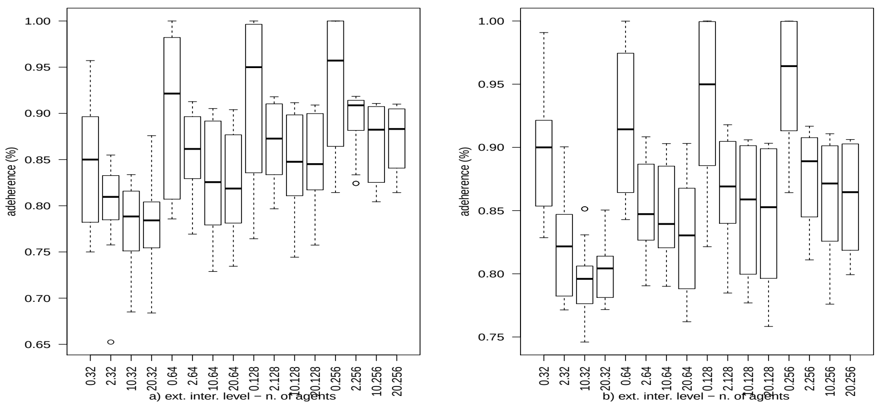


Figure 9.7: Two-way ANOVA boxplot

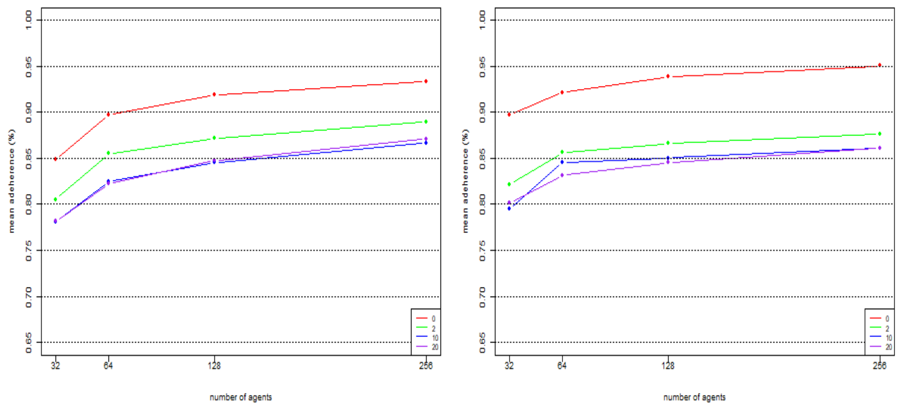


Figure 9.8: Knowledge average end points as function of number of agents

Table 9.3: Regression models 5-8 output

	Parameter	Estimate	Srd. Error	t value	Pr(|t|)	RSE
mod 5	a	0.959807	0.008471	113.308	1.52e-06 ***	0.01231
	b	4.374830	0.649387	6.737	0.00668 **	
mod 6	a	0.895795	0.006736	132.99	9.37e-07 ***	0.009768
	b	4.548499	0.556739	8.170	0.00384 **	
mod 7	a	0.963486	0.005339	180.465	3.75e-07 ***	0.007904
	b	2.550909	0.381057	6.694	0.0068 **	
mod 8	a	0.884566	0.006624	133.530	9.26e-07 ***	0.009739
	b	3.213421	0.528068	6.085	0.00891 **	

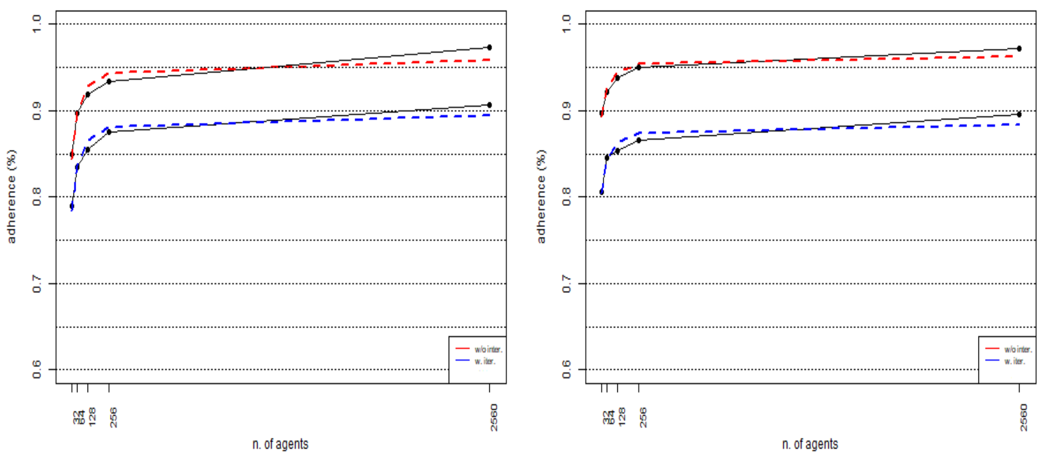


Figure 9.9: Models 5-8 trend

all, ii) increasing the number of agents, the knowledge adherence tends to an asymptote and the number of agents has lower and lower effects and iii) increasing the number of agents, the effect of learning rate decreases. Indeed the organization could not reach knowledge level above 100% and if the mean value tends to a fixed point, necessarily the standard deviation of points at learning rate of 0.1, 0.5 and 0.9 tends to zero. At this point, there is no difference among the different learning rates.

Hence, regardless the learning rate of the agents, a big organization tends to reach the same value of knowledge. And not only does not the organization knowledge depend on the learning rate of agents but beyond a certain size of the organization the knowledge is no more dependant on the size of the organization itself. This result somehow contradicts one of the result March derived from his model. The interesting thing in the proposed model is that with a small number of individuals the dynamics of the system is completely super-imposable with March one. Picture 9.10 gives a zoom of figure 9.2 considering only the case of 64 agents. It is clear that the pattern is always descendent: slow learning (low learning rate) performs better at the equilibrium and fast learning (high learning rate) perform worse. The figure shows that learning rate equal to 0.1 (slow learning) has always end point values grater than learning rate equal to 0.5 (medium learning) and learning rate equal to 0.5 has always values greater than learning rate equal to 0.9 (fast learning).

But, if the end points are intended to reach the maximum value increasing

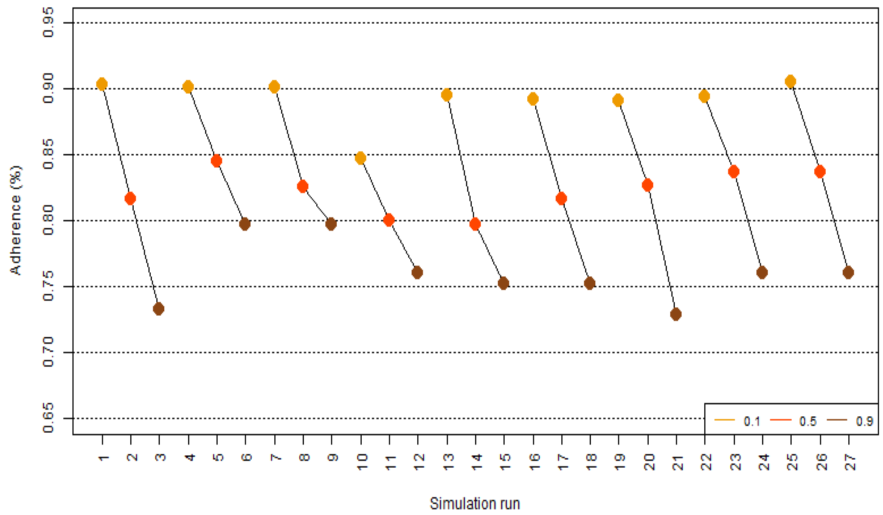


Figure 9.10: End point value as functinon of learning rate

the number of agents, the underlying hypothesis starts to crumble. What emerges from the data is the idea that for small communities individuals have a huge impact on its destiny but, increasing the communities, the effects of the single individuals tend to vanish. There is somewhere the transition from individuals to organisation in shaping the characteristics of the community. This is not the only clue supporting the idea, as explained in the next section.

9.2 Network analysis

The proposed model has the ability to create a snapshot of the connections network on timely basis. This feature allows to monitor how the structure of the network and its properties change over time, or during the evolution of the system. Considering the output of the previous section a dedicated simulation session has been created to deeply understand how network changes with different combination of learning rate, number of agents and level of external interaction.

The first point to discuss is how the network density changes during the evolution and which are its characteristics. Picture 9.11 shows a typical trend of density in the current model. The network snapshot is taken every 25 iterations so a simulation of 5000 iterations gives a 200 points density

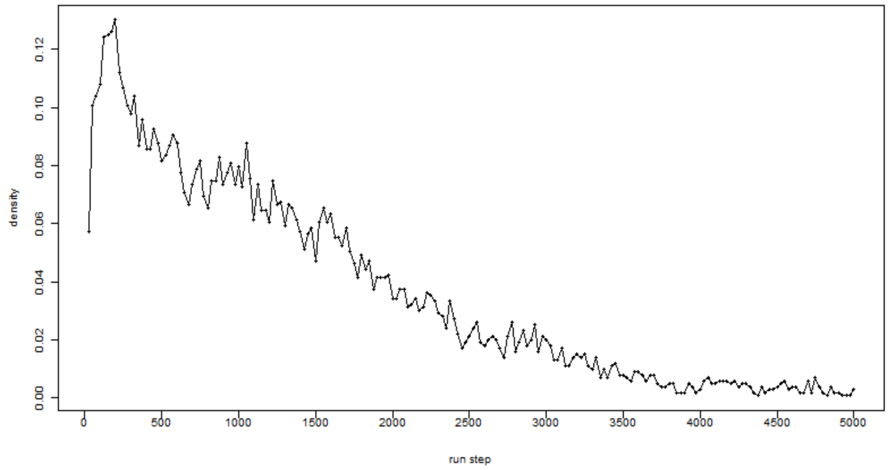


Figure 9.11: Example of network density evolution

plot.

The trend could be explained in this way. At the beginning of the simulation there is no preconstituted network, hence the density is zero. Agents seek for information and then create the connections, the density increases as the adherence to external reality increases. Once the adherence has reached the asymptote, there is no more knowledge to exchange and then the experience of the agents are all negative and the stack is empted. So, the tendency to decrease over time is linked to the structure of the model. But the most interesting part is the first one that is where the network is built and till the asymptote of adherence is reached.

Given the shape of density curve, two are the parameters to investigate: the maximum value and the time required to reach it. Picture 9.12 explores all the 832 maximum values of the simulations. Red points are the simulations without the knowledge distortion whereas the black ones are with knowledge distortion. Knowledge distortion seems not to influence the structure of the network in terms of number of contacts. The densities are almost the same.

Knowledge distortion only slightly delays the maximum attainment as figure 9.13 reports.

Moreover, network density reaches the peak always before the asymptote as figure 9.14 shows. An hypothesis could be that the organization keeps only the needed connection in order to reach the maximum adherence and only for the needed time. A couple of motivations could be found: first,

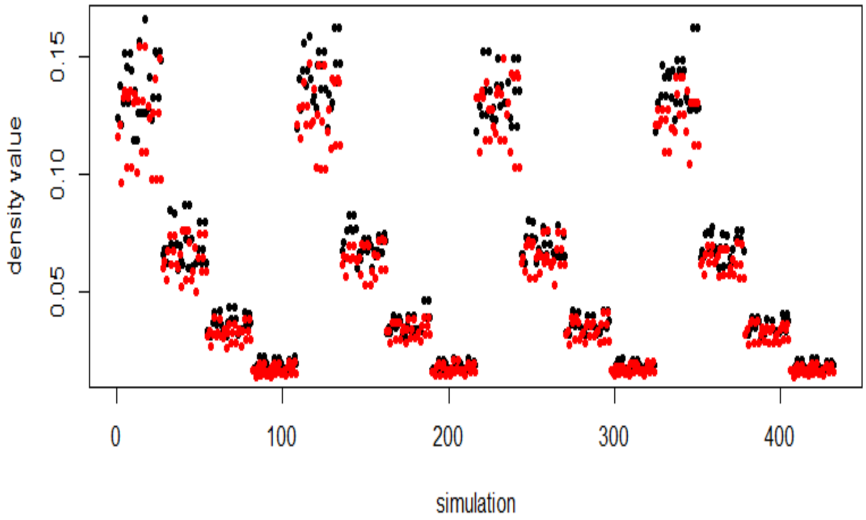


Figure 9.12: Density maximum value for the 832 simulation

the organization has a sort of inertia that is overcome by a dense network but, in a second moment, the network could become sparser once knowledge sharing boosts. Second, when the density maximum is reached, the network rearranges itself in a sort of hierarchical structure based on competence and knowledge and only the most knowledgeable agents are responsible to keep the organization to the adherence end points.

Picture 9.15 reports the complete density curves trend split by agents as alternative view of the cited clustering. The trend of all the curves is the same: all the curves increase at the beginning, reach the peak and then decrease. Of course, all the densities related to simulation with external interaction do not tend to zero since there is always little information to exchange also when the asymptote is reached. Interaction with external world always brings a piece of new information, and since it could be exchanges, the agents' stacks do not empty. Therefore the network density does not reach zero.

There are is a couple of interesting points to stress. First, the values of densities are low considering the possible number of connections. This implies that the stack capacity does not influence the density network, otherwise the density would have been much greater. If for example the level of stack was

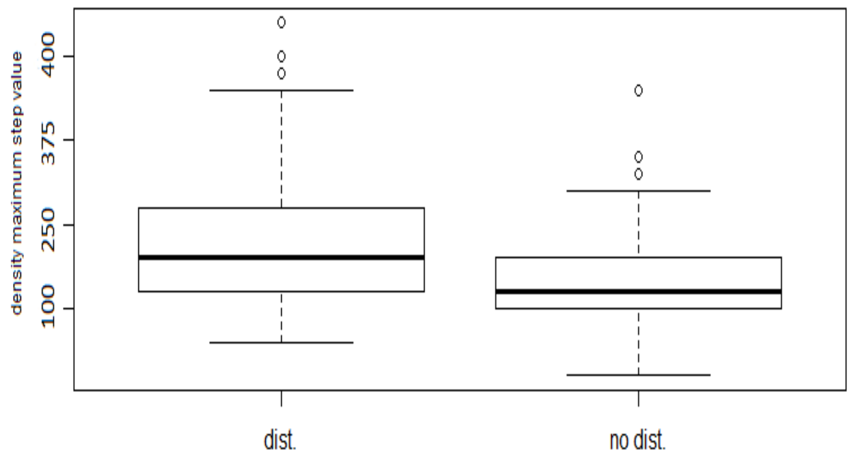


Figure 9.13: Density maximum onset for the 832 simulation

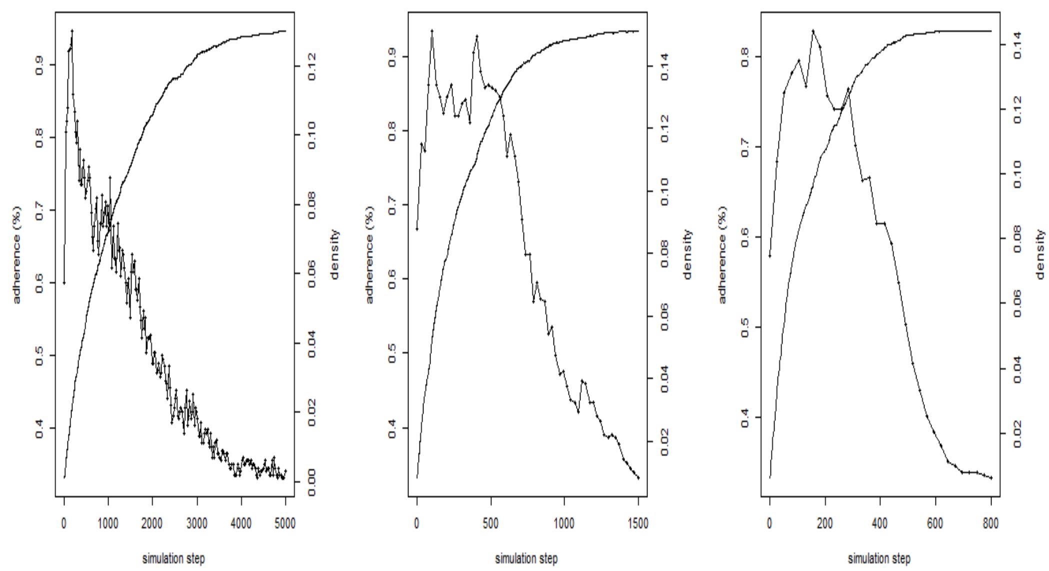


Figure 9.14: Adherence (crescent curve) plotted with density (de-crescent curve)

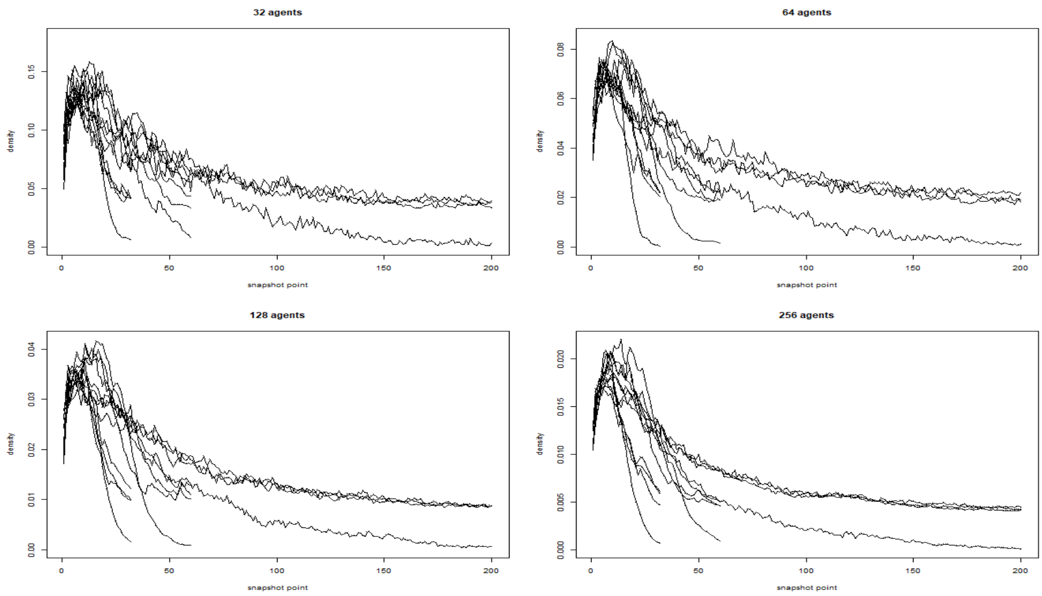


Figure 9.15: All density curves split by number of agents.

half of the number of agents and if the density was 0.5 it could be thought that the limiting factor is the length of agent stack. But since the values are much more smaller than 0.5 the conclusion is that the length of the stack is not a limiting factor.

Second, grouping the values by learning rate and number of agents lead to an interesting point. Picture 9.16 shows two kinds of boxplots. Focusing on the first row of boxplots, the first important point to highlight is that learning rates are not able to separate the density curves. Again, the simulations suggest that even the network structure is not impacted by the learning rate of the agents rather by the community structure. Indeed, looking at the second boxplot, it is clear that the network density seems to be function of the number of agents.

Moreover, the density maximum values seem to respect a low of doubling as the number of agents. If the density median value, when the agents are 32, is multiplied by 32 (the number of agents), the median value at 64 agents by 64, and so on, a pretty constant number is obtained. The complete output is shown in table 9.4 (pag. 156).

This result suggests that there is a certain amount of knowledge in the organization at the beginning of the simulation and a network free to set up finds a suitable configuration to hold it. The hypothesis is that the network finds the minimum necessary density also to minimize the cost of terms of

Table 9.4: Density values multiplied by agents

	10 dimensions		20 dimensions		30 dimensions	
n. of agents	median	result	median	result	median	result
32	0.1290	4.13	0.1421	4.55	0.1572	5.03
64	0.0667	4.27	0.0750	4.80	0.0788	5.05
128	0.0325	4.16	0.0368	4.72	0.0419	5.37
256	0.0180	4.62	0.0189	4.84	0.0207	5.31
2560	0.0019	4.88	0.0020	5.13	0.0023	6.08
mean		4.41		4.81		5.36

energy. To reinforce this result, three groups of simulations are carried out changing the structure of the knowledge. In the first group, the knowledge has 7 topic of 10 dimensions each: this means 70 dimensions in total. The second group has a more complex knowledge since every topic (always 7) has 20 dimension for a total o 140 dimensions. The third group has a heavier knowledge with 30 dimensions for every topic, for a total of 210 dimensions. Table 9.4 (pag. 156) reports the main density peaks for the different number of agents and the result of the multiplication by the number of agents.

What appears clear is that the result of the multiplication is still constant within the group of homogeneous knowledge structure. The median is preferred to the mean since the boxplots are skewed.

If the hypothesis holds there is a gap to bridge. From one side the network arranges itself in a way to display the information in the most proper way and on the other side, agents with their learning rate could or could not take advantage of that. It seems that the organization has an intrinsic property that is the ability to manage the knowledge and an extrinsic property, related to the characteristics of the agents. But this would be true only for small networks and with simple rules since according to linear models 5, 6, 7 and 8 (table 9.3), the ability of the agents to shape the behaviour of the network seems to mild increasing the number of agents.

The natural following step is to check whether or not the underlying network has some specific properties. The impact of knowledge exchange on the network configuration is still a piece of theory to discover as also underlined by Mueller et al.: *"it is astonishing that research on how knowledge diffusion processes affect existing network topologies is still rather scarce"* (Mueller, Bogner, and Buchmann 2017, 614). In their paper, the authors analyse with an AB model *"how and why the degree of distribution within a network can be harmful for network performance"* (Mueller, Bogner, and Buchmann 2017, 614). They substantially evolve a knowledge diffusion model on a Erdös

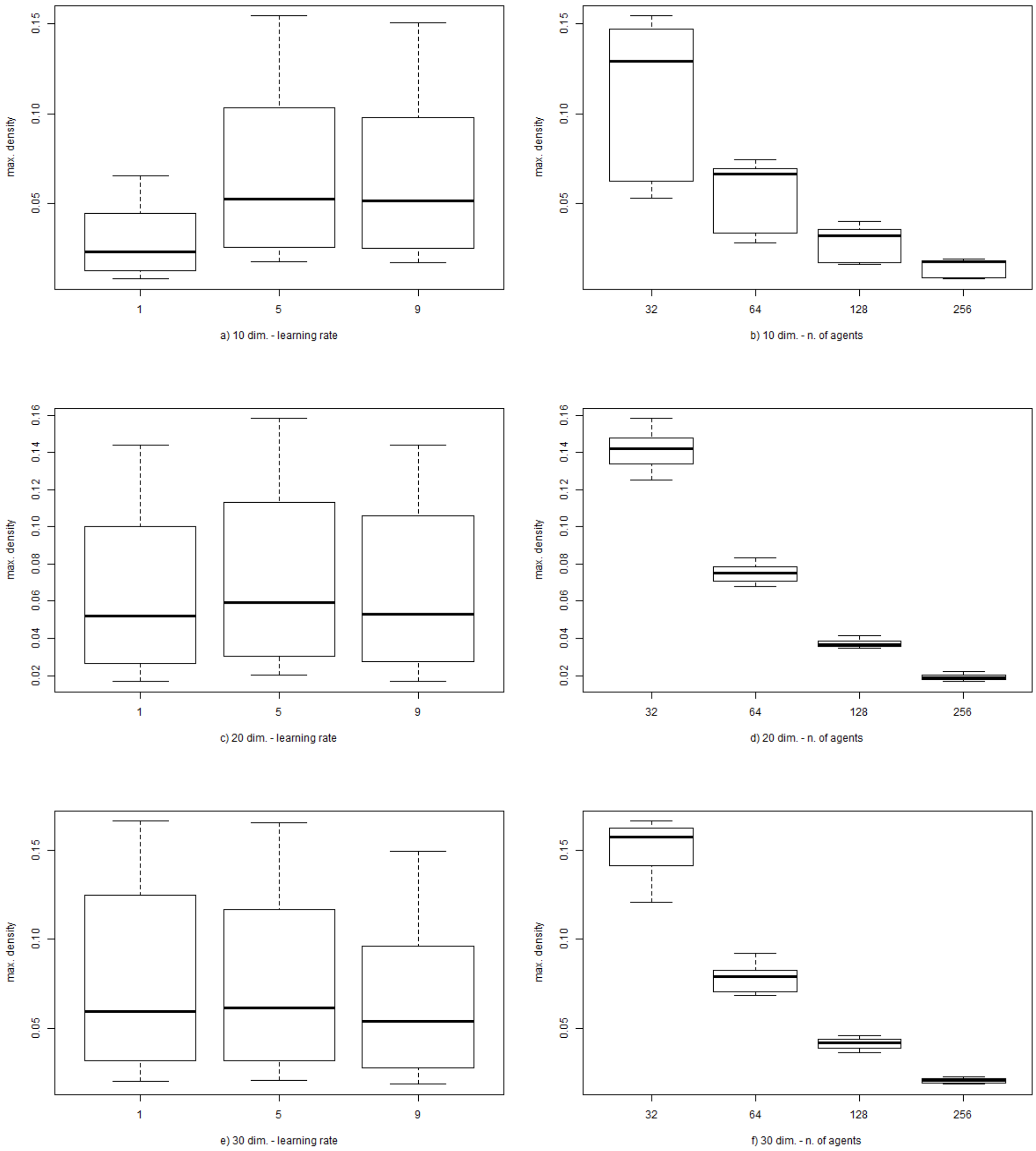


Figure 9.16: Maxima of densities grouped by learning rate and n. of agents (a-b: 10 dimensions, c-d: 20 dimensions, e-f 30: dimensions)

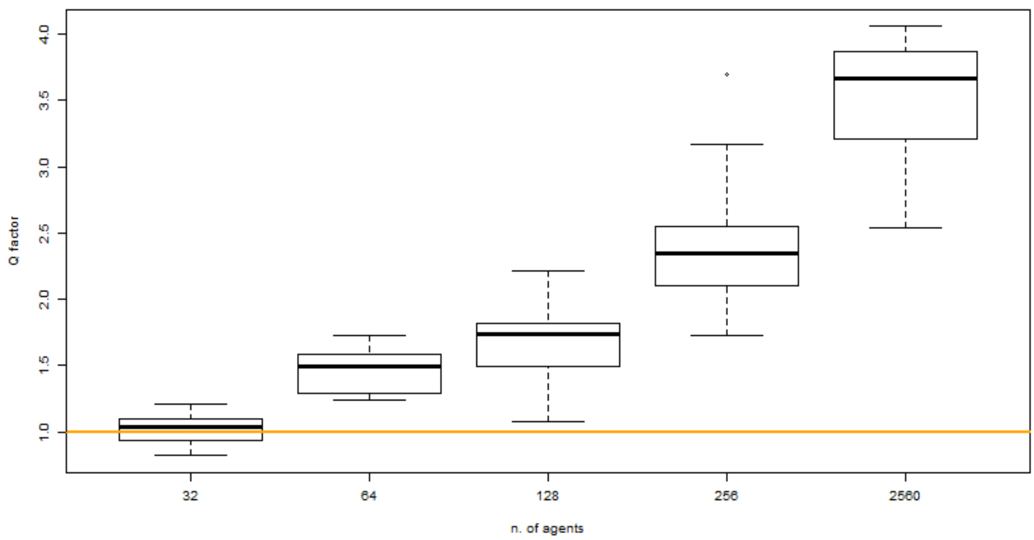


Figure 9.17: Q factors boxplot split by number of agents

and Rényi random network, on a Barabasi and Albert scale free network, on a Watts and Strogatz small world network and on a evolutionary network. The output is that *"Watts Strogatz networks perform best followed by Erdös Rényi networks, Barabasi Albert networks, and evolutionary networks"* (Mueller, Bogner, and Buchmann 2017, 622).

The small world configuration is one of these properties to be firstly assessed. According to Watts and Strogatz (Watts and Strogatz 1998) the Q index is a good indicator of the property of small world. Considering the aforementioned simulation, all the networks at the maximum of density are captured and than studied in their structural properties. There is a surprising output, as show in figure 9.17. The boxplot highlights the huge difference in the Q factors for networks of 32 agents, 64, 128, 256 and 2560 agents.

The bigger the size of the network, the more likely the network presents a small world effect.

As previously said, looking at table 9.4, the values suggest a configuration of energy minimum in transferring energy. Graph theorists have tried to characterise graphs (and then networks) studying their energetic properties. Recalling Gutman (Gutman 1978) it is possible to define the energy E_G of a graph as:

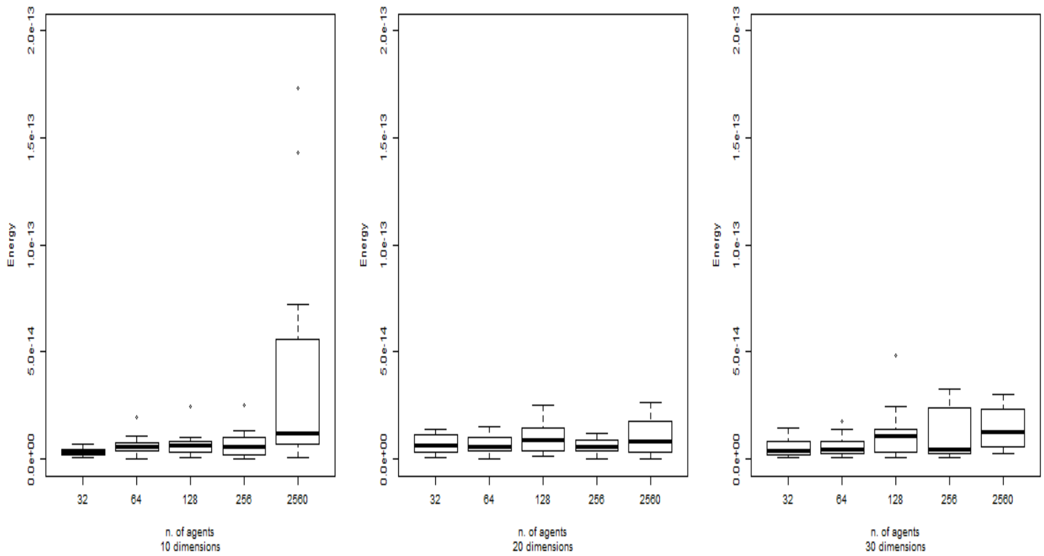


Figure 9.18: Graph energy

$$E_G = \sum_{i=0}^{n} |\lambda_i| \tag{9.2}$$

that is the sum of the module of all the eigenvalues of the graph.

The energy of the networks at the maximum of density for all the simulations is calculated and reported in figure 9.18. The output is pretty clear: all the networks have more or less the same energy regardless the number of nodes (agents) and the regardless the complexity of the structure of the knowledge.

Summarising the main output of this chapter analysis it could be said that there is somewhere a transition from a small network where the contribution and the characteristics of the agents are important toward a big network where the identity of the agents are transparent to it. Everything becomes insignificant and the adherence is crushed against the asymptote. It seems that along the increasing of the size, the networks organise themselves to become independent to whom the actors inside are, minimising the transmission energy and taking efficient configuration (small world).

Chapter 10

Cloisters

This chapter discusses the results of simulations dealing with propinquity channel. According to a wide literature, propinquity has a big effect on the flow of information within the organizations. The presented model could shape the propinquity effect during the simulation in a pretty sophisticated way (refer to section 5.7 and 6.3.14). First of all it is possible to select whether or not agents seek for a better knowledge or seek for knowledge per se. This option could model two different behaviours expressed by the agents. In the first case, when superiority of knowledge is selected, the individual accepts knowledge from the neighbours only when it is better than the owned one. But this is not always the case considering that the neighbour could not be the expert in the field. In this case individuals should discard the information and seek somewhere else. Individuals are not always keen to make the effort to seek the information elsewhere, accepting the one provided. This is exactly the second option in the model: when superior knowledge is not selected, the knowledge transaction takes place regardless the reciprocal levels. This splits the simulation in two different big scenarios according to the type of transaction selected for the propinquity. Indeed those two scenarios are the extreme representations of an organization, in the real world the strategy may be different from individual to individual and even the same individual could behaves differently during the time. In any case, the reality has two boundaries, the first with all the individuals seeking for superior knowledge and the second with all the individuals happy with the knowledge the neighbours provide.

The second macro parameter in the model is the distance of influence of the propinquity action. It is possible to set the boundaries within which the agents could ask for knowledge. Hence if the limit is set equal to 2, the individual could ask for knowledge at maximum to colleagues in the next offices. Set at 6 means a larger freedom to seek, and this could represents the idea of seeking knowledge in the corridors offices.

The simulation architecture is similar to the one used in chapter 9, as reported in figure 10.1.

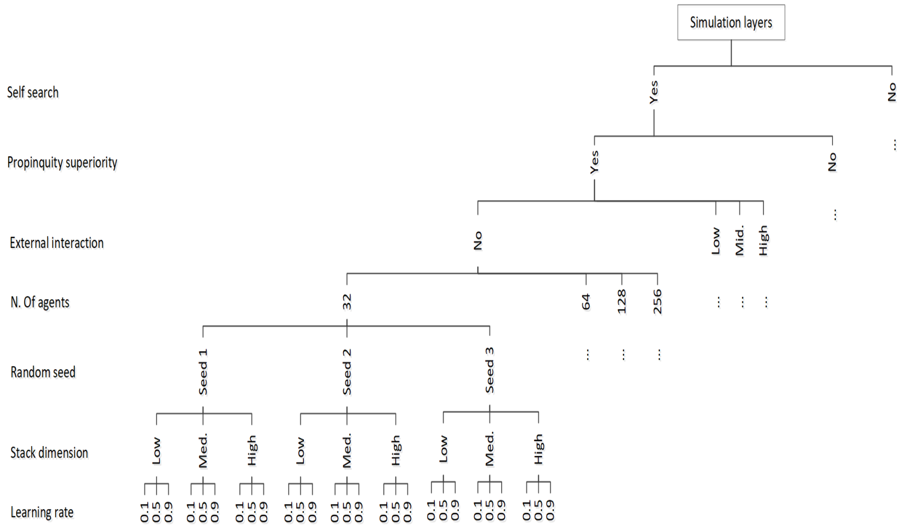


Figure 10.1: Simulation diagram

Table 10.1 reports the high level blocks configuration. Every block named with a letter has the structure as in figure 10.1.

The core of simulations is the same as proposed in chapter 9 and it is replicated considering the different combinations of the presence of self search and knowledge superiority. This means that there are four combinations[1]:

1. knowledge superiority **not** activated and self search **not** activated
2. knowledge superiority **not** activated and self search activated
3. knowledge superiority activated and self search **not** activated
4. knowledge superiority activated and self search activated

Pictures 10.2, 10.3, 10.4 and 10.5 show the end points values in the four different cases.

The discussion of the results could start from figure 10.2, where neither self search and knowledge superiority are activated. There is still a trend in

[1]This four combinations are replicated also for the influence parameter set to wide.

Table 10.1: Propinquity simulations blocks

Block	Sub block	Self search	Propinquity	Short range	Long range	Know. sup.	Ext. dist.
A	$1 \rightarrow 108$		x	x			no
	$108 \rightarrow 216$		x	x		x	no
B	$1 \rightarrow 108$		x	x			low
	$108 \rightarrow 216$		x	x		x	low
C	$1 \rightarrow 108$		x	x			medium
	$108 \rightarrow 216$		x	x		x	medium
D	$1 \rightarrow 108$		x	x			high
	$108 \rightarrow 216$		x	x		x	high
E	$1 \rightarrow 108$		x		x		no
	$108 \rightarrow 216$		x		x	x	no
F	$1 \rightarrow 108$		x		x		low
	$108 \rightarrow 216$		x		x	x	low
G	$1 \rightarrow 108$		x		x		medium
	$108 \rightarrow 216$		x		x	x	medium
H	$1 \rightarrow 108$		x		x		high
	$108 \rightarrow 216$		x		x	x	high
I	$1 \rightarrow 108$	x	x	x			no
	$108 \rightarrow 216$	x	x	x		x	no
J	$1 \rightarrow 108$	x	x	x			low
	$108 \rightarrow 216$	x	x	x		x	low
K	$1 \rightarrow 108$	x	x	x			medium
	$108 \rightarrow 216$	x	x	x		x	medium
L	$1 \rightarrow 108$	x	x	x			high
	$108 \rightarrow 216$	x	x	x		x	high
M	$1 \rightarrow 108$	x	x		x		no
	$108 \rightarrow 216$	x	x		x	x	no
N	$1 \rightarrow 108$	x	x		x		low
	$108 \rightarrow 216$	x	x		x	x	low
O	$1 \rightarrow 108$	x	x		x		medium
	$108 \rightarrow 216$	x	x		x	x	medium
P	$1 \rightarrow 108$	x	x		x		high
	$108 \rightarrow 216$	x	x		x	x	high

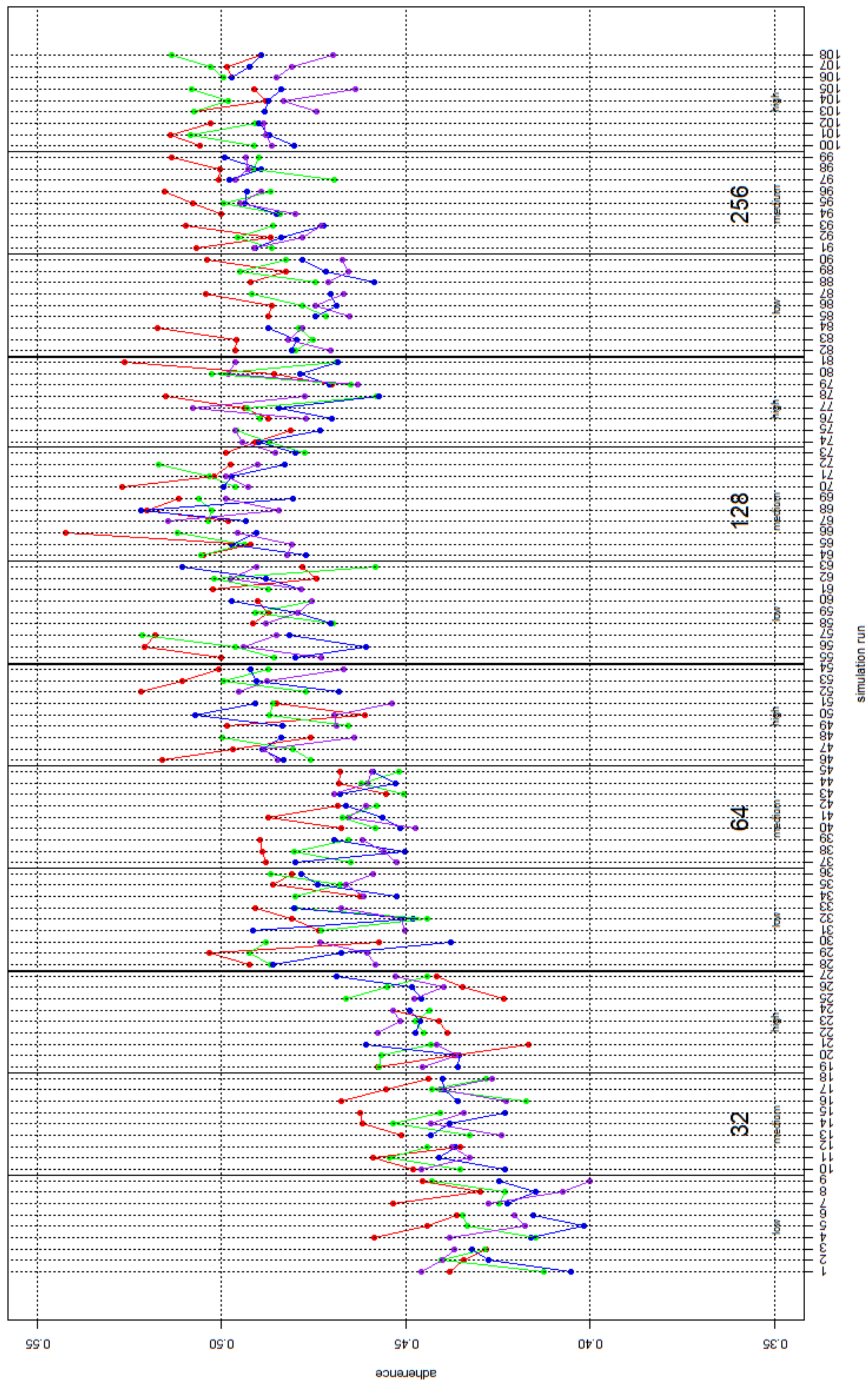


Figure 10.2: End points without knowledge superiority and without self search
(n. dist. dim: red = 0, green = 2, blue = 10, purple = 20)

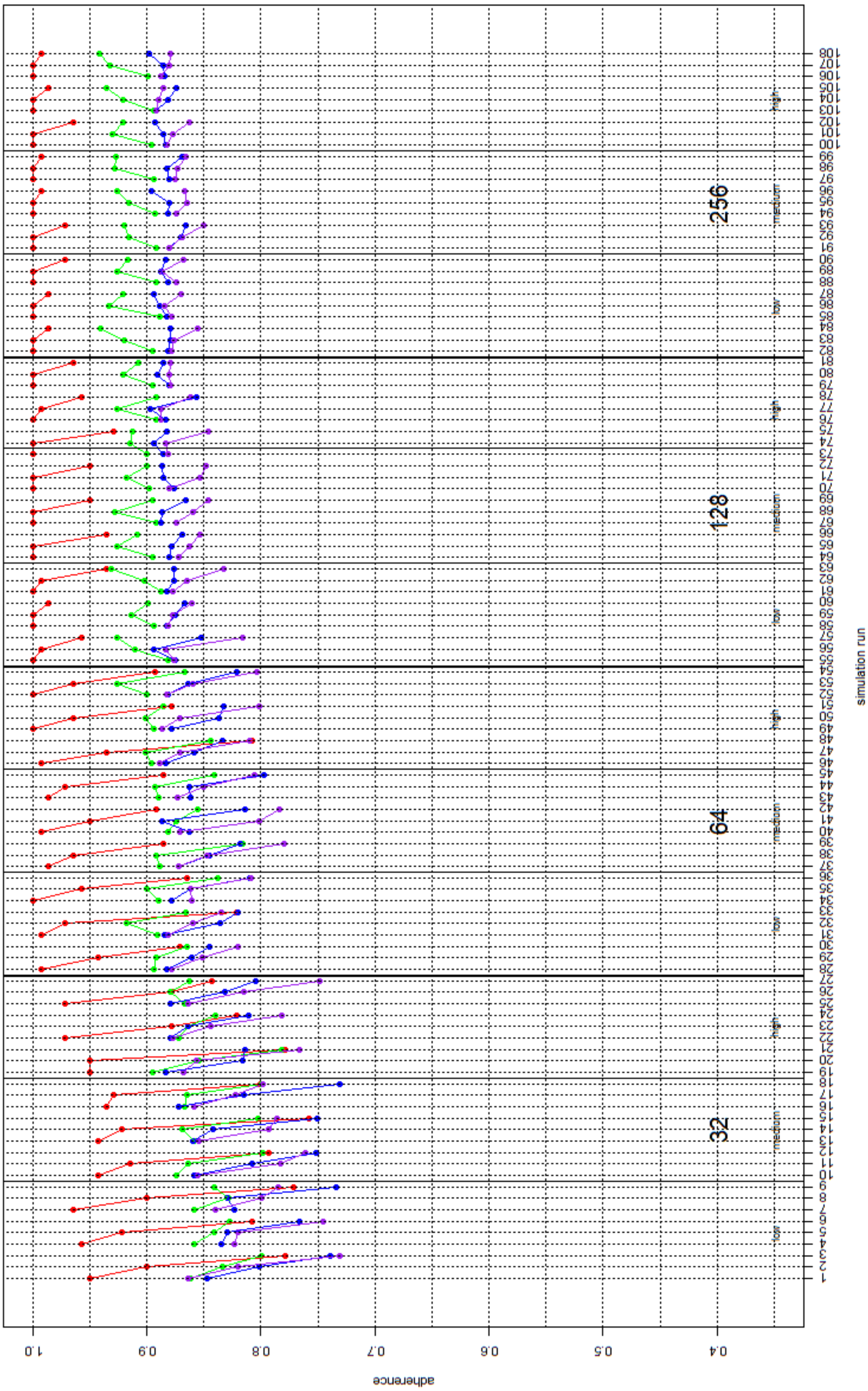


Figure 10.3: End points without knowledge superiority and with self search (n. dist. dim: red = 0, green = 2, blue = 10, purple = 20)

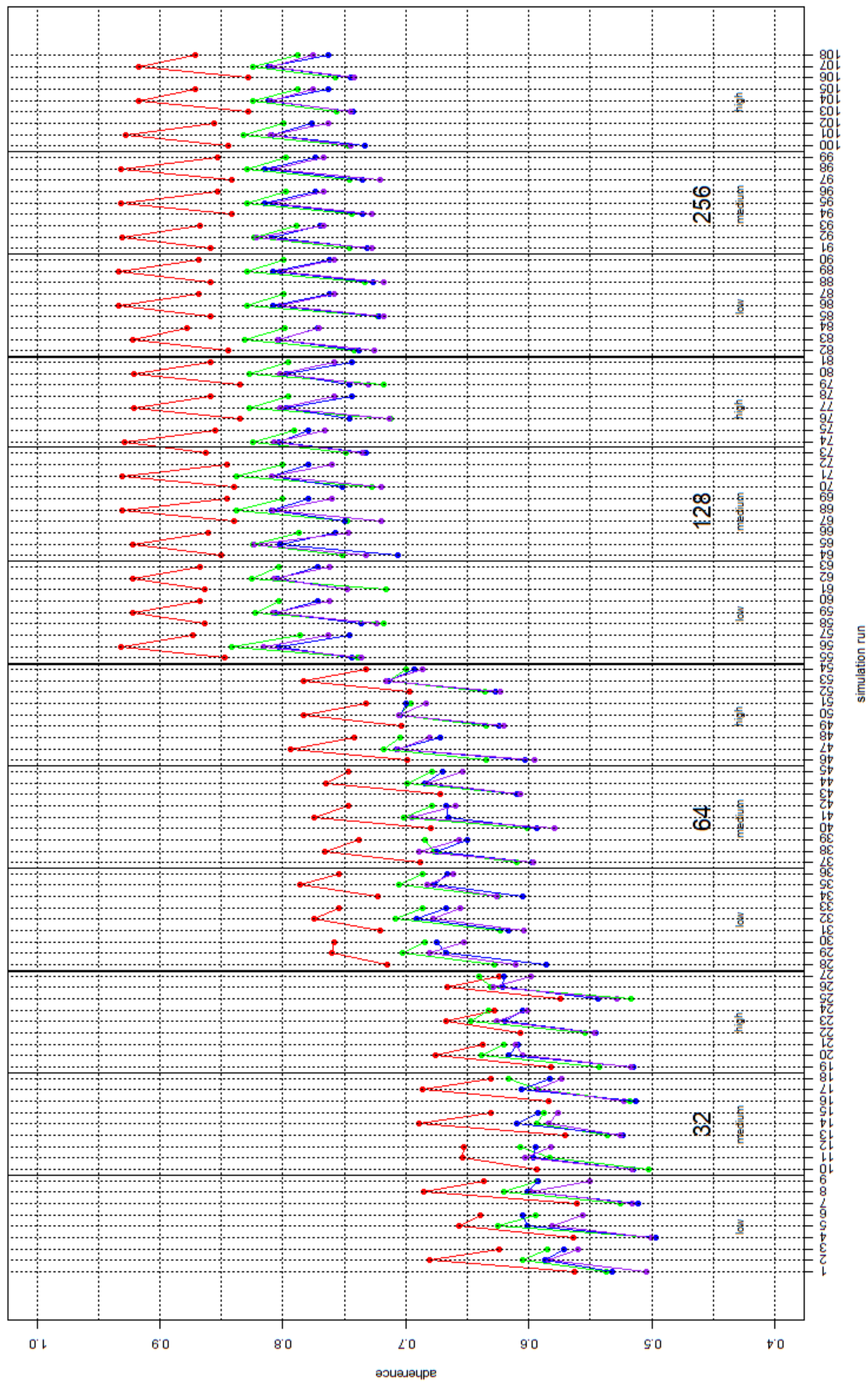


Figure 10.4: End points with knowledge superiority and without self search (n. dist. dim: red = 0, green = 2, blue = 10, purple = 20)

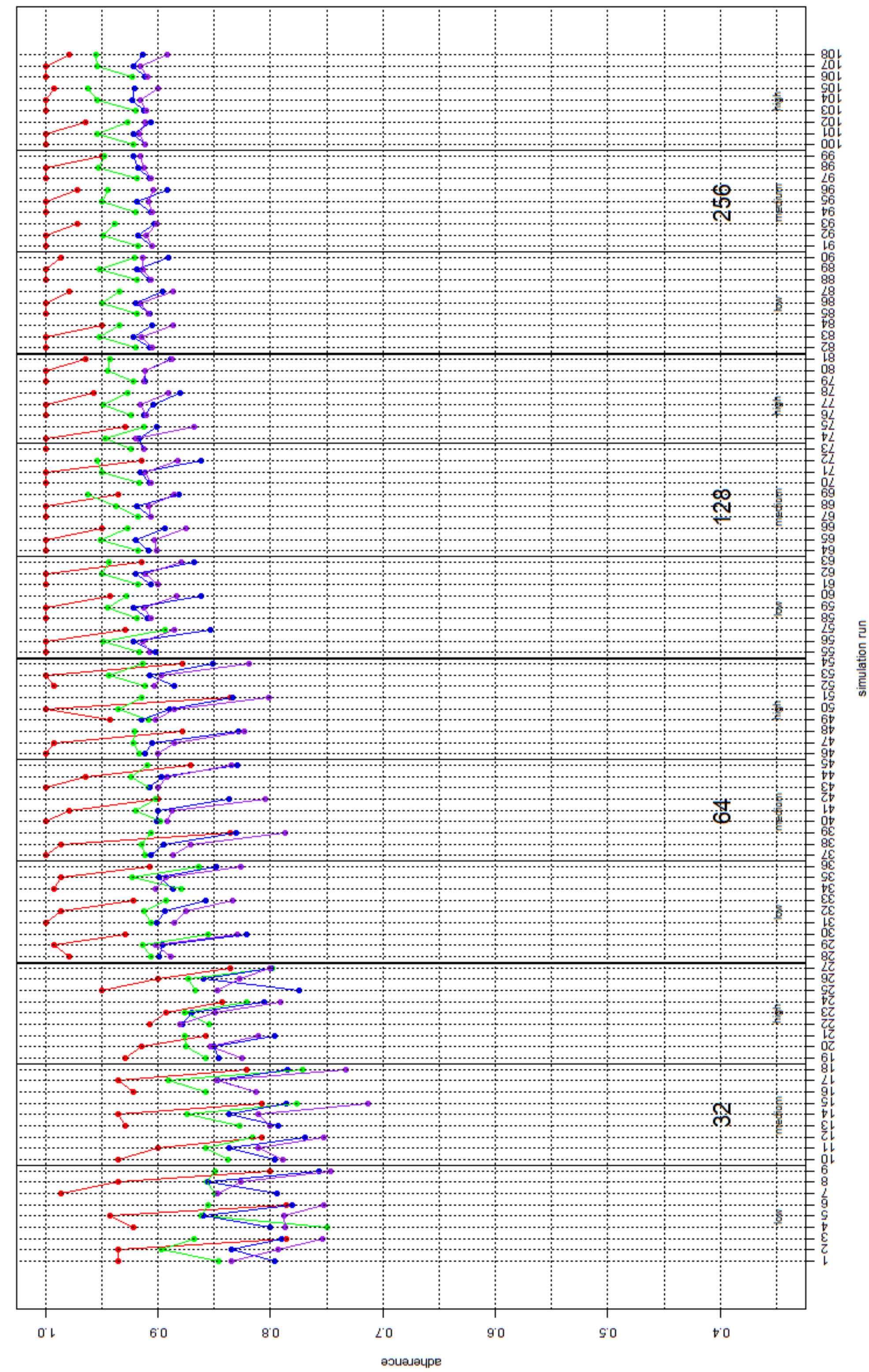


Figure 10.5: End points with knowledge superiority and self search
(n. dist. dim: red = 0, green = 2, blue = 10, purple = 20)

the mean values: organizations with 32 and 64 agents are less performant than organizations with larger number of agents. But, organizations with 128 and 256 agents seem to perform equally. This is somehow related to the degree of isolation agents feel in the simulation. Indeed, in the simulation scenario, the smallest the network the more isolated are the agents. Therefore there is less chance to share knowledge.

Things become more interesting looking at figure 10.3: the addition of propinquity to self search has a healthy effect, the end point values seem to be higher than the ones reported in figure 9.2.

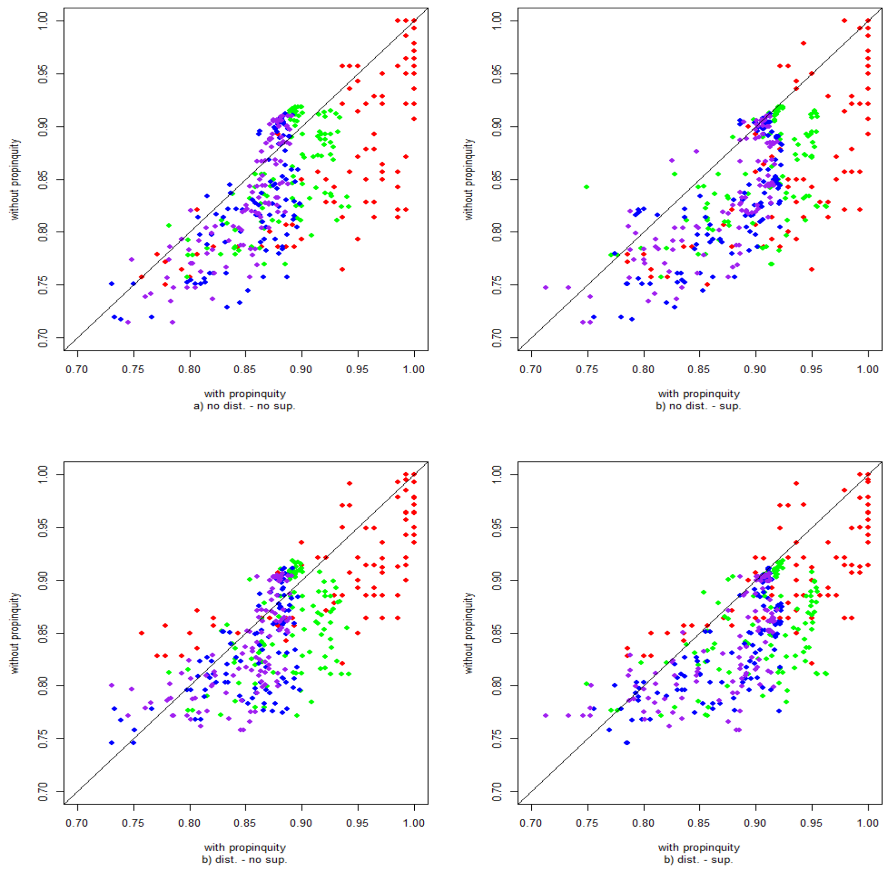


Figure 10.6: End points differences
(n. dist. dim: red = 0, green = 2, blue = 10, purple = 20)

Picture 10.6 shows a scatter plot with the end point values calculated in the four different situations:

a) self search does not have external interaction on, propinquity is activated without knowledge superiority

b) self search does not have external interaction on, propinquity is activated with knowledge superiority

c) self search has external interaction on, propinquity is activated without knowledge superiority

d) self search has external interaction on, propinquity is activated with knowledge superiority

If no effect was visible, all the points should stay on the black line but it is clearly visible that most of the points lay in the lower region, that is the end point values with two channels are higher than the equivalent with only self search on. A most quantitative analysis is reported in table 10.2 where a paired t test is reported (upper part of the table).

All the results are statistically significant, the two channels on average perform better than single channel. This is well evident from figure 10.7 where the activation of the second channel let the end point values increase in a range from 2% to 5%. The figure reports a boxplot of the differences among simulations with and without propinquity channel activated, in the four different combinations.

Considering that the t test is so significant, a further check is needed to clear any doubt about a possible use of overpowered simulations. Secchi and Seri (Secchi and Seri 2017) argue that statistical elaboration of data coming from AMB simulations could hide the trap of type-II error when the number of run is too high. In hypothesis test, a big number of simulations runs could lead to a satisfactory result minimizing type-I error but, consequently, being prone to type-II error. The authors provide an empirical formula to determine the number of run to perform avoiding overpower. For $\alpha = 0.01$ and $\beta = 0.05$ the proposed formula is:

$$N = J \cdot n(J, ES) \sim 10.091 \cdot J^{-0.640} \cdot SE^{-1.986} \tag{10.1}$$

where SE is the side effect and J is the number of configurations. Considering that every t test is performed with 108 data and the run are replicated three times changing the seed number, the number J of configuration is 36. The lower part of table 10.2, shows the estimation of SE, $n(J, SE)$ and then N. The derived numbers are around the numbers of runs made and it is possible to state that the t test is not affected by overpowered simulations.

A further remark is needed here: the increase is not an effect due to the chance of exchanging knowledge. It might be thought that the activation of

Table 10.2: Paired t test

		without knowledge superiority			with knowledge superiority		
	ext lev.	t	p-value	mean diff.	t	p-value	mean diff.
With ext. interaction	0	-10.2730	$\leq$ 2.2e-16	-0.051	-10.7730	$\leq$ 2.2e-16	-0.054
	2	-7.7370	2.994e-12	-0.032	-11.2040	$\leq$ 2.2e-16	-0.051
	10	-7.0761	8.058e-11	-0.028	-12.5310	$\leq$ 2.2e-16	-0.047
	20	-6.1694	6.220e-09	-0.019	-11.2040	$\leq$ 2.2e-16	-0.039
Without ext. interaction	0	-5.6520	6.635e-8	-0.024	-7.0284	1.018e-10	-0.026
	2	-7.8367	1.813e-12	-0.032	-11.5640	$\leq$ 2.2e-16	-0.051
	10	-5.3393	2.640e-07	-0.020	-11.6280	$\leq$ 2.2e-16	-0.038
	20	-4.6098	5.601e-06	-0.015	-9.1275	2.400e-15	-0.034
	ext lev.	SE	$n(J, SE)$	N	SE	$n(J, SE)$	N
With ext. interaction	0	0.628	3.582	$\sim$ 128	0.665	3.197	$\sim$ 115
	2	0.625	3.616	$\sim$ 130	0.996	1.433	$\sim$ 51
	10	0.488	5.911	$\sim$ 212	0.819	2.114	$\sim$ 76
	20	0.339	12.188	$\sim$ 438	0.696	2.920	$\sim$ 105
Without ext. interaction	0	0.440	7.261	$\sim$ 261	0.477	6.185	$\sim$ 222
	2	0.747	2.5538	$\sim$ 91	1.191	1.004	$\sim$ 36
	10	0.431	7.565	$\sim$ 272	0.818	2.119	$\sim$ 76
	20	0.330	12.857	$\sim$ 462	0.745	2.551	$\sim$ 92

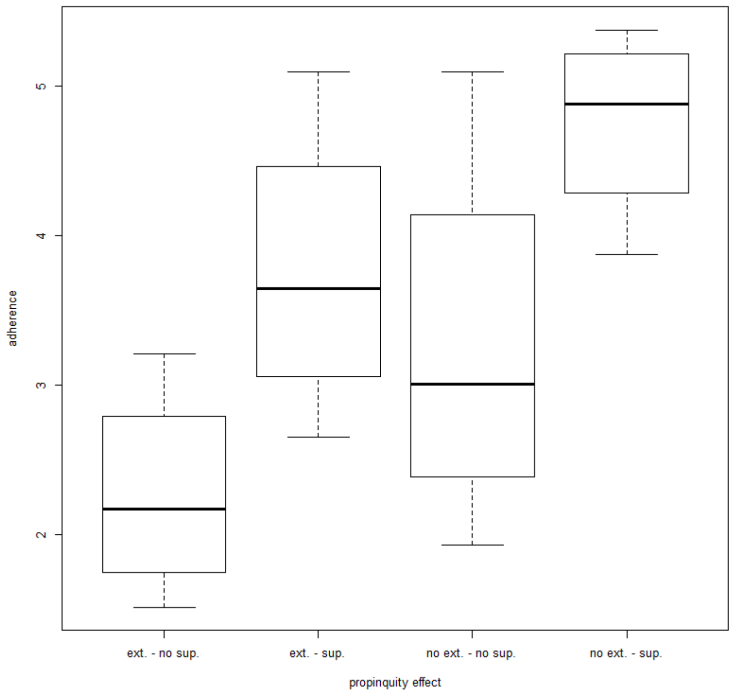


Figure 10.7: Mean differences boxplot

the second channel actually increases the chance agents have to meet each others and to exchange knowledge. The model is designed to prevent this effect. At the beginning of the simulation, the agents' time allocated to autonomous search[2] is 100%, and within this time the selection of the channel is made randomly with equal probability. Therefore, the probability to be assigned to a self search or a neighbourhood channel is the same. More important is the fact that for every simulation cycle, an agent could take one and only one channel. Hence, if an agent is assigned to neighbourhood search, she could not be assigned to other channels in the same simulation cycle. The overall probability to meet remains unvaried and, by consequence, the increase in the end point values is not a probabilistic effect. A first important result highlights the fact that the activation of neighbourhood exchange channel in addition to self search increases on average the performance of the organizations. The end points are on average higher.

[2]Autonomous search includes self search, neighbourhood search and friendship search.

Another interesting consideration could be made looking at end point values figures[3]: the output March found holds as long as the entire organization seeks always for better information. This is particularly evident in figure 10.2 where there is no structure in the output and there is no relationship between the agents learning rate and the end point values, due to the lack of knowledge superiority mechanism. This was an assumption in March model deducible in the model construction: *"the organizational code adapts to the beliefs of those individuals whose beliefs correspond with reality on more dimensions than does the code"* (March 1991, 74). Moreover, another constraint March paved the model on is that all the agents are involved in the mechanism and *"individuals modify their beliefs continuously as a consequence of socialization into the organization and education into its code of belief"* (March 1991, 74). In the simulation of propinquity this is not always true since individuals located in part of the building far from the colleagues have less opportunity to use the neighbour channel to transfer the knowledge. This means that the offices layout necessarily creates an asymmetry of opportunity in transferring the knowledge.

Surprising is figure 10.4 where the self search is not present and superior knowledge is active. In this case, through offices, agents exchange knowledge only if the donor has greater knowledge. The pattern followed by the output presents a regularity: if the agents have a medium learning rate, the output is better. So, slow learners are not able to perform better. The result could be explained by figure 10.8 where all the possible interactions in the 64 agents simulation are reported[4]. Isolated nodes apart that do not contribute to the final increase in the overall knowledge, there are only small groups of agents not greater than 14 nodes. Quite presumably, most of the exchange happens in the two big clusters and in the two smaller ones (4 and 5 nodes). Hence if March's output does not hold it means that under a certain number of individuals the mechanism breaks.

A warning of that effect was already visible also in figures 9.2 and 9.3 where in 32 agents section it visible a divergence from March's trend.

The simulations with propinquity channel suggest three reasons why March output could fail. First, the presence of more than one communication channel changes the way agents interacts and, boosting the output, it confuses the learning rate path. Second, it could fail when the number of agents involved is small.

A dedicated short simulation of self search with 16 agents (subgroup GA)

[3]Picture 10.2, 10.3, 10.4 and 10.5.

[4]Actually, the edges reported in the figure are not all the possible connections for the 64 agent propinquity simulation: these are all the possible ones with propinquity short range activated ($d_{i,j} \leq 2$), as explained in section 5.7.

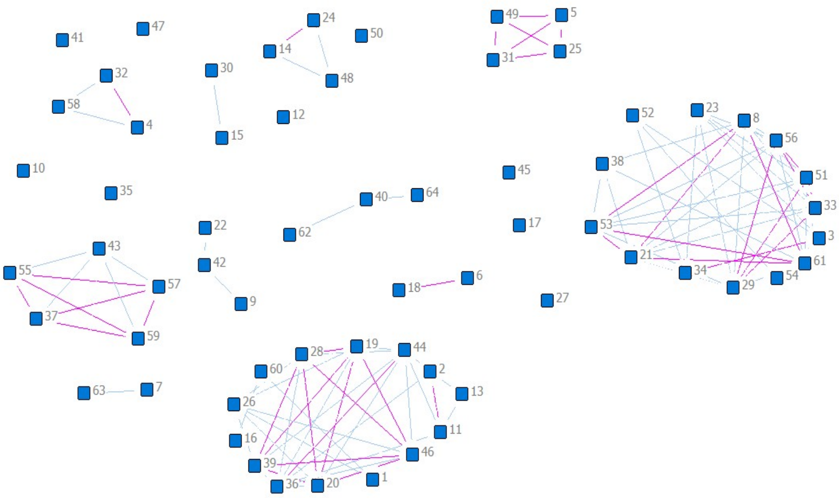


Figure 10.8: Isolation in short range 64 agents simulation

could clarify whether these two hypothesis are reasonable. The output is replicated in figure 10.9 where the trend of three replications are depicted[5].

It is extremely visible the break of the typical pattern March found in his paper. Hence it is not necessary to have a second channel to break the trend, but a small number of agents could. It is possible to conclude that March's effect is limited in a particular range of agents: there is a minimum limit under which the mechanism breaks and there is a maximum limit over which the learning rate effect is negligible (as discussed in previous chapters).

Third when knowledge superiority is not active. There is a profound implication in the last point, knowledge superiority entails that agents could recognize that alter knowledge is better than the owned one. And, since superiority is against external reality, the agent must be able to determine that her representation of external reality is lower than another agent's one. If this ability lacks, agents could exchange knowledge only because is different.

This suggest that the mechanism by which slow learners keep the organization at higher level in March's model lies in the code update step where the best performer are found. Best performers are non other than the agents with superior knowledge since *"at the same time, the organizational code adapts to the beliefs of those individuals whose beliefs correspond with reality on mode dimensions than does the code"* (March 1991, 74).

Hence to replicate the slow learners effect, a mechanism to seek superior

[5]The three replications have three different random seed numbers.

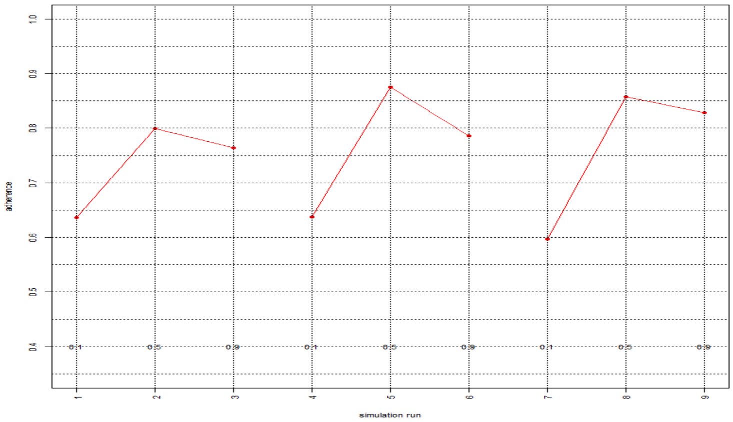


Figure 10.9: End points of simulations with 16 agents

knowledge must be present and it must be strong enough.

10.1 Wider propinquity

Understanding the impact of propinquity on network structure is an important point. As previously touched on, the model permits to modulate the influence of propinquity, allowing agents to search for knowledge in smaller or wider areas near their offices. The same experimental design reported in figure 10.1 has been replicated also with the propinquity width set to the maximum[6]. The comparison is the following step, starting from the figures 10.10, 10.11, 10.12 and 10.13 which capture the end point trends for the simulation with wider propinquity influence.

The influence of long range could be elaborated in a similar way as done before and a scatter plot of long range versus short range is reported in figure 10.14. All values lie on the bisector line showing an equivalent behaviour of the organization in the two scenarios. Only in the subplot A there is a slight spread related to simulations of figure 10.10 with 32 and 64 agents. This is due to the attenuation of the clustering effect obtained with a wider propinquity influence.

Again, a paired t test could be performed between the short range and long range simulations as reported in table 10.3. The table results are pretty evident, there is no clear predominance of long range over short range. The test, where significant, show a minimal difference, around 1% and not always

[6]The parameter is then set to 6.

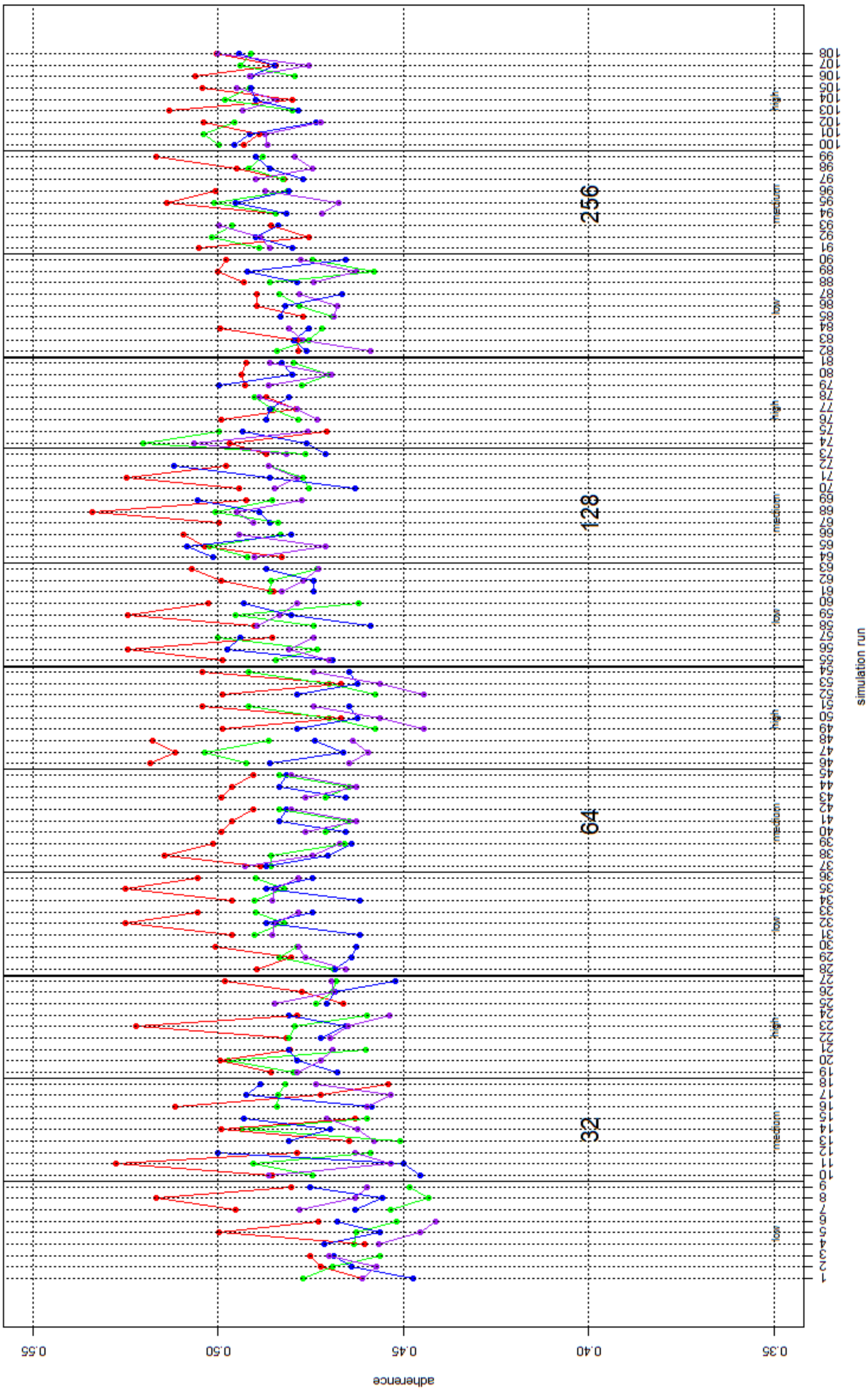


Figure 10.10: End points with long range but without self search and knowledge superiority
(n. dist. dim: red = 0, green = 2, blue = 10, purple = 20)

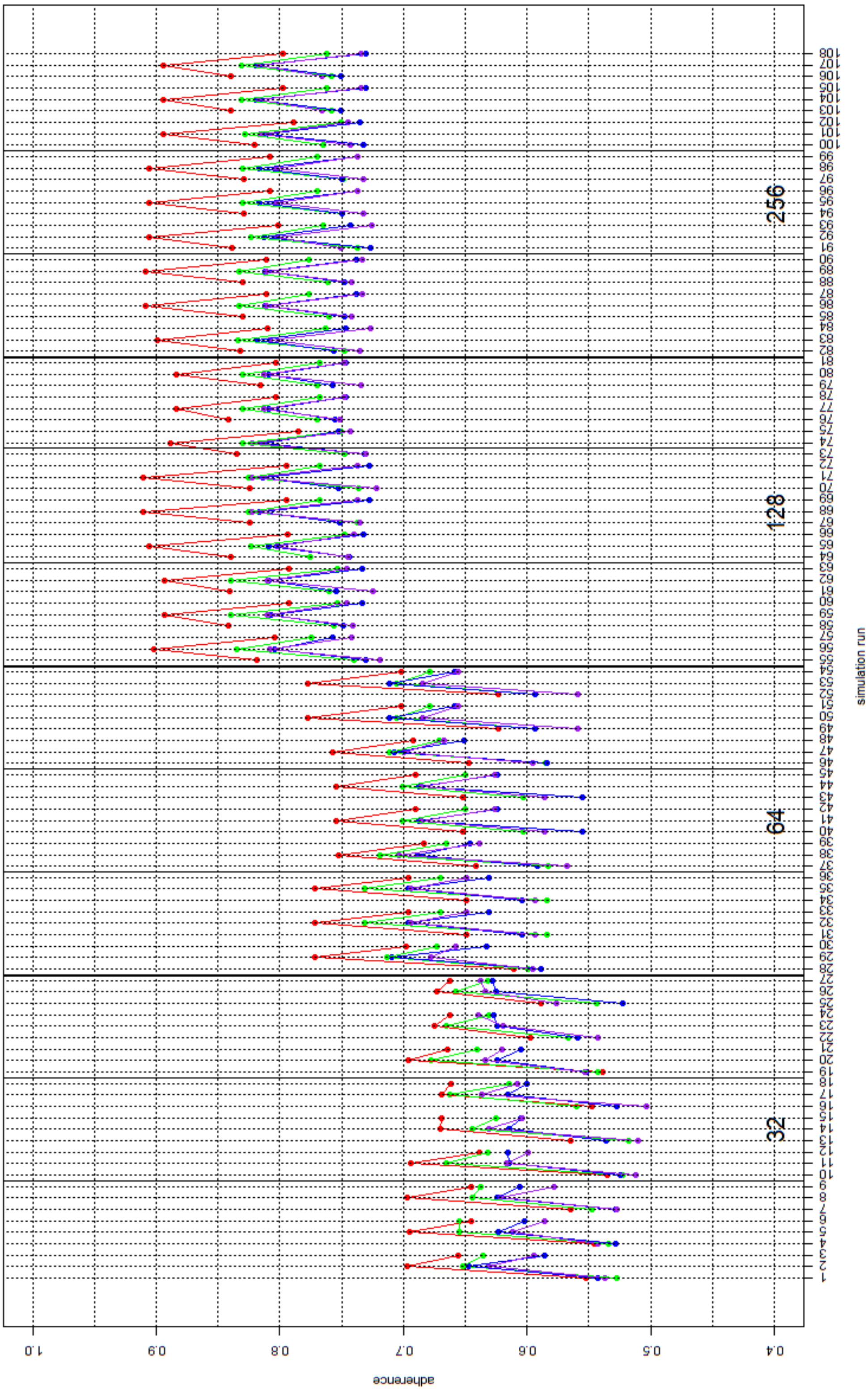


Figure 10.11: End points with long range and knowledge superiority but without self search
(n. dist. dim: red = 0, green = 2, blue = 10, purple = 20)

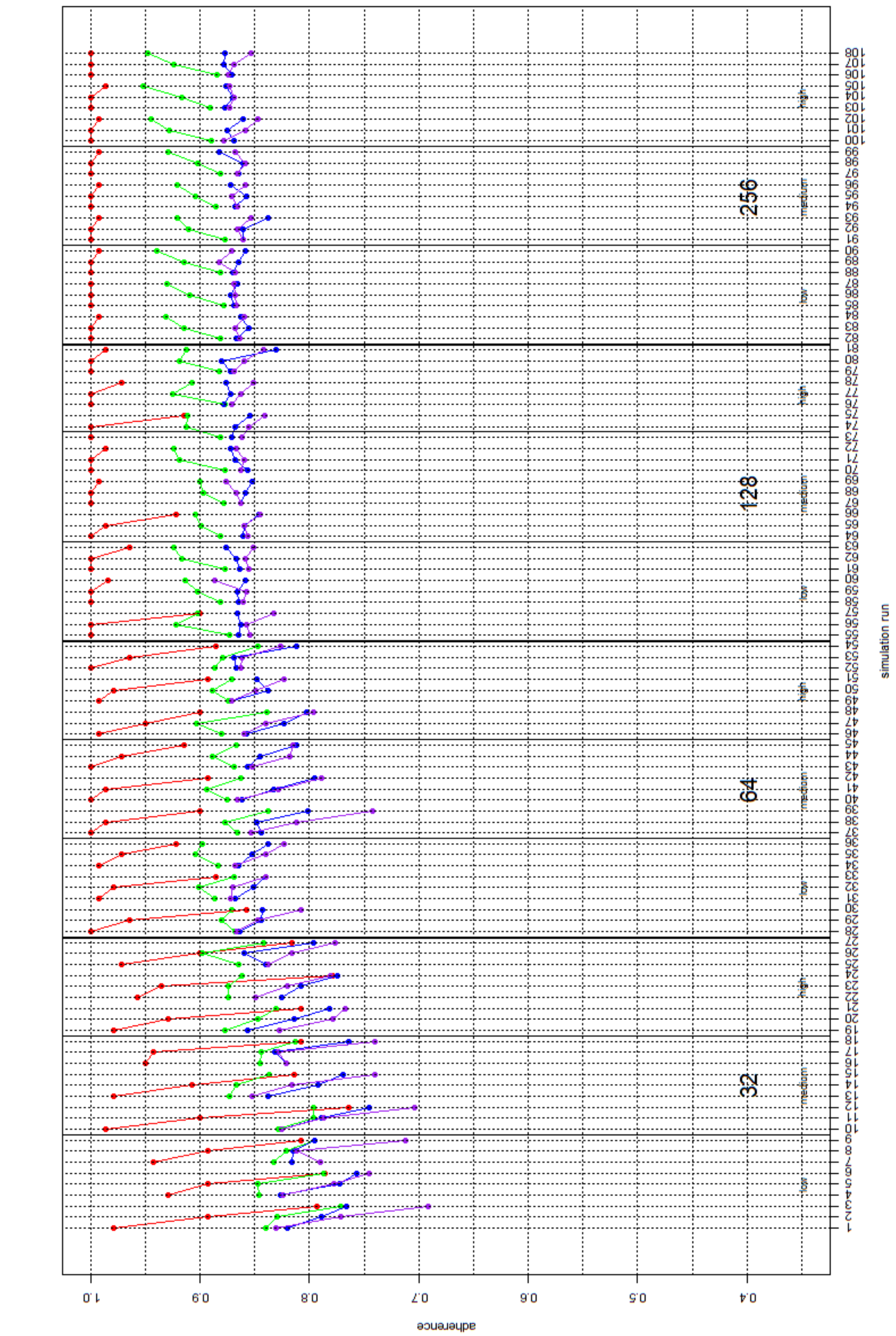


Figure 10.12: End points with self search, long range but without knowledge superiority
(n. dist. dim: red = 0, green = 2, blue = 10, purple = 20)

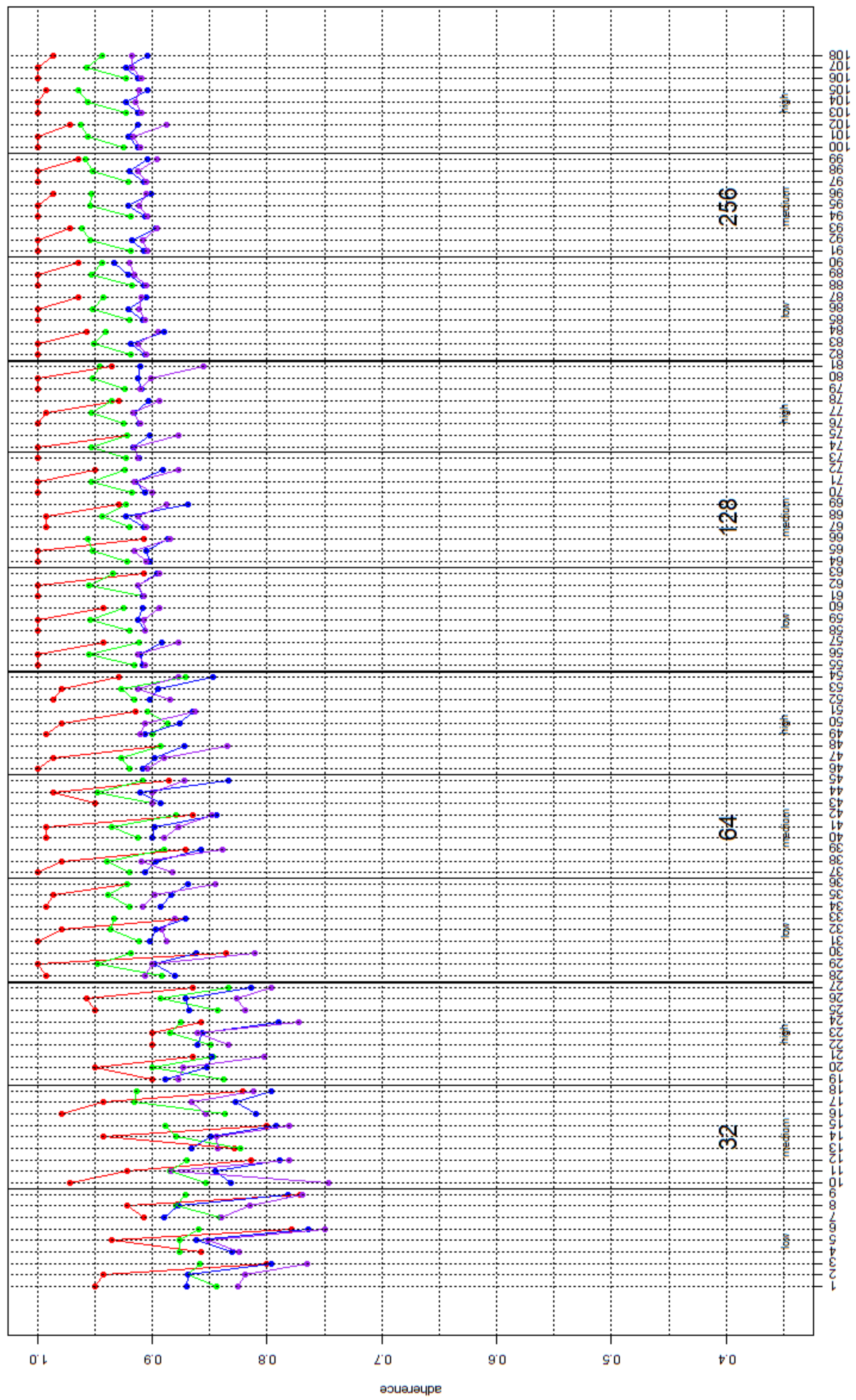


Figure 10.13: End points with self search, knowledge superiority and long range
(n. dist. dim: red = 0, green = 2, blue = 10, purple = 20)

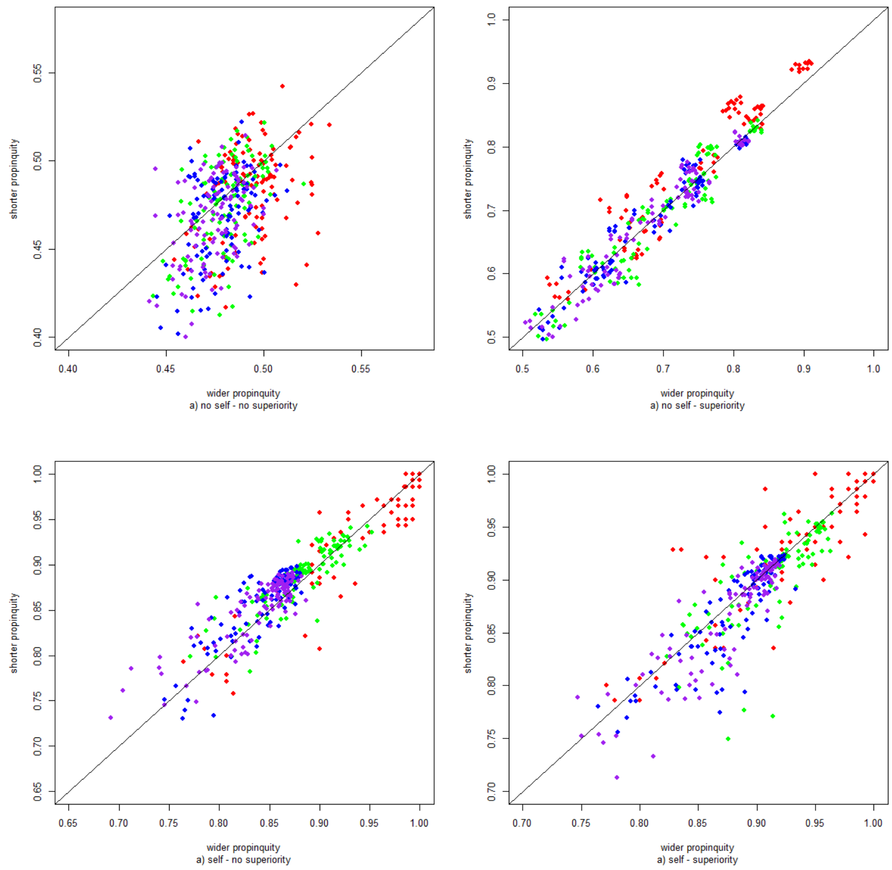


Figure 10.14: End points difference between short and long range (n. dist. dim: red = 0, green = 2, blue = 10, purple = 20)

in the same direction. Therefore a propinquity increase does not assure a clear increase of the end point values. Propinquity seems to work as an enabler. This results could in some way echoed by one of the finding in Levine and Prietula work (Levine and Prietula 2012) which assert that exchange pattern on local search perform better.

A final remark here worths, supporting the concept. Looking at figure 10.11 is still possible to note the inverse U-shape for end point values. It could be thought that widening the propinquity the number of agents in the offices clusters increases, moving in the region where the March effect is possible. Actually, the effect of widening the propinquity is not linked to the increase of the number of agents rather it is linked to the increase of relationships within the offices clusters. This could be noticed comparing figures 10.8 and

Table 10.3: Paired t test for long range

		without knowledge superiority				with knowledge superiority			
	ext lev.	t	p-value	mean diff.	H_1	t	p-value	mean diff.	H_1
no self	0	-4.6303	5.159e-06	-0.012	L	8.7943	1.349e-14	0.026	G
	2	-3.0553	0.0014	-0.006	L	-0.4271	0.3351	-0.001	L
	10	-4.1585	3.241e-5	-0.009	L	0.9569	0.170	0.002	G
	20	-3.6856	1.800e-04	-0.007	L	0.0525	0.4791	-0.000	G
self	0	-3.0596	0.0014	-0.006	L	0.3492	0.3638	0.001	L
	2	3.1643	0.0010	0.006	G	-3.0892	0.0012	-0.008	L
	10	7.3032	2.627e-11	0.014	G	-3.8353	0.0001	-0.008	L
	20	6.1542	6.679e-09	0.012	G	-3.7912	0.0001	-0.008	L

10.15.

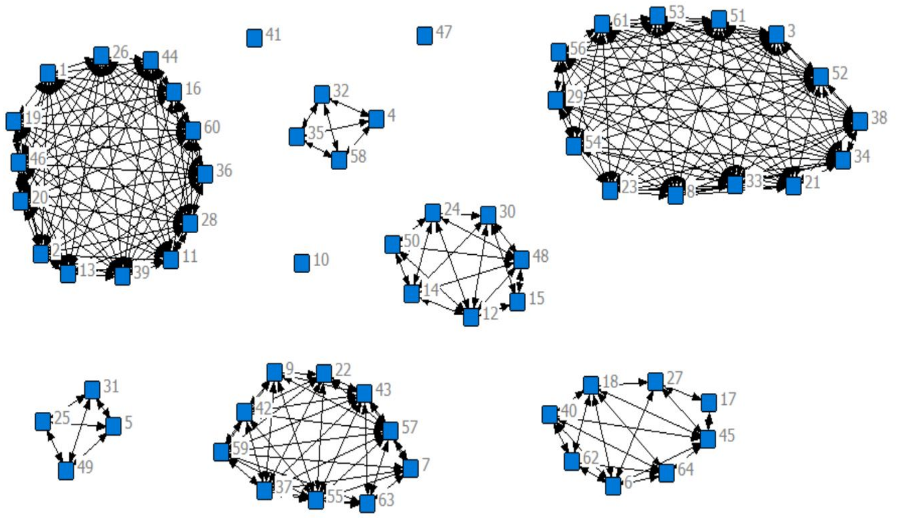


Figure 10.15: Isolation in wide range 64 agents simulation

The number of agents does not increase increasing the propinquity influence, only the number of possible links within the clusters are greater. This result is in line with the space syntax theory and support the idea that offices layout has a paramount effect on the knowledge transmission.

10.2 Network analysis

The analysis of network structures starts from the study of the density. As first step, the simulations involving the self search combined with propinquity are considered. Two subgroups are studied: the simulations where knowledge superiority is present in the propinquity channel and the simulations where knowledge superiority is suppressed. Each subgroup is then split into 4 replicates with different level of interaction with the external world.

Picture 10.16 shows the density curves for the first subgroup of simulation where the interaction with the external world is not present and superiority is switched off and figure 10.17 shows the density curves for simulations with superiority considered (coloured by stack level). The reading of these figures shows that there is a clustering effect that deserves some clarifications.

To understand better the role of the number of agents, a boxplot is presented for both groups, with and without superiority (picture 10.18 a and b).

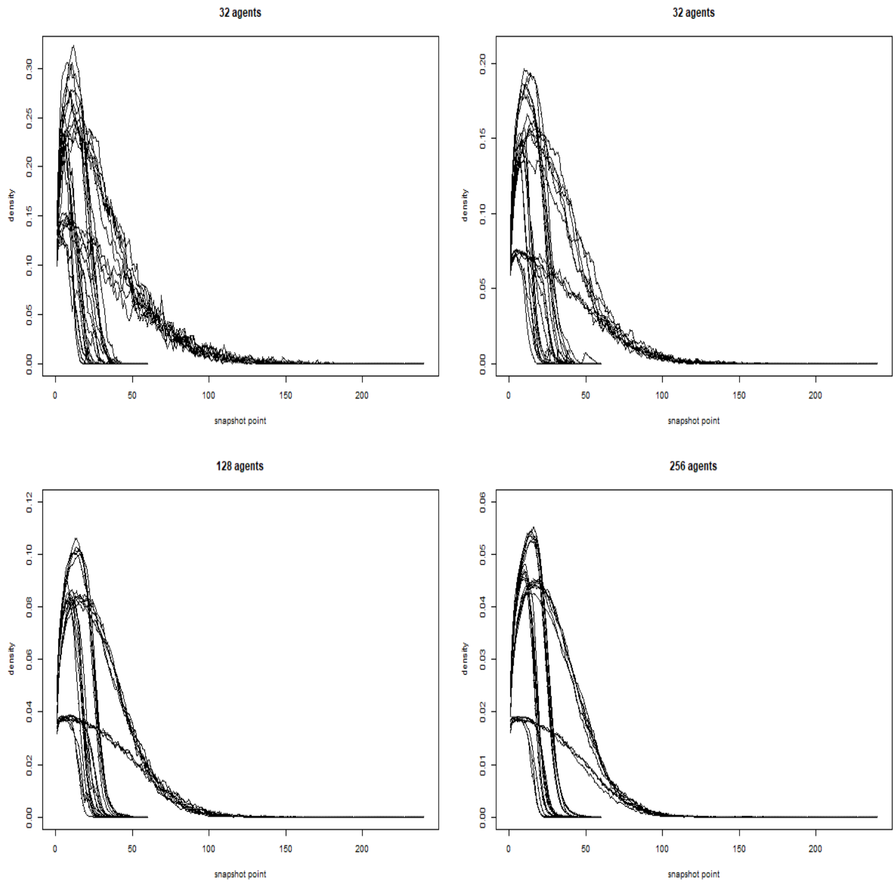


Figure 10.16: Density curves at different number of agents, without superiority

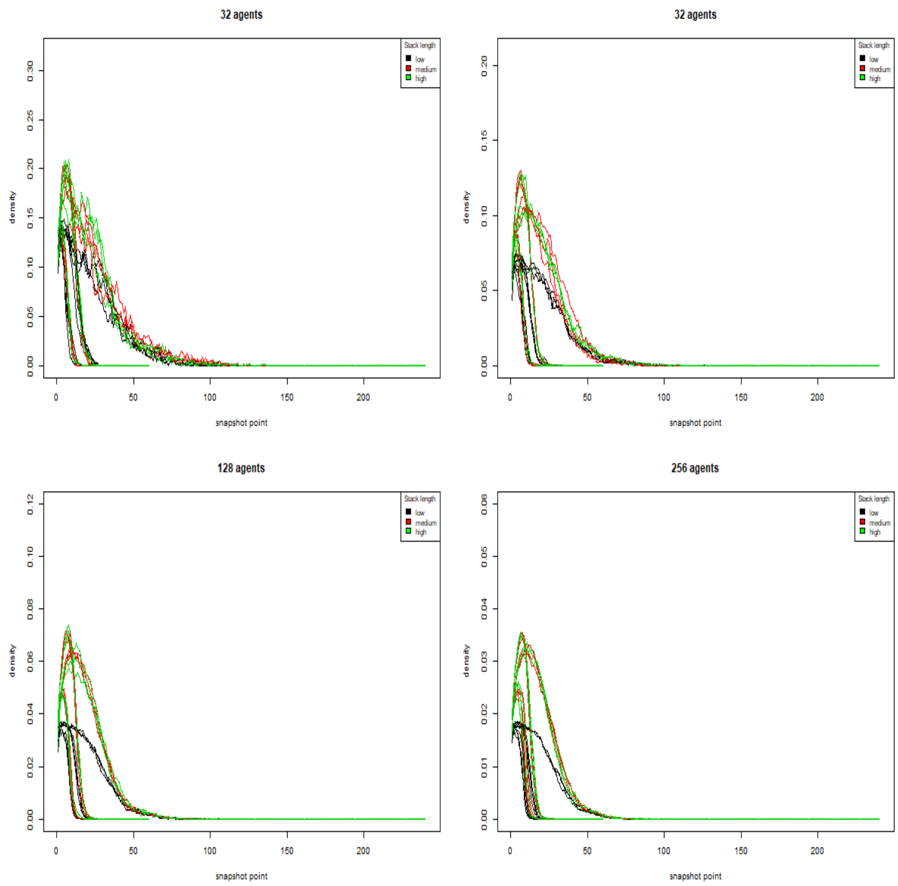


Figure 10.17: Density curves at different number of agents, with superiority (coloured by stack level)

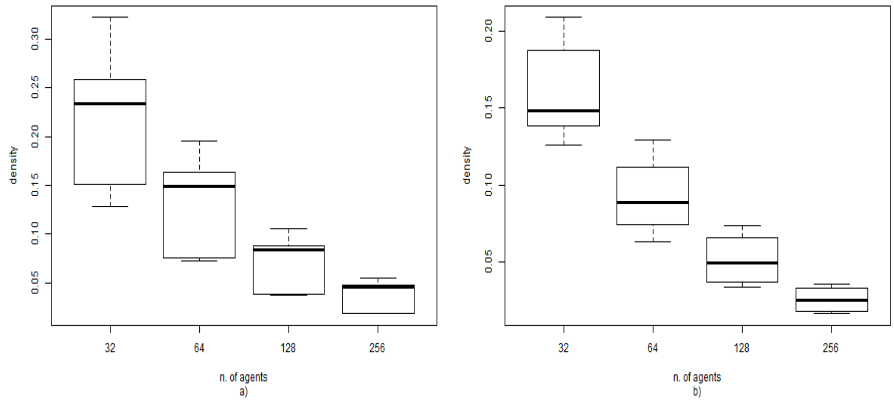


Figure 10.18: Density boxplots (without and with knowledge superiority)

The boxplot reports the distribution of maximum values of density for the different simulations. Indeed the number of agents is a significant factor and the values of the density are pretty different changing the number of agents. A deeper analysis of the boxplot reveals two main topics to explain. First, the density at the peak is greater in the simulation involving propinquity other than self search respect to the simulations where only self search is taken into account. Referring to tables 9.4 and 10.4 it is possible to appreciate the values whereas figure 10.19 offers a graphical representation.

The four dotted lines are at the density median values found for self search only simulations[7] (ref. table 9.4). Black line is for 32 agents, red is for 64 agents, blue for 128 and green for 256 agents. The points represent the median for the simulation with both channels active and they have the same colour coding as the lines. In the upper part of the plot is also reported the block number as in table 10.4. It is evident that the the points are always above the respective lines.

Some considerations arise: first, propinquity channel has the average effect to increase the density of the network, all points are above the corresponding lines. This means that the self search alone requires less dense networks. Second, knowledge superiority tends to create less dense networks despite the fact that the end points values are generally higher than the relative simulations without knowledge superiority (refer to table 10.4). This is evident appreciating that blocks 1, 3, 5, 7 have higher values than blocks

[7]This comparison is made with the knowledge at 20 dimensions per topic. All the simulations in table 10.4 have the same knowledge construct.

Table 10.4: Simulations density medians

			Knowledge sup.			
Sim. group	Agents	Block	No	Yes	Block	Range
A,B,C,D	32	1	0.2167	0.1522	2	Short
	64		0.1481	0.1060		
	128		0.0726	0.0543		
	256		0.0455	0.0295		
E,F,G,H	32	3	0.2077	0.1698	4	Wide
	64		0.1297	0.1081		
	128		0.1376	0.1123		
	256		0.0706	0.0557		
I,J,K,lL	32	5	0.2732	0.2041	6	Short
	64		0.1650	0.1205		
	128		0.0897	0.0668		
	256		0.0477	0.0337		
M,N,O,P	32	7	0.2727	0.2112	8	Wide
	64		0.1724	0.1277		
	128		0.1416	0.0964		
	256		0.0715	0.0483		

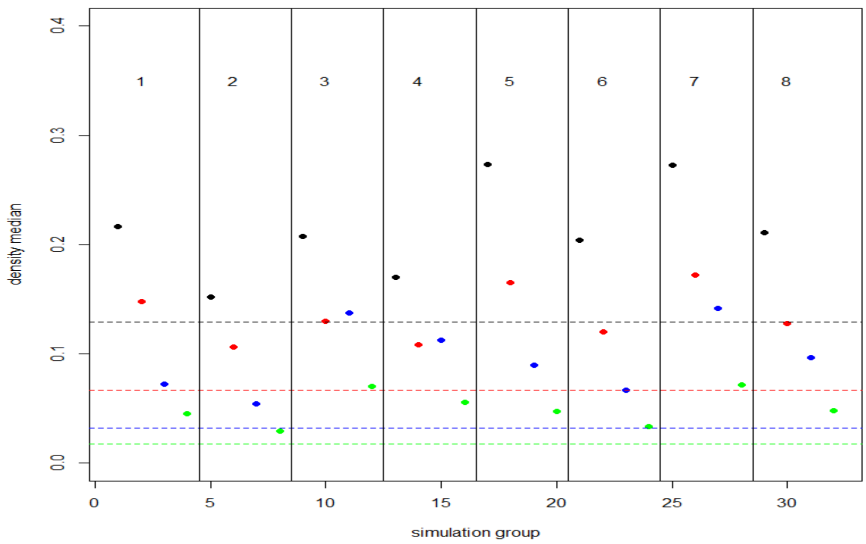


Figure 10.19: Density medians

2, 4, 6 and 8. Third, wider range of propinquity influence tends to create (slightly) denser networks, blocks 5, 6, 7, 8 have higher values than blocks 1, 2, 3 and 4.

It could be appreciated that the median increases, increasing the complexity of the task. Indeed the simulations with the propinquity have an higher complexity respect to the simulations with only the self search. But, between the two, the most complex simulations (without superiority) have greater density. Then it could be hypothesised that the network is denser, the more complex the task to perform. Complexity in not only in terms of number of channel activated in the agents' task but also in the energy to be spent to accomplish the task. This because the simulations without the superiority actually propagate not optimized knowledge in terms of the objective of the whole organization job, that is maximizing the adherence to external reality. Making a parallel with physics it seems that *ceteris paribus* a more entropic task is related to a denser network.

The second point to investigate is the anomaly in the density curves. A comparison of the boxplot in figure 10.18 with figure 9.16 (d) shows a less sharp division among the groups of data and the distribution are highly skewed, in an anomalous way. Moreover, analysing figure 10.16 results evident that there is something more in action.

Picture 10.20 shows the same output of figure 10.16 but coloured by stack level, and it is clear that in each plot there are roughly two different behaviours: the black curves and the red-green ones. The black curves belong to simulations with low length of agent stack whereas the red and the green belong to medium and high length of the stack.

Taking as reference the simulation without superiority[8] and removing temporarily the density profiles with low level of stack and plotting again the data, the boxplots of figure 10.21 are obtained. The output is clearer as the division by number of agents. This point is studied in the next section.

10.3 Stack level and energy

There is still the need to understand what happens in the simulations with a low level of stack. It is possible to discuss around the counter-intuitive result of the density maximum considering for sake of simplicity only the case of 32 agents without superiority[9]. For this specific elaboration further simulation (subgroups CQ and CR) were run with 32 agents covering the range of stack length from the minimum to the maximum. This means that

[8]The simulations with superiority exhibit the same behaviour.

[9]Without lack of completeness.

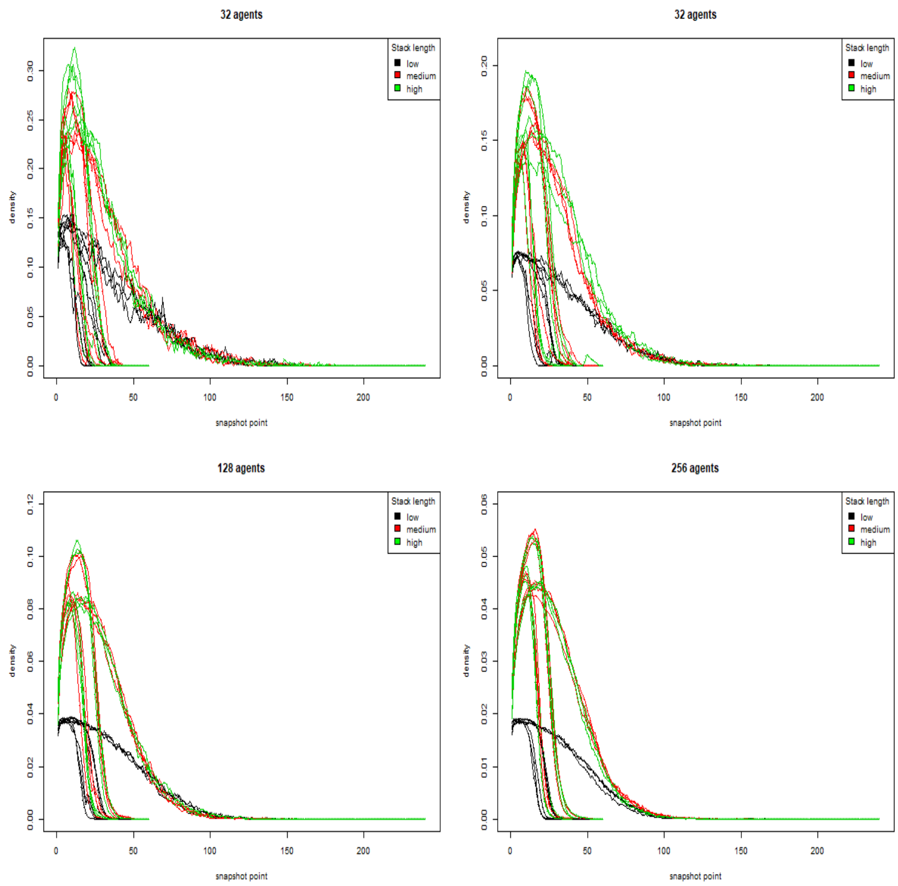


Figure 10.20: Density curves at different number of agents coloured by stack level, without superiority

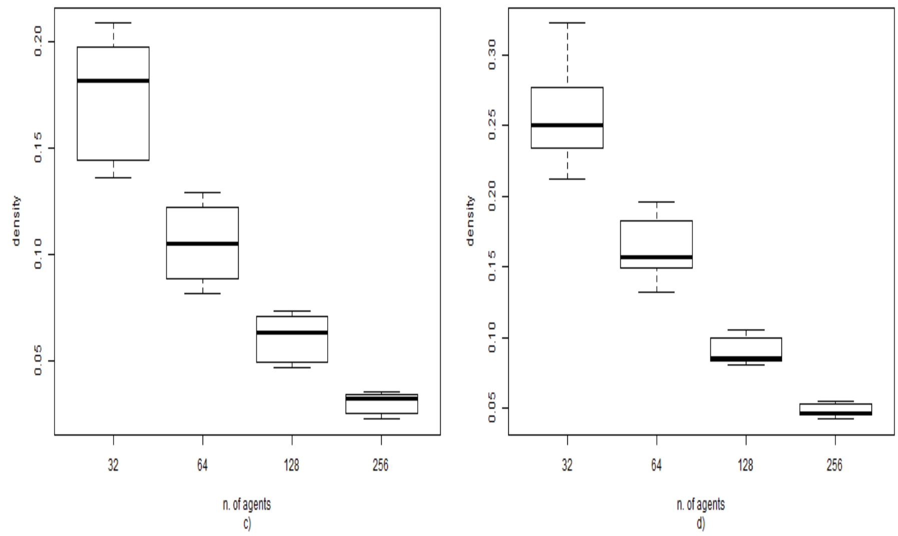


Figure 10.21: Density boxplots without low level of stack (with and without knowledge superiority)

simulations were run with stack length equal to 5, 8, 11, 14, 16, 20, 25 32. Before entering in the results discussion it could be useful to recall that the length of the stack limits the number of ties an individual could remember. By consequence, the maximum number of ties a simulation could have is given by the formula:

$$L_{max} = \frac{1}{2}\sum_{i=1}^{N_{ag}} S_i^L \tag{10.2}$$

where S_i^L represents the maximum number of links for each agent. Recalling the formula of network density and substituting the term L with the one found, the maximum density is given by the formula

$$\Delta_{max} = \frac{1}{N_{ag}(N_{ag}-1)} \cdot \sum_{i=1}^{N_{ag}} S_i^L \tag{10.3}$$

This equation is the more general one and, if all the agents have stacks with the same length, it could be simplify in the following one:

$$\Delta_{max} = \frac{S^L}{(N_{ag} - 1)} \tag{10.4}$$

The obtained result could be used to estimate the maximum density of the network during the simulation. Three are the levels of stack in the present experimental design: low, medium and high. Usually the level high is equal to $N_{ag} - 1$ and it allows all the individuals to potentially connect themselves with all the others in the organization. The medium level is usually set at $N_{ag}/2$, that is an agent could be connected maximum to half of the organization without discarding owned connections to create a new one. Minimum level is set to few units. As mentioned, for this study, further points were added. Hence, considering 32 agents, the formula 10.4 gives the results reported in table 10.5.

Table 10.5: Maximum obtainable density at different stack levels (32 agents)

	5 (low)	8	11	14	16 (medium)	20	25	1 (high)
Δ_{max}	0.16	0.26	0.35	0.45	0.52	0.64	0.81	1

Rearranging the density boxplot by stack level the output (figure 10.22) reveals an interesting point. The values increases at the increase of the stack level but till a certain point. After that point the density seems to stay constant regardless the level of the stack.

Picture 10.22 suggests two implications: first, the network reaches its density maximum due to the constraint of the stack level, trying to accomplish the task. This phenomenon was not evident in the self search alone scenario since the complexity of the task was manageable with a lower density. In propinquity case, since the task is more demanding, the system tries to establish more connections to face it but it is limited by the agents' memory. Second, above a certain level of stack the density does not change: the network has reached the energy level necessary to manage the task. Moreover, stack limitation gives a way to demonstrate that the network reaches the configuration to fulfil the task that minimizes the energy.

Before demonstrating this hypothesis it should be explained why there is no dependency between the net density and the end point value as reported in figure 10.23.

It is clear the absence of any pattern since the end point values are distributed uniformly with network density. This could seems an anomaly considering the constraint brought by the stack level, but the result could be

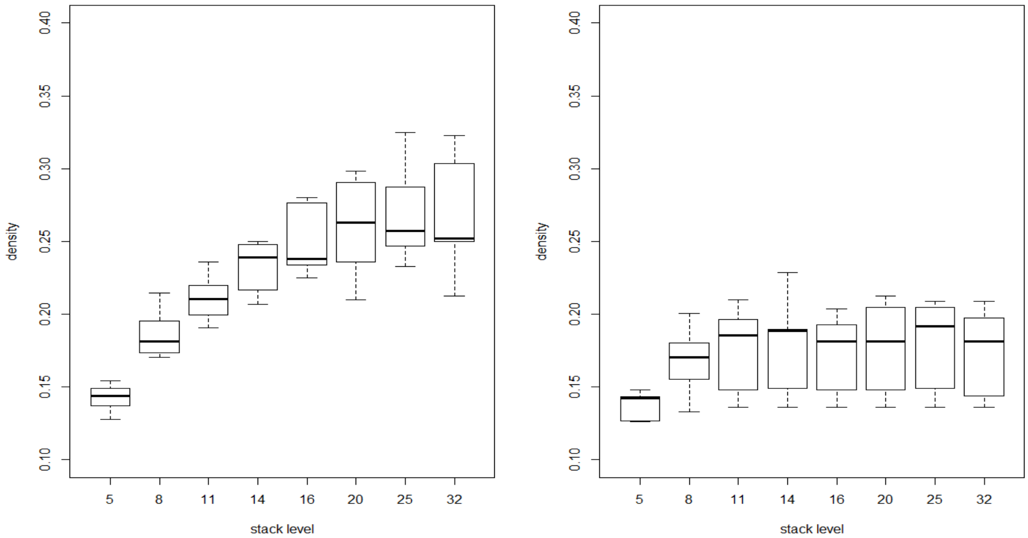


Figure 10.22: Density boxplots without low level of stack (with and without knowledge superiority)

explained by the inner mechanism of propinquity channel. As shown in figure 6.3, every time an agent needs a new connection but her stack is full, the least frequent link is removed to generate the space for the new connection. This means that a new connection, functional to the knowledge exchange is always set, without jeopardizing the overall end point reaching. This is the reason why end point is not affected by the limitation of stack level.

From the comparison among figure 10.22 and values in table 10.5 it is evident that the density could be potentially higher than the values actually found. For example with stack level equal to 5 the network is virtually at the maximum level (0.16) but, already at stack level 8, the maximum value is 0.26 when the simulations give values well below 0.20. But even if the density is not close to the theoretical one the values are still limited by the stack level.

Picture 10.24 insights into this phenomenon: the red lines are the maximum number of out degree admissible considering the stack level and the boxes are the displacement of out degree values at different stack levels. When the stack level limits the network, the distribution of out degree is limited on the side near the limit. This means that agents which potentially need more contacts actually cannot have them. This limits the density.

Hence, the density increases with the increasing of stack level (when the

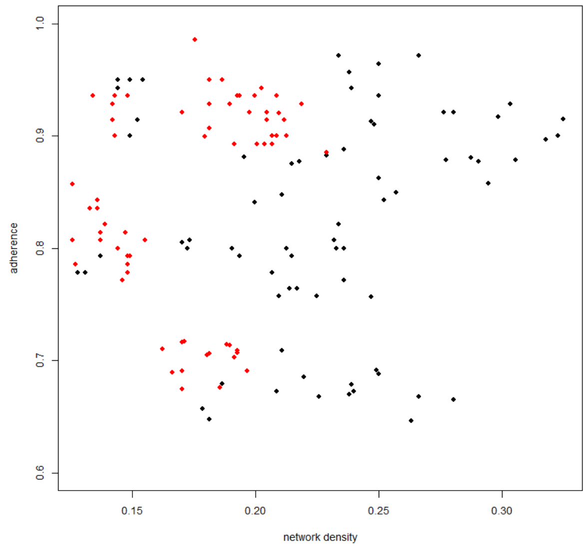


Figure 10.23: End point as function of network density (Red = with superiority, black = without superiority)

stack level is limiting) and in this part of the trend the density values are not the "natural" ones. At a certain point the stack level is no more limiting and hence the density tends to an asymptotic value (that is the "natural" one). The curve is then composed by two linear trend: the increasing part and the constant part. The intersection of these two trends marks the minimum density to accomplish the task in terms of stack level.

If the minimum did not exist, the intersection point should be near the maximum value of the stack level. Indeed, this would be indication that the density is actually still climbing without reaching a plateau. Vice versa, an intersection value far from the maximum is indication that with increasing freedom, the network is not greedy and always stays with the same value of density.

Picture 10.25 reports on the left the intersection points for the two groups of simulation (with and without superiority). The intersection points are at 8.7 for the red line and 17.3 for the black one. Both are far from 32.

Density limitation could be seen as a communication constraint: less ties are present in the network, so the individual has less interlocutors to meet with and to involve in a knowledge exchange. Puranam and Swamy (Puranam and Swamy 2016) studied the implications of limited communication in a coupled learning process. Their work is focused on the importance of

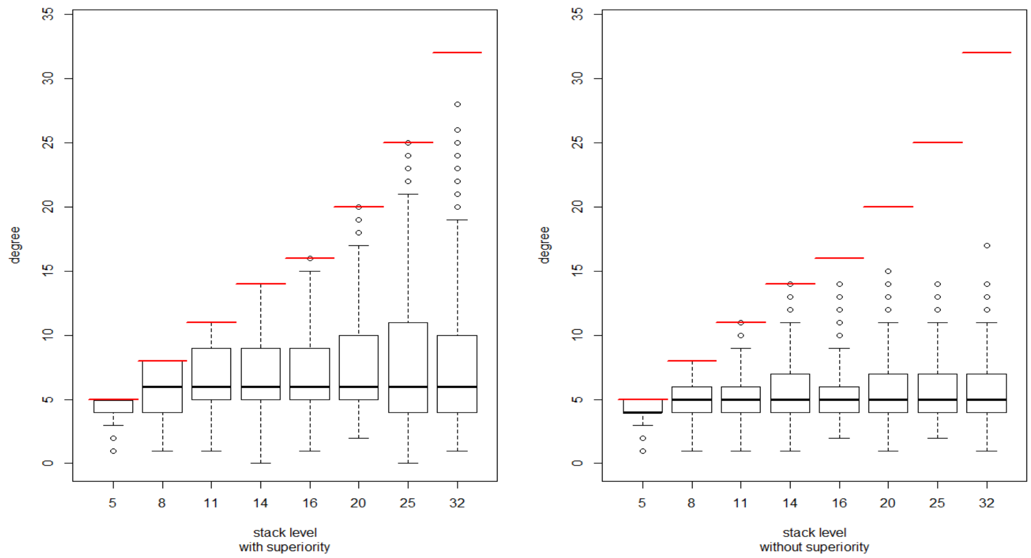


Figure 10.24: Out degrees as function of stack level

initial representation of the task when the communication among two specialist is impaired (or not) by the environment. The interesting finding is linked to the extreme importance of the representation agents have at the beginning of the task assigned. In a scenario where the communication is limited, the worse representation they have the better result at the end they get. A parallel could be done with the obtained results. At the beginning of the simulations agents are equipped with a random representation of the external reality that, on average, is correct for about a 30%[10]. This is a pretty bad representation owned by individuals but it could be functional to a good result when the stack level limits the number of possible ties.

Energy is then related to density since both are derived from the adjacency matrix. And if also energy values reach an asymptote, then it could said that the system reach a configuration of minimal energy to perform the task as is evident from the right part of figure 10.25.

This aspect could also be considered from another perspective. If the task requires more energy, then the cost of the task is more. This, in turn, means that the knowledge transfer with two channels costs more than the transfer with only one channel.

[10]The possible values for the individual beliefs are 1, -1 and 0.

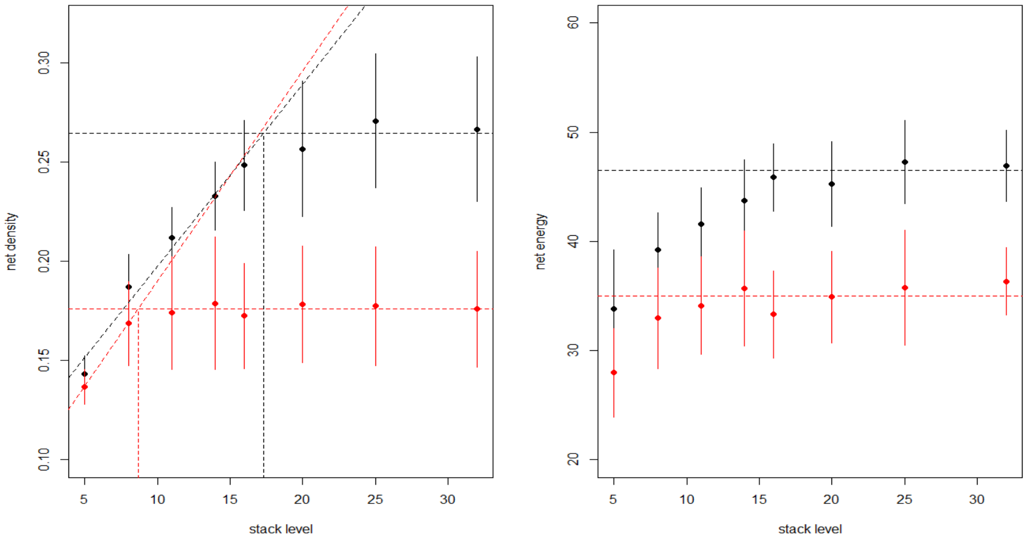


Figure 10.25: Network density and energy models

10.4 Office footprint

As seen in section 9.2 and, more precisely, in figure 9.17 there is the tendency by network to adopt a small world configuration when the size increases. Now, considering the propinquity aspect, there could be the interest in understanding what happens to the shape of the network. This at least for two reasons: first, in the simulation there are two different paths for knowledge diffusion. Not only do agents have the chance to gain knowledge by searching autonomously but they could also rely on their neighbours. Second, according to Hilier (Hillier 1996), the space configuration should have an important role in the way people experience and live the available space. Moving this concept into the exploration exploitation problem, the physical layout could in some way shape the way people talk to each other by constraining or fostering interactions. Doing that, the physical layout changes the probability two individuals meet and only certain relationships are favourite. Bafna argued that *"the demarcation of boundaries allows particular relationships of access or visibility to emerge among the component spaces, and this in turn generates probabilistic patterns of movement and encounter within the population being housed"* (Bafna 2003, 18). Boutillier et al. in 2008 (Boutellier et al. 2008, 373) supported this opinion writing: *"we defend the opinion that performance is influenced by communication but indirectly through knowl-*

edge sharing and creation". Strong of this support by literature, the Q factor (Watts and Strogatz 1998) is calculated for the simulations with self search and propinquity. Picture 10.26 reports the output, on the left, for propinquity without knowledge superiority and, on the right, for propinquity with knowledge superiority. Immediately evident are the values of the Q factor, smaller that the values obtained for self search only (cf. figure 9.17).

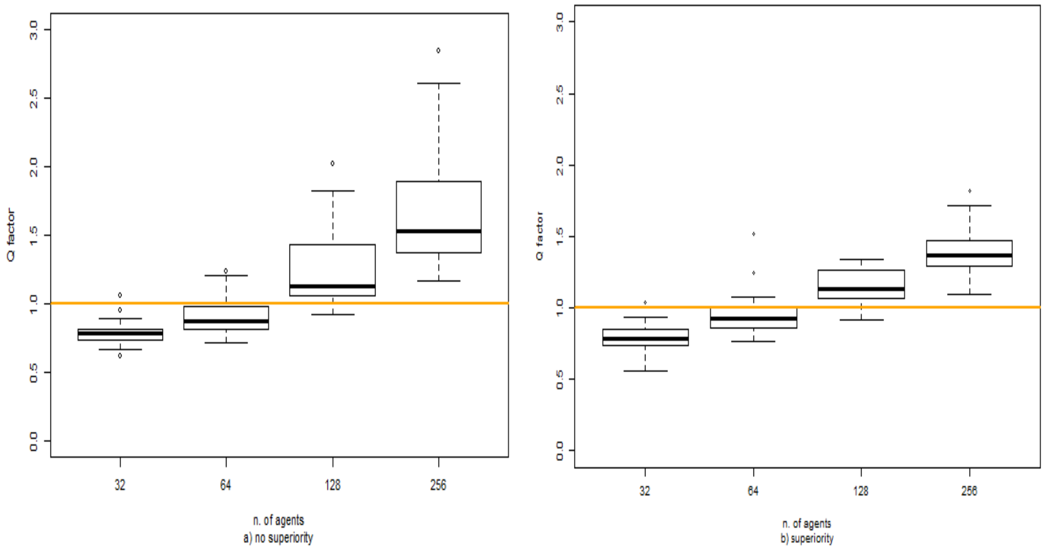


Figure 10.26: Q index for simulation with self search and propinquity

The result suggests the loss of the small world status when propinquity is switched on. It is plausible to associate this loss with the propinquity channel. Hence there is the need to elaborate more on propinquity concept.

Propinquity level is a parameter of the model. This implies that it is possible to create simulations with different level of propinquity. It is possible to change the influence of neighbours and study the effect. A second group of simulations is performed setting the parameter to the maximum level permitted. With this group of simulations there are three different scenarios to be compared: i) self search alone, ii) self search with short propinquity and iii) self search with long propinquity. The effect of propinquity could be assessed using the QAP test (Krackardt 1987, Krackhardt 1988) which gives the correlation between two different networks.

The propinquity network (pic. 5.7) is taken as reference and all the 256 agents networks of the three groups of simulations are assessed against it with the QAP test.

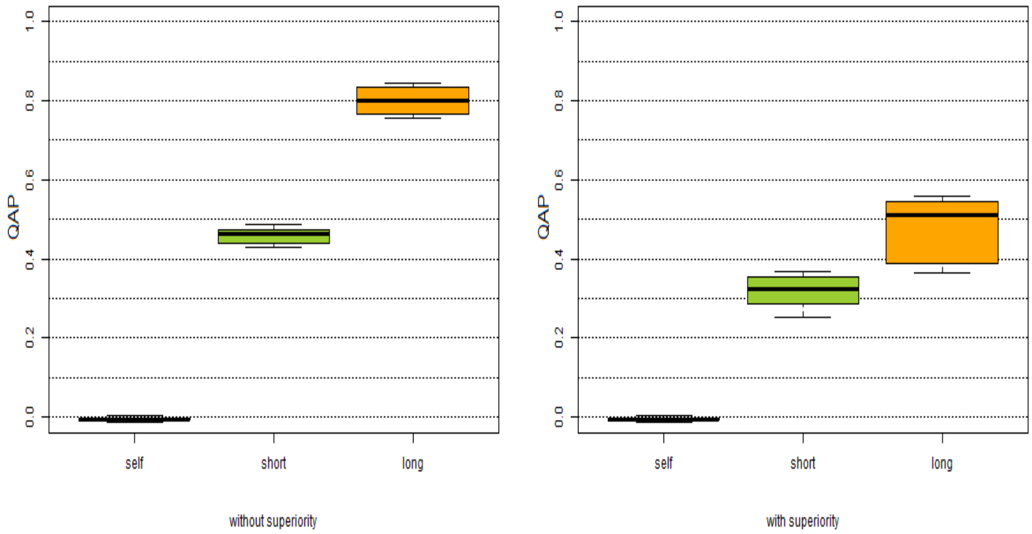


Figure 10.27: QAP correlation output

Picture 10.27 shows the results of QAP correlation test for the 256 agents networks in boxplots. On the left the propinquity without superiority is considered and on the right the propinquity with superiority is shown.

The trend is clear: the lone self search has very limited correlation with the propinquity network whereas the simulations with propinquity have higher correlation with the propinquity network. It is worth to remember that the propinquity network is the office configuration. It gives a representation of the possible neighbours interactions. A high correlation with this network could imply that the final interaction network at the end of the simulation is like the propinquity network, or at least is similar. Hence it could be said that the office layout is able to influence the connection formation during the time. Moreover, looking the boxplots a couple of further considerations could be made: the short propinquity shapes less than the long propinquity and, the knowledge superiority shape less than the exchange based on knowledge "as is" (that is knowledge superiority switched off).

The superimposition of more channels leads to higher values for the end point. Mueller et al. studied different networks configurations[11] and found that *"the better performing networks are characterised by a less asymmetric degree distribution. The worst performing networks indeed have a more asym-*

[11] Authors tried Erdös-Renyi, Barabasi-Albert, Watts-Strogatz and evolutionary networks.

metric degree distribution than the better performing networks" (Mueller, Bogner, and Buchmann 2017, 624).

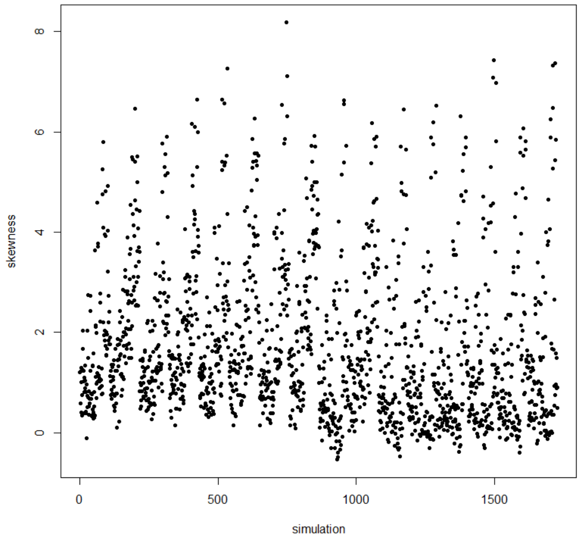


Figure 10.28: Networks degree distribution skewness

Picture 10.28 shows all the skewness values[12] for the simulations involving the propinquity channel plus self search channel[13]. Every point represents the skewness of the single simulation distribution. All the distributions have a pronounced skewness. Nonetheless, when self search is present, the combination of the two channels perform better than the self search channel alone, even if the distributions are asymmetric. This could imply that the mechanism is more complex than the one found by Mueller et al. (Mueller, Bogner, and Buchmann 2017) when the multilayer scenario is present.

10.5 End points convergence

One of the most interesting point emerged in chapter 9 is the convergence of the end point values increasing the number of agents involved in the organization. As depicted, with larger and larger organization, the end points of the adherence tend to converge to an asymptotic value regardless the learning rates of the agents. Introducing the propinquity channel, the scenario is

[12]The formula used for the skewness is $G_1 = g_1\sqrt{n(n-1)}/(n-2)$.

[13]Groups CI, CJ, CK, CL, CM, CN, CO, CP for a total of 1728 simulations.

more complex and the question could be whether or not this phenomenon is still present. In the present simulation chunk, the self search has to compete with another communication channel: the propinquity. Agents are now able to retrieve information from the neighbours or with a self search (as in chapter 9).

A look to end point values trends (figures 10.2, 10.3, 10.4, 10.5) suggest a sort of dependence on the number of agents, asking to check if a convergence is in some way present. Each coloured line in the figures represents the end points values of a particular simulation held at three different values of learning rate (0.1, 0.5, 0.9). It is then possible to estimate the range of each line, that is the difference between the maximum and the minimum value. If the convergence holds, the range should decrease with the number of agents.

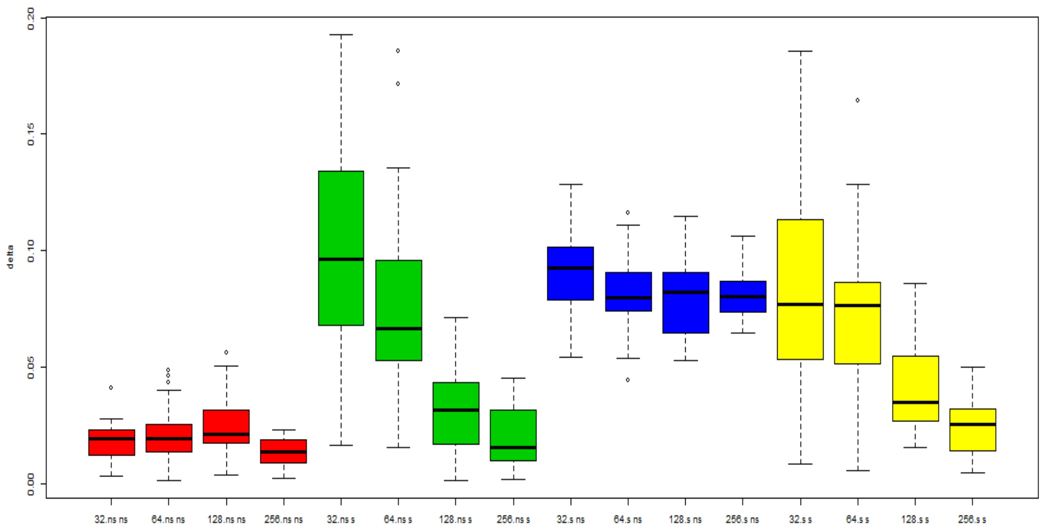


Figure 10.29: End points ranges

All the ranges are the collected and plotted in the boxplot 10.29. Four different colours are used to distinguish the different combinations in the simulations: (red) self search and superiority not activated, (green) self search activated and superiority deactivated, (blue) self search deactivated and superiority activated and (yellow) both activated.

It is quite evident that there are two distinct behaviours: the red-blue one and the green-yellow one. In the first pair the self search is not activated whereas in the second self search is activated. A further element is needed to better investigate the behaviour. In section 10.1 the idea of influence sphere of propinquity was introduced and the end point of that simulation could be also included in the analysis. The plot 10.30 has two more groups (gray and white) which are the same simulations of the green and white groups only with a larger influence of the propinquity.

Two regression models are proposed, as reported in table 10.6. The first considers all 6 groups of data with 4 predictors: the number of agents, the self search channel presence (dummy variable), the knowledge superiority option presence (dummy variable) and the longer propinquity influence presence (dummy variable)[14].

The table provides all the variables significant and a comment is worth: the fact that the variable superiority presence and longer propinquity are significant is in some way related to the presence of the red and blue boxes. Model 2 is created focusing only on simulation where the self search is present (green, gray, yellow and white) and the output confirms that the only significant variable is the number of agents. Hence, when the self search is active, no matter whether or not agents seek for superior knowledge or they consider a wider neighbourhood, the number of agents tend to crush the range. Again, even if the propinquity is present, the self search is able to dominate it.

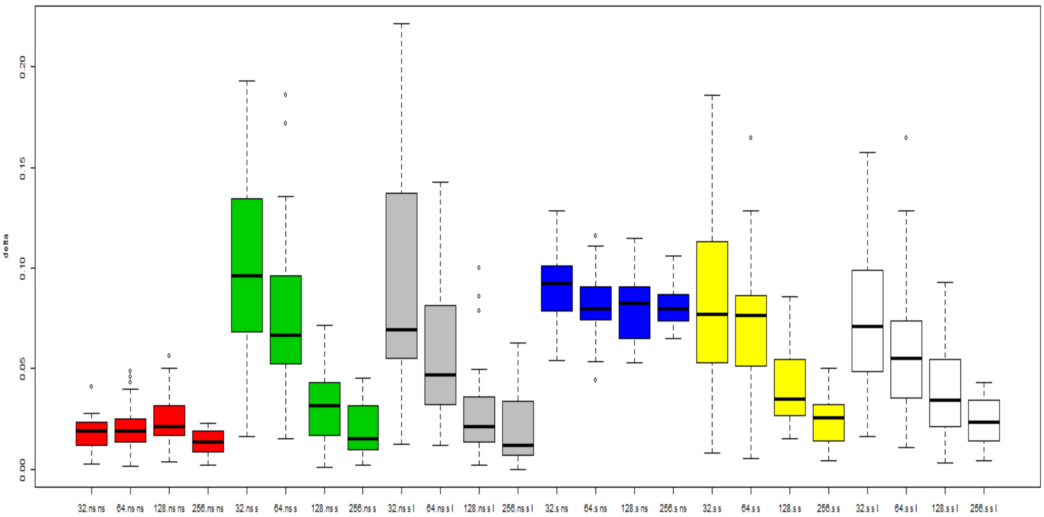


Figure 10.30: End points ranges with longer propinquity included

[14] All the dummy variables are set to 1 if the channel is present or the option activated, 0 elsewhere.

Table 10.6: Regression models 1 and 2 for range convergence

	Dependent variable: end point range	
	(1)	(2)
n. of agent	−0.0002 (0.00001)	−0.0003 (0.00002)
	$p = 0.000^{***}$	$p = 0.000^{***}$
self search (dummy)	0.008 (0.003)	
	$p = 0.004^{***}$	
kn. superiority (dummy)	11.171	
	$p = 0.000^{***}$	$p = 0.565$
longer prop. (dummy)	−0.022 (0.004)	−0.005 (0.004)
	$p = 0.000^{***}$	$p = 0.230$
Constant	0.060 (0.003)	0.086 (0.003)
	$p = 0.000^{***}$	$p = 0.000^{***}$
Adjusted R^2	0.285	0.336
Residual Std. Error	0.033 (df = 859)	0.033 (df = 572)
F Statistic	86.811*** (df = 4; 859) (p = 0.000)	97.998*** (df = 3; 572) (p = 0.000)

Note: *p<0.1; **p<0.05; ***p<0.01

Chapter 11

Gatherings

Meeting channel is the third main way through which knowledge could be exchanged. Meeting channel simulations framework is designed considering four different dimensions: type of meeting, number of meeting, time allocated to meeting, physical limitation. The combinations of these dimensions gives the high level structure of the simulations package, in addition, the presence of self search, propinquity channel, the number of agents, and learning rate are considered.

The proposed model has two different kinds of meetings: department meeting and project meeting. The existence of these two different types of meetings tries to cope with the typical scenario in an organization. Every organizational unit typically holds internal meetings to discuss about it own objectives, governance, best practices or about the knowledge it creates. The model captures this option with the first type of meeting where the potential attendees are selected within the department and the topic discussed is the one belonging to the department[1].

Similarly, the second type of meeting (called project meeting) encompasses all the cross-unit, cross-department meetings. In these meetings, attendees could belong to different departments and the topic treated could be heterogeneous, according to the necessity of the moment. Those meetings could represent project meetings where the work status is discussed, where lessons learnt are shared, where the project plan is discussed or where a particular technical aspect is discussed in order to find an agreement.

[1] The external reality representation for this model has N_{ER}^{cat} categories (cf. 6.3.3). The number N_{ER}^{cat} of categories must be equal to the number of departments present. In this way every department has its own knowledge category. In the simulation run, the number of categories (and departments) is set to 7.

		Number of meetings	
		Low	High
Time allocated by agents	High	Sporadic of meetings with basically a higher number of attendees	Frequent meetings with basically a higher number of attendees
	Low	Sporadic meetings with few attendees	Frequent meetings with few attendees

Figure 11.1: Meeting simulation combinations

The flexibility of the model allows also to mix the different meetings during a simulation. The simulations are run with only department meetings, only project meeting and with an equal number of both.

Two parameters rule the meetings activities in the model: the number of meetings and the time allocated to meetings[2]. The number of meetings tells the model how many meetings need to be simulated in each simulation cycle. The time allocated represents the daily time an agent could spend in meetings. This feature allows to simulate different agents' roles in the organization which could be differently involved in meetings. For example, a laboratory technician hardly could attend project meetings, rather she will be involved in the department meetings where the internal knowledge is discussed. Otherwise, a project team member will be mostly part of project meetings sharing knowledge with agents of different organization units.

Setting the allocated time to low means that agents have less probability to be involved in meetings during the simulation cycle. Hence the number of potential attendees is low. Vice versa, setting it to high means that agents could likely be part of the meeting exchange channel. This means that the potential number of attendees for each simulation cycle is high(er). For the performed simulations, all the agents have the same values of allocated time[3]. Modulating also the number of meeting per simulation cycle it could be possible to create four different combinations as reported in figure 11.1.

[2]These parameters are doubled: number of meetings and time allocated for department meetings (N_m^d and T_{dep}) and number of meetings and time allocated for project meetings (N_m^p and T_{prj}).

[3]The model allows every agent to have her own profile as discussed in chapter 13.

These four combinations are the core of the simulations upon which all the other parameters are changed. Again, the idea is to respect as much as possible the same simulation architecture used for the self search and the final structure of meeting simulation is reported in figure 11.2. It is important to highlight that the diagram has half of the simulations with self search and meeting alone and half of the simulations with self search, propinquity[4] and meeting active.

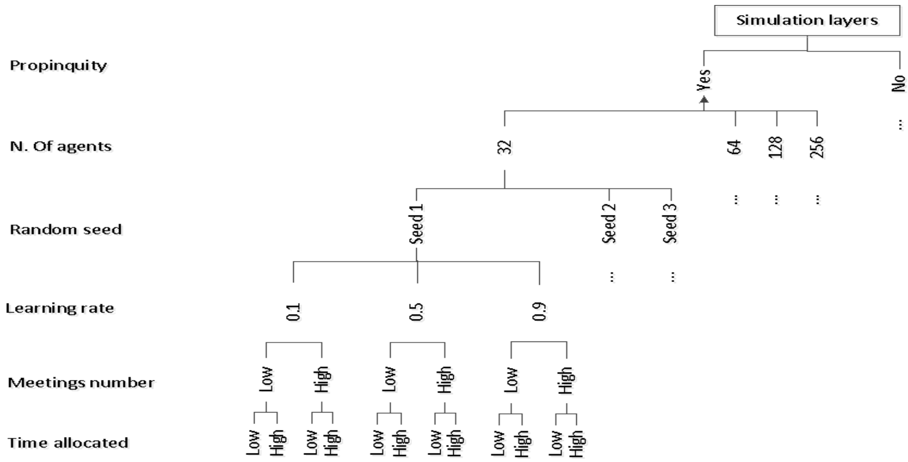


Figure 11.2: Simulations diagram

To better study the effect of physical layout on governance channel, subgroups DA, DB and DC are replicated changing the selection mechanism used for meeting attendees list definition. First three subgroups do not have any physical constraint in selecting the number of attendees (unrestricted mechanism) whereas the second three subgroups (DJ, DK, DL) have limitation in the selection of the number of attendees (restricted mechanism). The detailed mechanisms are reported in section 6.3.12. Therefore, the framework presented in figure 11.2 is replicated with the two different selection mechanisms.

The analysis of the result starts from the simulations with unrestricted selection mechanism. Figures 11.3, 11.4, 11.5 show the results of only department meeting, only project meeting and both types of meetings respectively. The figure proposed follows this structure: on the y-axis the mean value of the adherence in shown, in the x-axis the number of the simulation. The total number of simulations reported is 288 split in 4 groups of 72 simulations. This is linked with figure 11.1 which shows the four different combinations.

[4]For these simulation the knowledge superiority is always active.

Each combination is assigned to a colour: red for number of meetings low - time low combination, green for number of meetings low - time high, blue for number of meetings high - time low and purple for number of meetings high and time high.

The 72 points x-axis is split into 2 subgroups: 1-36 for simulations without propinquity channel and 37-72 for simulations with propinquity channel. Each of these two groups is spit into 4 groups according to the number of agents: 32, 64, 128, 256. Each final group of 9 x-axis's points is made by 3 replicates[5] of the final combination learning rates - meeting low/high values. This means that every combination of figure 11.1 is replicated at the three different learning rates (0.1, 0.5, 0.9).

Focusing on figure 11.3, the left part of the output is quite similar to the output found in self search (cf. chapter 9) where there is a slight dependence on the number of agents and the path linked to learning rate is almost everywhere respected. A recognizable pattern is evident when propinquity is active because the clustering effect breaks March's learning rate path (right part). Again, this is a further confirmation of the results found in chapter 10. Moreover regardless the combination between number of meetings and time allocated to them, the output are almost super-imposable. This means that meetings in the various knowledge clusters do not perturb the system so much.

A comparison could be made with the corresponding end point trends coming from self search and propinquity analysis. Picture 11.6 shows the differences in the end point values.

Plot A shows the difference between self search and self search plus department meeting simulations, Plot B reports the difference between self search plus propinquity and self search plus propinquity plus department meeting . The differences are slightly above 0, meaning that the contribution of department meeting is positive even if it does not influence so much the output of the other channels. Table 11.1 reports the average values of the differences. What could noted is that the time allocated to meeting is the most influencing parameter whereas the number of meeting give a practically insignificant contribution.

A first result could be derived: the department meeting tends to be transparent to the exploration-exploitation phenomenon in terms of final adherence.

But, sharing knowledge within homogeneous cluster tends to have an adverse effect on the time required to reach the end point. Figure 11.7

[5]Changing the random seed

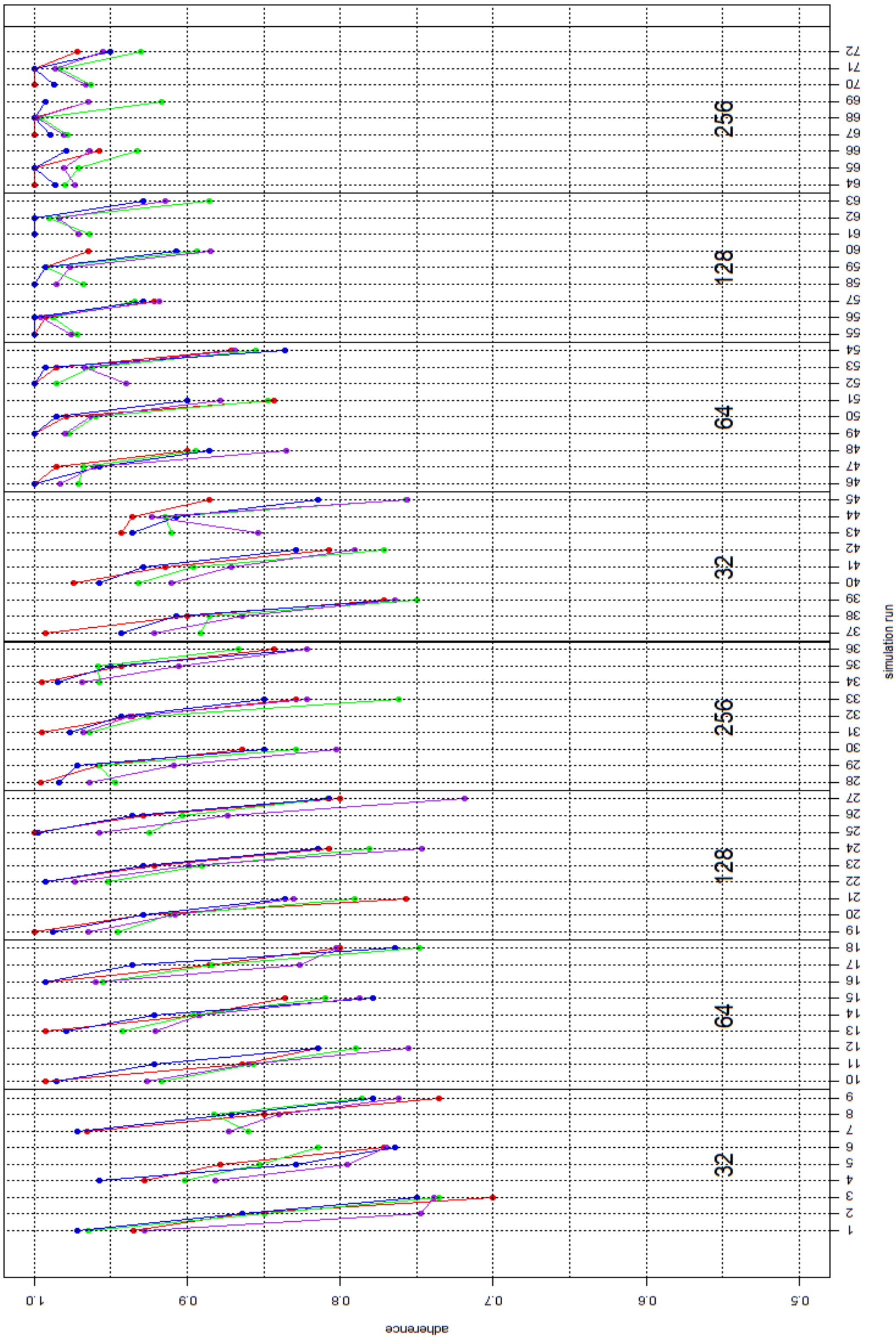


Figure 11.3: Knowledge end points for department meetings only (w/o physical constraints)
(parameter combination: red = few meetings few time, green = few meetings - lot of time, blue = many meetings - few time, purple = many meetings - lot of time)

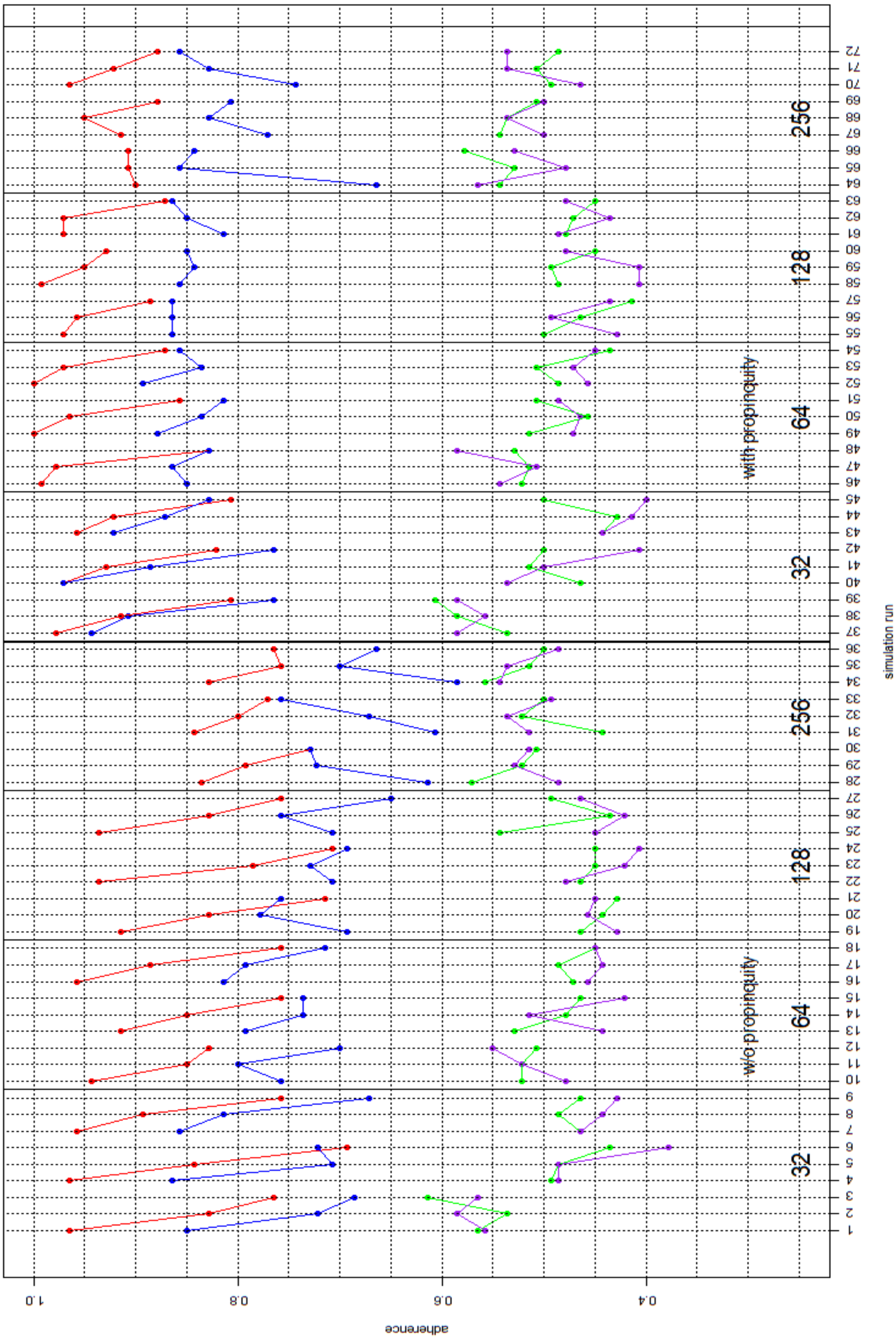


Figure 11.4: Knowledge end points for project meetings only (w/o physical constraints)
(parameter combination: red = few meetings few time, green = few meetings - lot of time, blue = many meetings - few time, purple = many meetings - lot of time)

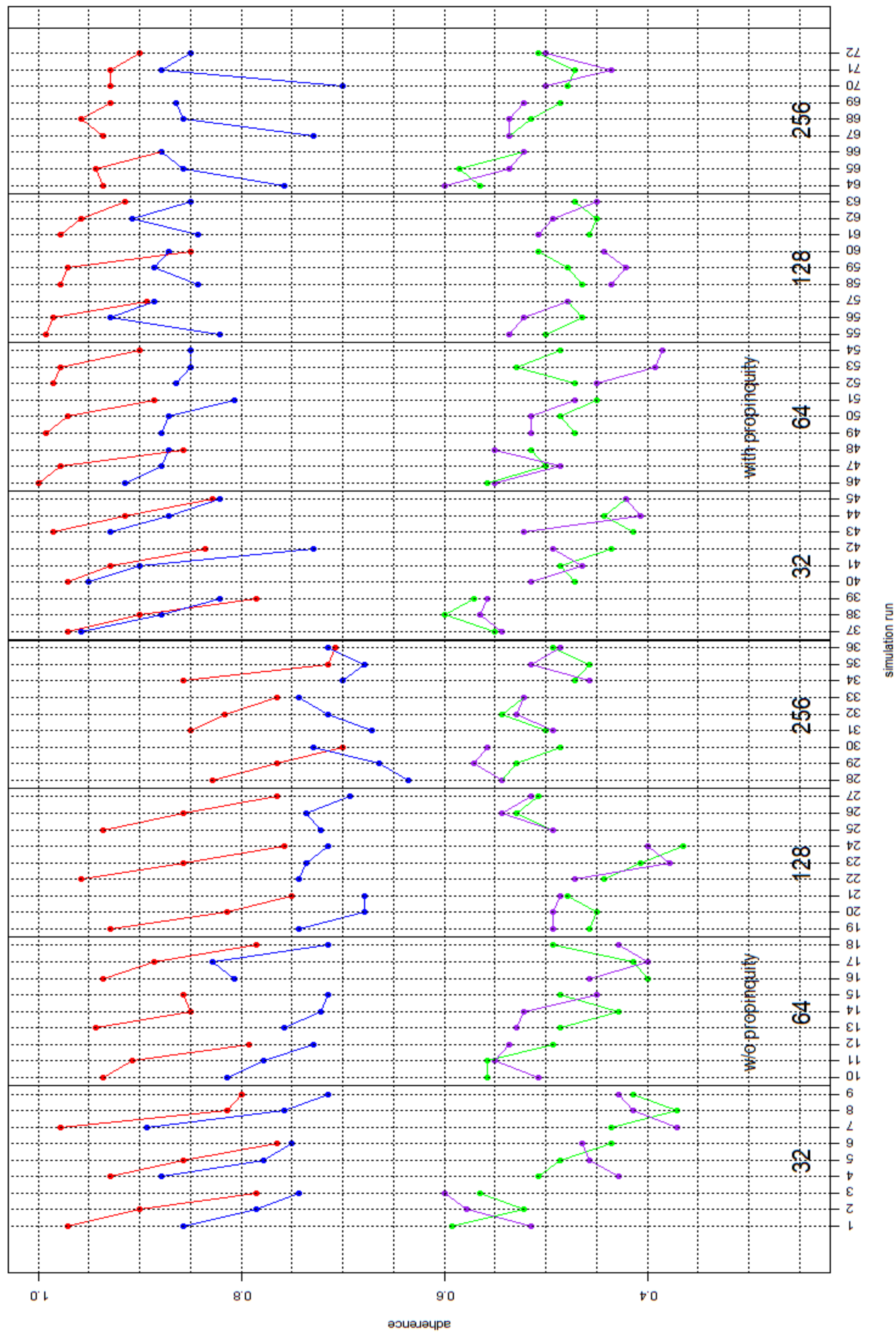


Figure 11.5: Knowledge end points for mixed meetings (w/o physical constraints)
(parameter combination: red = few meetings few time, green = few meetings - lot of time, blue = many meetings - few time, purple = many meetings - lot of time)

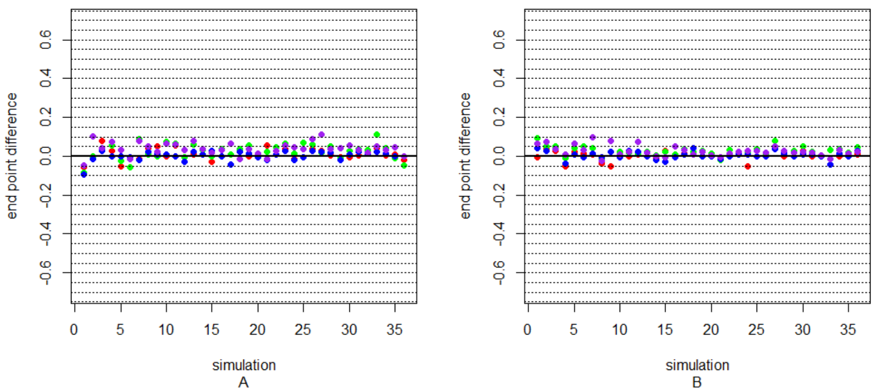


Figure 11.6: End point delta
(parameter combination: red = few meetings few time, green = few meetings - lot of time, blue = many meetings - few time, purple = many meetings - lot of time)

Table 11.1: Differences between simulations

Meeting combinations		Channels	
Number of meetings	Time allocated	Self search	Self search and propinquity
low	low	0.007	0.000
low	high	0.026	0.028
high	low	0.001	0.005
high	high	0.040	0.027

shows that, with meeting channel, the end point is reached with longer time[6] (simulations with more than 40000 run steps are needed). If a comparison is made between the self search channel and self search plus propinquity channels and their corresponding channels with department meeting active the output is pretty clear. Figure 11.8 depicts the delta, clearly supporting the idea that meeting channel is detrimental in terms of time required to increase knowledge.

Also in this case, the effect is mainly led by the time allocated to meeting. The suggestion from figure 11.8 seems to be that the most people are involved in meeting the longer it takes to increment the adherence. But in some way is also in line with the reading that department meetings actually foster clustering and then impede the equalization of knowledge in the whole organization. This results in longer time to achieve a common belief.

[6]The reported output is the adherence profile of run DF004.

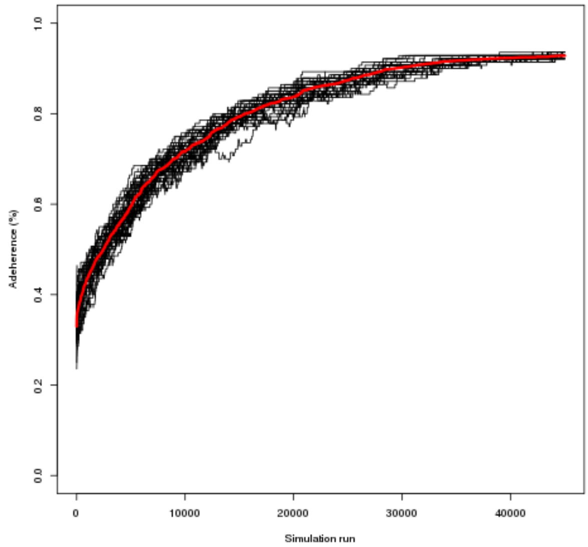


Figure 11.7: Adherence profile with department meeting active (High High combination)

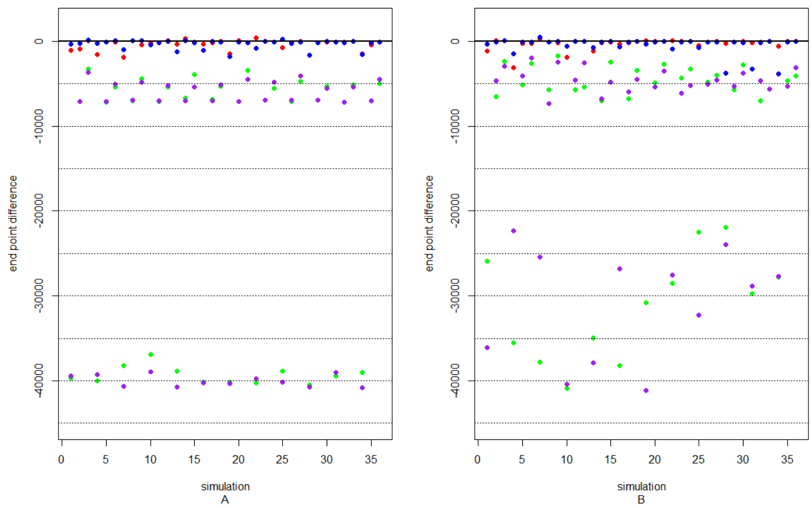


Figure 11.8: Onset point delta
(parameter combination: red = few meetings few time, green = few meetings - lot of time, blue = many meetings - few time, purple = many meetings - lot of time)

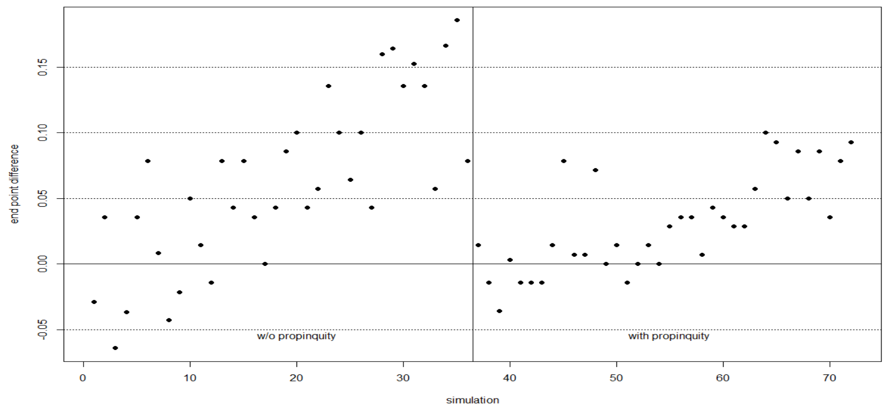


Figure 11.9: End point delta

The result related to department meeting channel is in line with the work of Clement and Puranam (Clement and Puranam 2018) since a formal structure of interaction leads to a (slightly) better performance.

Highly counter intuitive is the output of figure 11.4 where project meeting channel output is reported. Few considerations could be derived: first, considering the best results (red lines - low low combination) and comparing them with the same result obtained by simulations with department meeting channel, the final adherence is lower. This is visible in figure 11.9 where the difference between department and project meeting output are predominantly positive.

Almost all the delta are positive supporting the idea that the department meeting simulations have higher end point values. This means that even in the mildest configuration (small number of meetings and few daily time dedicated) the project meeting seems to be detrimental to the organization performance.

Second, the time allocated to meeting is able to separate the behaviour of project meeting (as it did in department channel). Here the effect is far more important than before practically splitting the chart in two families (gree and purple lines and blue and red lines). The number of meeting is responsible for the secondary division of red and blue lines. The lower the number of meetings the better the adherence.

Third March's effect is corrupted almost in every combination, project meetings tend to annihilate the effect of learning rate.

Fourth, project meeting acts as a short circuit in the organization. This is visible in figure 11.10 where the differences of the onset time between

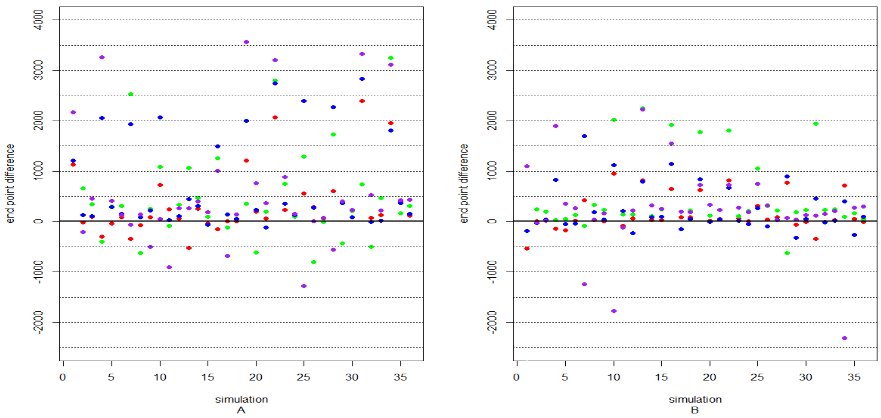


Figure 11.10: Onset value delta for project channel (Colours as per figures 11.3, and 11.4)

the self search channel and self search plus propinquity channels and their corresponding channels with project meeting active are show.

Since the values are positive means that the scenarios where the project meetings are disabled the time required to reach the end point is much longer. Combining figures 11.4 and 11.10 a hint of the mechanism could be sought. Project meeting, gathering together individuals from different departments and dealing with diverse topics, boosts the sharing of knowledge destroying the clustering effect. The exchange in a meeting is extremely effective since it is equal to $A_{meet} - 1$ exchanges. So, when department channel is active meeting exchange pushes toward clustering and when meeting channel is active it pushes toward unification.

Last consideration could be made looking at figure 11.5: the meeting exchange mechanism is more powerful in projects meeting than in department meetings. The figure shows the output of simulations where the number of department meetings is equal to the number of project meetings and the output is very similar to project channel alone (figure 11.4).

This result is instead not aligned with Clements and Puranam (Clement and Puranam 2018): a possible explanation lays in the fact that in their work the exchange does not directly impact the duration of the simulations.

Beyond the name given to the type of meetings, the most interesting thing is that exactly the same meeting mechanism gives extremely different output just changing the attendees. In the first group of simulations, only attendees belonging to the same department are allowed whilst in the second group attendees belonging to different departments are allowed. In the second part of simulations the impact on meetings channels by the physical layout is studied. To understand the impact the attendees selection mechanism needs to be changed.

11.1 Impact of physical layout

In the simulation with unrestricted selection mechanism at the beginning of every cycle (i.e. working day) every agent is assigned to a particular knowledge channel with a certain probability. For the rest of the cycle, the assigned agents to meeting channel form the panel of potential attendees at the different meetings. Hence, if for every cycle the number of meeting is N (specified by N_m^d and N_m^p), for N times the attendees are selected from the panel.

In this first group of simulations all the agents have equal chance to be assigned to the meeting channel. This leads to the binomial distribution as the driving law beneath the assignment[7]. The expected values is then Np, this means that for a simulation with 128 agents and a probability of assignment of 20%, on average the number of agents in the panel is 25. Within those 25 agents, all the attendees for the different meetings will be found.

At this point, the model decides the number of attendees of every meeting considering all the possible values from 3 to the number of potential attendees with the same probability and drawing a number out[8].

This gives the same chance to have small meetings and large meetings and large meetings do happen with a relatively high frequency. The result is that in large meetings many agents could change toward the consensus and with many large meetings, the organization could rapidly converge to a common belief. This distribution of attendees across the meeting could be thought as unrealistic.

To better understand the phenomenon, a second selection mechanism is presented (henceforth restricted mechanism). Still considering the effect of physical layout on daily work, the mechanism has been linked to the real constraint in having the meetings: the capacity of the meeting rooms. Now it is not possible to host a meeting with more attendees than the capacity of the meeting room and a list of all the characteristics of the rooms has been implemented (as explained in 6.3.12).

[7]Considering as n the number of assignment trials and considering p the probability to be assigned, the overall probability is: $P(k) = \binom{n}{k} p^k q^{n-k}$.

[8]The model samples from the uniform distribution with density: $f(x) = \frac{1}{N_i^{panel} - 3}$ for $x \in [3 \,..\, N_i^{panel}]$ where N_i^{panel} is the number of agents in the panel for the i-th cycle.

The restricted attendees selection mechanism output could be seen in figures 11.11 and 11.12 where the simulations took the same parameters as in previously studied subgroups[9].

A direct comparison of the effect on department and project meetings could be seen in figure 11.13 where the final adherence is reported for both groups of simulations. It is pretty evident that the mechanism has no effect on the department meetings but it has heavy impact on project meetings. The reason behind this discrepancy could be found in the fact that attendees to department meetings are also constrained by the affiliation. This means that the attendees are selected from the members of departments while for project meeting this does not happen. Here a couple of hypothesis could be made: first the number of attendees belonging to a particular department and assigned to department meeting is comparable to the capacity of meeting rooms. By consequence, meeting room capacity is not a big issue for department meetings. Project meetings experiences the opposite situation, the potential number of available attendees is on average higher than the rooms capacity.

Second, clustering effect brought by department meeting is stronger than physical constraints.

Anyway, once again, physical layout has a strong effect on exploration exploitation problem. Most importantly, not only autonomous search is impacted by physical layout, also governance is impacted by it.

Interesting to report is also the comparison of the onset values for meeting projects (figures 11.14). Not only the office layout does impact the end point values but also the onset points.

The restricted selection makes one group of simulation faster (few meetings - lot of time), one slightly faster (many meetings - lot of time) and the other two slower (few meetings - few time and many meetings - few time). Hence, summarising, the restricted selection mechanism is able to interact with the time dedicated to meetings. Anyway, its overall effect is still present.

11.2 Network analysis

Network analysis for meeting channel starts with the evidence that density profiles behave differently from the density profiles in the autonomous channels already studied (self search and propinquity)[10]. Picture 11.15 summaries

[9] The simulations group with both meeting types has not been run since considered not fundamental for the output analysis

[10] For network analysis the subgroups with only department meetings and only project meetings are chosen. Both are with restricted selection mechanism. This selection is made

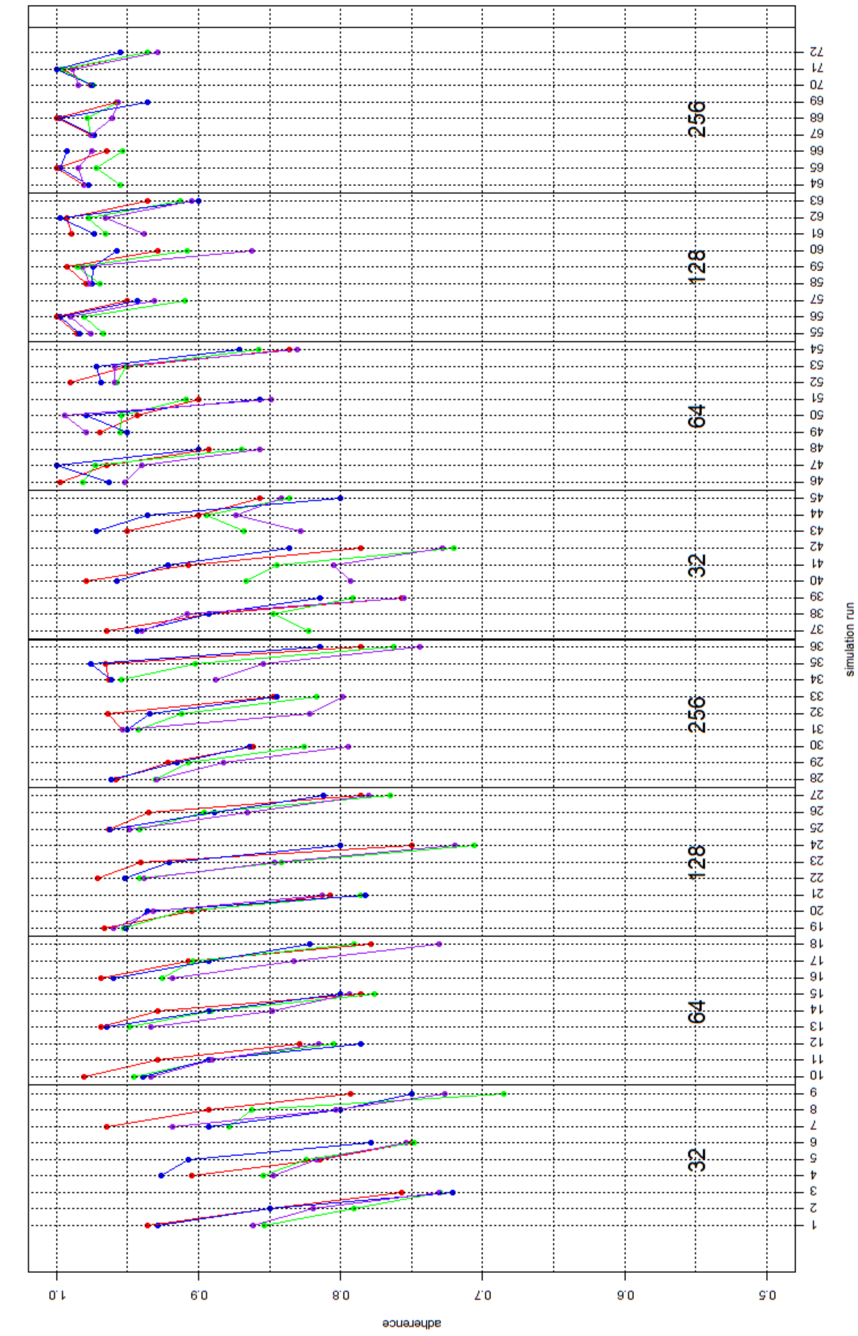


Figure 11.11: Knowledge end points for department meetings only (restricted mechanism)
(parameter combination: red = few meetings few time, green = few meetings - lot of time, blue = many meetings - few time, purple = many meetings - lot of time)

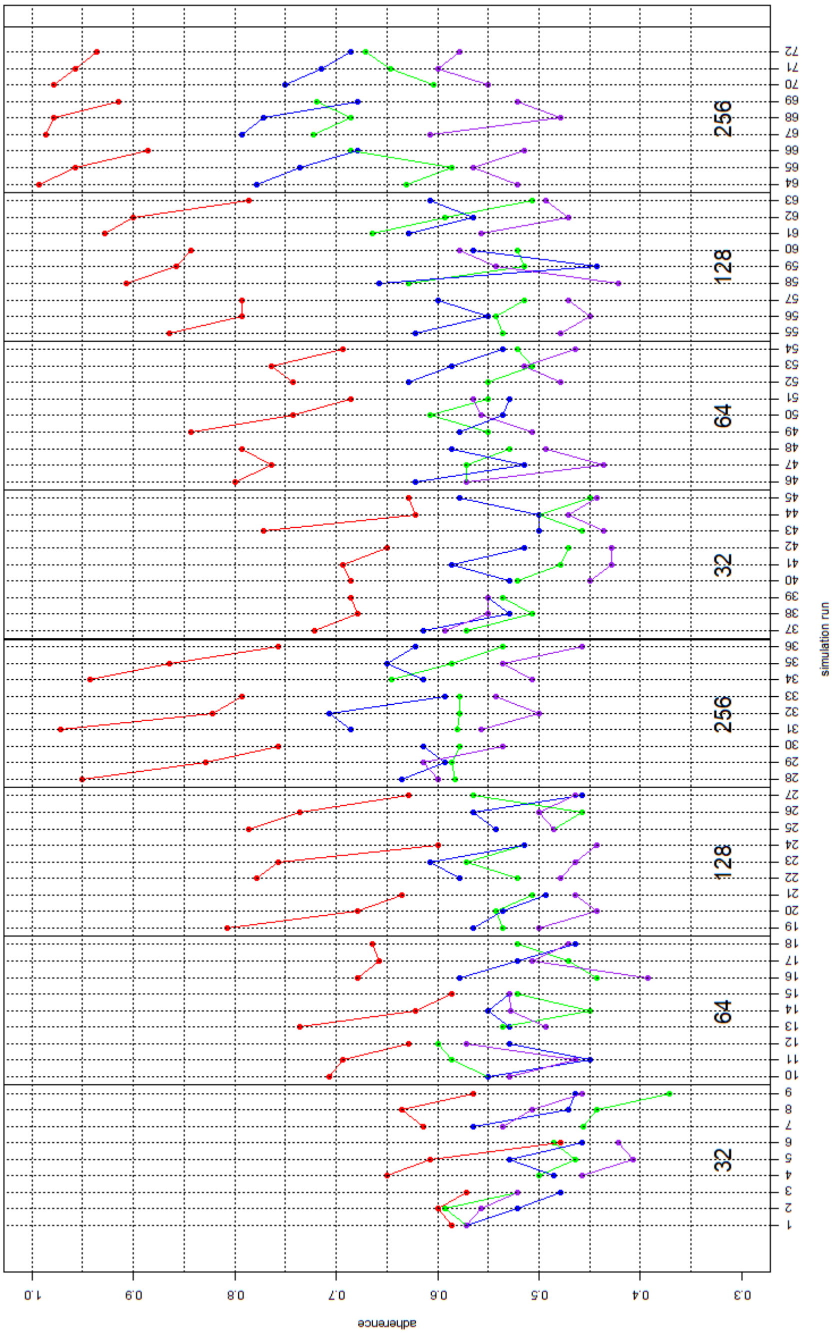


Figure 11.12: Knowledge end points for project meetings only (restricted mechanism)
(parameter combination: red = few meetings few time, green = few meetings - lot of time, blue = many meetings - few time, purple = many meetings - lot of time)

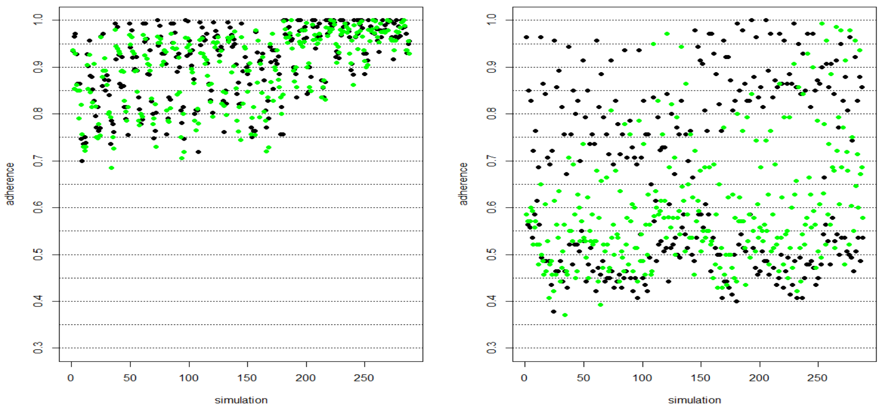


Figure 11.13: New selection mechanism impact
(black = unrestricted mechanism, green = restricted mechanism)

all the profiles for the department meetings and figure 11.16 reproduces all the profiles for the project meetings. Both figures have 8 different plots according to table 11.2.

Table 11.2: Characteristics of density subplots

Subplot	N. of meetings	Time allocated	Propinquity
A	low	low	no
B	low	low	yes
C	low	high	no
D	low	high	yes
E	high	low	no
F	high	low	yes
G	high	high	no
H	high	high	yes

Figures are clustered by the combination among time allocated to meeting, number of meetings and presence of propinquity. Each figure (A-H) encompasses the trends at different learning rates. Simulations with different learning rates show different run steps: when fast learners are present the run is shorter whereas when the slow learners are present the run is longer. This justifies the presence of curves of different lengths in the figures. Focusing, just for clarification, on figure 11.16-C, the green curves are shortest

without any loss of information.

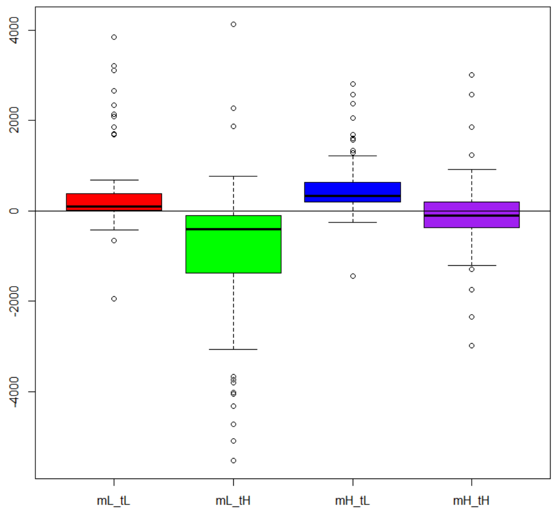


Figure 11.14: Onset values comparison
(parameter combination: red = few meetings few time, green = few meetings - lot of time, blue = many meetings - few time, purple = many meetings - lot of time)

since are the ones with fast learner, the black ones are the longest since slow learners are present. This is applicable to both the figures 11.15 and 11.16.

From the first look it is extremely evident that the two types of meetings have different influence on the network density trend. Generally speaking, meeting channel has the tendency to extend the overall density over time, during the simulation. This is, in some way, expected considering the meeting mechanism. The big difference among the self searches and the meeting relies exactly in the contrast among autonomy and governance. In the self search, the agent is the promoter of knowledge exchange whereas in the meeting channel the agents are called to share the knowledge. Indeed they are not forced to share but surely they are invited to interact with people whom they would not necessarily contact in a self search. Meetings are one of the most evident realization of governance, so it could be said that governance tends to keep agents linked longer than in self search.

But this is not the only difference. The two types of meetings have a different influence also on the maximum value of the density. Department meetings do not change the values of the maximum. A comparison of maximum values in figure 11.15 with figure 9.12 reveals that they are comparable. This means that the first type of meeting does not contribute heavily in the density of the network.

Vice versa, project meeting has a dramatic effect on network density. The peak is much more higher than in other configuration reaching, in some cases, 0.7[11]. Cross-functional meeting where different topics are discussed seems to created denser networks.

A possible explanation could be the following: in department meeting scenario the self search is the only way for an agent to meet people of different departments. For topics not belonging to agent department, she has to rely on autonomous search. When dealing with their own knowledge topic, they meet with internal SME[12] and they learn from them. Increasing the number of internal meeting, the frequency of meeting with internal SME obviously increases. By consequence all agents tend to link to the SME and then the density si low.

When dealing with a cross-functional project meeting agents meet all the SMEs in the organization because they could be involved in meeting with topics of any kind. This increases the chance to create "forced" linked they would never create in the autonomous search. This forced relationship faster the transfer letting the system to reach the equilibrium earlier and then at lower level but keeping more links.

It seems that department meeting acts as a slow learning entity and project meeting as a fast learning entity.

A general consideration could be that cross functional meetings have denser networks but they perform less better than intra department meetings. Kunz,in 2011, argued that *"explorative and exploitative processes on the group level are significantly affected by the possibility of interactions between groups"* (Kunz 2011, 6.2). In particular she also found that *"in case of high complexity and broad knowledge pools, within-group exploration performs better than between-group exploration that increases its performance level very slowly"* (Kunz 2011, 5.13).

Another interest point to highlight is the ability of department meeting to preserve the density clustering by number of agents. By contrast, project meeting could cluster by learning rate. Figure 11.17 shows the maximum values of the different network densities split as shown in table 11.2. One the y-axis the density maximum value is reported and in the x-axis the 12 groups of data are reported. These groups are formed by the 4 values of number of agents and the three values of learning rates[13]. What is visible is that the lower the number of agents, the higher the maximum value. This is true for

[11]Picture 11.16-F.

[12]Subject matter expert.

[13]The value 64_5 means 64 agents and learning rate equal to 0.5

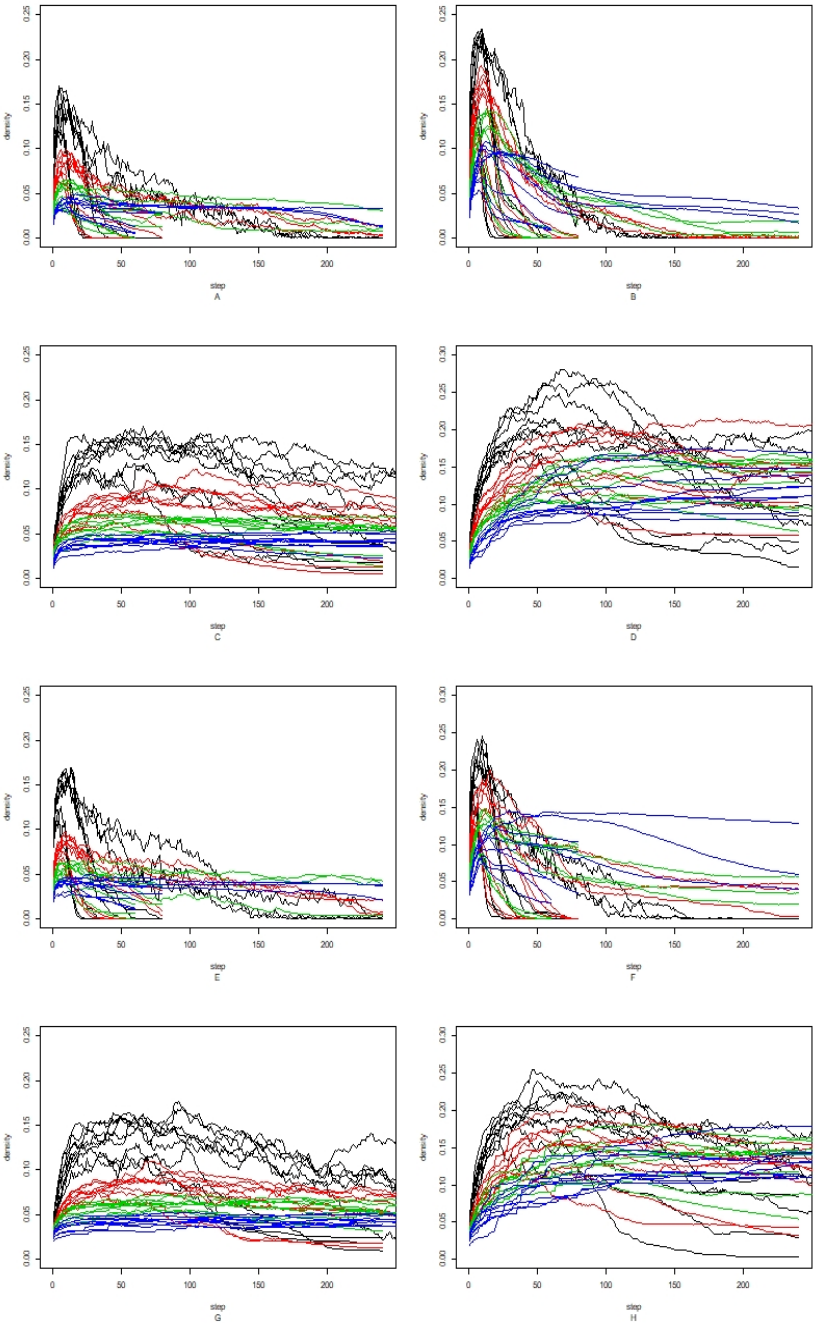


Figure 11.15: Department meeting densities
(Number of agents: black = 32, red = 64, green = 128, blue = 256)

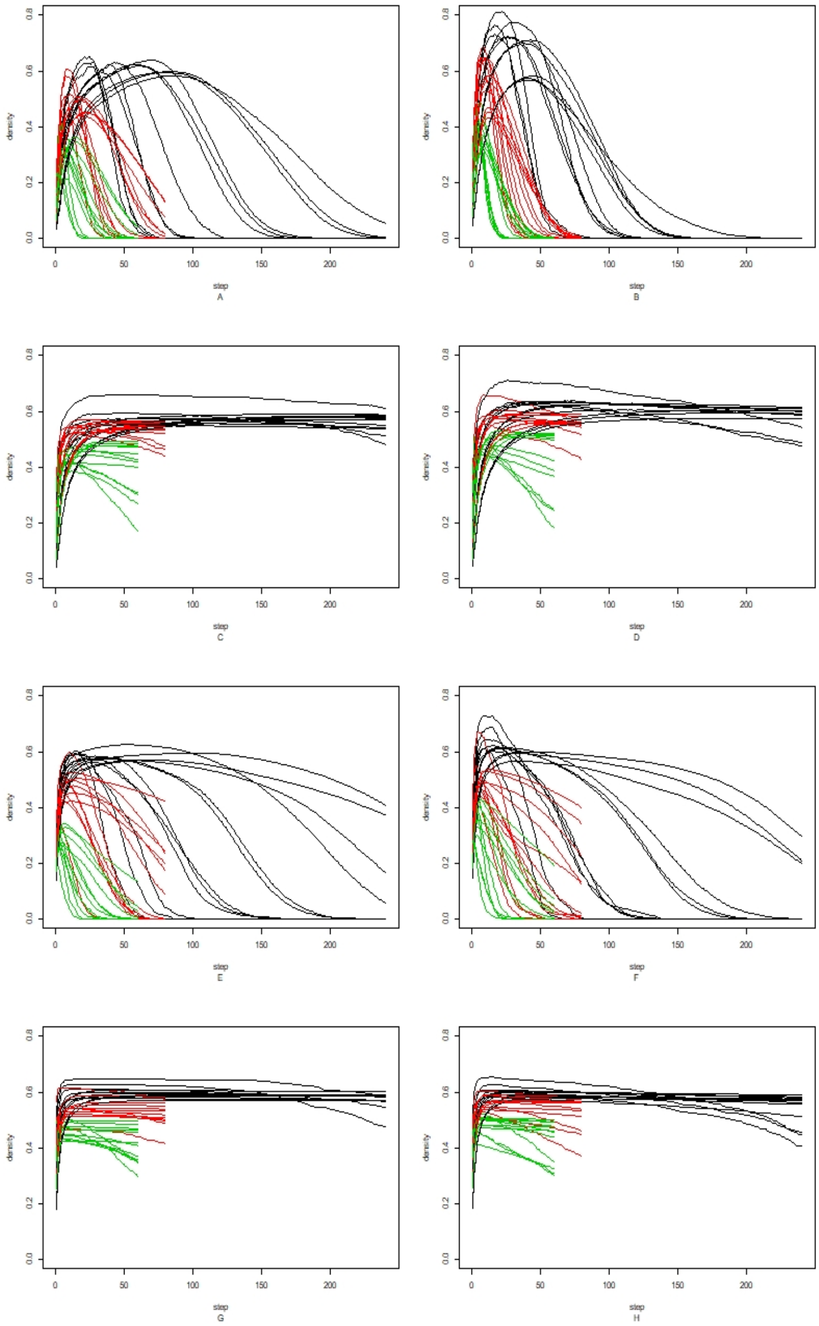


Figure 11.16: Project meeting densities
(Learning rates: black = 0.1, red = 0.5, green = 0.9)

subplot A, D, E, G where propinquity is not active. When propinquity is active (subplot B, D, F, H) the trend is less marked but still presents.

Figure 11.18 has the same structure but is shows the output of the project meeting simulations. Three are the colours according to the learning rate used in through the runs. Here the upper boxes are belong to 0.1 learning rage, the middle layer to 0.5 and the lower layer to 0.9 learning rate.

While the department meeting is present, the structure of the network is more similar to the ones found in the previous chapters (refer to chapters 9 and 10) whereas when project meeting channel enter in the simulation, the structure changes and also the cluster effect changes.

The results on network density could be complemented referring to Clement and Puranam work on formal organization design (Clement and Puranam 2018). The authors argue about the benefit in deploying a formal organization with a top-down approach rather than formalizing the spontaneous connections among individuals as a bottom-up process. Spontaneous connections could suffer of two distinct issues: error of omission and error of commission. In the first case individuals does not connect with the interlocutor she depends on and in the second case individual does connect with the wrong interlocutor. Hence the main finding of the work is that imposing a formal structure the number of connections are continuously recovered[14]. Hence, as the authors argued, the formal organization *"provide a common frame of reference, which helps agents regenerate the interconnection"* (Clement and Puranam 2018, 3887). The output found in the present work partially agree with this finding. A formal structure keep the interactions alive and this could be sought in the extremely high values and slow decay of network densities. By contrary, the performance in case of a formal structure such project meeting does not lead to a performance comparable to situations where there is no formal structure.

The aforementioned effect on the network structure should also be appreciable and figure 11.19 gives an insight on that showing the Q indexes[15] of the subgroups of simulations. The subplot A represents the Q values for 8 groups made by the combinations of number of agents and presence of propinquity[16]. The subplot is then split into 2 parts: the leftmost one (groups 1-4) representing the simulations with department meeting, self search but with no propinquity and rightmost one (groups 5-8) representing the sim-

[14]In the Clement and Puranam's model (Clement and Puranam 2018), the connections break at the injection of a new project and the authors study the level of recovery changing the way formal structure is imposed.

[15]As calculated in section 9.2.

[16]64_n means 64 agents an no propinquity, 128_p means 128 agents and presence of propinquity.

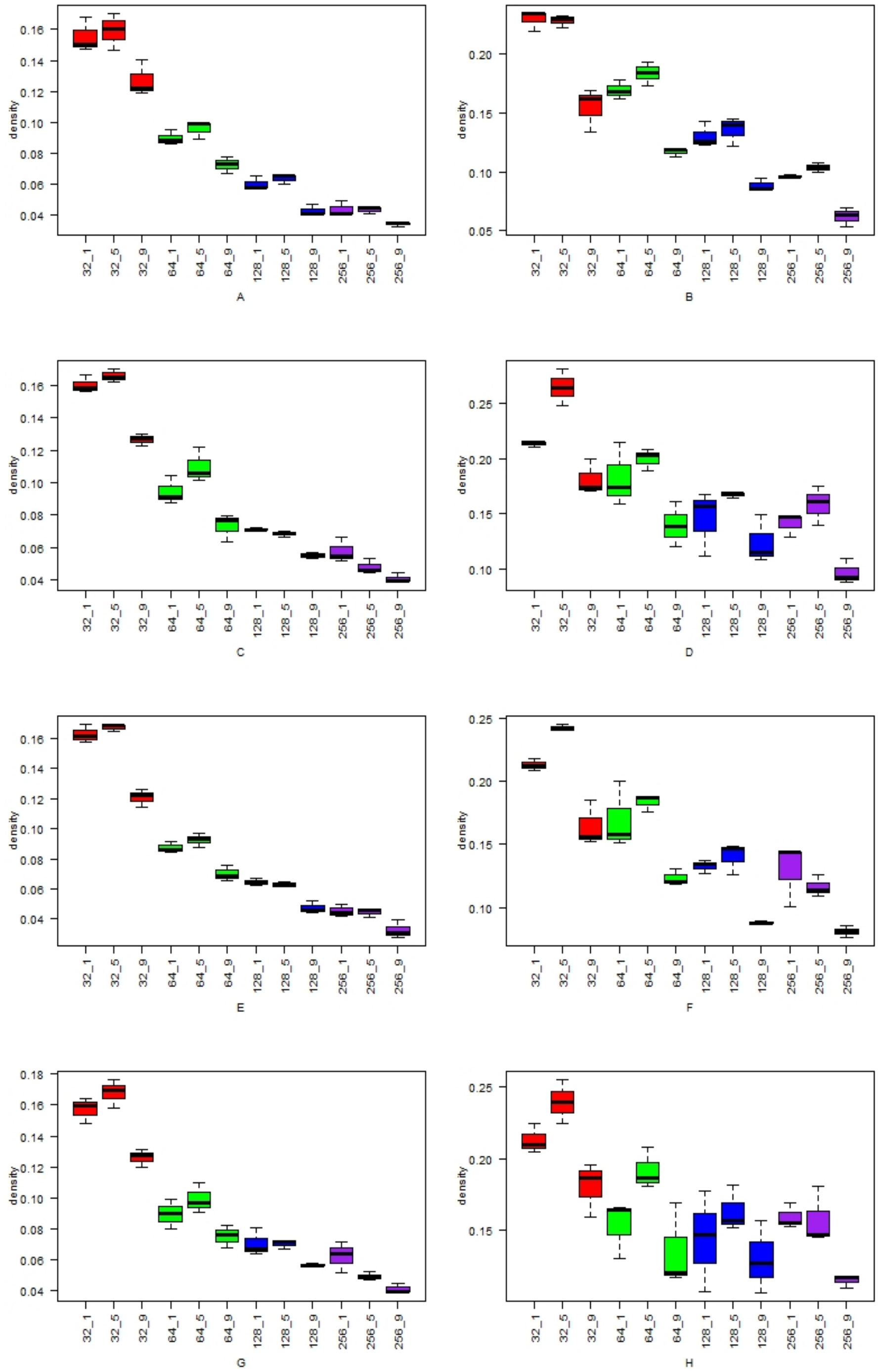


Figure 11.17: Department meeting densities
(Red = 32, green = 64, blue = 128, purple = 256)

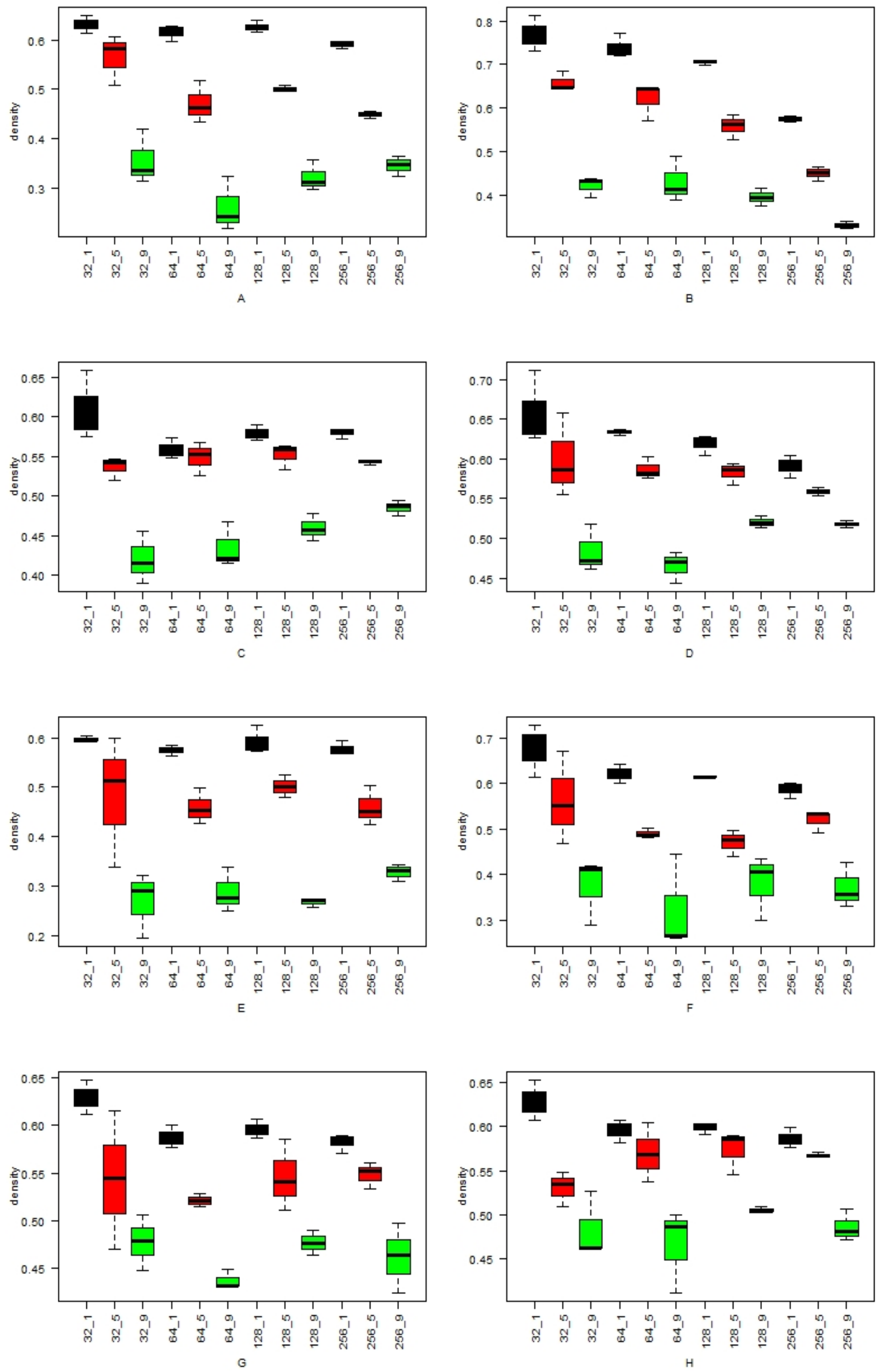


Figure 11.18: Project meeting densities
(Black = 0.1, red = 0.5, green = 0.9)

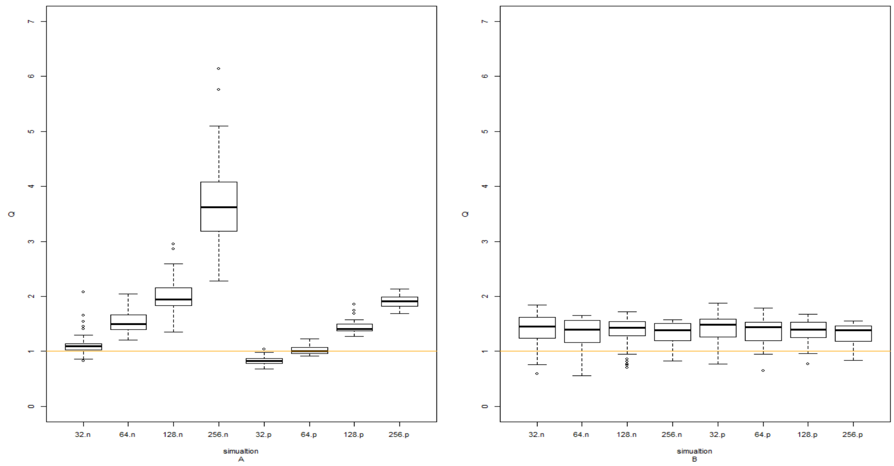


Figure 11.19: Q index for meetings

ulations with department department meeting, self search and presence of propinquity.

Two considerations could be done: first the effect of propinquity dampens the cluster effect as consequently the small world effect. Second, comparing this figure with figure 9.17 it is possible to state that the small world effect is enhanced thanks to the department meetings. This is particularly true considering that the 256 agents' simulations have a Q index even greater than the 2560 agents simulations with solely self search (cf. figure 9.17).

Indeed this is a confirmation that the department meeting stresses the clustering and then the small world effect, if it is combined with the self search.

Different is the situation in subplot B. Here, project meeting effect is depicted. The small world structure is negligible, and the increasing effect with number of agents is absent since all the simulations have the same Q index values around 1. This is a confirmation that cross functional teams tend to destroy the small world configuration.

This is also evident form figure 11.20, comparing the two distributions of in-degree values for a simulation with only department meeting and a simulation with only project meeting[17] (left side department meetings - right side project meetings).

The distribution of simulations accounting department meeting is extremely skewed on the left with the vast majority of agents owning a small

[17]The two simulations have all the other parameters equal.

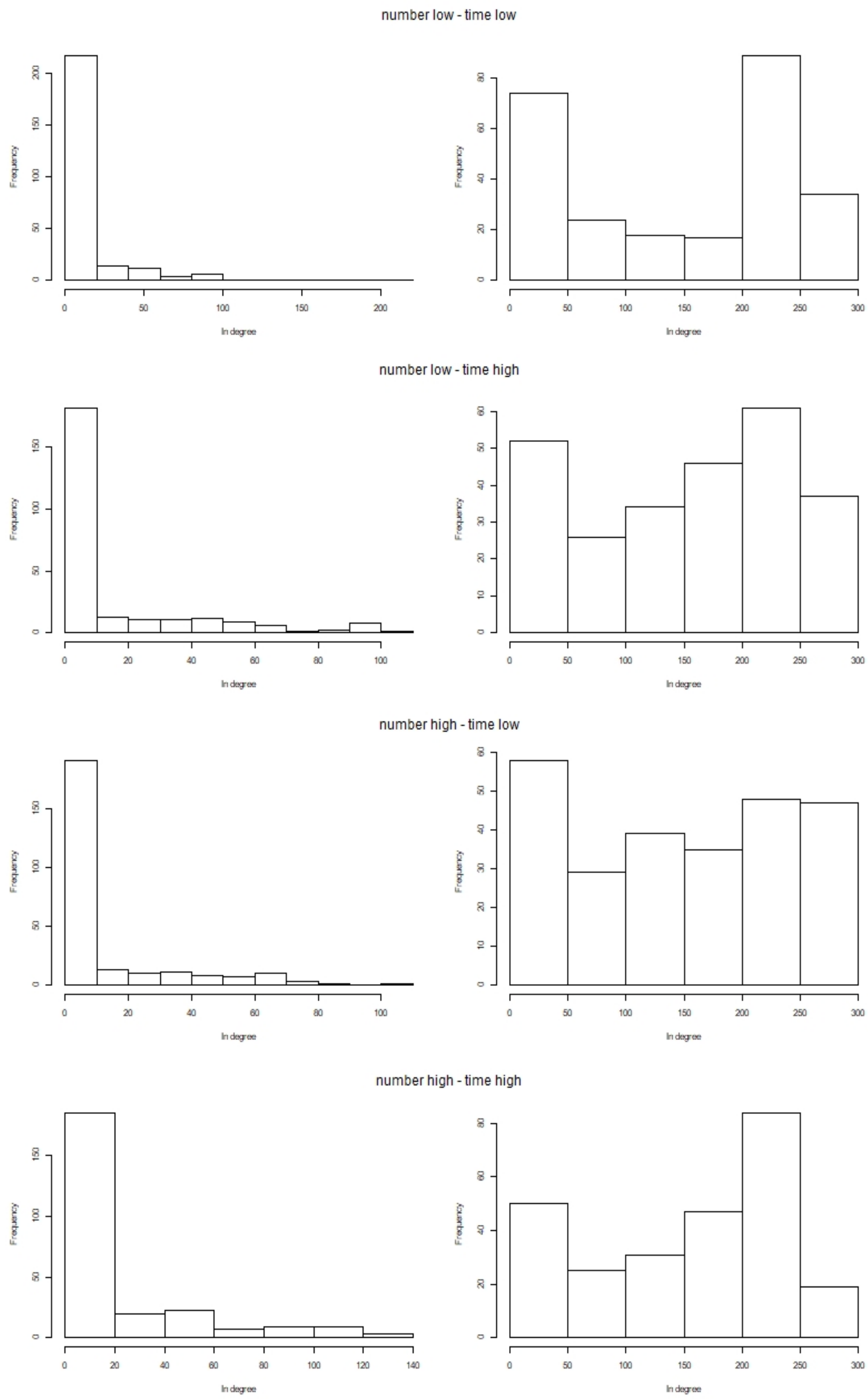


Figure 11.20: In degree values distribution
(left: department meeting, right: project meeting)

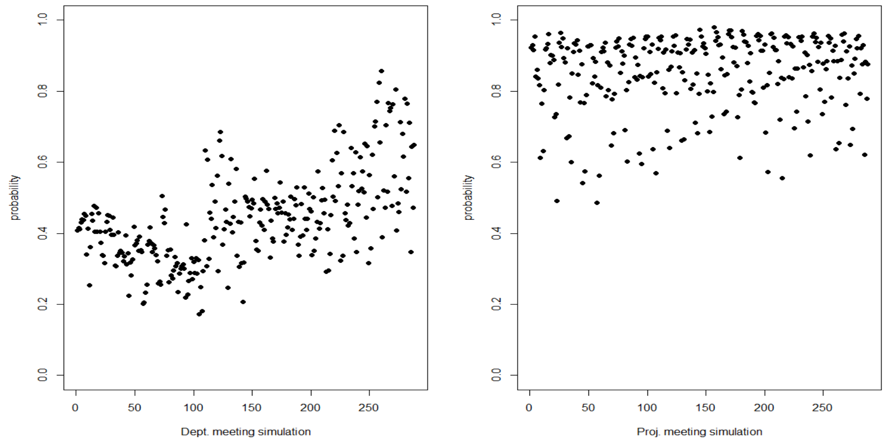


Figure 11.21: Cluster coefficient

number of incoming contact and few agents owning a large number. The corresponding distribution for project meeting is more flat. Moreover, the effect is present regardless the combination between number of meetings and time allocated; it seems an effect of the type of meeting, not coming from the intensity of its execution.

Another confirmation comes from cluster coefficient figure 11.21[18]. The figure is usually split in two: on the left the department meeting channel graph and on the right the project meeting channel part. Each subplot shows the 288 coefficient of the related subgroups and the y-axis shows the probability that adjacent vertices of a vertex are connected: the higher the probability the lower the clustering effect. Again, the department meeting has higher probability to create cluster than project meeting.

A word is worth on the mixed situation. Looking at figures 11.4 and 11.5 it is possible to highlight that the effect of project meeting is stronger than the effect of department meeting since the results are extremely similar. This is also corroborated by figure 11.22 where the in-degree distribution and the cluster index are reported for the simulations having both meeting channels active. The in-degree distribution and the clustering coefficient values are very similar to the values owned by the project meeting simulations.

A final interesting aspect is related to network energy (figure 11.23) when department and project meetings channels are involved. Generally speaking, meeting channel highly increases the energy if the values are compared with

[18]Coefficient cluster is estimated with CC_L index.

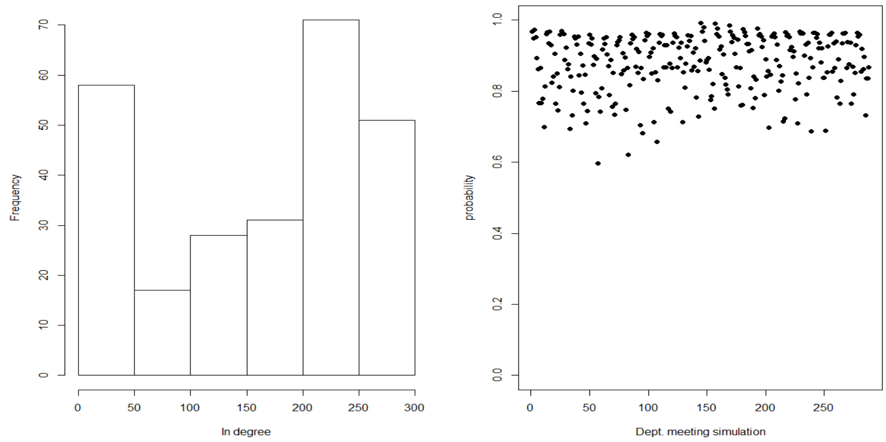


Figure 11.22: Effect of mixed meetings on network

those of solely self search active (figure 9.18). This implies that meeting channel is responsible for increasing the energy needed to perform the task. Furthermore, a dependency to the number of agents is introduced, dependency not present in the self search alone scenario. Probably this dependency is linked to the fact that increasing the number of agents also the number of exchanges through meetings increases. Also, the presence of propinquity increases the energy required demonstrating that adding channels the cost of the task increases. Finally, meeting project channel has significantly higher values in energy. After all, maintaining a denser network for longer time needs higher energy.

Levine and Prietula (Levine and Prietula 2012) argued about the cost of exchanging knowledge modelling different paths: self learning, embedded exchanges, performative ties and market exchange. The similarity with the presented work lays in the presence of an exchange trough social network of colleagues (local search) and in the presence of a global search through not intimate colleagues. But the presented work in some way contrasts their output which stresses that the performances are low when self search is involved and are higher in performative ties and even higher in embedded exchanges. There are a couple of important differences among the models that could explain this discrepancy: first, performance is differently defined. In Levine and Prietula the cost (linked to performance) is defined more as a delay in achieving a task rather than as knowledge acquired (and relative cost in terms of energy to acquire it). Second, in Levine and Prietula's work the channels are not modelled together, each channel is run separately. The concept of

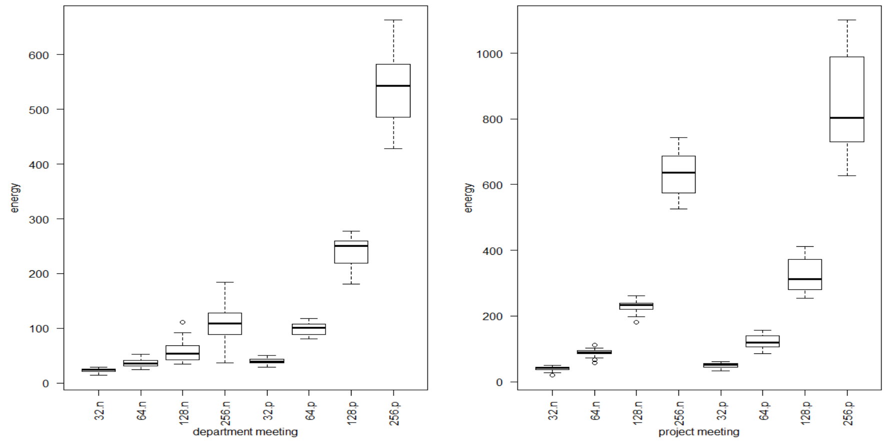


Figure 11.23: Meeting network energy

energy/cost associated to more complex task to achieve is not applicable.

11.3 Hierarchy

Beyond popularity, also hierarchy could lead meeting in a real organization. All the simulations analysed before do not have hierarchy activated. The meeting belief is then created starting from the popularity of the attendees, as explained in sections 5.5 and 6.3.12. In this part of the simulations, the hierarchy is introduced and it comes before popularity in the creation of meeting belief. Hierarchy could overcome popularity when a meeting output need to be found. Hierarchy represents here any agents who has the power to force an output of a meeting: she could be the manager or she could be the project manager. In the simulation some agents are randomly promoted to higher rank[19] and they could influence the output of a meeting.

The simulation architecture is the same as in the previous section and figure 11.24 reports the end points obtained for the department meetings. Figure 11.25 shows the end points for project meetings. It is noteworthy a similar behaviour as the simulations without hierarchy: the two figures show similarities with figures 11.3 and 11.4.

This is even more evident in figure 11.26 where the differences among simulations with and without hierarchy are reported. It could be said that,

[19]4 in 32 agents simulations, 8 in 64 agents simulations, 16 in 128 agents simulations and 32 in 256 agents simulations equal to 12.5% of the population.

on average, the presence of hierarchy does not influence so much the end point of the organization since values are randomly displaced around 0.

An interesting effect could be noted for the combination low time - low number of meetings in project meeting scenario (red lines in figure 11.25). Increasing the number of agents, the overall adherence increases. In this case the hierarchy seems to be beneficial.

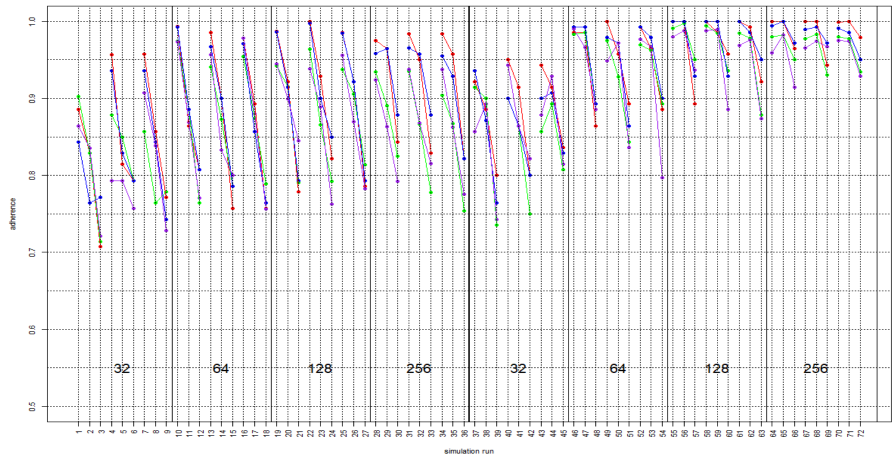


Figure 11.24: Department meeting end points (Colours as per figures 11.3, and 11.4)

Both department meeting and project meeting do not behave much differently in terms of network structure, when hierarchy is present. Figure 11.27 reports the differences in density values for both channels at different combinations of number of agents and presence of propinquity. Hierarchy has actually a tenuous effect on network density as reported in left subplot. In particular, when only department meetings are active, the difference among simulation with and without hierarchy is slightly positive then the presence of hierarchy creates denser network. This was in some way expected considering that department meeting tends to cluster and the hierarchy could introduce additional points of aggregation within every department. Vice versa in the project scenario networks are slightly sparser and around 0. An interpretation could be that the presence of hierarchy tends to rationalize the presence of SME (Subject Matter Expert) and to create less interactions. Of course, the effect is minimal but it is anyway interesting to appreciate the two different trends.

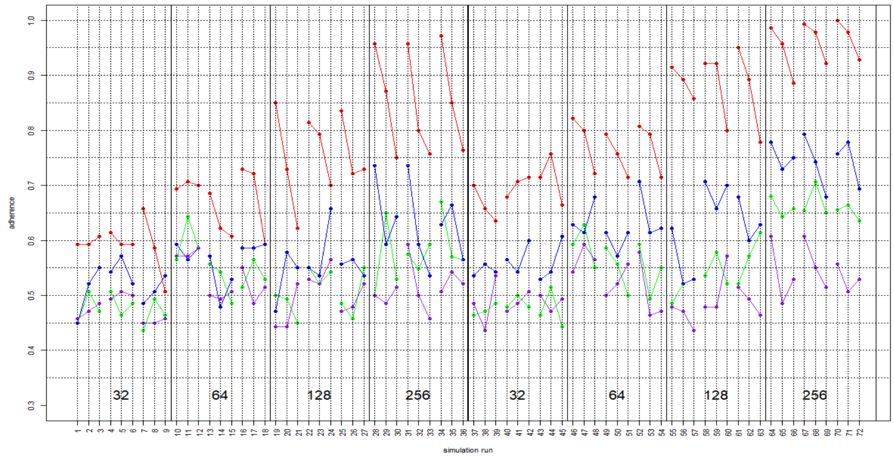


Figure 11.25: Project meeting end points (Colours as per figures 11.3, and 11.4)

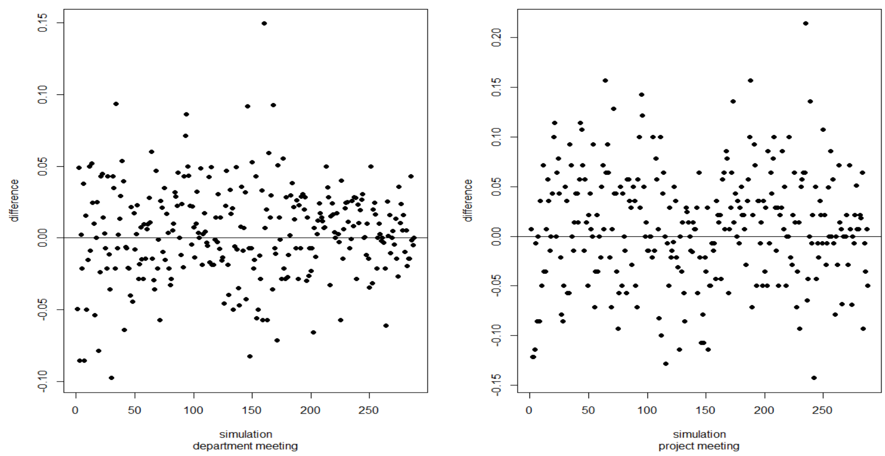


Figure 11.26: End point difference

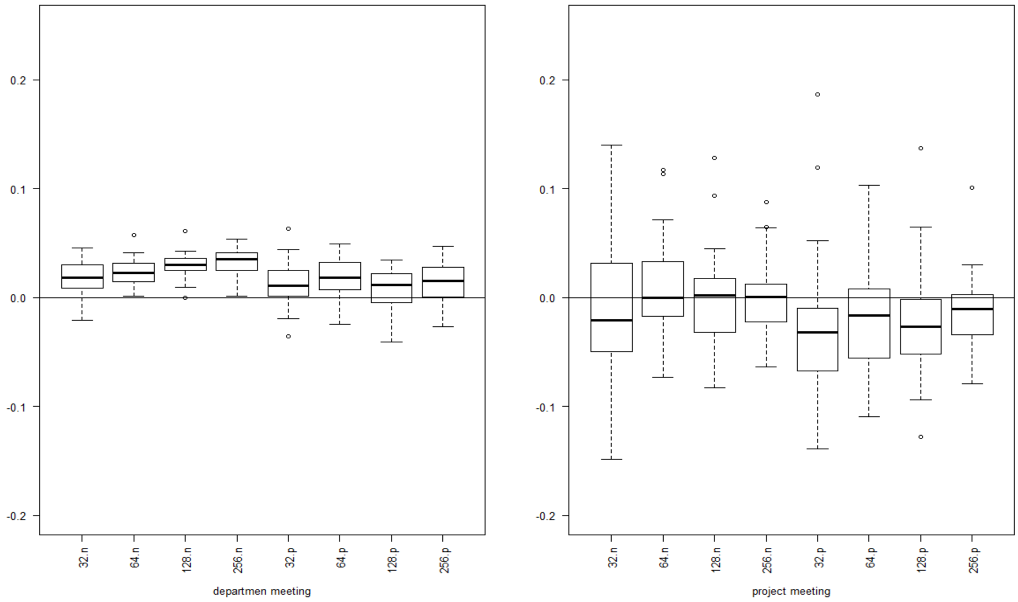


Figure 11.27: Network density boxplot

Chapter 12

Friends

Affective channel is the third and last autonomous way agents have to seek knowledge. As already mentioned in the general description of the model, informal channel is the only part which relies on a preconstituted network of relationships[1]. Individuals belonging to the organization used as prototype were asked to list their colleagues considered friends. The information was retrieved through a dedicated survey. The output has been transformed into an adjacency matrix to be used as network. The resulting network is used in the simulations as giving all the possible links an agents could reach through the informal way. Picture 12.1 indicates the output of the request done to the individuals.

Different scholars studied the structure of the friendship networks. For example, Amaral et al. (Amaral et al. 2000) studied different classes of small world networks and they found that friendship is one of them. Some years later, Singer et al. (Singer, Singer, and Herrmann 2009) tried to develop in silico the formation of a friendship network finding that it naturally tends to a small world structure. Hence, before starting with the simulations involving informal channel a preliminary study of network characteristics was performed. The already mentioned Q factor[2] for the network is equal to 4.64 (CCr = 6.50 and PLr = 1.09), suggesting a small world configuration. The network built for the model confirms what found in literature.

Friendship is simulated following the diagram in figure 12.2 and it is replicated two times. In the first case only informal channel is simulated whereas in the second also self search is added (mimicking the approach used in chapter 10). Only self search is activated presuming a small effect of

[1]Office layout network is not considered a proper relational network.
[2]cf. Uzzi and Spiro 2005.

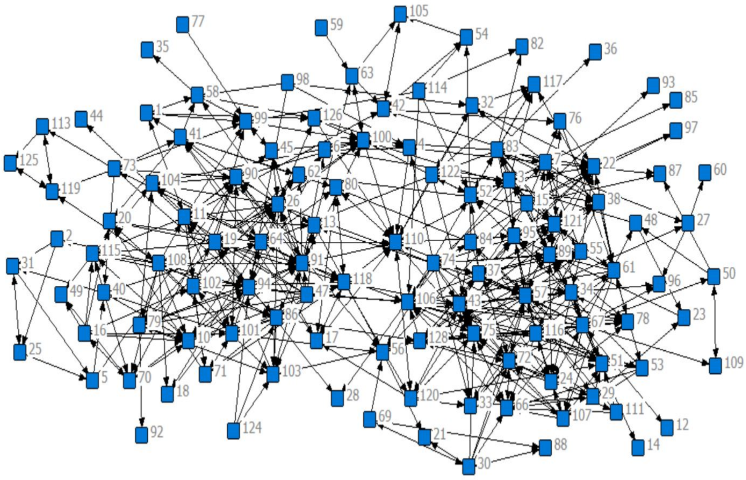


Figure 12.1: Friendship network

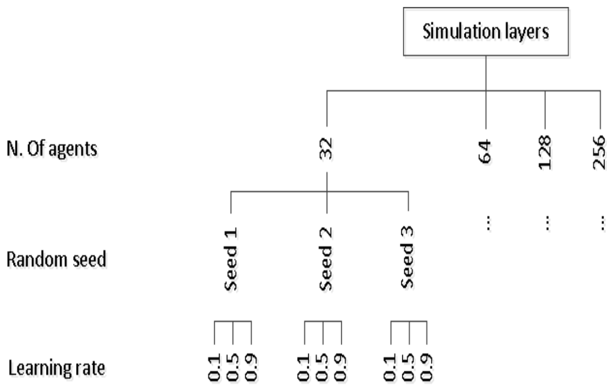


Figure 12.2: Friendship simulation diagram

informal channel. To be sure to detect the output propinquity and governance channels are switched off. The simulation schema is simpler than the ones used for propinquity and meeting channels and aims to study the effect at the usual three learning rate levels (0.1, 0.5 and 0.9), with four different number of agents (32, 64, 128 and 256). Everything replicated with three different random seeds. This entails 72 simulations.

The complete series of end point is reported in figure 12.3. The figure is split in two parts according to the simulation replicates but the structure is the same as already presented in previous chapters. It is pretty evident the different behaviour when self search is active or not: if only friendship is switched on, the end point values are low, when also self search is activated the end point values are high. The explanation could be found in two points: the absence of knowledge superiority mechanism in the friendship channel and the sparse connections among individuals. This hypothesis is supported by the output discussed in chapter 10.

Even if the output seems fairly plain, a couple of comparison could be made. The first obvious comparison is against the self search alone results in order to understand whether or not a second channel like friendship boosts the end point values as happens with propinquity. The second comparison is against propinquity simulation with superiority not active, to further discuss about network structure impact on exploration exploitation phenomenon.

The first comparison is shown in figure 12.4 and it is visible the weak contribution given by friendship. The two series of end points are well overlapped and the only effect that could be mentioned is the March's effect slightly sullied by the superimposition of the second channel. Just for comparison also the effect of propinquity on self search is reported (blue lines).

Friendship channel seems to have an effect on the onset of plateau of end point values as reported figure 12.5. The subplot A reports a scatter plot among self search onset values and self search plus friendship onset values. The interesting effect is that the second channel seems to harmonize the time needed by the organization to reach the end point. It slows down the fast and mid learner organization and it accelerates the slow learner organization. This is also shown in table 12.1 where it is visible the shift of the three means and the spread of the standard deviation.

Since friendship does not have knowledge superiority active during exchange it might be thought that the effect is caused by this absence. To assess this point the same comparison with propinquity channel is done. Part of the simulations were made without knowledge superiority mechanism[3] and subplot B (figure 12.5) shows the scatter plot of self search with

[3]subgroup CI, refer to table 7.1.

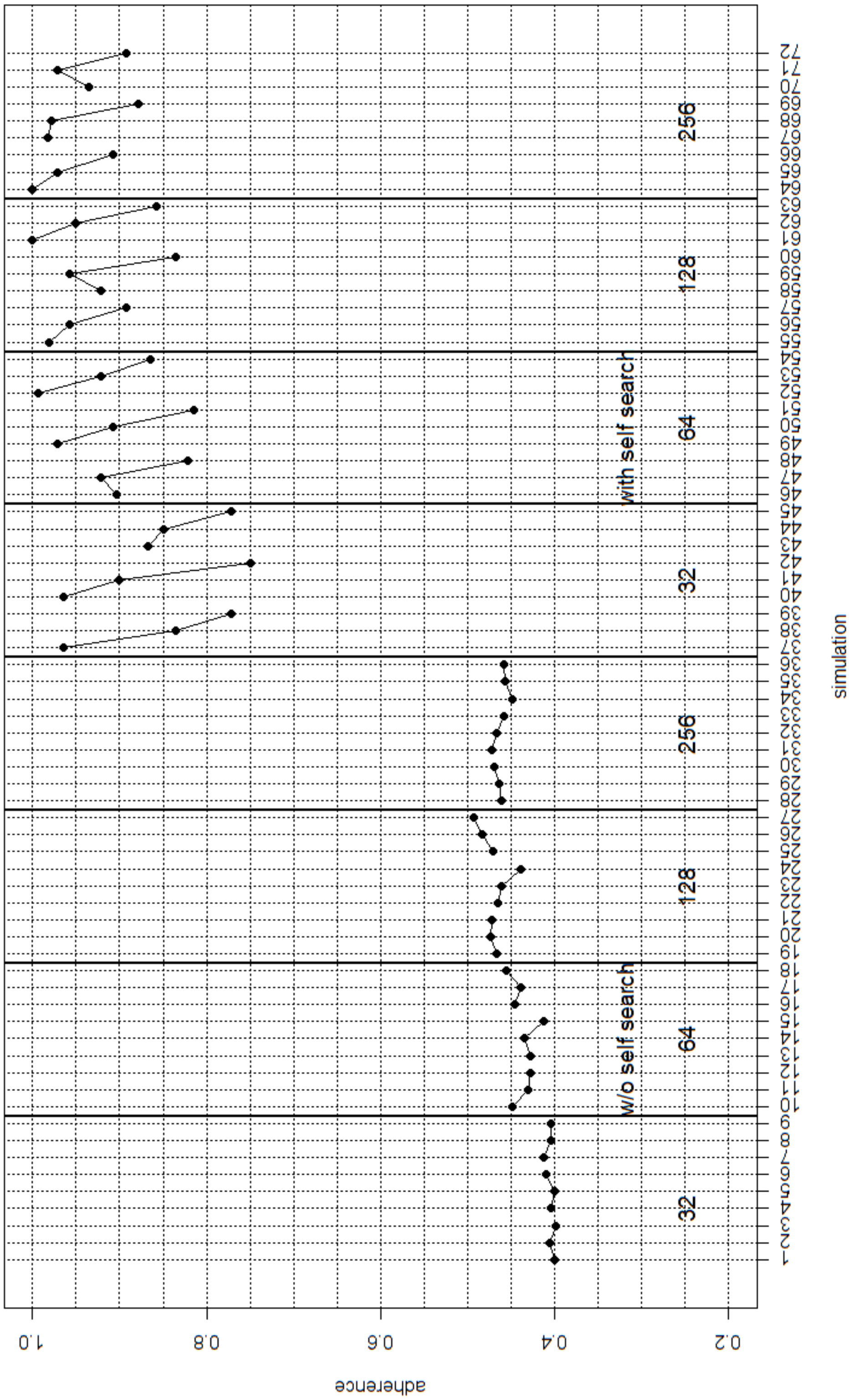


Figure 12.3: End points for friendship channel

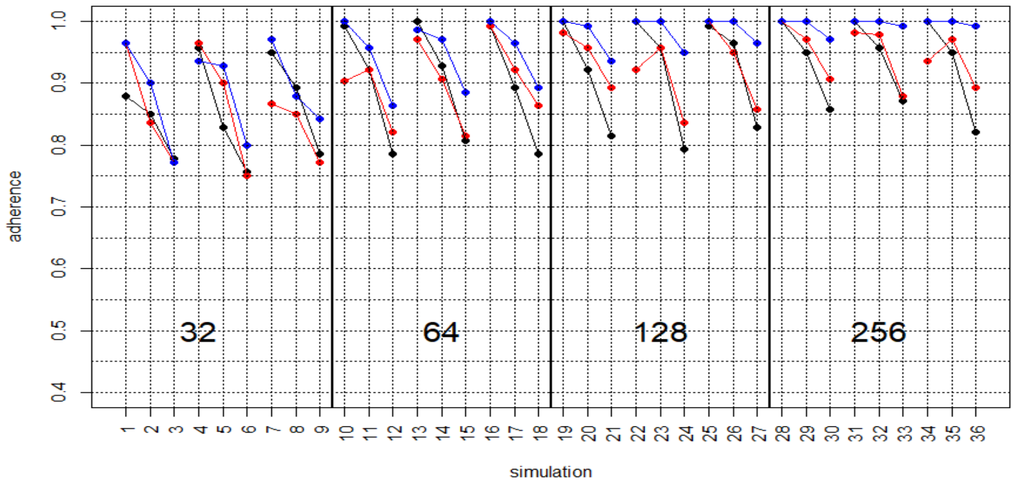


Figure 12.4: End points for self search with and without friendship channel (Simulation: black = without friendship, red = with friendship, blue = with propinquity)

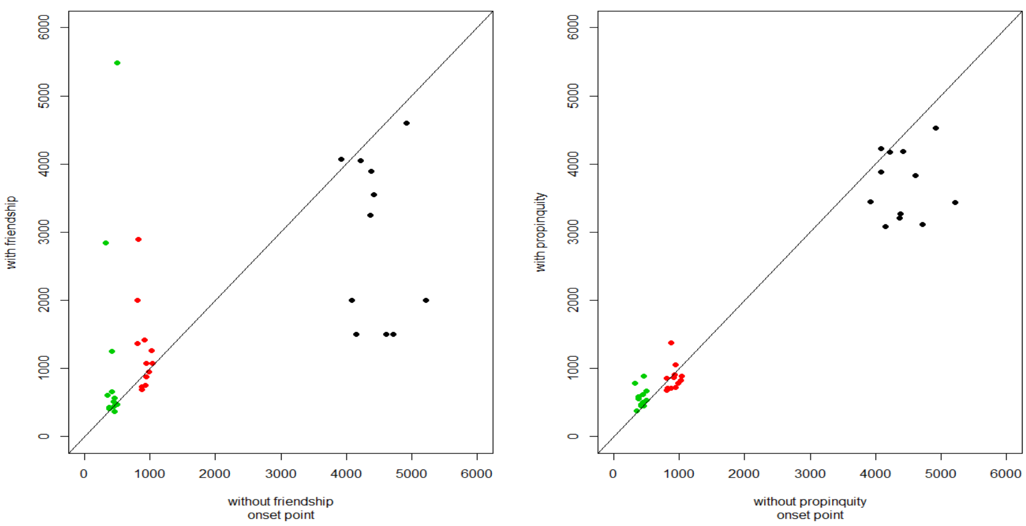


Figure 12.5: A (left) - self search with and w/o friendship end points
B (right) - self search with and w/o propinquity end points
(Learning rate: black = 0.1, red = 0.5, green = 0.9)

Table 12.1: Mean and standard deviation change

	Self search		Self search and friendship	
Learning rate	mean	st. dev.	mean	st. dev.
0.1	4423	381	2825	1183
0.5	913	80	1258	633
0.9	422	54	1171	1524

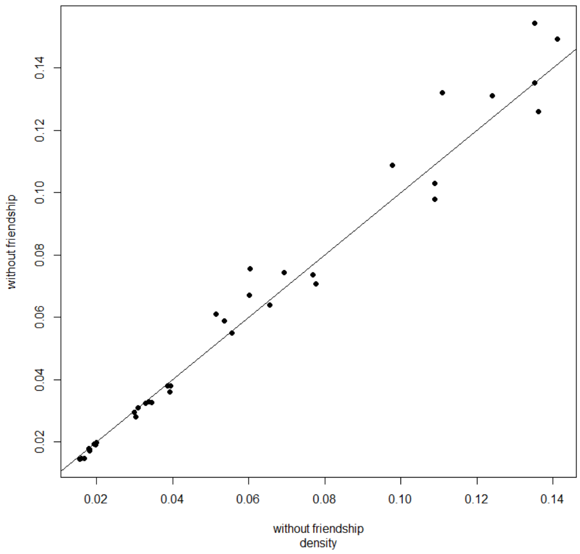


Figure 12.6: Density maximum scatter plot

and without propinquity and the effect is by far less important. The impact is not linked to the exchange mechanism rather it could be imputable to the nature of the network itself. Propinquity network does not present any small world properties as friendship network does and the hypothesis could be that superimposing small world networks the effect is the onset values spreading.

Interesting is the fact that although the onset point changed, the network density peaks do not as depicted in figure 12.6. The scatter plot among density maxima of self search alone and self search and friendship reports values well displaced along the bisector line witnessing an absence of difference.

This means that adding friendship on the self search does not affect the density of the resulting networks. But, since also self search simulations exhibit a small world structure (cf. section 9.2), the composition of the two channels could shape the structure differently.

The result confirms what found in chapter 10: not necessary a small world has higher performance than other networks structures. On exploration exploitation problem, it seems that a more clustered and isolated network such as the propinquity one[4] could perform better. Summarising: friendship network does not have the ability to remarkably impact the exploration exploitation problem in terms of adherence. Instead, a possible effect is to harmonize the onset values.

[4] Q index = 0.01175.

Chapter 13

Complexity and emergence

The exchange mechanism used in the present research is a simple one, well known and with the intrinsic characteristic to tend to an equilibrium. Since the knowledge exchange could take place only when superior knowledge is owned by the donor, the interactions naturally stop when all the agents have the same representation of the external reality: there is no more superior knowledge to share.

So far, the single communications layers and simple combinations have been studied with the aim to understand the impact of different parameters on the final output. In these scenarios, the exchange mechanism has the power to lead the system to an equilibrium and to show a crescent trend in the agents' and average adherences.

Taking some examples from the simulations output it is possible to appreciate this effect. Figure 13.1 shows typical outputs obtained from the simulations. Trends AA001 and AD001 are an example of self search activity, CI216 is an example from propinquity simulations and DA039 from governance simulations groups.

All four trends are crescent with an equilibrium point beyond that the exchange stops. Only the AD001 figure shows a certain burden at the end of the simulation: this is the effect of the interaction with the external world (or unlearning effect). Anyway the average level of adherence is stable, the equilibrium is reached.

But, things start becoming more interesting when the complexity of the simulation increases. Next sections discuss about some behaviours obtained as output that are not foreseen by the exchange mechanism. These results could be read in two directions: first, as anticipated the organization is a complex system and complex systems show emergent phenomena. Exploration exploitation is not an exception. Second, although the model uses a simple exchange mechanisms, sometime criticized for the lack of richness in

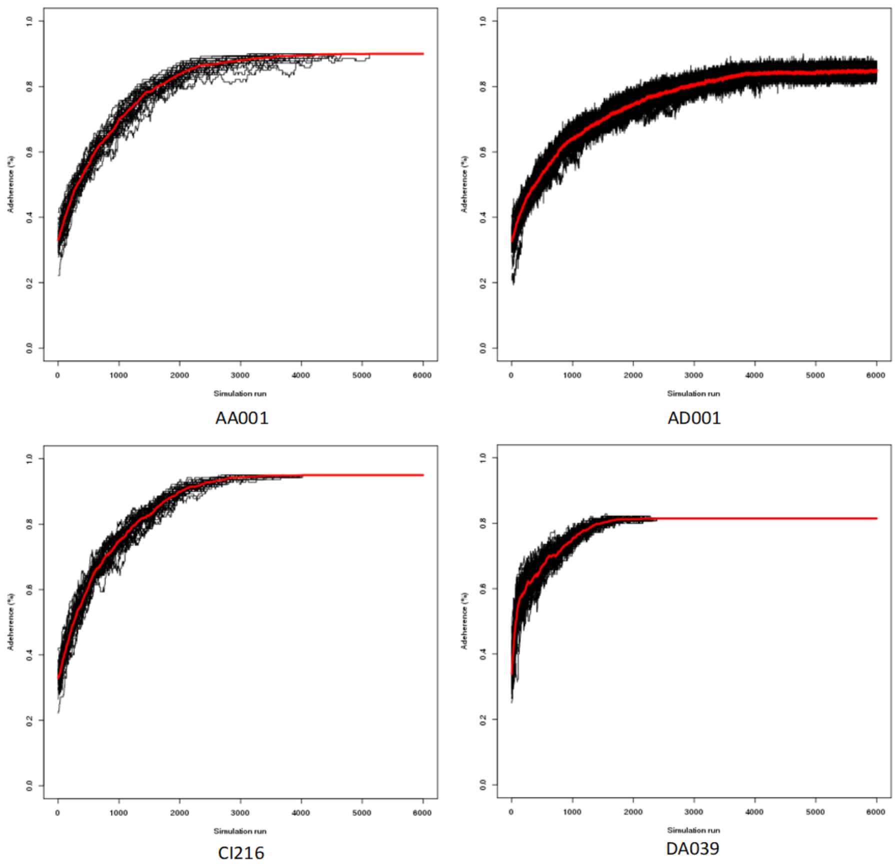


Figure 13.1: AA001 simulation trend

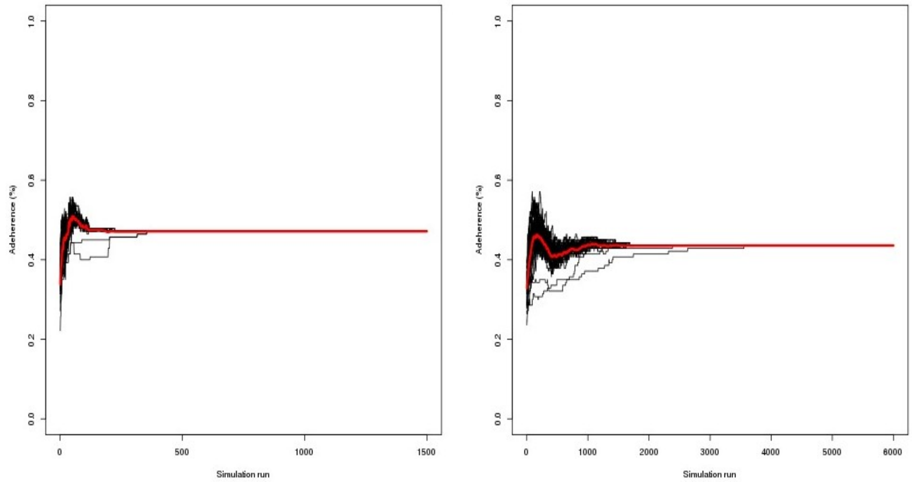


Figure 13.2: Unlearning effect

describing the nuances of knowledge transfer, it is still able to surprise. This could be seen as a further demonstration that ignoring the multiple facets of an environment could lead to miss some important and interesting phenomena. These phenomena deserves a dedicated further analysis but it is beyond the scope of the present work.

13.1 Unlearning

A first example of interesting effects comes from the simulations with meeting channel active. Picture 13.2 reports a couples of simulations where it is visible the unlearning effect. The adherence curve is not always crescent rather, the average adherence reaches a peak and then it decreases to an end point value at lower value.

This effect was not reported in March papers, neither in the following works in literature. It is indeed an emergent behaviour, typical of a complex systems. Moreover, it is not systematic, it is present only in some simulations, as further demonstration of its emergence. This is a confirmation of the entire hypothesis made at the beginning of the present work: exploration exploitation problem needs to be considered as a complex system, not all the behaviours could be inferred from the basic rules or narrowing the focus on a particular part of the system. Indeed this last approach permits to understand better the different features of the system but the composition of the results could not show all the phenomena only catchable considering complexity.

A possible explanation could be found in the ascendancy of some individuals in the simulation who became extremely popular in the organization and they become able to force the organization to adopt their belief. If these individuals do not change their beliefs according to the external reality, the entire organization tends to a lower level or adherence. Picture 13.2 shows two slightly different situations. The leftmost figures depicts an organization not capable to realign to external reality and hence it gains a lower level of adherence as end point. In this scenario, pivotal individuals are stronger than the organization. The rightmost figure introduces a further step, the organization at a certain time is able to change the trend and to gain again adherence to the external reality. In any case the realignment is made too late as the final end point is still lower than the peak reached in the first part of the simulation. Anyway, it is interesting to note that the organization regains the ability to follow the external reality.

Unlearning it not a new concept in literature. In early 80's Hedberg, (Hedberg 1981) argued about how organizations learn and unlearn and, more recently, Cegarra-Navarro et al. (Gabriel Cegarra-Navarro Gabriel, Sánchez-Vidal, and Cegarra-Leiva 2011) tried to study the unlearn context in the Spanish metal industry posing that unlearning stage is necessary to achieve an appropriate balance among exploration and exploitation since persons have to forget outdated knowledge before learning updated one.

Anyway, unlearning term should be used with care since Howells et al. (Howells and Scholderer 2016) strongly challenged Hedberg conclusions stating that from the review of literature there is no empirical evidence of his definition. Often, under the term unlearning fall different mechanisms as theory-change or the setting aside of owned understanding when new facts lead to a different one. They concluded that in all examined cases simpler concepts could be adopted instead of unlearning.

13.2 Unstability and divergence

The subgroup MB encompasses all the communications layers together, interaction with external world and complete heterogeneity of agents. In these simulations, all agents have different learning rate and stack level whereas in the previous subgroups the values of learning rates and stack levels were equal for all the agents since their effect was under study.

In MB group the heterogeneity of agents is the broader available. Stack levels and learning rates values for each agent are randomly sampled from the uniform distributions $\big(U(5, S^L), U(0.1, 0.9)\big)$. Times allocated to meetings

Table 13.1: Time allocated to different channels for different job roles

Job role	T_{sel}	T_{dep}	T_{prj}
1	15	25	60
2	40	60	0
3	30	20	50
4	15	25	60
5	20	40	40
6	30	50	20
7	30	20	50
8	30	10	60
9	15	15	70
10	40	30	30
11	30	20	50

and self search ($T_{sel}, T_{dep}, T_{prj}$) are derived from time allocations estimation based on real data. Table 13.1 shows the entire values used in the simulations for the 11 job roles and figure 13.3 reports the distribution of job roles in the simulations with 255 agents. Moreover, also the numbers of meetings are randomly chosen every run from distributions $U(1, N_m^p)$ and $U(1, N_m^d)$.

From figure 13.4 it is possible to appreciate that the typical trend shown in figure 13.1 is lost. There is no more a crescent trend and, most important, there is no more an equilibrium. The organization is stuck in a sort of periodic learning-unlearning cycle without any end. The right side subplot is showing also another effect highlighted in figure 13.5. Agents' trends are no more compacted as in typical trends as in figure 13.1 where, although with some noise, all black lines (agents' trends) move in a coordinated way toward the equilibrium.

In the figure 13.5 a subgroup of agents diverges in terms of adherence to external reality. In MB019 the divergence starts at the beginning and fades after a certain number of runs. In the MB014 case, a group of agents exits from the group of all the agents and stays diverted from the rest of the simulations. This support the idea of a segregation and isolation of a part of the organization, suggesting that some agents interact among themselves and they reach a local equilibrium. As long as the organization is not able to reabsorb this divergence, this group will stay isolated.

A possible interpretation suggests that a group of agents with low amount of time dedicated to project meetings (T_{prj}) stays isolated and have an internal exchange of knowledge (as, for example, job role 2 and 6). Without a prominence of project meeting channel in their job role an effective ho-

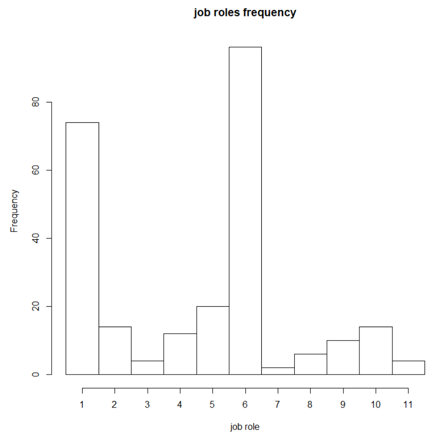


Figure 13.3: Job roles' distribution

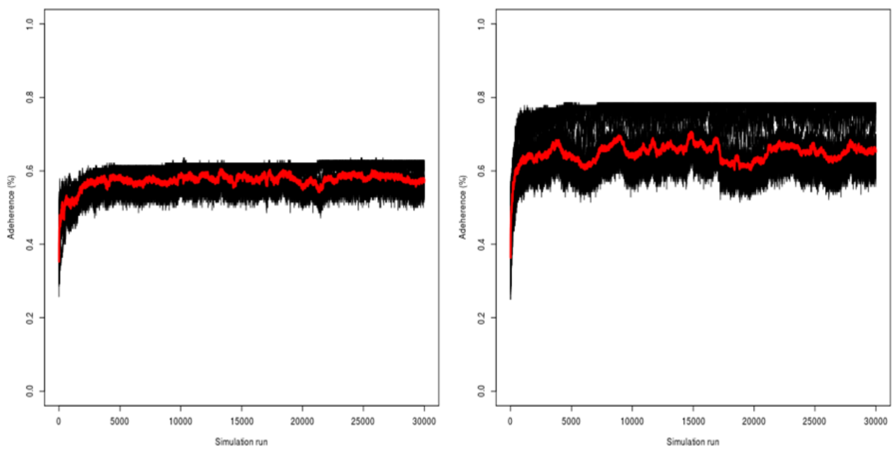


Figure 13.4: Unstability effect in MB001 and MB016

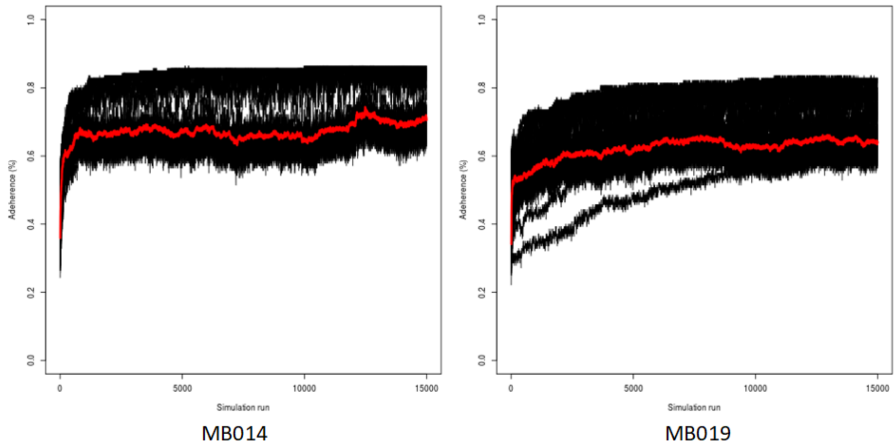


Figure 13.5: MB014 and MB019 simulation trend

mogenizer is missing. Also self search channel is not extremely high (T_{sel}) supporting the tendency to department isolation.

Chapter 14

Summary and contribution

This research is hinged to the idea that exploration exploitation problem could not be interpreted using limited perspective. Indeed March's merit was to begin a discussion about the importance of exploration and exploitation dilemma and the importance of having slow learner in the organizations. However, following literature focused the study on agents only or, at least, stressed this point. Undoubtedly various subsequent papers argued around the initial model investigating particular features or explored new aspects but always in a finite vision. From the output of the simulations done the importance of the environment emerges very strong. Exploration exploitation could not be studied only from the the agents' point of view, there is the need to consider at the same level also the environment as a whole and the underneath networks of relationships.

To support this view, many simulations were run, many configurations were tested, many data were elaborated[1]: it is time to summarize the output of the presented work.

14.1 Environment and March's effect

The first three propositions of the research question could have a common answer: environment matters. The output of this research is clear: individuals and environment compete to conquer the supremacy. In some situations

[1]R version 3.5.1. "Feather Spray" R Core Team 2013 was used to build entire model, to run the simulations and to perform data analysis. The following additional R packages were used: igraph (v.1.2.2), dplyr (v. 0.7.8), tidyr (v. 0.8.2), stringr (v 1.3.1), readxl (v 1.1.0), grDevices (v 3.5.1), e1071 (v 1.7-0.1), stargazer (v 5.2.2).

individuals prevail and their characteristics shape the output, in other situations the environment overcomes individuals but often the output reveals different tendencies simultaneously in act. There is no clear cut in the sphere of influence, it is a sort of continuum with all the combinations possible. The effect is not always intuitive neither trivial and the mix of these ingredients lead to a scenario with many facets and nuances that strongly challenges the output March and subsequent scholars found, paving the road for new questions. When focusing on a particular aspect the interplay seems clear but the when we look up the situation is far more intricate. The main output are now summarised by points.

The output suggests the necessity to consider a dual part of agents that is the environment. Environment in a broad meaning since it could include governance as meeting, physical layout and formal and informal channels. The evidence is clear from chapter 8, in particular figure 8.1 documents the effect of environment on the end point values. The effect is twofold: the absolute value of the end point is strongly impacted and also the ability of learning rate to shape the end points is not always granted. The environment is capable to condition the level of adherence the organization tends to and also it is able to drown the distinction among slow and fast learner performance.

The environment influences the mutual learning mechanism and it could ease the knowledge transmission toward a higher adherence or it could impede it, forcing the organization at lower level of adherence. Governance channel is a good example: department meetings tend to focus the organization on knowledge exchange, clustering it into groups of homogeneous topics and this benefits the self search, leading to high adherence values. Vice versa, project meetings keep the organization fully connected but it has and adverse effect on the mean value of the organization adherence.

Every communication channel shows an interesting result. First, in the self search channel the influence of learning rate on the final organization adherence in not homogeneous. March (March 1991) did find that a slow learners organization performs better than a fast learners one and this result was achieved every time. What emerges from the current simulations is that the ability of learning rate to differentiate the final adherence has a non linear effect. The ability seems to be a function of the number of agents present in the organization and it holds only in a limited range of number of agents. More precisely, the curve is an inverted U-shape curve as shown by figure 14.1.

The lost of this ability by learning rate is not due for the same phe-

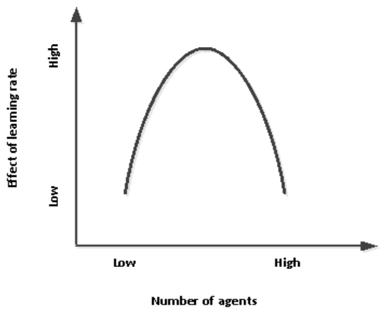


Figure 14.1: Influence of learning rate

nomenon when the number of agent is low and high. In the first case, when the number of agents is low, the lost seems to be linked to the absence of critical mass in the organization. The dynamic is too poor to let the learning rate to be able to have an effect. When the number of agents is high the inability is related to the convergence of the organization to an asymptotic value regardless the composition of the organization. It seems that when the organization increases its size, the contribution of the single agents diminishes in favour of an organization effect. Second, the phenomenon seems to be quite general since it could be seen also when propinquity is active.

Propinquity tends to destroy the effect of learning rate as per March's output exactly because it tends to cluster the agents. Indeed, large organizations are penalised by physical layout since it is more difficult to host them in a single place. This effect is particularly clear in figure 10.4 when only propinquity channel is active. Of course, the output is the pure effect of isolations but a closer look to figure 10.5 highlights the same tendency also when the self search is active. Moreover, even modulating the propinquity, the results do not change dramatically, as reported in figure 10.14. Then the propinquity effect seems to be dichotomous: when is active, the extension of propinquity influence does not change the intensity of the effect. Friendship has a mild effect on exploration exploitation, milder than expected. A possible reason lies in the similarities owned by self search and friendship networks. Both exhibit small world properties and the superimposition may not be able to change dramatically the end point values. But, friendship has an interesting effect: it seems to harmonize the onset values. This effect could be explained as a sort of short circuit in the network which bypasses the superior knowledge mechanism and tends to uniform the speed of diffusion.

The superimposition of different channels is not always negative. The activation of a second channel could improve the organization ability to adhere to external reality. This effect is particularly evident when propinquity is activated ad clearly figure 10.6 shows the results. All the values are in the propinquity regions meaning that the second channel boosts the organization. A similar effect could be noted comparing figure 9.2 with figure 11.11. Department meetings focuses the organization and the performances increase. A counter example is given by the governance and in particular by project meetings: cross functional meetings have denser networks but they perform less better than intra department meetings. And, in presence of both types of meetings, project meetings have the ability to shape the adherence to knowledge and the organization network characteristics, overriding department meeting effect. To attenuate the project meeting effect, maybe the number of held meetings should be lower slowing down the acceleration to knowledge homogenization (and therefore the knowledge exchange) and giving to the organization a chance to achieve a higher adherence.

These findings well support the first three propositions of the research question: different communication channels shape the exploration and exploitation output in a very strong way and also the superimposition of different channels lead to different adherence outputs. Moreover, March's effect has a validity in a narrower territory than expected.

14.2 Emergence in exploration exploitation dilemma

Results confirm that exploration exploitation problem shows all the typical treats of complex systems and it exhibits emergent phenomena. Beside these phenomena there are also other output that could share some characteristics with complexity, although they could not be completely ascribable to emergence. They are not obvious, they refuge the linearity and they give counter intuitive output. Below a more detailed description of the output supporting proposition 4.

Indeed unlearning period showed by the organization is an emergent phenomenon (as reported in figure 13.2). This behaviour is maybe the most evident proof of emergent properties owned by the system. It has an erratic presence since it depends on particular values of the model parameters rather than on macro feature of the simulation (that is, for example, the activation of a communication channel). The organization could sometime react to this effect and regain some of the lost knowledge but this is not grated. Moreover, even if the knowledge superiority mechanism is active, the organization unlearns deviating from the inner mechanism of mutual learning. The unlearning effect is not present in the rules suggesting that it is an effect of the interaction among individuals. Indeed, neither March (March 1991) and subsequent scholars reported this phenomenon as output of their models.

Isolation is a second potential emergence (figure 13.5). Part of the organization deviates from the remaining whole and pursues a different strategy in the knowledge exchange. This deviation could be reabsorbed by the organization or it could persist. Therefore, the ability of the organization to re-compact the individuals is not always granted. Why part of the organization diverges and why the organization does not always prevails over the clustering are questions which answers are not hard coded in the model.

When heterogeneity is pushed to the maximum allowed by the model and all the communications channels are active, the organization shows a new behaviour. There is the loss of a crescent adherence and the organization floats between learning and unlearning in a sort of periodic alternation. None of this was expected by the model neither from previous literature.

The impact of the communication layers calls for more attention. Intriguing is the subtle effect of friendship on the equilibrium onset and the almost imperceptible contribution on adherence. Propinquity, when added to self search, boosts the performances while project governance seems to be stronger than department governance.

The interaction with external world presents an interesting effect: it does not bring noise in a linear way, it seems to act as an enabler as discussed in chapter 9.

Also networks show clues of complexity. Network density is always clustered by the number of agents except when project governance is active. In this case, the density is clustered by learning rate.

14.3 The importance of networks

Networks deserve a dedicated space in the exploration and exploitation context and could not be considered a supporting actor on the stage. The results of the simulations strongly highlight that networks are a dynamic part of the exploration and exploitation and could give profound insight if adequately considered. A merit of the present work is the effort to link the effect of exploration and exploitation context on the relative networks configuration in a systematic way and to extract information from their properties. Besides the precious insights on the studied mechanism, the analysis of the networks suggests new indexes that could enrich the discussion giving new glasses to look to exploration and exploitation problem. Propositions 5 and 6 are well supported by data and below are described the main achievements.

The organization has the ability to change its structure of relationships and to adapt to the situation. Organization is alive and it changes over time. This is well depicted by figure 9.11: the density of the network changes over time. This is a direct proof of mutable nature of the organization. But organization changes also between configurations, different active channels or different scenarios make it adopt different structures. Pictures 10.16, 11.15 and 11.16 demonstrate this, showing different density shapes. Density changing in shapes is a signal of different organization structures. But there are other important clues about this: figures 9.17 and 10.26 show a changing values in the Q factor witnessing a different structure in the networks and figure 10.27 shows how physical layout models the agents connections, reporting different correlation values to office network. Again, it is evident how physical layout is important in the exploration-exploitation balance.

Compelling is the fact that network density seems to be function of the number of agents and not function of learning rate. Pictures 9.12, 10.18 and 11.17 demonstrate that, as long as the governance is absent or is transparent to exploration exploitation problem (as department meetings), the number of agents shapes the density. Again, the counter example is offered by figure 11.18: when project meetings are active, density is led by the learning rate and not by the number of agents. It seems that when the organization could take the desired configuration, the structure is chosen minimizing the energy and by consequence the density is correlated with the number of agents. But, when a certain type of governance which forces the organization to keep unnecessary links the idea of energy minimum is lost. It this case, the different learning rates could come out and shape the density since fast learners needs less links compared to slow learners.

The idea of energy could be associated to the organization and an intriguing result is the clue that the organization tends to keep itself in a configuration which could minimize the energy. This effect is particularly evident in figure 10.25. The relationships network is susceptible to the different environment agents experiment. Adding channel has the effect to create denser networks and by consequence more energetic networks. Hence, the more complex the task to perform, the denser the relative network. Complexity is not only in terms of number of channels activated in the agents tasks but also in terms of the energy to be spent to accomplish the task. *Ceteris paribus* a more entropic task is related to a denser network. The concept of energy could also be applied to knowledge, as reported in table 9.4. Looking carefully it is possible to appreciate that increasing the length of the knowledge, the network density increases. Hence, the more complex the knowledge to manage, the denser (and then the mode energetic) the network. Undeniably there is a strict connection between the environment and the energy to balance the exploration exploitation in it.

Picture 9.14 manifests another engaging behaviours of the organization. Network density reaches the peak well before the organization reaches the adherence end point. There is no need for the organization to keep all the connections for all the time in order to reach the adherence end point. As already discussed, a first explanation could be that the organization has a sort of inertia that could be overcome by a dense network but, once the organization starts heavily sharing knowledge, the network could become sparser (decreasing the associated energy). A second hypothesis is that once the density maximum is reached, the network rearranges itself in a sort of hierarchical structure based on competence and knowledge and only the most knowledgeable agents are responsible to keep the organization to the adherence end points.

All the results on network analysis justify the research question propositions 5 and 6: the context is able to call for different topological structures. These structures evolve over time and adapt themselves to the contexts showing interesting regularities or intriguing aspects. Also, interesting to highlight is that not always the best performance in terms of adherence is achieved by a small-world network.

14.4 Contribution

Although exploration exploitation dilemma has been on scholars' agenda since many years, the proposed work tries to contribute in a still florid topic in which many scholars gave insight. As many of them highlighted[2], many are the points still open and there is still the need for new results. The proposed work attempts to add a small tile in the big picture either from the academic perspective and from managerial perspective.

The novelty is in the tentative to integrate in the exploration exploitation problem many different facets in a model which has an empirical relevance. Albeit rooted in the well established agent based tradition, the present model embraces the new approach in which networks play a pivotal role. The model is designed to be aligned to the most recent modelling techniques and on top of that, it has the distinctive feature of considering the exploration exploitation problem as a multilayer system. The effect of the presence of different communication layers is one of the main remit of the model. Self search, propinquity, friendship and governance were never studied together[3].

[2]See for example Sachdeva 2013, Almahendra and Ambos 2015 and Wilden et al. 2018.
[3]As per literature review made in may 2019.

In the design phase, there was the need to use as much as possible information coming from the real world appreciating that most of the models considered in literature are isolated as Windrum et al. argued (Windrum, Fagiolo, and Moneta 2007). Indeed the benefit of this approach is to capture the essence of a particular mechanism or feature of the social context to study but it is rather aseptic and partial. In this scenario, propinquity is studied explicitly for the first time, not using a condensed and rare representation of the reality but using in the simulation a real layout of a company offices building. To do that, elements of the space syntax theory were withdrawn and transposed to the exploration exploitation arena. Again, friendship is studied based on real affective network, using data coming from a survey. Also the characterization of the agents withdraws from real data: job role, time allocation and ranking are base on real data.

Another pretty new result is the possible association of energy to the task. Exploration exploitation exercise has a cost which must be contemplated. Exploration exploitation balance received attention from the organization perspective without stressing the fact that any piece of the game has a potential energetic cost.

Generative social science framework has been adopted considering the organization as a complex system and giving it the ability to grow without any constraints in term of relationships between agents. This approach is quite new considering that the extant literature often starts the exploration exploitation study from preconfigured relationship networks. Then the aim is to understand which configuration perform better (as, for example in Mueller, Bogner, and Buchmann 2017). This study starts a step earlier, it gives the system the ability to freely grow and to assume the relevant network configuration. Apart from the freshness of the approach, the interesting thing is that the result complies with the literature. An example could be again taken from Mueller et al. (Mueller, Bogner, and Buchmann 2017) where the supremacy of small world network was testified. In the presented simulation, when free from constraints the organization tends to adopt a small world configuration (figures 9.17 and 11.19).

Converging to a final remark, it could be said that all the obtained results suggest the need to consider the exploration exploitation as a whole. Focusing on particular aspects ignoring the surrounding could indeed gives insight but the generalization power is reduced. The environment is extremely effective in changing the dynamic of the exploration exploitation problem, showing new aspects and behaviours.

Also managers could benefit from this study. Starting exactly from the last point just mentioned, the study gives managers the clue to consider the exploration exploitation problem in a broader way. It is indeed important to consider the learning rate but physical layout, governance, informal channels and the composition of the organization could dramatically change the results. Every aspect of the organization life is important in the seek of the desired balance.

Manager could also find the concept of energy interesting: an organization could not add tasks indefinitely. There is a limit beyond that it is no more possible to perform the task. This point was made clear by the work of Hearter et al. (Haerter, Jamtveit, and Mathiesen 2012) stressing the fact that network structure and dynamics are strictly intertwined. As they wrote, *"limitations on the processing capacities of nodes and links have a profound impact on the flow of information in [...] communication networks"* (Haerter, Jamtveit, and Mathiesen 2012, 1). Knowing the impact of the task in terms of energy could give manager a hint on how she is distant from the organization limit in processing the task.

Chapter 15

Limitations and possible expansions

Any work presents limitations and the present one is not an exception. Indeed during the creation of the study, the coding of relative model and data analysis, many assumptions, simplifications and limitations have been embraced.

Most of them are a by product effect of the research design. Simulation, as seen before, is not exempt from limitations. The first point is related to the concreteness of the model framework that is the level the reality modellable. Being a model, necessarily some assumptions need to be made as also some simplifications.

The organization is an idealization of a real one: the agents exhibit only features strictly relevant to the study. For all the other characteristics all the agents are treated in the same way or there is no explicit evidence of them in the simulations. For example there is no distinction by age, seniority or sex in the model, all the agents are modelled as owning the same treats. Other features are not modelled: agents fatigue is not considered in the simulations. Fatigue could impair the knowledge sharing or increment the knowledge distortion during the real life. The assumption during the design phase was to keep the model focused on the most relevant aspects for multilayer exploration exploitation problem. This assumption was made considering that, encompassing different channels and governance, the model would have become in any case articulated. In order to keep the model manageable the reduction has been applied in the *"second order"* features. Also the daily activity of agents has been simplified, excluding variation as vacations or job transfer. Hence all the agents are present and active for all the simulation.

Although external interaction is considered in the simulation, such interaction is a simplification of the real interaction between organizations. The interaction is modelled choosing randomly agents and changing their beliefs. Indeed in the real world, the interaction is not a duty of all the agents and different agents could interact with colleagues of different external organizations. On top of that, the number of external organizations is huge considering either the internal company organizations and all the potential external third parties which could interact with the organization object of the study. The external organizations could have different propensity to exploration or exploitation and this could influence differently parts of the studied organization.

The model does not consider all the potential constraints, laws or requirements an organization has to satisfy: as in all the literature works the exploration exploitation problem is considered abstract and free from limits.

Knowledge has been modelled according to March's initial model but many other way exists to represent knowledge, as Guechtouli argued (Guechtouli 2014). The complexity of different topics has been modelled increasing the length of knowledge. A further simplification was made avoiding knowledge overlapping: topics are considered separated to keep calculation time low. This because overlapping is supposed extremely impacting.

Also knowledge exchange mechanism is kept as in March's paper. Of course the mechanism *per se* could be challenged because considered too simplistic. Scholars have already attempted to consider this point, in particular Coinet et al. argued that *"even very basic yet credible variations between usual knowledge transmission mechanisms and realistic ones may yield sensibly distinct outputs"* (Coinet and Roth 2007, 6.1). Hence, different mechanisms could, in principle, exhibit different output.

The path assignment mechanism is not extremely sophisticated: individuals are assigned randomly according to a predefined probability. The selection of the communication channel is indeed more complex: if governance channel is chosen as example it is clear that peoples attend meeting according to projects they are involved in. This aspect is theoretically similar to department meeting and the the potential effect is captured by department meeting channel, for this reason it was not modelled. Moreover project meetings quite often have attendees not belonging to the same organization. This mixture of organization belonging could only increase the noise and it could be captured by the external interaction channel. Hence it was not modelled.

There are two other typical aspects that are outside the scope of the present work: organizational changes and individuals turnover. Changes in the organization are not easy to be modelled due to the extreme variability in deployment time, breadth and depth. Moreover it is not strictly necessary that an organization has to change over the time. In all the simulations, where relevant, the organization does not change. Of course changes in job profiles could happens but they are not so frequent and moreover they are not massive as when an organization change takes place. Hence job profile changes are not considered. But, in addition to organizational changes there is the fact that different organizations typically have different structures. In the model we tested only one organizational structure.

Turnover and external turbulence could be considered an important factor also in light of March's output, but it was judged as out of scope for the present set of simulations. Of course the model is already able to simulate the turnover but the run and subsequent analysis would have taken too long to be completed in the available time frame.

Propinquity has been modelled but with a couple of simplifications. First no changes in offices layout or desk assignments are taken into account. Employees could sometime change offices during the permanence in a company with an impact on the neighbours' network. This phenomenon was not considered too impacting to be included. Analogous approach was taken about offices layout changes. Hence the propinquity network does not evolve during the simulation. Second, offices are modelled through the Space Syntax dictates, using only the basic tools. Space Syntax provides also more sophisticated tools that, in principle, could improve the accuracy of the simulation.

Technological communication tools (email, instant messaging, and phone) are not considered as a distinct relationship path. This relationship is split into formal and informal channels because the scope of the research deals with an organization which has not geographical issues. Being all the activities done within the same place, formal and informal connections could be considered a good proxy for technological communications.

The last point to discuss is about the degree of generalization obtainable by the present research. It is evident that the methodology presented and the strong link with the field leads to a scarce generalization. However, although rooted in a real organization, the hypotheses originated by the simulation about the underpinning mechanisms could be considered quite general.

* * *

This study posits interesting cues for future studies. There are different directions toward which new studies could be addressed.

First, many combinations the model allows are still to be simulated. There is the need to build a wider map of the obtainable world. This discovery is in two dimensions: width and depth. Considering the width, new insight could come from spanning the model parameters ranges. For example, studying more on the impact of knowledge sharing channels addition and composition would be beneficial. Which is the impact of different organization structures or which is the impact of different physical layouts could clarify better the output seen in the present simulation. Dealing with depth, there is need to simulate around some interesting output. Unlearning effect is an emergent phenomenon extremely interesting which deserves a dedicated study. Again, propinquity should be further investigated using more sophisticated tools taken from Space Syntax theory.

Second, the connection between environment and underlying network structure is not fully understood. Why the organization network assumes different shapes in different situation needs further clarification.

Third, strictly related to networks, there is the need to further understand the idea of energy. It seems a promising idea with many potential results, but more simulations and analysis are necessary.

Fourth, agents features could be better shaped to obtain a more realistic description of an organization employee.

Fifth, turnover and organizational changes could be an interesting topic to add to the present output.

Sixth, external interaction needs for sure a deeper discussion. In the model the representation is extremely idealised and simplified. A richer conceptualization could improve the understanding of this important part of the daily life of an organization. Again, in the present work we were not able to simulate the impact of external reality changes. This aspect is extremely important and interesting also in light of network dynamics.

Seventh, knowledge representation and transfer mechanisms should be further challenged. A more sophisticated sensitivity analysis could worth as moving from March's knowledge representation and testing a more elaborated one.

Eight, it could be interesting to elaborate more on governance introducing new features. Rodan (Rodan 2005), for example, considered also policy makers and promotions. The inclusion of new governance tools could enrich also the output of the model.

Ninth, friendship impact may need further attention. The effect on onset values as the impact (or potential lack of) on end point values should be better investigated in more complex configurations.

Moreover, emergent phenomena recorded deserve a new campaign of simulations and a dedicated analysis to explain the possible causes.

Last, it could be interesting to gather experimental data to confirm some the output presented.

Chapter 16

Appendix A: regression diagnostics

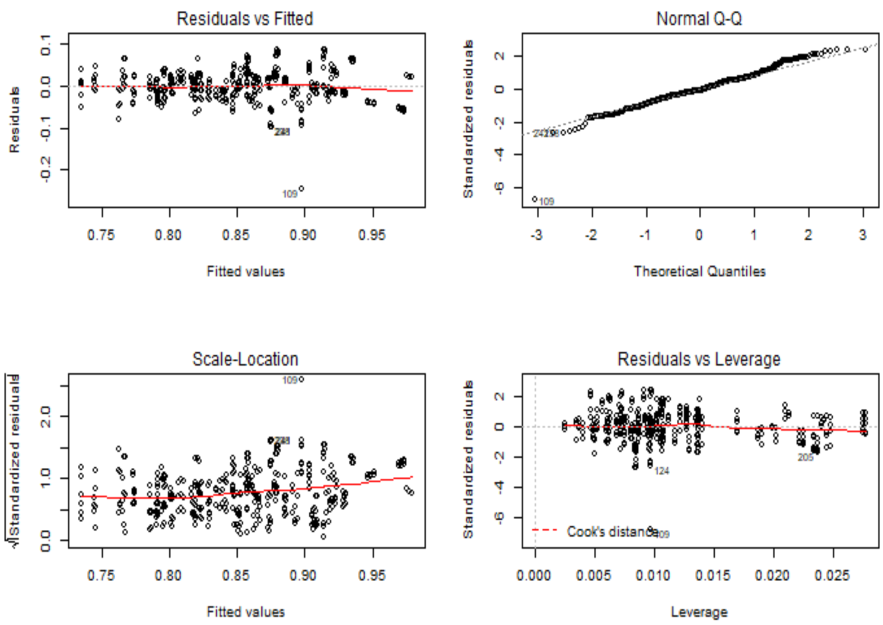


Figure 16.1: Model 1 regression diagnostics

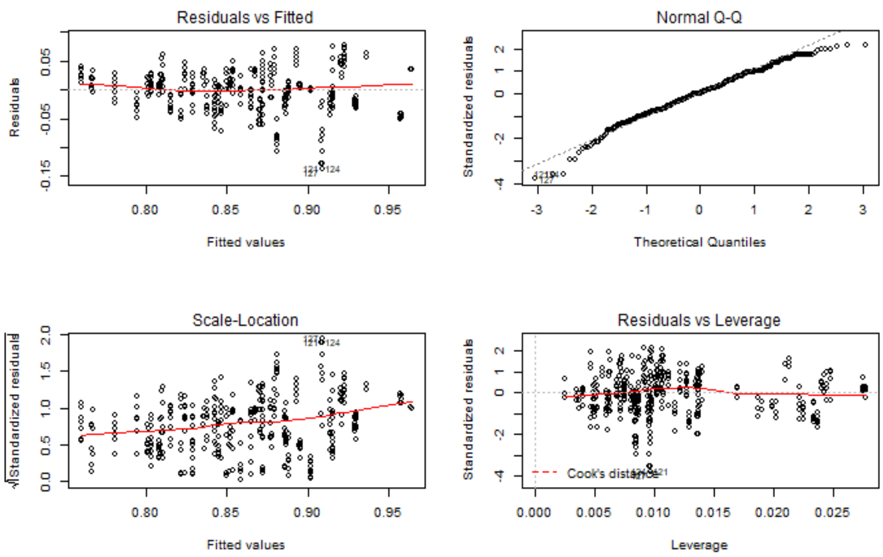


Figure 16.2: Model 2 regression diagnostics

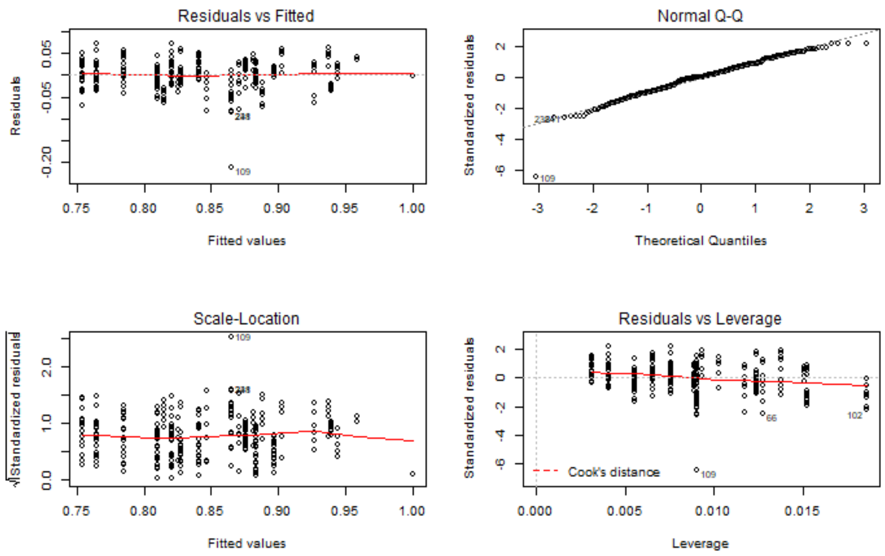


Figure 16.3: Model 3 regression diagnostics

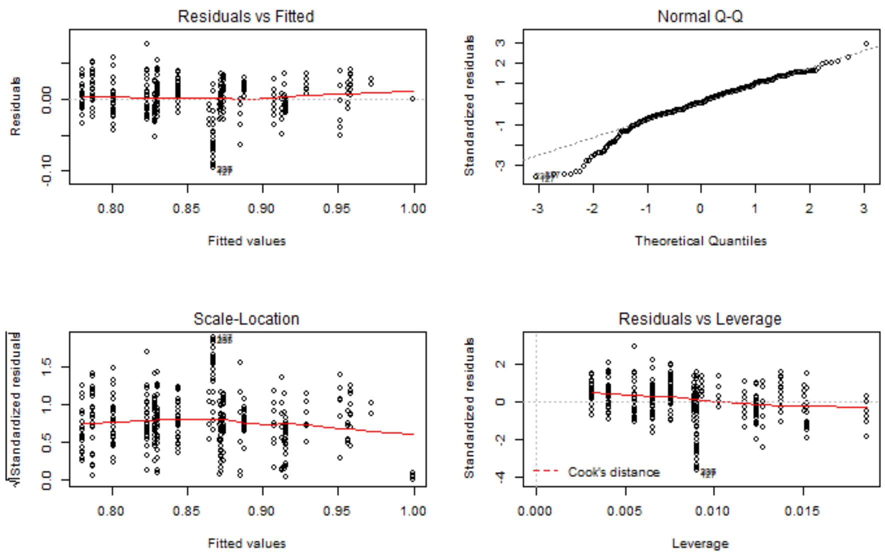


Figure 16.4: Model 4 regression diagnostics

Acknowledgements

This work could not have been done without the help and support of different persons.

My first though goes to my company Chiesi Farmaceutici for the opportunity to gain a PhD. In particular I'd like to thank dr. Rossello Simona for the support since the beginning of my adventure.

I need to acknowledge the University of Bologna for having accepted me with all the exceptions I brought. My gratitude is for prof. Boari Cristina who shed light on dark Tuesday afternoon and for prof. Torrisi Salvatore who introduced me into March's world.

A huge gratitude to my supervisors, the 3Gs: prof. Soda Giuseppe, prof. Morandin Gabriele and prof. Fioretti Guido for the support, advocacy, guidance and friendship. I really enjoyed the discussions we had. A special thank to prof. Morandin Gabriele for being my Guardian Angel during the journey.

A big thank to dr. Sala Irene, dr. Bellesia Francesca, dr. Baskakova Ekaterina and dr. Alinsunurin Jason for the help during these 4 years and for patience in dealing with a ghost classmate. You made me feel as part of the group.

I would also like to thank my colleagues dr. Benassi Andrea, dr. Gaspari Roberto and dr. Ghiretti Alessandro for all the discussions about physics and statistics related to this study.

But my deepest gratitude is to my family, especially to my wife and my children for all their support, love, patience and for all the time I did not spend with them during these years.

I really learnt a lot during the path, from technical stuff to personal emotions, but what I understood the most is the concept of resilience. In 2015 I made a bet and I sowed a seed. The bet is won but I still do not know where the seed fall.

ἀλλὰ γὰρ ἤδη ὥρα ἀπιέναι, ἐμοὶ μὲν ἀποϑανουμένῳ, ὑμῖν δὲ βιωσομένοις: ὁπότεροι δὲ ἡμῶν ἔρχονται ἐπὶ ἄμεινον πρᾶγμα, ἄδηλον παντὶ πλὴν ἢ τῷ ϑεῷ.

Nomenclature

$\hat{S}$	General stack
τ_{er}	Frequency for external reality change
τ_{ewi}	Frequency for interaction with external world
A	Agent's belief
d	Geodesic distances among agents
D_T	Agents' department belonging
E	External reality array
E_G	Networks energy
F^{ape}	Type of external reality change: periodic or aperiodic
F^{dpt}	Activation or deactivation of department meeting channel
F^{erc}	Activation or deactivation of external reality change during simulation
F^{ewi}	Activation or deactivation of interaction with external world
F^{fri}	Activation or deactivation of informal channel
F^{kd}	Activation or deactivation of knowledge distortion during exchange
F^{ksp}	Activation or deactivation of knowledge superiority in propinquity channel
F^{prj}	Activation or deactivation of project meeting channel
F^{pro}	Activation or deactivation of propinquity channel
F^{rnd}	Activation or deactivation of fixed random seed

F^{sel} Activation or deactivation of self search channel

G_{max} Maximum distance between agents in propinquity channel

G_{min} Minimum distance between agents in propinquity channel

H Agent's adherence to external reality

L Agents learning rate array

N_{ed}^{ag} Number of agents per step to involve in the interaction with external world

N_{ER}^{cat} Number of external reality knowledge categories

N_{ER}^{dim} Dimension of each knowledge category

N_m^d Number of department meetings per step

N^{fri} Agents' friendship connections

N_m^p Number of project meetings per step

N_{ag} Number of agents in the simulation run

N_{dd} Portion of knowledge to involve in the distortion during exchange

N_{edd} Portion of knowledge to involve in the interaction with the external world

N_{kd} Activation or deactivation of knowledge distortion during exchange

N_{run} Number of simulation steps

Q Small world index. It is the ratio between clustering coefficient ratio (CC_L) and path length ratio (PL_r)

R^{ag} Agents' organizational rank

S Agent's stack

S^L Agents stack dimension

S_{rnd} the seed number, if F^{rnd} is set to 1

T_{aut} Time dedicated to autonomous search

T_{dep} Time dedicated to department meetings

T_{prj} Time dedicated to projects meetings

References

Abelson, Robert P, and Alex Bernstein. 1963. "A computer simulation model of community referendum controversies". *Public Opinion Quarterly* 27 (1): 93–122.

Allen, James, Andrew D James, and Phil Gamlen. 2007. "Formal versus informal knowledge networks in R&D: a case study using social network analysis". *R&d Management* 37 (3): 179–196.

Allen, Thomas J. 1977. "Managing the flow of technology". *MIT Press Books* 1.

Almahendra, Rangga, and Björn Ambos. 2015. "Exploration and exploitation: a 20-year review of evolution and reconceptualisation". *International Journal of Innovation Management* 19 (01): 1550008.

Amaral, Luıs A Nunes, et al. 2000. "Classes of small-world networks". *Proceedings of the national academy of sciences* 97 (21): 11149–11152.

Aven, Brandy, and Evelyn Ying Zhang. 2016. "Social Distance and Knowledge Transformation: The Effects of Social Network Distance on Organizational Learning". *Sociological Science* 3:1103–1131.

Axelrod, Robert. 1997. "The dissemination of culture: A model with local convergence and global polarization". *Journal of conflict resolution* 41 (2): 203–226.

— . 1984. *The evolution of cooperation.* Basic books, New York.

Bafna, Sonit. 2003. "Space syntax: A brief introduction to its logic and analytical techniques". *Environment and Behavior* 35 (1): 17–29.

Barabási, Albert-László, and Réka Albert. 1999. "Emergence of scaling in random networks". *science* 286 (5439): 509–512.

Baum, Joel AC, Stan Xiao Li, and John M Usher. 2000. "Making the next move: How experiential and vicarious learning shape the locations of chains' acquisitions". *Administrative Science Quarterly* 45 (4): 766–801.

Bedau, Mark A. 2013. "Weak emergence and computer simulation". In *Models, simulations, and representations*, 109–132. Routledge.

Bell, Geoffrey G, and Akbar Zaheer. 2007. "Geography, networks, and knowledge flow". *Organization Science* 18 (6): 955–972.

Benner, Mary J, and Michael Tushman. 2002. "Process management and technological innovation: A longitudinal study of the photography and paint industries". *Administrative science quarterly* 47 (4): 676–707.

Benner, Mary J, and Michael L Tushman. 2003. "Exploitation, exploration, and process management: The productivity dilemma revisited". *Academy of management review* 28 (2): 238–256.

Blaschke, Steffen, and Dennis Schoeneborn. 2006. "The forgotten function of forgetting: Revisiting exploration and exploitation in organizational learning". *Soziale Systeme* 12 (1): 100–120.

Bocanet, Anca, and Cristina Ponsiglione. 2012. "Balancing exploration and exploitation in complex environments". *Vine* 42 (1): 15–35.

Boero, Riccardo, and Flaminio Squazzoni. 2005. "Does empirical embeddedness matter? Methodological issues on agent-based models for analytical social science". *Journal of artificial societies and social simulation* 8 (4).

Bonabeau, Eric. 2002. "Agent-based modeling: Methods and techniques for simulating human systems". *Proceedings of the National Academy of Sciences* 99 (suppl 3): 7280–7287.

Boutellier, Roman, et al. 2008. "Impact of office layout on communication in a science-driven business". *R&d Management* 38 (4): 372–391.

Bunge, Mario. 2004. "How does it work? The search for explanatory mechanisms". *Philosophy of the social sciences* 34 (2): 182–210.

Burgelman, Robert A. 1991. "Intraorganizational ecology of strategy making and organizational adaptation: Theory and field research". *Organization science* 2 (3): 239–262.

— . 2002. "Strategy as vector and the inertia of coevolutionary lock-in". *Administrative science quarterly* 47 (2): 325–357.

Burnes, Bernard. 2005. "Complexity theories and organizational change". *International journal of management reviews* 7 (2): 73–90.

Burns, Tom, and George M Stalker. 1961. "The management of innovation. London". *Tavistock Publishing. Cited in Hurley, RF and Hult, GTM (1998). Innovation, Market Orientation, and Organisational Learning: An Integration and Empirical Examination. Journal of Marketing* 62:42–54.

Burton, Richard M, and Børge Obel. 1995. "The validity of computational models in organization science: From model realism to purpose of the model". *Computational & Mathematical Organization Theory* 1 (1): 57–71.

Caligiuri, Paula. 2014. "Many moving parts: Factors influencing the effectiveness of HRM practices designed to improve knowledge transfer within MNCs". *Journal of International Business Studies* 45 (1): 63–72.

Camerer, Colin F. 2011. *Behavioral game theory: Experiments in strategic interaction.* Princeton University Press.

Carley, Kathleen M. 2009. "Computational modeling for reasoning about the social behavior of humans". *Computational and Mathematical Organization Theory* 15 (1): 47–59.

Chanda, Sasanka Sekhar. 2017. "Inferring final organizational outcomes from intermediate outcomes of exploration and exploitation: the complexity link". *Computational and Mathematical Organization Theory* 23 (1): 61–93.

Chanda, Sasanka Sekhar, and Sougata Ray. 2015. "Optimal exploration and exploitation: the managerial intentionality perspective". *Computational and Mathematical Organization Theory* 21 (3): 247–273.

Chanda, Sasanka Sekhar, Sougata Ray, and Bill Mckelvey. 2018. "The Continuum Conception of Exploration and Exploitation: An Update to March's Theory." *M@ n@ gement* 21 (3).

Christensen, Clayton. 2013. *The innovator's dilemma: when new technologies cause great firms to fail.* Harvard Business Review Press.

Clement, Julien, and Phanish Puranam. 2018. "Searching for structure: Formal organization design as a guide to network evolution". *Management Science* 64 (8): 3879–3895.

Coen, Corinne. 2009. "Simple but not simpler". *Computational and Mathematical Organization Theory* 15 (1): 1.

Cohen, Wesley M, and Daniel A Levinthal. 1990. "Absorptive capacity: A new perspective on learning and innovation". *Administrative science quarterly* 35 (1): 128–152.

Coinet, Jean-Philippe, and Camille Roth. 2007. "How realistic should knowledge diffusion models be?" *Journal of Artificial Societies and Social Simulation* 10 (3): 1–11.

Coleman, James S. 1990. *Foundations of social theory.* Harvard university press.

Cowan, Robin, and Nicolas Jonard. 2004. "Network structure and the diffusion of knowledge". *Journal of economic Dynamics and Control* 28 (8): 1557–1575.

Cowan, Robin, Nicolas Jonard, and J-B Zimmermann. 2007. "Evolving networks of inventors". In *Innovation, industrial dynamics and structural transformation*, 129–148. Springer.

Cunningham, Bryon. 2001. "The reemergence of'emergence'". *Philosophy of science* 68 (S3): S62–S75.

Cyert, Richard M, and J. G. March. 1963. *March, A behavioral theory of the firm.* Englewood Cliffs, NJ: Prentice-Hall.

Dahl, Michael S, and Christian ØR Pedersen. 2004. "Knowledge flows through informal contacts in industrial clusters: myth or reality?" *Research policy* 33 (10): 1673–1686.

Davidsson, Paul, and Harko Verhagen. 2013. "Types of simulation". In *Simulating Social Complexity*, 23–36. Springer.

Davis, Gerald F, and Christopher Marquis. 2005. "Prospects for organization theory in the early twenty-first century: Institutional fields and mechanisms". *Organization Science* 16 (4): 332–343.

Edmonds, Bruce. 2007. "Simplicity is not truth-indicative". *Gershenson, C. et al.(2007), Philosophy and Complexity. World Scientific*: 65–80.

Edmonds, Bruce, and Ruth Meyer. 2015. *Simulating social complexity.* Springer.

Edmonds, Bruce, and Scott Moss. 2004. "From KISS to KIDS–an 'anti-simplistic'modelling approach". In *International workshop on multi-agent systems and agent-based simulation*, 130–144. Springer.

Epstein, Joshua M. 2006. *Generative social science: Studies in agent-based computational modeling.* Princeton University Press.

Epstein, Joshua M, and Robert Axtell. 1996. *Growing artificial societies: social science from the bottom up.* Brookings Institution Press.

Erev, Ido, and Alvin E Roth. 1998. "Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria". *American economic review*: 848–881.

Fang, Christina, Jeho Lee, and Melissa A Schilling. 2010. "Balancing exploration and exploitation through structural design: The isolation of subgroups and organizational learning". *Organization Science* 21 (3): 625–642.

Fiol, C Marlene, and Marjorie A Lyles. 1985. "Organizational learning". *Academy of management review* 10 (4): 803–813.

Fioretti, Guido. 2013. "Agent-based simulation models in organization science". *Organizational Research Methods* 16 (2): 227–242.

Forrester, Jay Wright. 1961. *Industrial dynamics.* MIT/Wright Allen, Cambridge, MA.

Gabriel Cegarra-Navarro Gabriel, Juan, M Eugenia Sánchez-Vidal, and David Cegarra-Leiva. 2011. "Balancing exploration and exploitation of knowledge through an unlearning context: An empirical investigation in SMEs". *Management Decision* 49 (7): 1099–1119.

Gardener, Martin. 1970. "MATHEMATICAL GAMES: The fantastic combinations of John Conway's new solitaire game" life,"". *Scientific American* 223:120–123.

Gibson, Cristina B, and Julian Birkinshaw. 2004. "The antecedents, consequences, and mediating role of organizational ambidexterity". *Academy of management Journal* 47 (2): 209–226.

Giere, Ronald N. 2004. "How models are used to represent reality". *Philosophy of science* 71 (5): 742–752.

Gilbert, Nigel. 2008. "Agent-Based Models. Quantitative applications in the social science 153". *London, Thousand Oaks.*

Gilbert, Nigel, and Pietro Terna. 2000. "How to build and use agent-based models in social science". *Mind & Society* 1 (1): 57–72.

Gilbert, Nigel, and Klaus Troitzsch. 2005. *Simulation for the social scientist.* McGraw-Hill Education (UK).

Grimm, Volker, et al. 2006. "A standard protocol for describing individual-based and agent-based models". *Ecological modelling* 198 (1-2): 115–126.

Gross, Thilo, and Bernd Blasius. 2008. "Adaptive coevolutionary networks: a review". *Journal of the Royal Society Interface* 5 (20): 259–271.

Grüne-Yanoff, Till, and Paul Weirich. 2010. "The philosophy and epistemology of simulation: A review". *Simulation & Gaming* 41 (1): 20–50.

Guechtouli, Widad. 2014. "Agent-based modeling of knowedge transfer within social networks". *IPAG working paper series* 148:1–22.

Gullahorn, John T. 1952. "Distance and friendship as factors in the gross interaction matrix". *Sociometry* 15 (1/2): 123–134.

Gupta, Anil K, Ken G Smith, and Christina E Shalley. 2006. "The interplay between exploration and exploitation". *Academy of management journal* 49 (4): 693–706.

Gutman, Ivan. 1978. "The energy of a graph". *Ber. Math.–Statist. Sekt. Forschungsz* 103:1–22.

Haerter, Jan O, Bjørn Jamtveit, and Joachim Mathiesen. 2012. "Communication dynamics in finite capacity social networks". *Physical review letters* 109 (16): 168701.

Hannan, Michael T, and John Freeman. 1987. "The ecology of organizational founding: American labor unions, 1836-1985". *American Journal of Sociology* 92 (4): 910–943.

Hatch, Mary Jo. 1987. "Physical barriers, task characteristics, and interaction activity in research and development firms". *Administrative Science Quarterly*: 387–399.

He, Zi-Lin, and Poh-Kam Wong. 2004. "Exploration vs. exploitation: An empirical test of the ambidexterity hypothesis". *Organization science* 15 (4): 481–494.

Hedberg, Bo. 1981. "How organizations learn and unlearn". *Handbook of Organizational Design (1)*: 3–27.

Hey, John D. 1982. "Search for rules for search". *Journal of Economic Behavior & Organization* 3 (1): 65–81.

Hillier, Bill. 1996. *Space is the machine: a configurational theory of architecture. London, UK: Space Syntax.*

Holland, John. 1975. "Adaptation in artificial and natural systems". *Ann Arbor: The University of Michigan Press.*

Holmqvist, Mikael. 2003. "A dynamic model of intra-and interorganizational learning". *Organization studies* 24 (1): 95–123.

Howells, John, and Joachim Scholderer. 2016. "Forget unlearning? How an empirically unwarranted concept from psychology was imported to flourish in management and organisation studies". *Management Learning* 47 (4): 443–463.

Humphreys, Paul. 2008. "Computational and conceptual emergence". *Philosophy of Science* 75 (5): 584–594.

— . 2004. *Extending ourselves: Computational science, empiricism, and scientific method.* Oxford University Press.

Kane, Gerald C, and Maryam Alavi. 2007. “Information technology and organizational learning: An investigation of exploration and exploitation processes”. *Organization Science* 18 (5): 796–812.

Katila, Riitta, and Gautam Ahuja. 2002. “Something old, something new: A longitudinal study of search behavior and new product introduction”. *Academy of management journal* 45 (6): 1183–1194.

Kim, Tohyun, and Mooweon Rhee. 2009. “Exploration and exploitation: Internal variety and environmental dynamism”. *Strategic Organization* 7 (1): 11–41.

Krackardt, David. 1987. “QAP partialling as a test of spuriousness”. *Social networks* 9 (2): 171–186.

Krackhardt, David. 1988. “Predicting with networks: Nonparametric multiple regression analysis of dyadic data”. *Social networks* 10 (4): 359–381.

Kunz, Jennifer. 2011. “Group-level exploration and exploitation: A computer simulation-based analysis”. *Journal of Artificial Societies and Social Simulation* 14 (4): 18.

Kuran, Timur. 1988. “The tenacious past: Theories of personal and collective conservatism”. *Journal of Economic Behavior & Organization* 10 (2): 143–171.

Lane, Peter J, Balaji R Koka, and Seemantini Pathak. 2006. “The reification of absorptive capacity: A critical review and rejuvenation of the construct”. *Academy of management review* 31 (4): 833–863.

Lavie, Dovev, and Lori Rosenkopf. 2006. “Balancing exploration and exploitation in alliance formation”. *Academy of management journal* 49 (4): 797–818.

Lawson, Benn, et al. 2009. “Knowledge sharing in interorganizational product development teams: The effect of formal and informal socialization mechanisms”. *Journal of Product Innovation Management* 26 (2): 156–172.

Lee, Sunkee. 2019. “Learning-by-Moving: Can Reconfiguring Spatial Proximity Between Organizational Members Promote Individual-level Exploration?” *Organization Science.*

Levine, Sheen S, and Michael J Prietula. 2012. “How knowledge transfer impacts performance: A multilevel model of benefits and liabilities”. *Organization Science* 23 (6): 1748–1766.

Levinthal, Daniel A. 1997. “Adaptation on rugged landscapes”. *Management science* 43 (7): 934–950.

Levinthal, Daniel A, and James G March. 1993. “The myopia of learning”. *Strategic management journal* 14 (S2): 95–112.

Levinthal, Daniel, and James G March. 1981. “A model of adaptive organizational search”. *Journal of Economic Behavior & Organization* 2 (4): 307–333.

Levitt, Barbara, and James G March. 1988. “Organizational learning”. *Annual review of sociology* 14 (1): 319–338.

Link, Albert N, Donald S Siegel, and Barry Bozeman. 2007. “An empirical analysis of the propensity of academics to engage in informal university technology transfer”. *Industrial and corporate change* 16 (4): 641–655.

Lorenz, E. 1993. *The essence of chaos.* UCL press, London.

Luhmann, Niklas. 1997. *Die Gesellschaft der Gesellschaft.* HeinOnline.

Luo, Shuangling, et al. 2015. “A study on coevolutionary dynamics of knowledge diffusion and social network structure”. *Expert Systems with Applications* 42 (7): 3619–3633.

Macal, Charles M, and Michael J North. 2006. “Tutorial on agent-based modelling and simulation part 2: how to model with agents”, 73–83. L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, / R. M. Fujimoto, eds.

Mäkelä, Kristiina, and Chris Brewster. 2009. “Interunit interaction contexts, interpersonal social capital, and the differing levels of knowledge sharing”. *Human Resource Management: Published in Cooperation with the School of Business Administration, The University of Michigan and in alliance with the Society of Human Resources Management* 48 (4): 591–613.

Mäki, Uskali. 2011. “Models and the locus of their truth”. *Synthese* 180 (1): 47–63.

Manson, Steven M. 2001. “Simplifying complexity: a review of complexity theory”. *Geoforum* 32 (3): 405–414.

March, James G. 1996. “Continuity and change in theories of organizational action”. *Administrative Science Quarterly*: 278–287.

— . 1991. “Exploration and exploitation in organizational learning”. *Organization science* 2 (1): 71–87.

March, James G, and Herbert Alexander Simon. 1958. *Organizations.* Wiley.

Miller, Kent D. 2015. “Agent-based modeling and organization studies: A critical realist perspective”. *Organization Studies* 36 (2): 175–196.

Miller, Kent D, and Dirk Martignoni. 2016. "Organizational learning with forgetting: Reconsidering the exploration–exploitation tradeoff". *Strategic Organization* 14 (1): 53–72.

Miller, Kent D, Meng Zhao, and Roger J Calantone. 2006. "Adding interpersonal learning and tacit knowledge to March's exploration-exploitation model". *Academy of Management Journal* 49 (4): 709–722.

Mitomi, Yuki, and Nobuo Takahashi. 2015. "A missing piece of mutual learning model of March (1991)". *Annals of Business Administrative Science* 14 (1): 35–51.

Montello, Daniel R. 2007. "The contribution of space syntax to a comprehensive theory of environmental psychology". In *Proceedings of the 6th International Space Syntax Symposium, _Istanbul, iv-1–12. Retrieved from http://www. spacesyntaxistanbul. itu. edu. tr/papers/invitedpapers/daniel_montello. pdf.*

Morone, Piergiuseppe, and Richard Taylor. 2004. "Knowledge diffusion dynamics and network properties of face-to-face interactions". *Journal of evolutionary economics* 14 (3): 327–351.

Mueller, Matthias, Kristina Bogner, and Tobias Buchmann. 2017. "The effect of structural disparities on knowledge diffusion in networks: an agent-based simulation model". *Journal of Economic Interaction and Coordination* 12 (3): 613–634.

Nelson, Richard R, and S Winter. 1982. *An evolutionary theory of economic change*. Boston, MA: Harvard University Press.

Oldham, Greg R, and Daniel J Brass. 1979. "Employee reactions to an open-plan office: A naturally occurring quasi-experiment". *Administrative Science Quarterly*: 267–284.

Oldham, Greg R, and Nancy L Rotchford. 1983. "Relationships between office characteristics and employee reactions: A study of the physical environment". *Administrative Science Quarterly*: 542–556.

Phelps, Corey, Ralph Heidl, and Anu Wadhwa. 2012. "Knowledge, networks, and knowledge networks: A review and research agenda". *Journal of management* 38 (4): 1115–1166.

Prietula, Michael, Kathleen Carley, and Les Gasser. 1998. *Simulating organizations: Computational models of institutions and groups*. Vol. 1. The MIT Press.

Puranam, Phanish, and Murali Swamy. 2016. "How initial representations shape coupled learning processes". *Organization Science* 27 (2): 323–335.

R Core Team. 2013. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. `http://www.R-project.org/`.

Reagans, Ray, and Bill McEvily. 2003. "Network structure and knowledge transfer: The effects of cohesion and range". *Administrative science quarterly* 48 (2): 240–267.

Reagans, Ray, and Ezra W Zuckerman. 2001. "Networks, diversity, and productivity: The social capital of corporate R&D teams". *Organization science* 12 (4): 502–517.

Rodan, Simon. 2005. "Exploration and exploitation revisited: Extending March's model of mutual learning". *Scandinavian Journal of Management* 21 (4): 407–428.

Rosenkopf, Lori, and Atul Nerkar. 2001. "Beyond local search: boundary-spanning, exploration, and impact in the optical disk industry". *Strategic Management Journal* 22 (4): 287–306.

Sachdeva, Megha. 2013. "Encounter with March's Organizational Learning Model". *Review of Integrative Business and Economics Research* 2 (2): 602.

Sadek, Ahmed Hassem, and Mardelle McCuskey Shepley. 2016. "Space Syntax Analysis: Tools for Augmenting the Precision of Healthcare Facility Spatial Analysis". *HERD: Health Environments Research & Design Journal* 10 (1): 114–129.

Sailer, Kerstin, and Ian McCulloh. 2012. "Social networks and spatial configuration—How office layouts drive social interaction". *Social networks* 34 (1): 47–58.

Sailer, Kerstin, et al. 2009. "Comparative studies of offices pre and post—how changing spatial configurations affect organisational behaviours". Royal Institute of Technology (KTH).

Schelling, Thomas C. 1971. "Dynamic models of segregation". *Journal of mathematical sociology* 1 (2): 143–186.

Schilling, Melissa A, and Christina Fang. 2014. "When hubs forget, lie, and play favorites: Interpersonal network structure, information distortion, and organizational learning". *Strategic Management Journal* 35 (7): 974–994.

Schüle, Michael, Rainer Herrler, and Franziska Klügl. 2004. "Coupling gis and multi-agent simulation–towards infrastructure for realistic simulation". In *German Conference on Multiagent System Technologies*, 228–242. Springer.

Schumpeter, Joseph A. 1934. "The Theory of Economic Development (translation of second German edition by Redvers Opie)". *Cambridge, MA, Harvard University.*

Schwandt, DR. 2008. "Adult learning". *Clegg, SR & Bailey, J., R.(eds.) International Encycolpedia of Organization Studies. Sage Publications Inc., United States of America.*

Secchi, Davide. 2015. "A case for agent-based models in organizational behavior and team research". *Team Performance Management* 21 (1/2): 37–50.

Secchi, Davide, and Martin Neumann. 2016. *Agent-Based Simulation of Organizational Behavior*. Springer.

Secchi, Davide, and Raffaello Seri. 2017. "Controlling for false negatives in agent-based models: a review of power analysis in organizational research". *Computational and Mathematical Organization Theory* 23 (1): 94–121.

Sidhu, Jatinder S, Harry R Commandeur, and Henk W Volberda. 2007. "The multifaceted nature of exploration and exploitation: Value of supply, demand, and spatial search for innovation". *Organization Science* 18 (1): 20–38.

Siggelkow, Nicolaj, and Daniel A Levinthal. 2003. "Temporarily divide to conquer: Centralized, decentralized, and reintegrated organizational approaches to exploration and adaptation". *Organization Science* 14 (6): 650–669.

Singer, Hermann M, Irina Singer, and Hans J Herrmann. 2009. "Agent-based model for friendship in social networks". *Physical Review E* 80 (2): 026113.

Soda, Giuseppe, and Akbar Zaheer. 2012. "A network perspective on organizational architecture: performance effects of the interplay of formal and informal organization". *Strategic Management Journal* 33 (6): 751–771.

Squazzoni, Flaminio, Wander Jager, and Bruce Edmonds. 2014. "Social simulation in the social sciences: A brief overview". *Social Science Computer Review* 32 (3): 279–294.

Stacey, Ralph D. 2007. *Strategic management and organisational dynamics: The challenge of complexity to ways of thinking about organisations*. Pearson education.

Stacey, Ralph D, Douglas Griffin, and Patricia Shaw. 2002. *Complexity and management*. Routledge.

Styhre, Alexander. 2002. "Non-linear change in organizations: organization change management informed by complexity theory". *Leadership & Organization Development Journal* 23 (6): 343–351.

Tetenbaum, Toby J. 1998. "Shifting paradigms: From Newton to chaos". *Organizational dynamics* 26 (4): 21–33.

Troitzsch, Klaus G. 2009. "Perspectives and challenges of agent-based simulation as a tool for economics and other social sciences". In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, 35–42. International Foundation for Autonomous Agents and Multiagent Systems.

— . 1997. "Social science simulation—origins, prospects, purposes". In *Simulating social phenomena*, 41–54. Springer.

Tushman, Michael L, and Charles A O'Reilly III. 1996. "Ambidextrous organizations: Managing evolutionary and revolutionary change". *California management review* 38 (4): 8–29.

Tushman, Michael L, and Elaine Romanelli. 1985. "Organizational evolution: A metamorphosis model of convergence and reorientation." *Research in organizational behavior*.

Uzzi, Brian, and Jarrett Spiro. 2005. "Collaboration and creativity: The small world problem". *American journal of sociology* 111 (2): 447–504.

Vermeulen, Freek, and Harry Barkema. 2001. "Learning through acquisitions". *Academy of Management journal* 44 (3): 457–476.

Wall, Friederike. 2016. "Agent-based modeling in managerial science: an illustrative survey and study". *Review of Managerial Science* 10 (1): 135–193.

Watts, Duncan J. 1999. *Small Worlds: the dynamics of networks between order and randomness*. Princeton University Press.

Watts, Duncan J, and Steven H Strogatz. 1998. "Collective dynamics of 'small-world'networks". *nature* 393 (6684): 440.

Weick, Karl E. 1976. "Educational organizations as loosely coupled systems". *Administrative science quarterly*: 1–19.

Weirich, Paul. 2011. "The explanatory power of models and simulations: A philosophical exploration". *Simulation & Gaming* 42 (2): 155–176.

Wheatley, Margaret J. 1994. *Leadership and the new science: Learning about organization from an orderly universe.* ERIC.

Wilden, Ralf, et al. 2018. "Revisiting James March (1991): Whither exploration and exploitation?" *Strategic Organization* 16, no. 3 (): 352–369.

Windrum, Paul, Giorgio Fagiolo, and Alessio Moneta. 2007. "Empirical validation of agent-based models: Alternatives and prospects". *Journal of Artificial Societies and Social Simulation* 10 (2): 8.

Wright, S. 1964. "Stochastic processes in evolution". *Stochastic models in medicine and biology* 25:199–241.

Wright, Sewall. 1932. "The roles of mutation, inbreeding, crossbreeding, and selection in evolution", 356–366. Proceedings of the VI International Congress Genetics. Ithaca, NY.

Xulvi-Brunet, Ramon, and Igor M Sokolov. 2002. "Evolving networks with disadvantaged long-range connections". *Physical Review E* 66 (2): 026118.