

# **GREQAM**

Groupement de Recherche en Economie  
Quantitative d'Aix-Marseille - UMR-CNRS 6579  
Ecole des Hautes Etudes en Sciences Sociales  
Universités d'Aix-Marseille II et III

**Document de Travail**  
**n°2006-45**

## **Learning in a community of practice: Complete vs. incomplete information**

**Widad GUECHTOULI**

**June 2006**

**DT-GREQAM**

# Learning in a community of practice: Complete vs. incomplete information

Widad Guechtouli

Ph.D. Student

GREQAM – University of Aix-Marseille III

[widad.guechtouli@univ-cezanne.fr](mailto:widad.guechtouli@univ-cezanne.fr)

**Abstract** This paper is a contribution to the modelling of interactions within social networks. We use an agent-based model to explore the learning process, in a particular form of social network: a community of practice. It is based on an empirical study, made in a research centre in France. We shed light on the importance of two parameters in learning in communities: the availability of the community members, and the information they have about their neighbourhood. The first parameter represents their willingness and engagement in the development of their community, whereas the second one gives us an idea on agents' ability to build a correct shared representation of the competencies of the other agents. We lead two sets of simulations, simulations with complete information and simulations with incomplete information.

**Key words:** learning, interaction, competencies, community of practice, agent-based model.

**JEL classification:** C63; C90; D83; Z13

## 1. Learning in a community of practice:

Because knowledge is considered as a key feature in innovation and since individuals rarely face complete information situations, social learning plays a very central part in any innovation process. Multi-agent simulations are considered as appropriate tools to investigate this field (Phan, 2003). In this paper, we will use an agent-based model to explore individual and collective learning process, in a particular form of social network: a community of practice (CoP). This concept is seen as one of the most efficient concepts to study the process of sharing of knowledge in groups (Lave and Wenger, 1991; Brown and Duguid, 2000). Some empirical studies using multi-agent modelling can be found in the literature (Diani and Muller, 2004; Dupouët *et al*, 2003; Dupouët and Yildizoglu, 2003), our approach differs in that we address the learning of agents through the raise of their skills, in a specific practice.

According to Wenger (2000), CoPs' success is essentially due to their informal and independent status, as well as to the voluntary engagement of their members and the knowledge created through their interactions. This implies a certain degree of trust shared by the community members. In this paper, we will shed light on the importance of two parameters in learning in communities. These two parameters are the availability of the members of the community, and the information they have about their neighbourhood. Therefore, we build an agent-based model, to describe the community and the behaviour of its agents. We will describe their interactions, and how they learn through these interactions.

Our model is based on an empirical study, made in June 2005 in the CIRAD (Centre de Coopération Internationale en Recherche Agronomique et Développement) in Montpellier, France. We met people belonging to a specific network: the Cormas network. This network

emerged in 1998, and one of its most important aims is the use and development of software named Cormas, created by members of the network. This network is composed of more than 385 agents, 3 of them are the creators of the software, and the others are users of this software. Most of the interactions between the community members are about the use of Cormas and how to solve problems experienced by some individual.

Our approach consists in observing how an individual behaves in order to raise the level of his own competency. We then consider that an agent learns through this network, if he can raise his competency in the use of the software, which represents the practice in this community.

## **2. Case description and methodology:**

We chose to study the Cormas network, belonging to the CIRAD. It's a network composed of people interested in natural resources management and multi-agent simulations. Cormas is software that was created by the members of this network; it can be downloaded off the Internet, free of charge and training courses are offered to teach people how to use it. One of the major goals of the community is the development of this specific tool. It was one of the several criteria that made us point at this network as a community of practice. In fact, the most important feature of a community of practice is the practice that binds the members of this community together. Here, this practice is the software they all use.

In addition to that, we can sum up what made us think of the Cormas network as a CoP in the following points:

- The status of the network: the Cormas network is a rather informal community, independent from the CIRAD and autonomous.
- It emerged in 1998 when people working in the same field first felt the need to work together and exchange their knowledge and experiences on the used of multi-agent simulations in natural resources management; this network was not created by a hierarchical authority.
- The voluntary engagement of the network members, on which depends the length of life of this community.
- The structure of the network: there are 3 very competent agents (the creators of the software) and 385 less competent, learning agents (the users of the software). This fits perfectly with the structure of a CoP. Indeed, according to Lave and Wenger (1991), the members of a CoP divide into two groups: the most competent ones will represent the *core* of the community, whereas the rest of the members will step at the periphery of the community.
- The free access to this community: people can join and leave the community free of charge. Anyone who shows interest for the use of multi-agent simulations in the natural resources field, and especially for the Cormas software can be a member of this community.

Data were collected in three ways: personal interviews, questionnaires sent through the Internet, and observations of interactions on the Cormas forum<sup>1</sup>. This let us build a multi-agent model, based on the structure of the Cormas network. This model will be described in the next section.

---

<sup>1</sup> <http://cormas.cirad.fr/fr/reseaux/forumCormas.htm>

### 3. The model:

The structure of the community we are modelling is based on data collected in the empirical study mentioned above. Because of issues of time and software constraints, we had to reduce the size of the population, keeping the same proportions for all types of agents. We have a population of 110 agents, divided in 2 populations: the *info-seekers population* and the *info-providers population*. The former represents 100 agents with no skills in the use of the software and looking for increasing their competencies through repeated interactions with more competent agents; the latter is composed with more competent agents, regarding the use of the software.

In the info-providers' population, we will have 1 agent with a competency equal to 1 and 9 agents with a competency equal to 0.75. We define an agent's competency as the percentage of questions about the software that he is able to give answers to. Therefore, the agent with a competency equal to 1 is able to answer 100% of the questions he may receive. Initially, this individual is the only agent in the core of the info-providers' population. As we said above, the core of this population is composed of the most competent agents in the community.

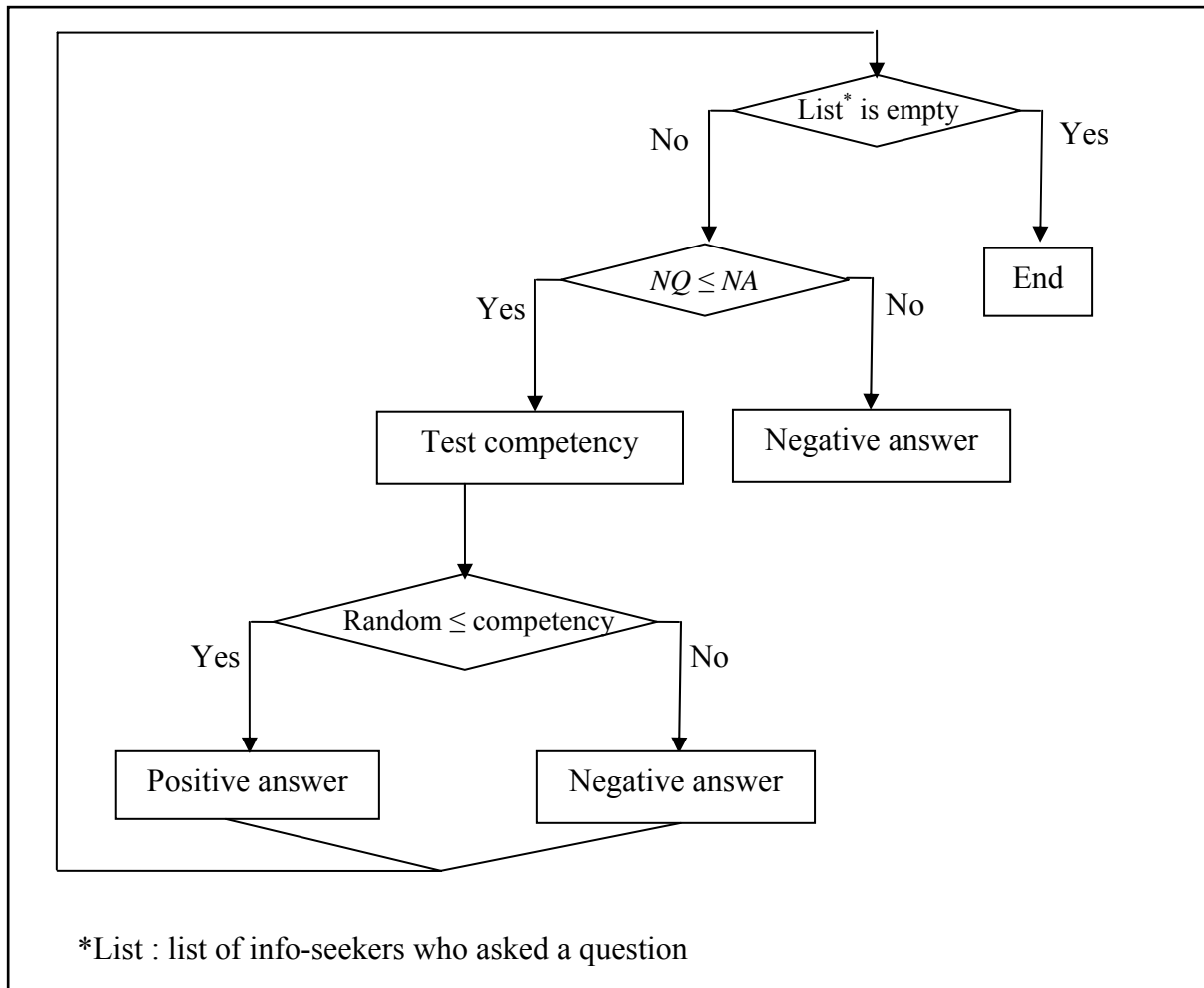
The 9 other agents are able to answer 75% of the questions they may be addressed, they belong to the periphery of the info-providers population.

The members of this population are subject to a constraint on the numbers of questions they can treat at a time. This constraint is a time constraint, and will take values between 1 and 10, considering that if this constraint is equal to 1, an agent can only treat one question at a time, and therefore, is not very available. On the contrary, if this constraint is equal to 10, this means that an agent can treat up to 10 questions at a time, and is therefore available and has enough time to answer those questions.

#### 3.1. Agents and interaction:

An agent is characterized by a name, a type, a competency and a set of info-providers. This set differs according to simulations and is composed of all info-providers that didn't give more than 3 negative answers to this specific agent.

Agents meet once every period of time, an info-seeker chooses randomly and info-provider among his set of info-providers, and asks one question. An info-provider can give a positive answer or a negative one. This will depend on two parameters: his competency and the number of questions he is allowed to answer. The decision process is shown in the figure below, where  $NQ$  is the number of questions received by an info-provider, and  $NA$  the number of answers he is allowed to give per time-step.



**Fig. 1** An info-provider's decision process

An info-seeker will accept 3 negative answers from an info-provider, before removing him from his set of info-providers. Once an info-seeker's set of info-providers is empty, he leaves the community.

### 3.2. The learning process

The learning process depends on the following rules:

- all agents with a competency smaller than 1 seek to increase it.
- An agent's competency increases by 0.01 each time he has a positive answer.
- Once an individual's competency reaches 0.75, other agents will be able to address him questions and he will answer them according to his competency.
- Once an info-seeker's competency is equal to 1, he will become an info-provider, and will stop asking questions.

We consider that learning occurs when an info-seeker is able to increase his competency. The evolution of learning will be measured by the evolution of the number of info-providers in the community.

## 4. Simulations:

We lead two types of simulations, simulations with complete information and simulations with incomplete information. In each set of simulations, the number of questions allowed per period takes values between 1 and 10. Simulations are run until the info-seekers population is empty, either because some of its members became info-providers and some left the community, or because all its members changed type and or left the community.

### 4.1. Simulations with complete information:

Considering that the goal of info-seekers is to increase their competencies in using software, all they need to know is the info-providers' competencies in that particular field. Hence, in complete information simulations, an info-seeker knows all info-providers and each individual's competency.

In this set of simulations, an info-seeker's set of info-providers is divided in 2 groups: agents with high competencies (equal to 1), and agents with average competencies. We will have then 2 subsets: the high competency subset, and the average competency one.

At each period, an info-seeker will ask a question to one of the most competent agents in his set of info-providers. This info-provider will be chosen randomly within the high competency subset. If this subset is empty, the info-seeker will choose randomly an agent within the average competency subset.

### 4.2. Simulations with incomplete information:

In these simulations, info-seekers know nothing about info-providers competencies. Therefore, they will use a new parameter to choose the info-provider they will address their questions to: the info-provider's reputation. It is calculated as follows:

$$rep_{t+1} = 0.4 rep_t + 0.6 coeff_{t+1} \quad (1)$$

with:

$$coeff_{t+1} = \frac{\text{Number of positive answers given}_{t+1}}{\text{Number of questions received}_{t+1}} \quad (2)$$

Here, an info-seeker's set of info-providers is composed of 3 subsets of agents with 3 levels of reputation. The high reputation subset contains agents with reputation equal to 1; the average reputation subset contains agents with reputations between 0.5 and 1, and the low reputation subset contains agents with reputations smaller than 0.5. At each period, an info-seeker asks a question to an info-provider chosen randomly within the high reputation subset. If this subset is empty, this agent will choose randomly an agent within the average reputation subset, and if this one is empty too, he will pick randomly an agent within the last subset, the low reputation subset.

### 4.3. Observations:

Observations will be on the three following points:

- a. **The length of simulations:** This parameter will give us an idea on how fast learning occurs in both types of simulations, according to info-providers' availability.
- b. **Global learning indicator:** We will focus on the evolution of the number of info-providers in the community. That is because we believe that if a community is composed of a large number of info-providers, i.e. very competent agents, there will be a large amount of positive interactions between these agents (by positive interactions, we mean interactions where info-seekers get positive answers to their questions, and increase their competencies). Agents will exchange knowledge through these interactions, and a more global knowledge will be created at the community level. The community will then be able to interact with other communities and ultimately, learn from these interactions.
- c. **The core of the info-providers' population:** the core of the info-providers' population is composed of the agents who received the largest number of questions, throughout every simulation. We will observe the number of questions received by each info-provider. We will then compare the core obtained in complete information simulations, and in simulations with incomplete information. This parameter will indicate whether agents were able to learn about each other's competencies through their interactions and build a correct and identical representation of the agents surrounding them, the correct representation being the structure in simulations with complete information.

## 5. Results:

### 5.1. Description of interactions:

Before studying the results of both types of simulations, let's see how info-seekers get access to the information needed to increase their competencies. To shed light on this question, we will observe info-seekers behaviour when information is very limited, i.e. when the number of question allowed is equal to 1. We will see how the first info-seeker to become info-provider behaves, and how the first info-seeker that left the community behaves.

In complete information simulation, agent 35, which was the first and only info-seeker to increase its competency and turn to info-provider, always asked agent 1 (the most competent agent in the community). It only got 3 negative answers despite the large number of questions addressed to agent 1. This is because agent 1 was removed from most info-seekers' set of info-providers as soon as the 5<sup>th</sup> step. Agent 35 was the only agent that didn't have more than 3 negative answers from agent 1 in the 5 first steps, and therefore from the 6<sup>th</sup> step on, it was the only agent asking questions to agent 1. Hence, it always got a positive answer.

In simulation with incomplete information, it is agent 46 that was the first and only info-seeker to turn to info-provider. Because it didn't know info-providers' competencies, it asked the ones with the highest reputations. But as time went by, it eliminated incompetent agents and as soon as the 46<sup>th</sup> step, it only asked agent 1, and from then on, it only got positive

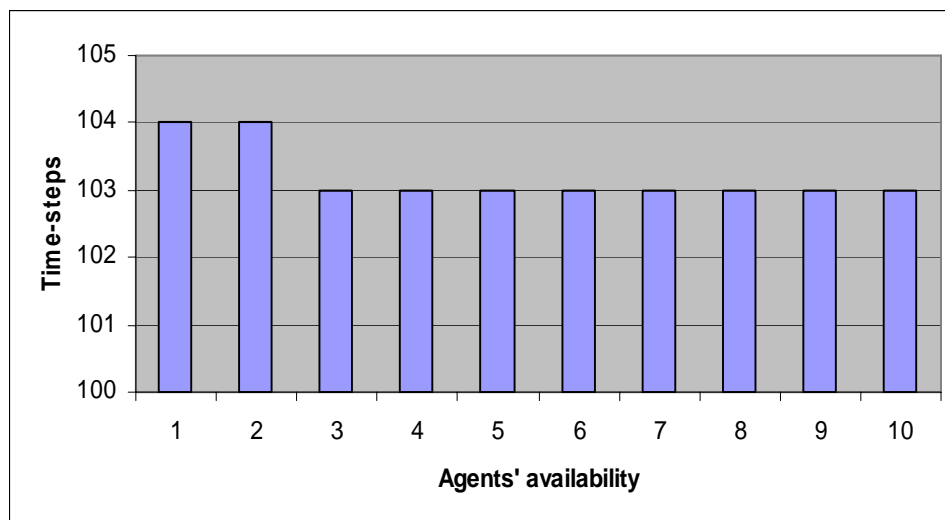
answers. It is the same moment where agent 1 was removed from all info-seekers' set of info-providers, except for agent 46.

In both simulations with complete information and simulations with incomplete information, agent 11 was among the first agents leaving the community. In fact, like 53 other info-seekers in complete information simulations and 28 other info-seekers in simulations with incomplete information, it left the community as soon as the 41<sup>st</sup> step, after asking all info-providers and having 4 negative answers from each info-provider. It seems then that if an agent is lucky enough to be the first to ask the most competent info-provider, then he will be able to increase his competency and learn faster than the other info-seekers.

## 5.2. Simulations with complete information:

### 5.2.1. Length of simulations:

From figure1, we can see that the length of simulations is almost the same no matter what values the number of questions allowed takes. Considering that simulations stop when the population of info-providers is empty, this means that info-seekers leave the info-seekers' population just the same no matter how available the info-providers are.



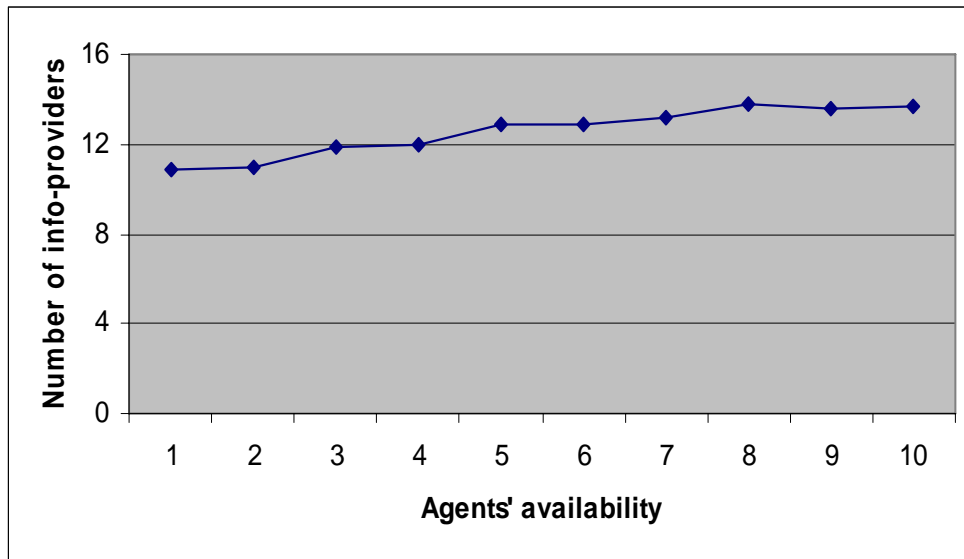
**Fig. 2** Length of simulations with complete information

Let's see now why this happens, i.e. why the info-seekers population gets empty. This occurs because some info-seekers, the ones that didn't leave the community and did get answers to their questions, must have changed type and become info-providers. This is shown next with the global learning indicator.

### 5.2.2. Global learning:

On the next figure, we can see the evolution of the number of info-providers in the community according to agents' availability. This is measured at the end of each set of simulations and is composed by initial info-providers, plus new comers (initial info-seekers who were able to increase their competencies and become info-providers).





**Fig. 3** Number of info-providers in simulations with complete information

One expected result: the more available info-providers are, the bigger the number of info-seekers becoming info-providers at the end of the simulations. However, this number remains a bit low considering that at the biggest value of info-providers' availability, only 4 info-seekers were able to turn into info-providers. This represents only 4% of the initial info-seekers population.

### 5.2.3. The core of the info-providers population:

Table 2 shows the agents composing the core of the info-providers population, the core being composed by the most asked agents, rated according to the number of questions they received during the simulation. Thus, it appears that agent1 is always the only info-provider in the core no matter what values the number of questions allowed takes.

These results are taken at the end of each set of simulations, and show that agent 1 is always the most asked agent in the info-providers population. This may seem a little strange. Indeed agent 1 is the most competent info-provider and all info-seekers know it. One would imagine that knowing that, all info-seekers would ask him at the same time. Then he would be eliminated from most info-seekers' set of info-providers and therefore shouldn't be part of the core.

Nevertheless, this prediction appeared to be wrong. All info-seekers asked agent 1 first, indeed. According to his availability, he was not able to give positive answers to all info-seekers and therefore, most of them didn't ask him anymore from the 5<sup>th</sup> step on. However, the few info-seekers who were lucky enough to get positive answers from agent 1 will keep on asking him questions and at the end of the simulations, he is the one with the highest number of questions received.

Number of questions allowed	Agents in the core	Number of questions received
1	1	542
2	1	642
3	1	636
4	1	836
5	1	936
6	1	936
7	1	1136
8	1	1236
9	1	1236
10	1	1436

Table 1: The core of the info-providers population in complete information simulations

### 5.3. Simulations with incomplete information:

#### 5.3.1. Length of simulations:

The first thing we can say about simulations with incomplete information is that they last longer than simulations with complete information. That is because info-seekers know nothing about info-providers' competencies, and therefore need more time to learn which agents are more competent than others and ask them questions. This is done by the mean of info-providers' reputations. Thus, some agents learn, some don't and leave the community. Simulations stop when the info-seekers community is empty.



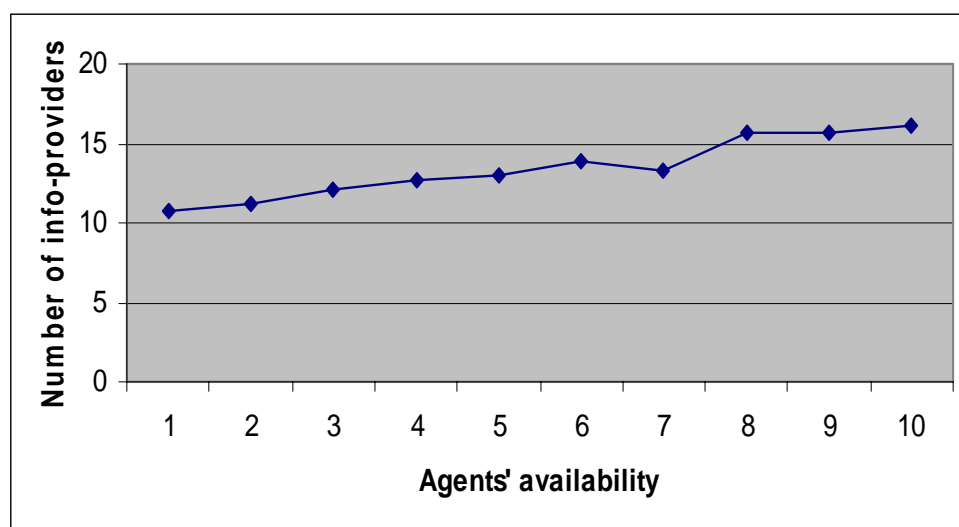
**Fig. 4** Length of simulations with incomplete information

So, unlike in complete information simulations, info-seekers' learning is slower and takes more time as the number of questions allowed is bigger. There is a difference of about 35

time-steps on average between lengths of simulations with complete information lengths of simulations with incomplete information.

### 5.3.2. Global learning:

We can see from figure 4 that there are more info-providers in the community when agents are allowed to answer a bigger number of questions. Considering that the level of learning is measured by the number of info-providers in the community, just like in the previous set of simulations, it seems that agents' availability fosters learning. However, even though agents are interacting without knowing anything about other agents' competencies, the level of learning in simulations with incomplete information is a bit higher than in simulations with complete information.



**Fig. 5** Number of info-providers in simulations with incomplete information

Indeed, 6 initial info-seekers became info-providers in simulations with incomplete information, whereas only 4 were able to change type in simulations with complete information. This difference is very small, yet it has to be noticed as it is not something we expected. This could be explained by the fact that, because agents don't know each other's competencies, they don't all ask the same agent at first (which was the case in the first type of simulations we led). Therefore, demands are spread in a more general way and agents are more likely to have answers to their questions. This stands at least at the beginning of the simulations. Let's see now what happens at the end of the simulations. Were info-seekers able to find out which info-provider was the most competent one?

### 5.3.3. The core of the info-providers' population:

We can see from table 2, that agent 1 is the one with the highest number of questions, and it's the only agent in the core, for all numbers of questions allowed. This agent also happens to be the most competent agent in the community. Just like in simulations with complete information, the core of the info-providers' population is composed of one unique agent, the most competent one. This clearly shows that, even though info-seekers knew nothing about

info-providers' competencies, they were able to find the most competent agent and ask him questions.

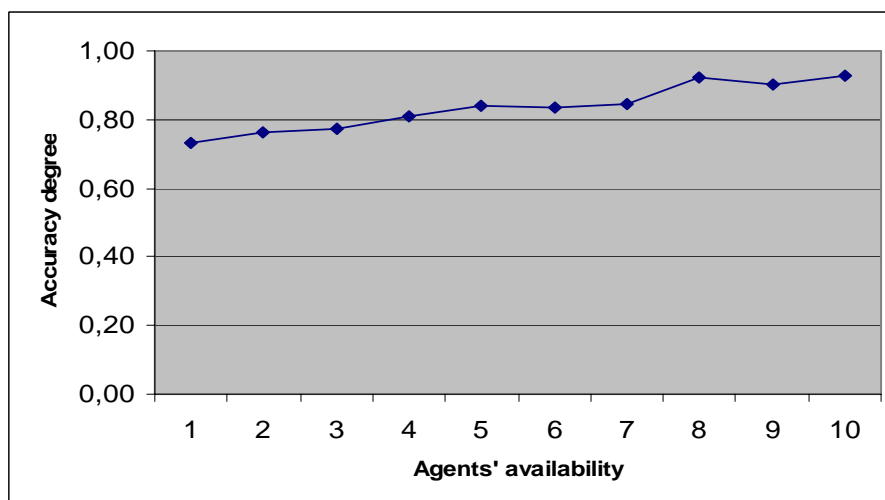
Number of questions allowed	Agents in the core	Reputation	Number of questions received
1	1	0,95	515
2	1	0,97	575
3	1	0,96	645
4	1	1,00	661
5	1	0,99	725
6	1	0,99	755
7	1	1,00	715
8	1	1,00	798
9	1	1,00	817
10	1	0,97	808

**Table 2:** The core of the info-providers population in simulation with incomplete information

What about the rest of the info-providers' population? Was info-seekers' representation of an info-provider' competency far from his real competency? To find out whether it is or not, we introduce a new indicator: the degree of accuracy of info-seekers' representation. It is calculated for each value of info-providers' availability as follows:

$$Accuracy\ degree = 1 - \frac{\sum_{i=1}^{10} (competency_i - reputation_i)}{10} \quad (3)$$

Results are shown in the next figure:



**Fig. 6** Accuracy of agent's representation of info-providers' competencies

This figure shows the accuracy of info-seekers' representation, according to info-providers' availability. We can see that the error made by info-seekers is relatively small. Its highest value is about 0.27 when agents' availability is equal to 1, whereas when the availability is 10, the mean error is only 0.07 on average for each info-provider. Considering these results, we can assume that the representation that info-seekers have of info-providers' competencies, which was built during the simulations, is correct indeed.

## 6. Discussion:

Communities of practice are considered as a very efficient tool in the sharing and capitalizing of knowledge. Organisations should therefore encourage such communities and respect their most important feature: their totally spontaneous and informal status (Wenger, 1999). In the model presented above, we considered interactions within a community, without taking into account the environment of this community, i.e. if this community belongs to a specific organisation. We wanted to shed light on two parameters considered as two important features of a community of practice: agents' availability and the knowledge they have about their environment (Wenger, 2000). To test agents' availability, they were allowed to answer a certain number of questions per time-step. We let this number take values between 1 and 10 in each set of simulations. What we could see was that the more available agents were, the faster learning happened. This was a quite obvious and expected result.

The interesting outcome here is that, after the learning process, and in both types of simulations, not all info-seekers were able to increase their competencies. Indeed, at the highest value of info-providers' availability and in simulations with complete information, only 10 info-seekers became info-providers. That is barely 10% of the info-seekers population. The rest of the agents decided to leave the community because they didn't get answers to their questions. Even though our agents are in a cooperative and non-strategic environment, not all of them can have access to information because of the info-providers unavailability. Therefore, there exists an implicit competition to access information. A congestion effect emerges. This is the second interesting and unexpected outcome. This kind of phenomenon is often observed in situations with strategic rationality, where there exists an explicit competition between agents looking for the same goal to reach. But, in this model, agents have a procedural rationality and interact in a non-strategic environment. Nevertheless, a congestion effect is observed during the simulations. Now, the question we are considering here is how can these results be interpreted in real life? What is the best strategy for individuals belonging to a social network where there is no explicit competition between them (such as a community of practice) in order to access information and learn faster?

One can think that the best way to do so is to ask the most competent agent, taking the risk to have negative answers considering it is the most asked agent in the community. And one can think that info-seekers should rather not ask the most competent agent if it is not available, and ask other competent agents in the community taking the risk that they cannot give a correct answer. Well, according to our model, agents asking the most competent one are those who are able to learn faster.

On the other hand, if we compare the results obtained from both types of simulations in terms of the constitution of the core of the info-providers' population, we'll find exactly the same agent in the core in both simulations, which is agent 1, the most competent agent in the community. This outcome reveals that info-seekers were able to judge info-providers'

competencies according to their reputation, as revealed by the accuracy indicator. Therefore, even if agents know nothing about others' competencies, they can still acquire that knowledge through the ability they have to wholly observe interactions happening inside the community. At the end of the simulations, they what agents were the most competent and therefore, they knew who was in the core of the community and who was not. Ultimately, interactions based on the reputation build by the info-seekers led them to share a correct representation of agents' competencies and the structure of the community.

Future steps of our work will be to explore more the parameters of this model. We started doing that in a paper that is still a work in progress where we studied the impact of info-seekers' motivation to stay in the community. We lead simulations with motivation between 1 and 10 and preliminary results showed that even at the highest values of info-seekers' motivation and info-providers' availability, still not all info-seekers were able to learn and change type. The congestion effect remained. The next step will consists in finding an algorithm to solve the congestion problem, so that all info-seekers can have access to the information needed and learn fast enough to become info-providers themselves. The new algorithm may be one such as info-providers would have a discrimination rule when deciding whether to answer a question or not, or info-seekers would be more patient and decide not to leave the community as soon as there is no agent they ask questions to, but rather wait for a new info-provider to emerge. To conclude, even though this model is based on data collected within the Cormas network, it is still a bit abstract and future development of the model should also include some new algorithms to better represent the Cormas network, such as a variable size of the community to represent the free access to the network; non-directed interactions (blackboard) to better represent the most common type of interactions in the Cormas network; differentiation of the several types of knowledge exchanged in the community.

## References

- Brown, J.S. and Duguid, P. (2000) *The Social Life of Information*. Harvard Business School Press, Boston.
- Diani M. and Muller P. (2004) *Dynamics of Trust within Communities of Practice: Goodwill, Reputation and Social Norms*. 4th Congress on Proximity Economics: Proximity, Networks and Co-ordination; Université de la Méditerranée, Marseille, June 17-18 2004.
- Dupouët O., M. Yildizoglu (2003) *Organizational performance in hierarchies and communities of practice*.
- Dupouët O., Yildizoglu M. and Cohendet P. (2003), *Morphogenèse de communautés de pratique*, 15ème édition du séminaire annuel pluridisciplinaire Organisation, Innovation, International (OI2)(UTC, Costech).
- Lave, J. and Wenger, E. (1991) *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge, New York, Melbourne.
- Nickols F. (2000), *Community of practice: overview*.
- Phan D. (2003), *From Agent-Based Computational Economics towards Cognitive*. In: Bourguine P. And Nadal J.-P. (Eds) *Cognitive Economics*, pp 1 – 23, Springer-Verlag, Copyright Springer, 2003.
- Rouchier J. (2005), *Les filières de légumes étudiées à travers un aller-retour entre terrain et modélisation multi-agents*, Les journées scientifiques francophones, 12 et 13 novembre 2005, Tokyo.

- Wenger E. (1999), Communities of practice: the key to knowledge strategy, *Knowledge Directions* 1(Fall): 1-16.
- Wenger E. (2000), *Communities of Practice and Social Learning Systems*. SAGE Volume 7(2): 225–246.