Learning in a community of practice: an agent-based model

Widad GUECHTOULI
Doctoral student GREQAM (CNRS) - Université d'Aix-Marseille III
15/19 Allée Claude Forbin
Bâtiment Austerlitz
13627 Aix-en-Provence - France
Email: widad.guechtouli@univ.u-3mrs.fr

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

1. Learning in a community of practice:

Because knowledge is considered as a key feature in innovation and since individuals rarely face perfect-information situations, individual learning and collective learning play a very central part in any innovation process. Multi-agent simulations are considered as appropriate tools to investigate this field (Phan, 2003). They allow us to model social structures, with independent computational agents that have the ability to communicate wit one another (Rouchier, 2005). 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 community members, and the information the agents have about their neighbourhood. The first parameter will represent the agents' willingness and engagement in the development of their community, whereas the second one will give us an idea of how agents trust information received from others. Therefore, we build an agent-based model, to describe the community and the behaviour of its agents. We will study 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. Among the 355 users, around 27 are considered as very competent users who can help solving problems which may experience some users in the use of the software. Most of the interactions among 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 its own competencies. We consider that an agent learns through the network studied here, if it can raise its skills in the use of the software.

2. The model:

The structure of the community we are modelling is based on data collected in the empirical study mentioned above; it contains 110 agents, divided on 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; the latter is composed with the most competent agents, regarding the use of the software.

We define an agent's competency as the percentage of questions about the software it is able to give answers to. Here, we will have 1 agent with a competency equal to 1 and 9 agents with a competency equal to 0.75. Therefore, the agent with a competency equal to 1, is able to answer 100% of the questions it may be addressed. This individual is the core of the info-providers population. 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 and have the ability to ask questions, in order to increase their competencies. 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. In this model, time constraint is the same for every agent, they all have the same availability.

2.1. 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. Agents meet once every period of time, an info-seeker chooses randomly an info-provider among its set of info-providers, and asks one question. An info-provider can choose to give a positive answer or a negative one. This will depend on two parameters: its competency and the number of questions it is allowed to answer. An info-seeker will accept 3 negative answers from an info-provider, before removing it from its set of info-providers. Once an info-seeker's set of info-providers is empty, it leaves the community.

2.2. The learning process

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

2.3. Simulations:

We lead two sets of simulations, simulations with perfect information and simulations with reputation. 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 all its members became info-providers, or because all its members left the community.

2.3.1. Simulations with perfect 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 perfect 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 (less than 1). We will have then 2 subsets: the high competency subset, and the average competency one.
At each period, an info-seeker will ask its question to one of the most competent agents in its 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.

2.3.2. Simulations with reputation:

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. This parameter is a way of modelling how agents trust others judgement on one's competency. It is calculated as the sum of positive answers given during the last ten periods, divided by the number of questions received on those periods.
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 or bigger than 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.

3. Results:

3.1. Simulations with perfect information:

3.1.1. Individual learning:

What we can see here, is that the smaller the number of questions allowed, the more agents leave the community. That is because when an info-provider can only treat one question at a time, info-seekers get more negative answers and tend to leave the community faster. From now on, each time we talk about info-providers, we mean the original population of info-providers, plus the info-seekers that increased their competencies and didn't leave the community.
Nevertheless, some info-seekers did increase their competencies and became info-providers. In fact, info-seekers competencies reach 0.75 around the 79th step, in all of the 10 simulations (see table1).

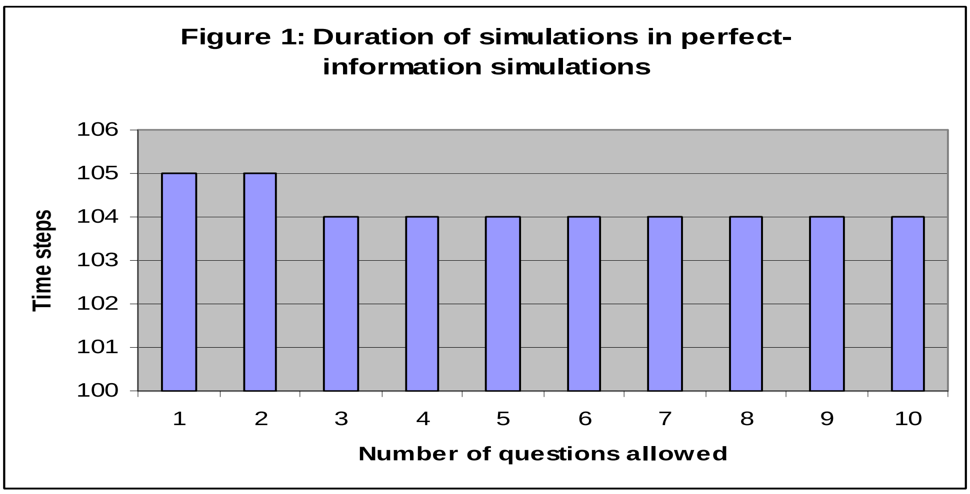
Number of questions allowed	Number of info-seekers leaving the community	Time-step when remaining info-seekers get answering abilities
1	99	80
2	98	80
3	98	79
4	96	79
5	95	79
6	95	79
7	93	79
8	92	79
9	92	79
10	90	79

Table 1: Individual learning and leaving agents

For the info-seekers that don't leave the community, they keep on increasing their competencies until they reach 1, then they leave the info-seekers' population for the info-providers' population. This is where collective learning happens.

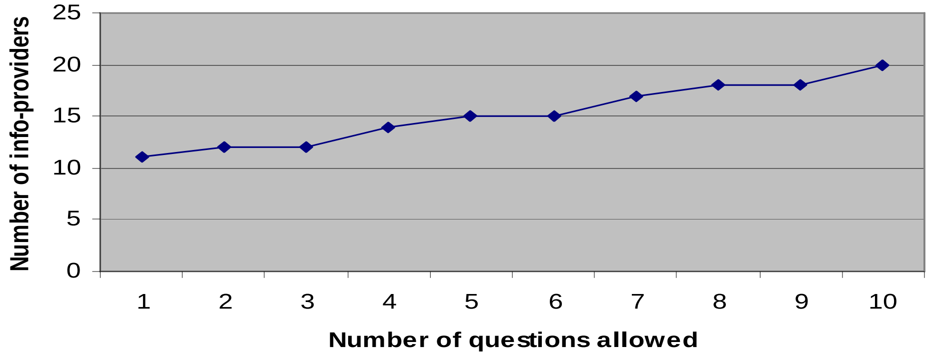
3.1.2. Collective learning:

From figure1, we can see that the duration 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, either because they became info-providers, or because they didn't have answers to their questions and decided to leave.



Let's see now why this happens, i.e. why the info-seekers population gets empty. From table1, we could see that there are a large number of info-seekers leaving the community. However, this number decreases as the availability of info-providers increases. This means that some info-seekers, the ones that didn't leave the community and don't belong to the info-seekers population anymore, must have changed type and become info-providers. Hence, individual and collective learning happened.
As we said above, an info-seeker becomes able to answer questions as soon as the 79th step in most simulations. The collective learning process though differs from one simulation to another. We measure the collective learning by the evolution of the number of info-providers in the community. We can see from figure 2, that this number is bigger as the number of questions allowed increases. This clearly shows how important is the availability of agents in the collective learning process.

Figure 2: Collective learning in perfect-information simulations



3.1.3. 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 most addressed agents, classified 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 considering that being the most competent agent in the community, all info-seekers will address their questions to this agent first. And they did indeed. However, it was removed from most info-seekers' set of info-providers, except for a few agents. These few agents will then continue addressing their questions to agent 1, and at the end, it is the agent that received the biggest number of questions.

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 2: The core of the info-providers population in perfect-information simulations

3.2. Simulations with reputation:

3.2.1. Individual learning:

The first thing we can say about simulations with reputation is that they last longer than simulations with perfect-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. Individual learning here starts later than in the pervious set of simulations and all info-seekers' competencies reach 0.75 around the 118th step. By info-seekers we mean those agents that didn't leave the community. After that, info-seekers increase their competencies, individual learning keeps going until all info-seekers' competencies reach 1. Here starts the collective learning process.
Simulations stop when the info-seekers community is empty. Results presented in table 3 are taken at that period of time, we can see the number of agents leaving the community in each set of simulations, and the evolution of the answering abilities for the agents staying in the community. We can see that there are a very large number of info-seekers that left the community. The rest of the agents turned to info-providers.

Number of questions allowed	Number of info-seekers leaving the community	Time-step when remaining info-seekers get answering abilities
1	99	117
2	98	118
3	98	118
4	97	119
5	96	118
6	95	118
7	93	118
8	92	118
9	92	117
10	91	117

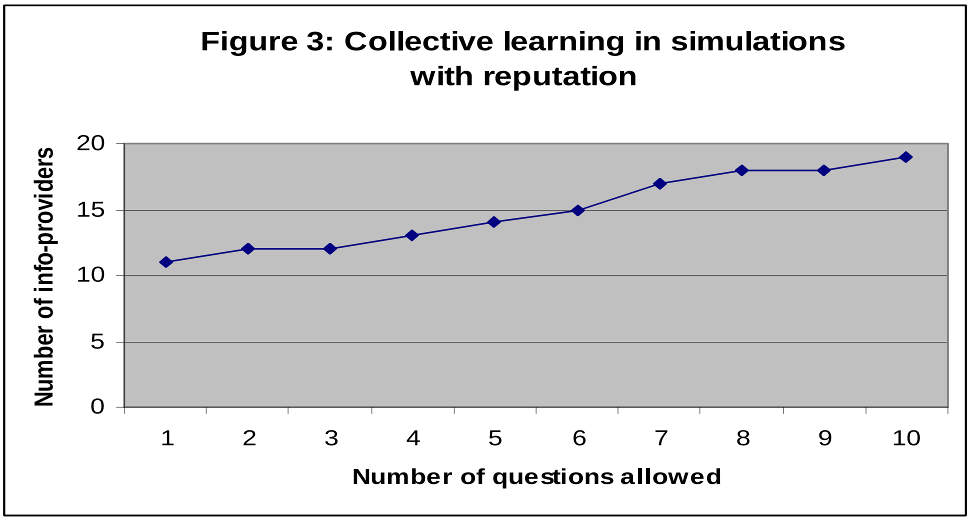
Table 3: Number of agents leaving the community and time steps where info-seekers get answering abilities

3.2.2. Collective learning:

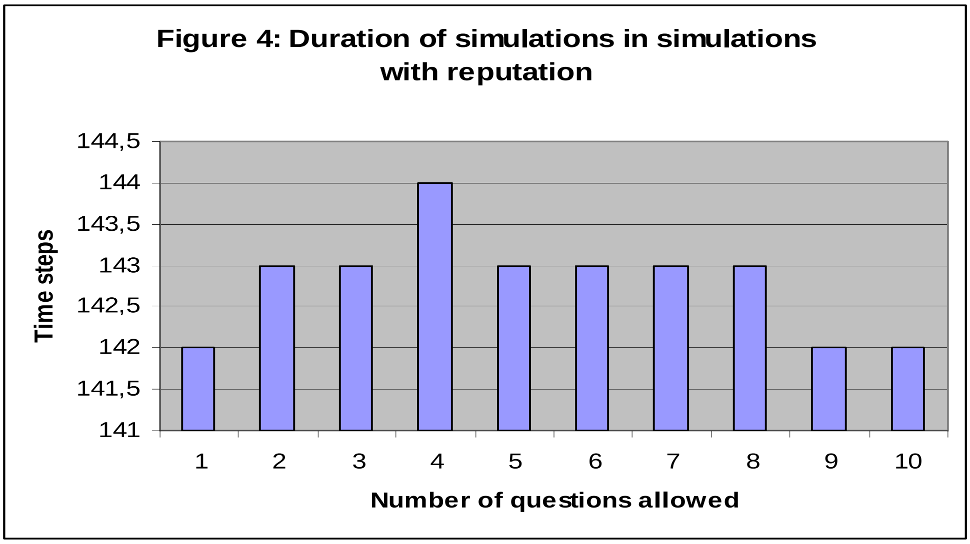
Table 3 shows that, unlike what happens in perfect-information simulations, info-seekers' learning is slower and takes more time as the number of questions allowed is bigger. This may be due to the fact that info-seekers are facing two types of learning:

- They gather information and learn to know their environment and neighbourhood.
- They learn from the answers they get and increase their competencies in the use of the software.

We can see from figure 3 that there are more info-providers in the community when agents are allowed to answer a bigger number of questions. Considering that collective learning is measured by the number of info-providers in the community, it becomes obvious that collective learning happens faster when agents are more available.



Looking at figure 4, one could wonder why does the duration of simulations tend to increase before the number of questions allowed get 5, then decrease when the number of questions allowed gets bigger? We can see that simulations last longer when the number of questions allowed increases until it reaches 4. This probably means that info-seekers need more time to know info-providers competencies and choose what agents they should ask. However, the number of info-seekers that got to increase their competencies to 1 and turned to info-providers is very small, between 1 and 3 (see figure 3). Collective learning happens, but in a rather slow rhythm.
When info-providers can treat more than 4 questions at a time, simulations tend to be shorter as info-providers are more available, and the number of info-seekers turning to info-providers is quite important (between 4 when the number of questions allowed is 5, and 9 when the number of questions allowed is 10). Collective learning is faster, here. Hence, we can easily assume that agents' availability plays a very central part not only in the happening of the collective learning process, but also in the speed of this process.



3.2.3. The core of the info-providers population:

We can see from table 4, 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 perfect-information, even though this agent has been removed form most info-seekers' set of info-providers, at the end of the simulations, i.e. when info-seekers' population is empty, it is still the agent that received the most questions during the simulations.
That is because, before being removed from most info-seekers' set of info-providers, all info-seekers addresses their questions to agent 1. Therefore it got more questions than any other agent in the community. Besides, it is the agent with the highest reputation over the last 10 steps as we can see in table 4. That is because all remaining

info-seekers addressed their questions to this agent and all other info-providers didn't get any question because they were removed from each info-seeker's set of info-providers long before the end of the simulations. An agent's reputation being calculated over the last 10 steps, agent 1 is the only agent that still receives questions from a chosen few agents. All other info-providers are being ignored by the remaining info-seekers.

Number of questions allowed	Agents in the core	Reputation	Number of questions received
1	1	1	574
2	1	1	682
3	1	1	657
4	1	1	778
5	1	1	863
6	1	1	971
7	1	1	1162
8	1	1	1280
9	1	1	1255
10	1	1	1403

Table 4: The core of the info-providers population in simulation with reputation

3.3. Access to information:

Let's see now how info-seekers get access to the information needed to increase their competencies. What is the best strategy that can allow agents to access information faster? 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. 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.
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 behave the first info-seeker to turn to info-provider, and how behave the first info-seeker that left the community.

In perfect-information simulations, agent 35 which was the first info-seeker and only 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 form most info-seekers' set of info-providers as soon as the 5th 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 6th step on, it was the only agent asking questions to agent 1. Hence, it always got a positive answer.

In simulations with reputation, 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 46th 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 perfect-information and simulations with reputation, agent 11 was among the first agents leaving the community. In fact, like 53 other info-seekers in perfect-information simulations and 28 other info-seekers in simulations with reputation, it left the community as soon as the 41st step, after asking all info-providers and having 4 negative answers from each info-provider. It seems then that if an agent wants to increase its competency, it should always ask the most competent one.

4. Conclusion:

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 trust (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 individual learning happened. This also had an impact on collective learning. In fact, individual learning went fast enough to let info-seekers' competencies, which started at 0, reach 1 after about 104 time steps in perfect-information simulations, and after about 143 time steps in simulations with reputation. Hence, the number of info-providers in the community increased and reached 20 agents in some simulations.

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 is a very interesting outcome. This means that info-seekers were able to judge info-providers' competencies according to their reputation. Therefore, even if agents know nothing about others' competencies, they can still acquire that knowledge through the repeating of interactions, by trusting the information they get from other info-seekers concerning info-providers' competencies. This implies a certain degree of trust between info-seekers.

Given these outcomes, in trust-based environment, the more available info-providers are, the faster info-seekers learn individually; and the faster collective learning occurs. This clearly shows that in a network where agents seek to increase their competencies in the use of a specific tool, trust and availability are very important conditions, for both individual and collective learning.

However, these results should also depend on the environment and the status of the community. According to Nickols (2000), members of a community of practice would lose the will to improve their practice if an organisation attempted to control this community, and the community of practice would soon disappear. Hence, the status of the community plays a great part in its success. Future development of this model would include an indicator to test the impact that would have the status of a community on the learning of its members.

Besides, in this model, we only treated info-providers availability. Further research should concern info-seekers' willingness to stay in the community as mutual and voluntary engagement is also considered as a key feature in the success of a community of practice. The number of negative answers that an agent is willing to accept from an info-provider before removing it from its set of info-providers could be a good indicator for testing this parameter.

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: Bourgine 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.