INTERACTIVE MULTI-AGENT LEARNING

Pedro Rafael Graça and Graça Gaspar
Department of Computer Science
University of Lisbon
Bloco C5 - Piso 1 - Campo Grande, 1700 Lisboa, Portugal
E-mail: prafael@di.fc.ul.pt, gg@di.fc.ul.pt

Abstract

The ability to learn is essential to intelligent agents that need to adapt to dynamic, non-deterministic environments. Traditional machine learning mechanisms can be used in a straightforward way in single agent individual learning; however, if agents are to learn as a team and benefit from each other's effort, these mechanisms must be extended towards a multi-agent learning perspective. To do so, additional ways of sharing and gathering information have to be considered so that the agents can interact during the learning process. Since a multi-agent system is a system of agent interaction, its inborn potential can be used to serve this purpose. In other words, rather than simply extending the traditional individual learning mechanisms with specific interaction protocols, the activity that concerns the intermediate steps of the collective learning process can use the structure of the multi-agent system, benefit from its potentialities, and follow the flow of agent interaction. This form of multi-agent learning is called interactive multi-agent learning and there are good reasons to expect that it can lead to better results in dealing with certain problems that appeal to collective learning.

1. Introduction

In this paper we address the recent field of study of interactive multi-agent learning (IML), proposing answers for two main questions:

- Why should interactivity be introduced in multi-agent learning?
- How can multi-agent learning become an interactive process?

Concerning the first question, we discuss the costs and benefits that can result from IML and compare it with other multi-agent learning perspectives. As for the second question, we discuss what kind of information should be exchanged between the agents and how should that exchange occur in an interactive way. Furthermore, we complement the theoretical discussion by showing how to modify and extend a case-based reasoning traditional individual learning mechanism so that it can be used as an IML mechanism.

IML is a field of study that, although considered interesting and promising, has not been broadly researched yet. Since its boundaries are still somehow undefined, we will begin by distinguishing it from other types of multi-agent learning. For that, we will consider the three categories proposed in [8] and present our own generic definition of IML.

We will use an example that helps to illustrate these three categories. For this effect, imagine a place where all meals can be bought through three different services: take away, snack bar or restaurant. Imagine also that this place, called "Three Orientals", serves traditional food from India, China and Japan, and that all the recipes used are exclusive and secret. Moreover, these recipes are very difficult to learn and take a long time to master.

1.1 Multiplied Learning

In this type of multi-agent learning, each agent has an individual learning mechanism and does not share his results. The learning effort is multiplied by every individual involved in it. Multiplied learning can be a solution when agents cannot share information (privacy, hostility) or when communication cannot be supported (communication too expensive, lack of resources).

As an example, consider that the "Three Orientals" has three isolated kitchens, one for each of the three different meal services. Imagine that three chefs are hired, one for each kitchen. Since they do not communicate with each other, each one of them has to learn how to cook all the secret recipes by himself. The learning effort is multiplied.

1.2 Divided Learning

When information exchange is possible, multiplied learning becomes redundant. The division of the learning task among the agents reduces this redundancy, but depends on several initial definitions:

- Which are the learning sub-tasks;
- Which agents execute each sub-task;
- When and how is the knowledge (acquired in each sub-task) transmitted and integrated.

These initial settings are the limits within which divided learning occurs, but such static definitions may not be possible (for example, the learning task may be too complex to be clearly divided and distributed) or desirable (for example, in an unstable or highly unpredictable environment, the transmission of information may become too difficult and that may hinder the sharing of acquired knowledge on preset occasions).

Consider now that the "Three Orientals" has only one kitchen. The learning effort can, for example, be divided among the three chefs in the following ways:

- one learns to prepare the Indian dishes, the second one the Chinese dishes and the third one the Japanese dishes;
- one learns how to prepare the ingredients, the second one how to cook and the third one how to gather the cooked food and arrange the dishes.

By dividing the learning task, the redundancy of the collective effort is greatly reduced. Notice however that such initial divisions are not always adjusted. For example, if all the services of the "Three Orientals" were reserved by the "Sushi Lovers Association" during one month, neither of the suggested divisions would be appropriate, because during that time only one of the chefs would be actively learning.

1.3 Interactive Learning

In interactive multi-agent learning there also exists a division of the learning task, but in this case the division is not initially set. The two basic principles of IML are:

- to let the succession of environmental events and social interactions guide the path of the learning process;
- to use the potentialities of multi-agent systems on behalf of the learning process.

The idea behind these principles is that, by considering the successive environmental and social states to interactively define well adjusted learning sub-tasks, by selecting agents to execute these sub-tasks and ways of sharing knowledge, and by using the means of social interaction provided by a multi-agent system to benefit distributed learning, the global results of the collective learning process can be enhanced. In IML, learning ceases to be a process which is isolated from other social processes. Instead, as it happens in human societies, it assumes its place inside the social system and follows its patterns of interaction.

In the shared kitchen of the "Three Orientals", the process of interactively learning how to cook the secret recipes takes place in an unpredictable way. The three chefs communicate frequently with each other and decide what is the best thing to do at each moment. They influence each other's learning path with advices and opinions. With this system they can, for example, identify relationships between the ways of preparing the same ingredients for different dishes; for instance, part of the knowledge acquired by one chef while cooking Chinese rice can help another chef that is learning how to cook Indian rice. When the "Sushi Lovers Association" reserves the services of the "Three Orientals", the three chefs adjust to the situation, divide the available tasks among them and are always able to keep learning.

2. Related work

The subject of multi-agent learning is addressed in a general way in [8], by proposing a distinction between multiplied, divided and interactive learning. To these authors, it is this last category that may "explain all the benefits of multi-agent learning".

In [2] we also find a general discussion about multi-agent learning. The authors consider interactive learning to describe "true distributed learning" and they add that "very little research exists in this area".

Some works ([3] and [7] are examples) address the subject of multi-agent reinforcement learning in ways that show some degree of agent interactivity.

Techniques that integrate case-based reasoning into multi-agent systems, allowing agents to interact and exchange learned cases, are proposed in [5]. More recently, the same authors introduced a case bartering protocol to improve multi-agent learning ([6]). Although their work addresses agent interaction over previously learned information, it does not discuss the exchange of data during the learning process.

The benefits of mutual agent interaction during the learning process are discussed in [4]. where an advice exchange technique that allows cooperative learning is introduced. Although we also propose advice exchange (among other ways of interaction) in our investigation, there are important differences in the way the exchange is made and in the kind of information exchanged.

The research reported in [1] shows that, on a dynamic multi-agent environment, the collective learning task can be optimised through simple agent interaction.

3. Interactivity and communication

Learning in an interactive way depends on the existence of a communication system that allows information exchange so that the learning path can be successively defined and the partial results of the learning effort can be efficiently shared. The type of information exchanged and the ways in which this exchange occurs must express this interactivity.

3.1 Type of Information Exchanged

We can distinguish between two types of information that can circulate among the agents in order to allow IML: information that concerns the learning process specifically, and information that describes the state of the environment and the agents.

3.1.1 Information about the learning task

The interactive circulation of specific knowledge that is related to the learning effort constitutes the core of IML, allowing agents to share information and complement their individual experience with external data gathered from other agents. The agents can share not only raw results (for example, a set of cases in case-based reasoning) and refined results (for example, a generalization of a set of cases), but also specific units of information that allow a richer interaction during the learning process. These units of information can be such as:

- a problem description;
- a solution: for example, the solution that an agent would select to solve a given a problem;
- a justification: raw or refined data that supports a proposed solution;

- an evaluation: a measure of how good a given solution is for solving a problem.

Such units can be interactively created by the agents (according to their current experience) and allow the exchange of intermediate results (hypothesis) of the learning process. Here are some examples of messages that can be used to regulate the interactive circulation of this information and allow the exchange of partial results at any intermediate stage of the learning process:

- Suggestion: A possible solution for a problem.
- Opinion: An evaluation of how adjusted a possible solution is for a problem.
- Help request: A request for advice from another agent that includes a problem description. It may also include a suggestion and its justification.
- Advice: The answer to a help request that may include a suggestion and its justification.

3.1.2 Information about the environment

As mentioned before, information that concerns the state of the environment and of the multi-agent system can also circulate in order to help the definition of sub-tasks and the selection of agents to execute them. Here are some ideas of how the flow of social interactions can help to define the path of the learning process:

- Learning sub-tasks: At each moment, the current state of the environment can be taken into account for the definition of sub-tasks that can be executed in an easier or more advantageous way;
- Selection of agents: The agents' individual state can be considered as a selection criteria (for example, the importance of the current task can be taken into consideration); the current relations between agents can be considered when forming teams to execute learning sub-tasks;
- Knowledge sharing: When in need, the agents can ask for the help of other agents; agents can form groups to debate different opinions on possible solutions for a problem; contacts established for other purposes can be used as a way of also sharing experience acquired through learning.

3.2 Ways of Information Exchange

The generic communication protocols that, in each multi-agent system, serve agent interaction, can be used to exchange information that regards the learning process. For instance, when agents communicate while executing a non-learning task, it may be possible to use this contact to also exchange learned data (for example, suggestions could be made).

In order to make better use of agent interaction on behalf of the learning process, it may be useful to introduce specific interaction protocols. These protocols may facilitate the transmission and exchange of specific data units (for example, providing the diffusion of help requests) and support specific forms of group discussion that address cooperative learning (for example, voting schemes).

4. Costs and advantages of IML

Not all problems that involve multi-agent learning appeal for interactivity. For instance, when there are strong restrictions to communication, when the agents do not cooperate or are hostile, or when the problem has a degree of predictability according to which a sequence of collective learning steps (divided learning) can be easily set, then IML may not be a viable or advisable option. The environments that appeal for interactivity are those that concern multi-agent learning on dynamic and unpredictable problems that can be solved through mutual cooperation. It may be adjusted to say that, in these cases, it is through IML that the potentialities of a multi-agent system can better serve the learning process; however, since the introduction of interactivity in learning has relevant costs, it is essential to discuss the reasons that lead us to believe that IML mechanisms can allow better results.

4.1 Costs of IML

The introduction of interactivity in multi-agent learning has important costs that must be considered. First, the ways in which agent interaction during the learning process occurs have to be planned, and this plan is generally more complex than a static initial definition (as it happens in multiplied or divided learning). Second, communication resources must be available and allow frequent information exchange. Third, agents have to be able to analyse and assimilate the information that circulates interactively during the learning process.

4.2 Advantages of IML

Comparing interactive and divided learning, we can say that the former extends the latter (the division of the learning effort becomes interactive) or that the latter is a simplified form of the former (interactivity is pre-determined). Regarding this, we discuss several advantages of IML using divided learning as a reference.

According to the divided learning perspective, results can only circulate after the learning sub-tasks are completed. When this circulation occurs at any intermediate steps of the learning process (IML perspective), useful information can be assimilated during the performance of sub-tasks, leading to the decrease of redundancy in the collective effort and allowing interactive adjustments to the path of current sub-tasks. This can also help to avoid excessive specialization (that can result from undertaking specific sub-tasks without intermediate interaction), and create a diversity of perspectives that allows different ways of addressing the same problems and the exploration of new solutions.

When the learning effort is divided, there exists an initial notion of which agents will execute which learning sub-tasks, and when will they execute them. Specially on dynamic and unpredictable environments, considering that specific conditions may temporarily hinder the execution of some of these sub-tasks, such initial settings may become misadjusted. In these cases, the possibility of interactively

defining sub-tasks and agents to execute them can allow selections that match the current conditions.

The division of the learning task creates dependencies between the agents. If an agent fails to execute his task, his fault can seriously delay or even disrupt the collective effort. In order to avoid this, the accomplishment of each critical duty must be assured. This may prevent the application of divided learning mechanisms to environments where agents are to have a considerable degree of autonomy and be able to decide to stop learning during a period of time (for example, to attend to more important immediate tasks) or permanently (for example, to attend to a new privacy policy), or in which faults in the communication system may occur and cause the isolation of agents for a significant time extent. The flexibility of IML allows the collective learning task to proceed in these cases. If necessary, the learning effort may temporarily assume a multiplied perspective (for example, when communication becomes unavailable) and reassume its interactivity when possible.

As a reference, imagine the cooperation between a group of ecologists and a group of travel agents over the management of a natural reserve. The ecologists wish to preserve the reserve but need money to take good care of it, and the travel agents want money but they need a well-kept reserve to attract tourists. In learning problems that appeal for this kind of cooperation, IML mechanisms can be specially well-adjusted, allowing the crucial interactive exchange of information between sets of agents that perform different but complementary tasks.

Attending to these potential advantages, there are good reasons to broaden the research of IML, for it may allow a faster collective learning with better results, provide better tolerance to agent or system faults, and address a wider scope of problems.

5. Cooperative case-based reasoning

At this stage of our research, we are focusing our attention on the adaptation of a case-based reasoning (CBR) traditional individual learning mechanism so that it can be used as an IML mechanism. We propose a cooperative CBR generic learning system that allows agent interaction during the collective learning effort. In doing this, we wish to illustrate how some of the ideas presented above can be applied.

5.1 Traditional and Multi-agent CBR

The traditional individual CBR generic learning process is based on the following four basic steps:

1. Retrieval of previously gathered cases (that concern problems somehow similar to the current problem);
2. Selection of a solution;
3. Revision of the solution (resolution of the problem and evaluation of the solution);
4. Storage of a new case.

In multi-agent CBR, this generic individual process has to be modified in order to include ways of information exchange between the agents.

The information used in the traditional CBR process is retained in a single case base, whose contents express, at each moment, the experience acquired. In multi-agent CBR, different types of case bases can be used: individual or collective, centralised or distributed. Our research focuses on the use of individual case bases, a scenario in which information exchange is crucial for the collective learning process, allowing in this sense a more deep and complete discussion concerning the introduction of interactivity. The use of separate case bases can be justified for reasons like the need for privacy (for example, the existence of a secrecy policy concerning some of the data retained in individual case bases) or the need for efficiency on data access or storage (for example, to avoid traffic bottleneck situations that can result from having a collective case base).

5.2 Cooperative CBR

Using the traditional mechanism as reference, we have identified three situations that are suitable for interactive information exchange. The first is during the process of consulting previously gathered cases and deciding which solution to select to solve the current problem: instead of simply consulting his individual case base, an agent may request the help of other agents, consider their advices and discuss possible solutions with them. The second situation is during the revision of the chosen solution: while trying to solve a problem, an agent may again consult other agents to exchange information and interactively adjust his procedure. The third situation is during the storage of a new case: the new case can be shared among several agents (possibly by those involved in the previous two situations). In Table 1, we extend the basic CBR steps in order to incorporate these information exchange substeps.

Table 1: Cooperative CBR learning process

Basic process	Extensions
I. Information retrieval	a) Retrieval of individual cases
	b) Access to external information (request help from other agents and gather advices)
II. Selection of a solution	a) Analysis of internal and external information
	b) Selection of a solution
III. Revision of the solution	a) Access to external information to adjust the solution while solving the problem
	b) Evaluation of the solution
IV. Storage of a new case	a) Creation of a new case
	b) Distribution of the new case for storage in different individual case bases

Depending on each agent's individual desire, the opportunities for interactive exchange of information can be used or ignored. The information that circulates whenever an agent decides to ask for help can be useful not only to himself but also to the other agents contacted. For instance,

during the process of selecting a solution for a problem, if the analysis of the information is done in group, then all the agents involved have the opportunity to retain useful knowledge that circulates during the discussion.

5.3 Agent interaction in cooperative CBR

The interactive information exchange between the learning agents involved in cooperative CBR appeals for specific processes of interaction. Considering the first of the three situations of interactive information exchange identified before, we propose two different protocols to regulate the process of accessing external information, analysing it and selecting a solution (steps I.b), II.a) and II.b) of the cooperative CBR learning process presented in Table 1). In these two protocols, the analysis of the information and the decision is centralised on one agent. This makes the process simpler and less expensive (regarding resource consumption); the distributed variants (ways of group discussion, voting schemes) are more complex and expensive, but offer powerful options for interaction. Choosing between centralised and distributed interaction depends upon the specific nature of the learning problem and the available resources.

In the tables ahead (tables 2 and 3) we show the details of the two protocols, specifying how each step corresponds to the substeps of the generic cooperative CBR learning process.

Let **A** be the agent that is searching for a solution and that decides to ask a group $\mathbf{G} = \{\mathbf{B}_{(1)}, \mathbf{B}_{(2)}, \ldots \mathbf{B}_{(n)}\}$ of **n** other agents for advice. In the first sequence (Table 2), agent **A** sends the same help request to all agents of **G**, gathers their advices, analyses the information available, and then selects a solution.

Table 2: Interactive information exchange: fixed help request

Fixed help request
Step 1 - Help request generation
Agent **A** consults his case base and generates a help request **Hreq** = [Problem, Suggestion, Justification] (corresponds to I.a)).
Step 2 - Information exchange
2a) Agent **A** sends **Hreq** to each agent of **G** (corresponds to the first part of I.b)) .
2b) Each agent $\mathbf{B}_{(i)}$ of **G** answers by sending his advice $\mathbf{Adv}_{(i)}$ = [$\text{Suggestion}_{(i)}$ + $\text{Justification}_{(i)}$]) to **A** (corresponds to the second part of I.b)).
Step 3 - Selection of a solution
Agent **A** analyses (corresponds to II.a)) the information available and selects a solution (corresponds to II.b)).

In the second proposed sequence (Table 3), agent **A** sends one help request at a time, and waits for each answer. In this case, agent **A** changes the contents of his help request whenever he receives a better suggestion, one that has a justification that is considered stronger. Considering that justifications are composed by sets of cases or its generalizations, a stronger justification may, for example, be one that:

- Contains the case that better matches the current problem and/or has a better solution (according to the evaluation criteria in use);
- Expresses a higher experience level.

Table 3: Interactive information exchange: variable help request

Iterative help request
Step 1 - Help request generation
Agent **A** consults his case base and generates a help request **Hreq** = [Problem, Suggestion, Justification] (corresponds to I.a)).
Step 2 - Information exchange
Agent A exchanges and analyses information (corresponds to I.b) and II.a)) according to the following algorithm: -Initialise **Hvar** = **Hreq** -For each agent $\mathbf{B}_{(i)}$ of **G** do: a) Send **Hvar** to $\mathbf{B}_{(i)}$ b) Receive $\mathbf{Adv}_{(i)}$ = [$\text{Suggestion}_{(i)}$ + $\text{Justification}_{(i)}$] c) If $\text{Justification}_{(i)}$ is stronger than Justification, then: Suggestion = $\text{Suggestion}_{(i)}$, Justification = $\text{Justification}_{(i)}$
Step 3 - Selection of a solution
Agent **A** analyses (corresponds to II.a)) the information available and selects a solution (corresponds to II.b)).

It is important to notice that the process of selecting a solution can consist in more than simply choosing one of the suggestions; instead, the information received can be used to compose solutions that combine aspects of different suggestions.

These sequences describe the process of interaction from the point of view of agent **A**. There is however another process involved in this interaction: the process of generating an advice that is performed by each agent $\mathbf{B}_{(i)}$ of **G**. In Table 4 we present a generic sequence of steps that describes this process. This process also involves information retrieval and the selection of a solution, and in this sense, some of its steps also correspond to some of the substeps of the generic cooperative CBR learning process.

Table 4: Process of advice generation

Advice generation
Step 1 - Information retrieval
1a) Agent $\mathbf{B}_{(i)}$ receives a help request **Hreq** from agent **A** (**Hreq** = [Problem, Suggestion, Justification]).
1b) $\mathbf{B}_{(i)}$ consults his case base and gathers information on how to solve the current Problem (corresponds to I.a)).
Step 2 - Composition of a suggestion
Considering the retrieved information (if there is any), $\mathbf{B}_{(i)}$ composes a $\text{Suggestion}_{(i)}$ and its $\text{Justification}_{(i)}$.
Step 3 - Creation of an advice
If $\mathbf{B}_{(i)}$ considers his $\text{Suggestion}_{(i)}$ useful (after comparing it with the Suggestion received from agent **A** and also considering both justifications; corresponds to II.a)), then he creates a new advice ($\mathbf{Adv}_{(i)}$ = [$\text{Suggestion}_{(i)}$ + $\text{Justification}_{(i)}$]) and sends it to agent **A** (corresponds to II.b)).

The second situation of possible agent interaction is during the resolution of a problem. At this time, an agent may again consult other agents to exchange information and interactively adjust his procedure. The processes that regulate this interaction are similar to the ones proposed for the first situation.
The third situation occurs during the storage of a new case, when it can be shared among several agents. The process associated to this third situation consists in simply performing a selection of a group of agents to receive the information and send the new case to them. The crucial point in this process is the selection criteria. One idea is to consider those agents involved in the previous two situations and analyze the information they shared (their advices) in order to choose good candidates (for example, an agent whose advice shows lack of experience could be considered a good candidate).

5.4 Example of agent interaction

The following example resumes the application of the processes of agent interaction previously described. Furthermore, it suggests the kind of motivations and situations that may trigger the interaction and define how each agent gets involved in it.
Suppose that after consulting his case base, agent **A** finds out that he has little experience on how to solve his current problem. Realizing that, he generates a help request (that includes a suggestion and its justification) and decides to use a variable help request protocol to send it to agents **B**, **C** and **D**. Agent **B** is the first to be contacted and answers that he has no experience in solving the current problem. Agent **C** is the second to receive **A**'s request and, after consulting his case base, he replies with an advice that suggests a different solution. Upon analysing **C**'s advice, **A** decides to modify his help request (replacing the current suggestion and its justification with the ones received from **C**) and then sends it to the third agent. Agent **D** analyses the request and concludes that he agrees with the current suggestion. In his advice he confirms the suggestion and adds his own justification. After receiving this last advice, agent **A** decides to accept the external suggestion to solve the current problem. During the resolution of the problem, he finds no need for external help (he skips the second situation of interaction). After solving the problem, agent **A** creates a new case, adds it to his case base, and, since agents **C** and **D** already have considerable experience in the subject, sends only one copy of the case to agent **B**.
It is important to notice that whenever the information that circulates during the interaction is considered useful it can be stored by the agents. For example, since agent **B** had no experience on the problem for which his help was requested, he could have considered useful to store the information present in the help request.

6. Conclusion

We have proposed a definition of IML, discussed its costs and advantages, and described ways for introducing interactivity in multi-agent learning. Using a generic CBR learning mechanism, we showed how the general principles of IML can be applied. Furthermore, we proposed specific processes to regulate agent interaction during the learning task and showed how the information can circulate in an interactive way.
Soon we will test our ideas on a simulation workbench that is currently under construction. A specific multi-agent learning problem (using CBR) will help us to obtain results that can further clarify the potentialities of IML.

References

[1] Graça, P. R. and Gaspar, G. *Using Cognition and Learning to Improve Agents' Reactions*. In E. Alonso, D. Kudenko and D. Kazakov (eds.), *Adaptive Agents and Multi-Agent Systems*. series LNAI 2636, pp. 239-259, Springer, 2003.

[2] Kazakov, D. and Kudenko, D. *Machine Learning and Inductive Logic Programming for Multi-agent Systems*. In M. Luck, V. Marik and O. Stepankova (eds.), *Multi-Agent Systems and Applications*, series LNAI 2086, pp. 246-270, Springer, 2001.

[3] Makar, R., Mahadevan, S., and Ghavamzadeh, M. *Hierarchical Multi-Agent Reinforcement Learning*. In Proceedings of the 5th Conference on Autonomous Agents (Agents'01), Montreal, 2001.

[4] Nunes, L. and Oliveira, E. C*ooperative Learning Using Advice Exchange*. In E. Alonso, D. Kudenko and D. Kazakov (eds.), *Adaptive Agents and Multi-Agent Systems*. series LNAI 2636, pp. 33-48, Springer, 2003.

[5] Ontañon, S. and Plaza, E. *Collaboration Policies for Case-based Reasoning Agents*. In Proceedings of the Workshop on Learning Agents Autonomous (Agents'01), Montreal, 2001.

[6] Ontañon, S. and Plaza, E. *A Bartering Approach to Improve Multiagent Learning*. In Proceedings of the 1st Conference on Autonomous Agents and Multiagent Systems (AAMAS 2002), Bologna, 2002.

[7] Turner, K., Agogino, A. K., and Wolpert, D. *Learning Sequences of Actions in Collectives of Autonomous Agents*. In Proceedings of the 1st Conference on Autonomous Agents and Multiagent Systems (AAMAS 2002), Bologna, 2002.

[8] Weiβ, G. and Dillenbourg, P. *What is 'Multi' in Multi-Agent Learning*. In P. Dillenbourg (ed.), *Collaborative Learning*, pp. 64-80, Pergamon Press, 1999.