Weiss, G.,& Dillenbourg P. (1999) . What is 'multi' in multi-agent learning? In P. Dillenbourg (Ed) *Collaborative-learning: Cognitive*

Chapter 4

What is ‘multi’ in multi-agent learning?

Gerhard Weiß
Institut für Informatik, Technische Universität München, Germany
weissg@informatik.tu-muenchen.de

Pierre Dillenbourg
Faculty of Education & Psychology, University of Geneva, Switzerland
Pierre.Dillenbourg@tecfa.unige.ch

Abstract

The importance of learning in multi-agent environments as a research and application area is widely acknowledged in artificial intelligence. Although there is a rapidly growing body of literature on multi-agent learning, almost nothing is known about the intrinsic nature of and requirements for this kind of learning. This observation is the starting point of this chapter which aims to provide a more general characterization of multi-agent learning. This is done in an interdisciplinary way from two different perspectives: the perspective of single-agent learning (the ‘machine learning perspective’) and the perspective of human-human collaborative learning (the ‘psychological perspective’). The former leads to a ‘positive’ characterization: three types of learning mechanisms - multiplication, division, and interaction - are identified and illustrated. These can occur in multi-agent but not in single-agent settings. The latter leads to a ‘negative’ characterization: several cognitive processes like mutual regulation and explanation are identified and discussed. These are essential to human-human collaborative learning, but have been largely ignored so far in the available multi-agent learning approaches. Misunderstanding among humans is identified as a major source of these processes, and its important role in the context of multi-agent systems is stressed. This chapter also offers a brief guide to agents and multi-agent systems as studied in artificial intelligence, and suggests directions for future research on multi-agent learning.

1. Introduction: What is intrinsic to multi-agent learning?

Learning in multi-agent environments[1] had been widely neglected in artificial intelligence until a few years ago. On the one hand, work in distributed artificial intelligence (DAI) mainly concentrated on developing multi-agent systems whose activity repertoires are more or less fixed; and on the other hand, work in machine learning (ML) mainly concentrated on learning techniques and methods for single-agent settings. This situation has changed considerably and today, learning in multi-agent environments constitutes a research and application area whose importance is broadly acknowledged in AI. This acknowledgment is largely based on the insight that multi-agent systems typically are very complex and hard to specify in their dynamics and behavior, and that they therefore should be equipped with the ability to self-improve their future performance. There is a rapidly growing body of work on particular algorithms and techniques for multi-agent learning. The available work, however, has almost nothing to say about the intrinsic nature of and the unique requirements for this kind of learning. This observation forms the starting point of this chapter which is intended to focus on the characteristics of multi-agent learning, and to take the first steps toward answering the question *What is 'multi' in multi-agent learning?*

Our approach to this question is interdisciplinary, and leads to a characterization of multi-agent learning from two different perspectives:

- single-agent learning (*'machine learning perspective'*) and
- human-human collaborative learning (*'psychological perspective'*).

The ML perspective is challenging and worth exploring because until some years ago learning in multi-agent and in single-agent environments have been studied independently of each other (exceptions simply prove the rule). As a consequence, only very little is known about their relationship. For instance, it is still unclear how exactly a computational system made of multiple agents learns differently compared to a system made of a single agent, and to what extent learning mechanisms in multi-agent systems differ from algorithms as they have been traditionally studied and applied in ML in the context of single-agent systems. The psychological perspective is challenging and worthy of exploration because research on multi-agent learning and research on human-human collaborative learning grew almost independently of each other. Despite the intuitively obvious correspondence between these two fields of research, it is unclear whether existing multi-agent learning schemes can be observed in human environments, and, conversely whether the mechanisms observed in human-human collaborative learning can be applied in artificial multi-agent systems.

[1] Throughout this text, the phrase 'multi-agent' refers to artificial agents as studied in artificial intelligence (as in e.g. 'multi-agent environment' or 'multi-agent learning') and does not refer to human agents. Human agents will be explicitly mentioned when comparing work on multi-agents with psychological research on collaborative learning.

The structure of this chapter is as follows. Section 2 provides a compact guide to the concepts of agents and multi-agent systems as studied in (D)AI. The purpose of this section is to establish a common basis for all subsequent considerations. This section may be skipped by the reader who is already familiar with the (D)AI point of view of these two concepts. Section 3 then characterizes multi-agent learning by focusing on characteristic differences in comparison with single-agent learning. Section 4 then characterizes multi-agent learning by contrasting it with human-human collaborative learning. Finally, section 5 concludes this chapter by briefly summarizing the main results and by suggesting potential and promising directions of future research on multi-learning.

2. **A brief guide to agents and multi-agent systems**

Before focusing on multi-agent learning, it is useful and necessary to say a few words about the context in which this kind of learning is studied in DAI. This section offers a brief guide to this context, and tries to give an idea of the AI point of view of agents and multi-agent systems. The term ‘agent’ has turned out to be central to major developments in AI and computer science, and today this term is used by many people working in different areas. It is therefore not surprising that there is considerable discussion on how to define this term precisely and in a computationally useful way. Generally, an agent is assumed to be composed of four *basic components*, usually called the sensor component, the motor component, the information base, and the reasoning engine. The sensor and motor components enable an agent to interact with its environment (e.g. by carrying out an action or by exchanging data with other agents). The information base contains the information an agent has about its environment (e.g. about environmental regularities and other agents' activities). The reasoning engine allows an agent to perform processes like inferring, planning and learning (e.g. by deducing new information, generating behavioral sequences, and increasing the efficiency of environmental interaction). Whereas this basic conception is commonly accepted, there is ongoing controversy on the characteristic properties that let an object like a program code or a robot be an agent. Following the considerations in (Wooldridge and Jennings, 1995b), a weak and a strong notion of agency can be distinguished. According to the *weak notion*, an agent enjoys the following properties:

- autonomy,
- reactivity, and
- pro-activeness.

According to the more specific and *strong notion*, additional properties or mental attitudes like

- belief, knowledge, etc. (describing information states),
- intention, commitment, etc. (describing deliberative states), and
- desire, goal, etc. (describing motivational states)

are used to characterize an agent. Examples of 'typical' agents considered in (D)AI are shown below. The reader interested in more details on AI research and application on agency is particularly referred to (Wooldridge & Jennings, 1995a; Wooldridge, Müller & Tambe, 1996; Müller, Wooldrige & Jennings, 1997).

In artificial intelligence recent years have shown a rapidly growing interest in systems composed of several interacting agents instead of just a single agent. There are three major reasons for this interest in multi-agent systems:

- they are applicable in many domains which cannot be handled by centralized systems;
- they reflect the insight gained in the past decade in disciplines like artificial intelligence, psychology and sociology that intelligence and interaction are deeply and inevitably coupled to each other;
- recently, a solid platform of computer and network technology for realizing complex multi-agent systems is available.

Many different multi-agent systems have been described in the literature. According to the basic conception and broadly accepted viewpoint, a multi-agent system consists of several interacting agents which are limited and differ in their motor, sensory and cognitive abilities as well as in their knowledge about their environment. We briefly illustrate some typical multi-agent scenarios.

One widespread and often used scenario is the transportation domain (e.g. Fischer et al., 1993). Here several transportation companies, each having several trucks, transport goods between cities. Depending on the level of modeling, an individual truck, coalitions of trucks from the same or different companies, or a whole company may correspond to and be modeled as an agent. The task that has to be solved by the agents is to complete customer jobs under certain time and cost constraints. Another example of a typical scenario is the manufacturing domain (e.g. Peceny, Weiß & Brauer, 1996). Here several machines have to manufacture products from raw material under certain predefined constraints like cost and time minimization. An individual machine as well as a group of machines (e.g. those machines responsible for the same production step) may be modeled as an agent. A further example is the loading dock domain (e.g. Müller & Pischel, 1994). Here several forklifts load and unload trucks according to some specific task requirements. Analogous to the other domains, the individual forklifts as well as groups of forklifts may correspond to the agents. As a final example the toy-world scenario known as the predator/prey domain should be mentioned (e.g. Benda, Jaganathan & Dodhiawala, 1986). Here the environment consists of a simple two-dimensional grid world, where each position is empty or occupied by a predator or prey. The predators and prey (which can move in a single-step modus from position to position) correspond to the agents, and the task to be solved by the predators is to catch the prey, for instance by occupying their adjacent positions. Common to these four examples is that the agents may differ in their abilities and their information about the domain, and that aspects like interaction, communication, cooperation, collaboration, and negotiation are of particular relevance to the solution of the task. These examples also give an impression of the three *key aspects* in which multi-agent systems studied in DAI differ from each other (Weiß, 1996):

- the environment occupied by the multi-agent system (e.g. with respect to diversity and uncertainty),
- the agent-agent and agent-environment interactions (e.g. with respect to frequency and variability), and
- the agents themselves.

The reader interested in available work on multi-agent systems is referred to the Proceedings of the First and Second International Conference on Multi-Agent Systems (1995, 1996, see references). Standard literature on DAI in general is e.g. (Bond & Gasser, 1988; Gasser & Huhns, 1989; Huhns, 1987; O'Hare & Jennings, 1996).

3. The ML perspective: Multi-agent vs. single-agent learning

As mentioned above, multi-agent learning is established as a relatively young but rapidly developing area of research and application in DAI (e.g. Imam, 1996; Sen, 1996; Weiß & Sen, 1996; Weiß, 1997). Whereas this area had been neglected until some years ago, today it is commonly agreed by the DAI and the ML communities that multi-agent learning deserves particular attention because of the inherent complexity of multi-agent systems (or of distributed intelligent systems in general). This section characterizes and illustrates three classes of mechanisms which make multi-agent learning different from single-agent learning: *multiplication, division* and *interaction*. This classification provides a *'positive' characterization* of multi-agent learning (we later present also a 'negative' classification). For each mechanism, a general 'definition' is provided and several representative multi-agent learning techniques are briefly described.

This coarse classification does not reflect the variety of approaches to multi-agent learning. Each class is an abstraction of a variety of algorithms. Most work on multi-agent learning combines these mechanisms, and therefore a unique assignment is not always possible. This also means that the three classes are not disjointed, but are of different complexity and partially subsume one another: divided learning includes elements of multi-plied learning, and interactive learning includes elements of divided learning. Our goal was not to compare all existing work in DAI, but to isolate what is very specific to multi-agent learning and has not been investigated in single-agent learning. There are of course a few exceptions, since there is for instance something 'multi' in research on parallel inductive learning (e.g. Chan & Stolfo, 1993; Provost & Aronis, 1995) or on multistrategy learning (e.g. Michalski & Teccuci, 1995).

3.1. Multiplication mechanisms

General. If each agent in a multi-agent system is given a learning algorithm, the whole system will learn. In the case of multi-plied learning there are several learners, but each of them learns independently of the others, that is, their interactions do not impact their individual learning processes. There may be interactions among the agents, but these interactions just provide input which may be used in the other agents' learning processes. The learning processes of agents but not the agents themselves are, so to speak, isolated from each other. The individual learners may use the same or a different learning algorithm. In the case of multi-plied learning each individual learner typically pursues its own learning goal without explicitly taking notice of the other agents' learning goals and without being guided by the wish or intention to support the others in achieving their goals. (The learning goals may mutually supplement each other, but this is more an 'emerging side effect' than the essence of multi-plied learning.) In the case of multi-plied learning an agent learns 'as if it were alone', and in this sense has to act as a *'generalist'* who is capable of carrying out *all* activities that as a whole constitute a learning process.

At the group level, the learning effects due to multiplication may be related to variables such as the number of agents or their heterogeneity. For instance, if the system includes several identical robots, it is less dramatic if one of them breaks down (technical robustness). If the systems includes agents with different background knowledge, it may work in a wider range of situations (applicability). The effects of multiplication are salient in sub-symbolic systems or systems being active below the knowledge level which include a large number of elementary learning units (which, however, can hardly be called agents). A good example of such systems are artificial neural networks composed of large numbers of neurons.

Examples of 'multi-plied learning'. Ohko, Hiraki and Anzai (1996) investigated the use of case-based learning as a method for reducing communication costs in multiagent systems. As an application domain a simulated environment occupied by several mobile robots is used. The robots have to transport objects between predefined locations, where they differ from each other in their transportation abilities. Each robot has its own case base, where a single case consists of the description of a transportation task and a solution together with its quality. Several agents may learn (i.e. retrieve cases and detect similarities among cases) concurrently, without influencing each other in their learning processes. Although the robots interact by communicating with each other in order to announce transportation jobs and to respond to job announcements, the learning processes of different agents do not interfere with each other.

Haynes, Lau and Sen (1996) used the predator/prey domain in order to study conflict resolution on the basis of case-based learning. Here two types of agents, predators and prey, are distinguished. According to this approach each predator can be thought of as handling its own case base, where a single case consists of the description of a particular predator/prey configuration and a recommended action to be taken in this configuration. It is clear that predators interact by occupying the same environment and by sensing each others' position. Despite this, however, a predator conducts its learning steps independently of other predators and their learning steps (even if the different learning processes may result in an increased overall coherence).

Vidal and Durfee (1995) developed an algorithm that allows an agent to predict other agents' actions on the basis of recursive models, that is, models of situations that are built on models of other situations in a recursive and nested way. The models of situations are developed on the basis of past experience and in such a way that costs for development are limited. Here learning is multi-plied in the sense that several agents may concurrently learn about each other according to the same algorithm without requiring interaction except mutual observation. There is no joint learning goal. The learning processes of multiple learners do not influence each other, and they occur independently of each other. Observation just serves as a ‘mechanism’ that provides the information upon which a learning process is based.

Terabe et al. (1997) dealt with the question of how organizational knowledge can influence the efficiency of task allocation in multiagent settings, i.e. to assign tasks to agents in such a way that the time for completing the tasks is reduced. Organizational knowledge is information about other agents' ability to solve tasks of different types. In the scenario described by Terabe and his colleagues several agents can learn independently of each other how specific tasks are best distributed across the overall system. There is no explicit information exchange among the agents, and the only way in which they interact is by observing each others' abilities. The observations made by an agent are used to improve its estimate of the abilities of the observed agent, but the improvement itself is done by the observer. Each agent responsible for distributing a task conducts the same learning scheme, and none of the responsible agents is influenced in its learning by the other agents.

Carmel and Markovitch (1996) investigated the learning of interaction strategies from a formal, game-theoretic perspective. Interaction is considered as a repeated game, and another agent's interaction strategy is represented as a finite state-action automata. Learning aims at modeling others' ‘interaction automata’ based on past observations such that predictions of future interactions become possible. As in the research described above, the learning of the individual agents is not influenced by the other agents. The agents interact by observing each others' behavior, and the observations then serve as an input for separate learning processes. There is also no shared learning goal; instead, each agent pursues its own goal, namely, to maximize its own ‘profit’. Related, game-theoretic work on multiagent learning was presented by Sandholm and Crites (1995) and Mor, Goldman and Rosenschein (1996).

Bazzan (1997) employed the evolutionary principles of mutation and selection as mechanisms for improving coordination strategies in multiagent environments. The problem considered was a simulated traffic flow scenario, where agents are located at the intersections of streets. A strategy corresponds to a signal plan, and evolution occurs by random mutation and fitness-oriented selection of the strategies. There are several other papers on learning in multi-agent systems that follow the principle of biological evolution; for instance, see (Bull & Forgarty, 1996; Grefenstette & Daley, 1996; Haynes & Sen, 1996). These are essentially examples of the multi-plied learning approach because the agents do not influence each other in their learning processes. It has to be stressed that a classification of evolution-based learning is difficult in so far as ‘evolution’, reduced to the application of operators like mutation and selection, can be also viewed as a centralized process. From this point of view the agents just act, whereas ‘learning by evolution’ is something like a meta-activity that takes place without being influenced by the agents themselves.

3.2. Division mechanisms

General. In the case of 'divided learning' a single-agent learning algorithm or a single learning task is divided among several agents. The division may be according to functional aspects of the algorithm and/or according to characteristics of the data to be processed in order to achieve the desired learning effects. As an example of a functional division one could consider the task of learning to optimize a manufacturing process: here different agents may concentrate on different manufacturing steps. As an example of a data-driven division one could think of the task of learning to interpret geographically distributed sensor data: here different agents may be responsible for different regions. The agents involved in divided learning have a shared overall learning goal. The division of the learning algorithm or task is typically done by the system designer, and is not a part of the learning process itself. Interaction is required for putting together the results achieved by the different agents, but as in the case of multiplied learning this interaction does only concern the input and output of the agents' learning activities. Moreover, the interaction does not emerge in the course of learning but is determined a priori and in detail by the designer. This is to say that there are interactions among the participants, but without a remarkable degree of freedom (i.e. it is known a priori what information has to be exchanged, when this exchange has to take place, and how the exchanged information has to be used). An individual agent involved in 'divided learning' acts as *'specialist'* who is just responsible for a specific subset of the activities that as a whole form the overall learning process.

Different benefits can be expected from a division of labour. The fact that each agent only computes a subset of the learning algorithm makes it simpler to design and reduces the computational load of each agent. The 'agent' metaphor insists on the autonomy of functions with respect to each other. A benefit expected from this increased modularity is that the different functions may be combined in many different ways. A further potential benefit is that a speed-up in learning may be achieved (provided that the time gained by parallelism is not weighted out by the time required for coordination). The main difference between division and multiplication is *redundancy*: in division, each agent performs a different subtask while in multiplication, each agent performs the same task. In other words, the rate of redundancy in agent processing would be 0% in 'pure' division mechanisms and 100% in 'pure' multiplication, the actual design of existing systems being of course somewhere in between these two extremes.

Examples of 'divided learning'. Sen, Sekaran and Hale (1994) concentrated on the problem of achieving coordination without sharing information. As an illustration application, the block pushing problem (here two agents cooperate in moving a block to a specific predefined position) was chosen. The individual agents learn to jointly push the block towards its goal position. The learning task, hence, was divided among two agents. Sen and his colleagues showed that the joint learning goal can be achieved even if the agents do not model each other or exchange information about the problem domain. Instead, each agent implicitly takes into consideration the other agent's activities by sensing the actual block position. In as far as the individual agents do not exchange information and learn according to the same algorithm (known as Q-learning), this work is also related to multiplied learning.

Plaza, Arcos and Martin (1997) applied multiagent case-based learning in the domain of protein purification. The task to be solved here is to recommend appropriate chromatography techniques to purify proteins from tissues and cultures. Each agent is assumed to have its own case base, where a case consists of a protein-technique pair, as well as its own methods for case indexing, retrieval and adaptation. Two different modes of 'case-based cooperation' are proposed: first, an agent's case base can be made accessible to other agents; and second, an agent can make its methods for handling cases available to other agents. These modes are applied when a single agent is not able to solve a given protein purification task, and both result in the division of a single learning task among several agents.

Parker (1993) concentrated on the question of how cooperative behavior can be learnt in multi-robot environments. Parker was particularly interested in the question of how coherent cooperation can be achieved without excessive communication. The main idea was to realize cooperative behavior by learning about ones own and others' abilities. The scenario consisted of heterogeneous robots that have to complete some predefined tasks as fast as possible, where the individual robots differ in the abilities to carry out this tasks. Learning is divided in as far as the robots collectively learn which robot should carry out which task in order to improve cooperation. In as far as this approach requires that the robots learn about each other (and, hence, to develop a model of the others) in order to achieve a shared goal, it is closely related to interactive learning. Other interesting research on multi-robot learning which falls into the category of divided learning is described in (Mataric, 1996).

Weiß (1995) dealt with the question of how several agents can form appropriate groups, where a group is considered as a set of agents performing compatible actions that are useful in achieving a common goal. As an application domain the multiagent blocks world environment was used. The task to be solved in this domain was to transform initial configurations of blocks into goal configurations, where the individual agents are specialized in different activities. The proposed algorithm provides mechanisms for the formation of new and the dissolution of old groups. Formation and dissolution is guided by the estimates of the usefulness of the individual agents and the groups, where the estimates are collectively learnt by the agents during the problem solving process. Learning is divided in as far as different agents are responsible for different blocks and their movements. This work is related to interactive learning in as far the agents and groups dynamically interact (by building new groups or dissolving old ones) during the learning process whenever they detect a decrease in their performance.

3.3. **Interaction mechanisms**

General. The multiplication and division mechanisms do not explain all the benefits of multi-agent learning. Other learning mechanisms are based on the fact that agents interact *during* learning. Some interaction also occurs in the two previous mechanisms, but it mainly concerns the input or output of the individual agents' learning processes. Moreover, in the case of multi-plied and divided learning, interaction typically occurs at the relatively simple and pre-specified level of 'pure' data exchange. In the case of interactive learning, the interaction is a more dynamic activity that concerns the intermediate steps of the learning process. Here interaction does not just serve the purpose of data exchange, but typically is in the spirit of a 'cooperative, negotiated search for a solution of the learning task'. Interaction, hence, is an essential part and ingredient of the learning process. An agent involved in interactive learning does not so much act as a generalist or a specialist, but as a *'regulator'* who influences the learning path and as an *'integrator'* who synthesises the conflicting perspectives of the different agents involved in the learning process. For an illustration of the differences between the three classes of learning, consider the concept learning scenario. In an only-multiplied system, two inductive agents could interact to compare the concepts they have built. In an only-divided system, a 'negative instance manager' could refute the concept proposed by the 'positive instances manager'. The resulting concept may be better but the reasoning of one agent is not directly affected by the others. In an interactive approach, two inducers could negotiate the generalization hierarchy that they use. The term 'interaction' covers a wide category of mechanisms with different potential cognitive effects such as explanation, negotiation, mutual regulation, and so forth. The complexity of these interactions makes up another difference between interactive learning on the one hand and multi-plied/divided learning on the other.

The difference between the interactions in the multiplication/division approaches and the interaction approaches is mainly a matter of granularity: agents using the 'multiplication/ division' mechanisms interact about input and/or output of the agents' learning processes, while the mechanisms classified under the 'interaction' label concern intermediate learning steps. Obviously, the difference between an input and an intermediate step is a matter of granularity, since intermediate data can be viewed as the input/output of sub-processes. The criterion is hence not fully operational. However, the interaction mechanisms (e.g. explanation) is also concerned with how the intermediate data are produced. We briefly illustrate this point in the context of case-based reasoning:

- *Multiplication:* case-based reasoning agents interact about the solution they have derived by analogy.
- *Division:* the 'case retrieval' agent asks the 'case adapter' agent for specifying the target case more precisely.
- *Interaction:* a 'case retrieval' agent asks a 'case adapter' agent for justification for the selection of a particular case or for naming a criterion that should be used for retrieval.

Examples of 'interactive learning'. Sugawara and Lesser (1993) investigated how coordination plans can be learnt in distributed environments. The diagnosis of network traffic was chosen as an application domain. By observing the network, each agent responsible for diagnosing some part of the network learns by observation a model of the network segment in which it is located and of the potential failures that may occur. Based on this learning, the agents collectively develop rules and procedures that allow improved coordinated behavior (e.g. by avoiding redundant activities) on the basis of diagnosis plans. The learning task is not divided a priori; instead, the agents dynamically contribute to the joint learning task depending on their actual knowledge.
Bui, Kieronska and Venkatesh (1996) studied mutual modeling of preferences as a method for reducing the need for time- and cost-consuming communication in

negotiation-intensive application domains. The idea was to learn from observation to predict others' preferences and, with that, to be able to predict others' answers. The work is centered around the concept of joint intentions and around the question of how such intentions could be incrementally refined. This incremental refinement requires conflict resolution during a negotiation process. As an application domain they chose a distributed meeting scheduling scenario in which the agents have different preferences with respect to the meeting date. Moreover, it is assumed that none of the agents posses complete information about the others, and it is taken into consideration that the privacy of personal schedules should be preserved.

In the context of automated system design, Nagendra Prasad, Lesser and Lander (1995) investigated the problem of how agents embedded in a multiagent environment can learn organizational roles. The problem to be solved by the agents is to assemble a (simplified) steam condenser. A role is considered as a set of operators an agent can apply to a composite solution. Each agent is assumed to be able to work on several composite solutions concurrently, and to play different roles concurrently. Three types of organizational roles were distinguished, namely, for initiating a solution, for extending a solution, and for criticizing a solution. It was shown that learning of organizational roles enables a more efficient, asynchronous and distributed search through the space of possible solution paths. Although role learning itself is done by an agent independent of the other agents, all agents are involved in an overall conflict detection and resolution process. (This work is also of interest from the point of view of ‘divided learning’ in so far as the organizational roles are predefined and the agents are responsible for different parts of the steam condenser (e.g. the ‘pump agent’ and the ‘shaft agent’).

In closely related work, Nagendra Prasad, Lesser and Lander (1996) investigated multiagent case-based learning in the context of automated system design. As in their work mentioned above, the problem of building a steam condenser was chosen as an illustrative application. Here the focus was not on organizational roles, but on negotiated case retrieval. Each individual agent is assumed to be responsible for integrating a particular component of the overall system, and to have its particular view of the design problem. Moreover, each agent is associated with its own local case base and with a set of constraints describing the component's relationships to other components, where a case consists of the partial description of a component configuration and the constraints associated with it. The agents negotiate and iteratively put together their local ‘subcases’ such that constraint violations are reduced and a consistent overall case that solves the design problem is achieved. With that, both the cases as well as the processes of case indexing, retrieval and adaptation are distributed over several agents, and learning requires intensive interaction among the individuals.

Tan (1993) dealt with the question of how cooperation and learning in multiagent systems are related to each other, where exchange of basic information about the environment and exchange of action selection policies are distinguished as different modes of cooperation. Experiments were conducted in the predator/prey domain. The results indicated that cooperation can considerably improve the learning result, even if it may slow down learning especially at the beginning. This work is closely related to divided learning in as far as Tan also considered situations in which learning and sensing is done by different agents, and it is related to multi-plied learning as far as Tan considered situations in which different predators learn completely independent of each other.

4. The psychological perspective: Multi-agent learning vs. human-human collaborative learning

This section characterizes available approaches to multi-agent learning from the perspective of human-human collaborative learning. As noted in (Weiß, 1996), in DAI and ML the term 'multi-agent learning' is used in two different meanings. First, in its stronger meaning, this term refers only to situations in which the interaction among several agents aims at and is required for achieving a common learning goal. Second, in its weaker meaning, this term additionally refers to situations in which interaction is required for achieving different and agent-specific learning goals. For a brief illustration of these meanings, consider the multi-agent scenarios sketched in the section 2. The transportation domain is a good example for multi-agent learning in its weaker meaning; here the interacting agents (trucks or companies) pursue different learning goals, namely, to maximize their own profit at the cost of the other agents' profit. The manufacturing domain is a good example for multi-agent learning in its stronger meaning, because here all agents or machines try to learn a schedule that allows the manufacturing of products in minimal time and with minimal costs. Obviously, it is multi-agent learning in its stronger meaning that is more related to human-human collaborative learning, and which is therefore chosen as a basis for all considerations in this section.

The three types of learning mechanisms introduced in the previous section allows us to contrast multi-agent and single-agent learning, but are not appropriate for comparing multi-agent to human-human collaborative learning. In human groups, multi-plied, divided and interactive learning mechanisms occur simultaneously. 'Multiplication' mechanisms occur. For instance, we observed several instances (Dillenbourg et al., 1997) in a synchronous written communication environment of peers producing the same sentence simultaneously, i.e. they where conducting in parallel the same reasoning on the same data. The 'division' mechanisms do occur in collaborative learning. We discriminate (see chapter 1, this volume) cooperative settings, in which the division of labour is regulated by more or less explicit rules, and collaborative learning, with no explicit division of labour. However, even in collaborative tasks, a spontaneous division of labour can be observed (Miyake, 1986), although the distribution of work is more flexible (it changes over time).

Some of the mechanisms which make collaborative learning effective (Dillenbourg & Schneider, 1995) relate to the first and second types of the learning mechanisms. However, the 'deep secret' of collaborative learning seem to lie in the cognitive processes through which humans progressively build a *shared understanding*. This shared understanding relates to the third type of mechanisms. It does not occur at once, but progressively emerges in the course of dialogue through processes like conflict resolution, mutual regulation, explanation, justification, grounding, and so forth. Collaborative dialogues imply a sequence of episodes, some of them being referred to as ‘explanations’, ‘justification’, or whatever, and they are all finally instrumental in the construction of a shared solution. The cognitive processes implied by these interactions are the focus of current studies on human-human collaborative learning - and they have been largely ignored so far in the available studies on multi-agent learning. In the following, we will therefore concentrate on three of these processes - conflict resolution, mutual regulation, and explanation - in more detail in order to show what current approaches to multi-agent learning do *not* realize. Additionally, a closer look is taken on the importance of misunderstanding, which usually is considered by (D)AI researchers as something that should be avoided in (multi-)agent contexts, but which plays an important role for human-human collaborative learning. With that, and in contrast to the previous section, this one provides a *‘negative’ characterization* of multi-agent learning which allows us to make suggestions for future research on this kind of learning.

We are speaking of ‘(different types of) mechanisms’ because this chapter concentrates on computational agents. It is clear, however, that human-human collaborative learning constitutes a whole process in which ‘multiplication’, ‘division’ and ‘interaction’ correspond to different aspects of the same phenomena. Therefore, in focusing on the processes of conflict resolution, regulation, and explanation, we illustrate these three regards.

4.1. **Conflict Resolution**

The notion of conflict, which is popular in DAI , is the basis of the socio-cognitive theory of human-human collaboration (Doise and Mugny, 1984). According to this theory, the benefits of collaborative learning are explained by the fact that two individuals will disagree at some point, that they will feel a social pressure to solve that conflict, and that the resolution of this conflict may lead one or both of them to change their viewpoint. This can be understood from the *‘multiplication’* perspective, because a conflict between two or several agents is originated by the multiplicity of knowledge. It appears, however, that learning is not initiated and generated by the conflict itself but by its resolution, that is, by the justifications, explanations, rephrasements, and so forth, that lead to a jointly accepted proposition. Sometimes, a slight misunderstanding may be enough to generate these explanations, verbalizations, and so forth. Hence, the real effectiveness of conflict has to be modelled through interactive mechanisms.

4.2. **Mutual regulation**

Collaboration sometimes leads to improved self-regulatory skills (Blaye, 1988). One interpretation of these results is that strategic decisions are precisely those for which there is a high probability of disagreement between partners. Metaknowledge is verbalized during conflict resolution and, hence, progressively internalized by the partners. Another interpretation attributes this effect to a reduced cognitive load: when one agent looks after the detailed task-level aspects, the other can devote his cognitive resources to the meta-cognitive level. This interpretation could be modelled through a *‘division’* mechanism. However, mutual regulation additionally requires that each partner maintains ‘some’ representation of what the other knows and understands about the task. This is not necessarily a detailed representation of all the knowledge and viewpoints of a partner, but a task-specific differential representation (‘With respect to that point my partner does have a different opinion to me’). The verbal interactions through which one maintains a representation of one’s partners knowledge have been partly studied in linguistics, but the cognitive effects of this mutual modelling process, constitutes an interesting item on the agenda of psychology research. Interestingly, this ability was proposed by Shoham (1990) as a criterion to define what an agent is and whether an agent is more than a simple function in a system

4.3. **Explanation**

The third phenomena considered here is learning by building an explanation to somebody else. Webb (1989) observed that subjects who provide elaborated explanations learn more during collaborative learning than those who provide straightforward explanations. Learning by explaining does even occur when the learner is alone: students who explain aloud worked-out examples or textbooks, spontaneously or on a experimenter's request, acquire new knowledge (Chi et al., 1989). The effect of explanation in collaborative learning could be reduced to a multiple self-explanation effect (see Ploetzner et al. - this volume). For instance, in a multi-agent system each agent could be given a ‘learning by explaining’ algorithm inspired by VanLehn's computational model of the self-explanation effect (VanLehn et al., 1992) or by some explanation-based learning algorithm (EBL) (Mitchell et al., 1986). The benefit of *‘multiplicity’* could result from the heterogeneity of outcomes: The generalization stage - the most delicate step in EBL - could be ‘naturally’ controlled as the least general abstraction of the proofs built by different agents.

An explainer has to simultaneously build the explanation and check for its consistency. In self-explanation, this self-regulation process is costly, while in collaboration, the cost of regulation can be shared. The cognitive load is lower when one checks the consistency of an explanation given by somebody else, than when one has to check one's own explanation. Hence, this interpretation mechanism can be viewed from the perspective of the second type of mechanisms, the division of labour. This idea was implemented in People Power (Dillenbourg & Self, 1992). An agent explains a decision by showing the trace of the rules fired to take that decision. Some of these rules were too general and learning consisted in specializing them progressively. When the second agent received an explanation, he tried to refute the argument made by the first agent (in order to show that it is too general). Actually both agents used exactly the same rules, hence learning was not related to multiplicity. It simply was too expensive for one agent to build a proof, to try to refute any step of its proof, to refute any step of this self-refutation, etc. In other words, the functions ‘produce an explanation’ and ‘verify an explanation’ were distributed over two agents, as in the *‘division’* perspective.

However, explanation is most relevant from the perspective of the *interaction* mechanisms. An explanation is not simply something which is generated and delivered by an agent to another (as it could be viewed in the case of multiple self-explanation). Explanation is, of course, interactive. Current studies on explanation view explanation as a mutual process: explanation results from a common effort to understand mutually

(Baker, 1992). The precise role of interactivity is not yet very clear. It has not been shown that interactive explanations generate higher cognitive effects. The situations in which the explainee may interact with the explainer do not lead to higher learning outcomes than those where the explainee is neutral (see Ploetzner et al. - this volume).

4.4. **The importance of misunderstanding**

The three mentioned mechanisms are modeled in some way in DAI. The main difference is that the computational models use formal languages while humans use natural language (and non-verbal communication). It is of course common sense to use a formal language between artificial agents, but the use of natural language creates a major difference which might be investigated in DAI without having to ‘buy’ completely the natural language idea: it creates *room for misunderstanding*.

Misunderstanding leads to rephrasement, explanation, and so forth, and hence to processes which push collaborative learners to deepen their own understanding of the words they use. As Schwartz (1995) mentioned, it is not the shared understanding per se which is important, but the efforts towards shared understanding. We see two basic conditions for the emergence of misunderstanding:

- Misunderstanding relies on semantic ambiguity. It is a natural concern of multi-agent system designers to specify communication protocols in which the same symbol conveys the same ‘meaning’ for each agent, simply because ‘misunderstandings’ are detrimental to system efficiency. With that, it turns out that there is a conflict between efficiency and learning in multi-agent systems: misunderstanding and semantic ambiguity increase the computational costs, but at the same time the potential of collaborative learning.
- Having room for misunderstanding also means having room for discrepancies in the knowledge and experience of interacting individuals. Such discrepancies may arise, for instance, because the individuals act in different (perhaps geographically distributed) environments, because they are equipped with different sensing abilities (although they might occupy the same environment), and because they use different reasoning and inference mechanisms. From this point of view it turns out that heterogeneity should be not only considered as a problem, but also as an opportunity for designing sophisticated multi-agent systems.

This room for misunderstanding implies that agents are (partly) *able to detect and repair misunderstanding*. Obviously, it is not sufficient to create space for misunderstanding, but the agents must be also capable of handling it. Any misunderstanding is a failure in itself, and it constitutes a learning opportunity only if it is noticed and corrected. The mechanisms for monitoring and repairing misunderstandings in human contexts are described extensively in this volume (Chapter 3, Baker et al.). What matters here from the more technical multi-agent perspective is that the monitoring and repairing of misunderstandings require the agent to maintain a representation of what it believes and what it’s partners believes. (see section 4.2)

These considerations seem to indicate that multi-agent learning necessarily has to be limited with regard to collaboration, as long as the phenomenon of misunderstanding is suppressed and ignored in multi-agent systems. A solution to this limitation would be to consider 'mis-communication' not always as an error that has to be avoided during design, but sometimes as an opportunity that agents may exploit for learning. The same distinction can be made with linguistic studies, where the concept of 'least collaborative effort' is used to describe how interlocutors achieve mutual understanding, while we prefer to refer to the 'optimal collaborative effort' to emphasize that the additional cognitive processes involved in repairing mutual understanding may constitute a source of learning (Traum & Dillenbourg, 1996).

Two additional remarks. Firstly, the above considerations may sound inappropriate since the notion of (mis)understanding is purely metaphorical when talking about artificial agents. However, one can understand the notion of 'meaning' as the relationship between a symbol and a reference set. Typically, for an inductive learning agent A1, a symbol X will be related to a set of examples (S1). Another agent A2 may connect X with another set of examples (S2), overlapping only partly with S1. Then, negotiation can actually occur between A1 and A2 as a structure of proofs and refutations regarding the differences between S1 and S2.

Secondly, in addressing the above question, we restricted ourselves to verbal interaction and dialogues, because this is most interesting from the point of view of artificial agents. In particular, we did not deal with non-verbal interaction (based on e.g. facial expressions and body language), even though it is an important factor in human-human conversation. However, as a first step in that direction, we are currently working on implicit negotiation in virtual spaces where artificial agents negotiate division of labour by movements: Agent-A watches which rooms are being investigated by Agent-B and acknowledges Agent-B's decision by selecting rooms which satisfy the inferred strategy (Dillenbourg et al, 1997). Finally, it should be noted that we did not deal with certain aspects of dialogues that are relevant to human-human collaborative learning, but are difficult to model in artificial contexts. Examples of such aspects are the relationship between verbalization and consciousness (humans sometimes only become aware of something when they articulate it), dialogic strategies (one purposefully says something one does not believe to test ones partner) and the internalization of concepts conveyed in dialogues (these points are addressed in chapter 7, Mephu Nguifo et al., this volume).

5. **Conclusions**

Summary. This chapter aimed at taking the first steps toward answering the question *What is 'multi' in multi-agent learning?* We attacked this question from both a ML and a psychological perspective. The ML perspective led to a 'positive' characterization of multi-agent learning. Three types of learning mechanisms - multiplication, division, and interaction - were distinguished that can occur in multi-agent but not in single-agent systems. This shows that multi-agent learning can take place at a qualitatively different level compared to single-agent learning as it has been traditionally studied in ML. Hence, multi-agent learning is more than just a simple magnification of single-agent learning. The psychological perspective led to a ‘negative’ characterization of multi-agent learning. Several processes like conflict resolution, mutual regulation and explanation were identified that play a more significant role in human-human collaborative learning than in multi-agent learning since they contribute to the elaboration of a shared understanding. The cognitive effort necessary to build this shared understanding , i.e. to continuously detect and repair misunderstanding, has not received enough attention in multi-agent learning where noise is a priori not treated as a desirable phenomenon. Hence, despite the fact that multi-agent learning and human-human collaborative learning constitute corresponding forms of learning in technical and in human systems, it is obvious that the available multi-agent learning approaches are of lower complexity.

Some Implications. There are many approaches to multi-agent learning that are best characterized as multiplication or division mechanisms, but less that are best characterized as interaction mechanisms. This indicates important potential directions for future research on multi-agent learning. What is needed are algorithms according to which several agents can interactively learn and, at the same time, influence - trigger, re-direct, accelerate, etc. - each other in their learning. From what is known about human-human collaborative learning, the development of these kinds of multi-agent algorithms may be facilitated by putting more emphasis on the implementation of negotiation processes among agents, triggered by the possibility of misunderstanding and oriented towards the emergence of a shared understanding.

Acknowledgments. We are grateful to the ESF programme ‘Learning in humans and machines’ for the opportunity to work together in the past years. This chapter summarizes results of our co-work.

References

Baker, M. (1992) The collaborative construction of explanations. Paper presented to "Deuxièmes Journées Explication du PRC-GDR-IA du CNRS", Sophia-Antipolis, June 17-19 1992.

Bazzan, A. (1997). Evolution of coordination as a methaphor for learning in multi-agent systems. In (Weiß, 1997, pp. 117-136).

Benda, M., Jaganathan, V., & Dodhiawala, R. (1986*). On optimal cooperation of knowledge sources.* Technical Report. Boing Advanced Technical Center, Boing Computer Services, Seattle, WA.

Blaye, A. (1988) *Confrontation socio-cognitive et resolution de problemes.* Doctoral dissertation, Centre de Recherche en Psychologie Cognitive, Université de Provence, 13261 Aix-en-Provence, France.

Bond, A.H., & Gasser, L. (Eds.) (1988*). Readings in distributed artificial intelligence*. Morgan Kaufmann.

Bui, H.H., Kieronska, D., & Venkatesh, S. (1996). Negotiating agents that learn about others' preferences. In (Sen, 1996, pp. 16-21).

Bull, L., & Fogarty, T. (1996). Evolutionary computing in cooperative multi-agent environments. In (Sen, 1996).

Carmel, D., & Markovitch, S. (1996). Opponent modeling in multi-agent systems. In (Weiß & Sen, 1996, pp. 40-52).

Chan, P.K., & Stolfo, S.J. (1993). Toward parallel and distributed learning by meta-learning. *Working Notes of the AAAI Workshop on Knowledge Discovery and Databases* (pp. 227-240).

Chi M.T.H., Bassok, M., Lewis, M.W., Reimann, P. & Glaser, R. (1989) Self-Explanations: How Students Study and Use Examples in Learning to Solve Problems. *Cognitive Science*, 13,145-182.

Dillenbourg P. & Schneider D. (1995) Mediating the mechanisms which make collaborative learning sometimes effective. *International Journal of Educational Telecommunications* , 1 (2-3), 131-146.

Dillenbourg, P., & Self, J.A. (1992) A computational approach to socially distributed cognition. *European Journal of Psychology of Education*, 3 (4), 353-372.

Dillenbourg, P., Jermann, P. , Buiu C., Traum , D. & Schneider D. (1997) The design of MOO agents: Implications from an empirical CSCW study. *Proceedings 8th World Conference on Artificial Intelligence in Education*, Kobe, Japan.

Doise, W. & Mugny, G. (1984) The social development of the intellect. Oxford: Pergamon Press.

Fischer, K., Kuhn, N., Müller, H.J., Müller, J.P., & Pischel, M. (1993). Sophisticated and distributed: The transportation domain. In *Proceedings of the Fifth European Workshop on Modelling Autonomous Agents in a Multi-Agent World* (MAAMAW-93).

Gasser, L., & Huhns, M.N. (Eds.) (1989*). Distributed artificial intelligence, Vol. 2.* Pitman.

Grefenstette, J., & Daley, R. (1996). Methods for competitive and cooperative co-evolution. In (Sen, 1996).

Haynes, T., & Sen, S. (1996). Evolving behavioral strategies in predators and prey. In (Weiß & Sen, 1996, pp. 113-126).

Haynes, T., Lau, K., Sen, S. (1996). Learning cases to compliment rules for conflict resolution in multiagent systems. In (Sen, 1996, pp. 51-56).

Huhns, M.N. (Ed.) (1987). *Distributed artificial intelligence*. Pitman.

Imam, I.F. (Ed.) (1996). Intelligent adaptive agents. Papers from the 1996 AAAI Workshop. Technical Report WS-96-04. AAAI Press.

Mataric, M.J. (1996). Learning in multi-robot systems. In (Weiß & Sen, 1996, pp. 152-163).

Michalski, R., & Tecuci, G. (Eds.) (1995*). Machine learning. A multistrategy approach*. Morgan Kaufmann.

Mitchell, T.M., Keller, R.M. & Kedar-Cabelli S.T. (1986) Explanation-Based Generalization: A Unifying View. *Machine Learning*, 1 (1), 47-80.

Miyake, N. (1986) Constructive Interaction and the Iterative Process of Understanding. *Cognitive Science*, 10, 151-177.

Mor, Y., Goldman, C.V., & Rosenschein, J.S. (1996). Learn your opponent's strategy (in polynomial time)! In (Weiß & Sen, 1996, pp. 164-176).

Müller, J., & Pischel, M. (1994). An architecture for dynamically interacting agents. *Journal of Intelligent and Cooperative Information Systems*, 3(1), 25-45.

Müller, J., Wooldridge, M., & Jennings, N. (Eds.) (1997). *Intelligent agents III*. Lecture Notes in Artificial Intelligence, Vol. 1193. Springer-Verlag.

Nagendra Prasad, M.V., Lesser, V.R., & Lander, S.E. (1995). *Learning organizational roles in a heterogeneous multi-agent system*. Technical Report 95-35. Computer Science Department, University of Massachussetts.

Nagendra Prasad, M.V., Lesser, V.R., & Lander, S. (1996). On reasoning and retrieval in distributed case bases. *Journal of Visual Communication and Image Representation*, Special Issueon Digital Libraries, 7, 1, 74-87.

O'Hare, G.M.P., & Jennings, N.R. (Eds.) (1996*). Foundations of distributed artificial intelligence*. John Wiley & Sons, Inc.

Ohko, T., Hiraki, K., & Anzai, Y. (1996). Learning to reduce communication cost on task negotiation among multiple autonomous mobile robots. In (Weiß & Sen, 1996, pp. 177-191).

Parker, L. (1993). Adaptive action selection for cooperative agent teams. In *Proceedings of the Second International Conference on Simulation of Adaptive Behavior* (pp. 442-450).

Peceny, M., Weiß, G., & Brauer, W. (1996*). Verteiltes maschinelles Lernen in Fertigungsumgebungen*. Report FKI-218-96. Institut für Informatik, Technische Universität München.

Plaza, E., Arcos, J.L., & Martin, F. (1997). Cooperative case-based reasoning. In (Weiß, 1997, pp. 180-201).

Proceedings of the First International Conference on Multi-Agent Systems (ICMAS-95, 1995). AAAI Press/MIT Press.

Proceedings of the Second International Conference on Multi-Agent Systems (ICMAS-96, 1996). AAAI Press/MIT Press.

Provost F.J., & Aronis, J.M. (1995). Scaling up inductive learning with massive parallelism. *Machine Learning*, 23, 33f.

Sandholm, T., & Crites, R. (1995). Multiagent reinforcement learning in the iterated prisoner's dilemma. *Biosystems*, 37, 147-166.

Schwartz, D.L. (1995). The emergence of abstract dyad representations in dyad problem solving. *The Journal of the Learning Sciences*, 4 (3), pp. 321-354.

Sen, S. (Ed.) (1996). *Adaptation, coevolution and learning in multiagent systems*. Papers from the 1996 AAAI Symposium. Technical Report SS-96-01. AAAI Press.

Sen, S., Sekaran, M., & Hale, J. (1994). Learning to coordinate without sharing information. In *Proceedings of the 12th National Conference on Artificial Intelligence* (pp. 426-431).

Shoham, Y. (1990) *Agent-Oriented Programming*. Report STAN-CS-90-1335. Computer Science Department, Stanford University.

Sugawara, T., & Lesser, V. (1993). *Learning coordination plans in distributed problem-solving environments*. Technical Report 93-27. Computer Science Department, University of Massachussetts.

Tan, M. (1993). Multi-agent reinforcement learning: Independent vs. cooperative agents. In *Proceedings of the Tenth International Conference on Machine Learning* (pp. 330-337).

Terabe, M., Washio, T., Katai, O., & Sawaragi, T. (1997). A study of organizational learning in multiagent systems. In (Weiß, 1997, pp. 168-179).

Traum, D. & Dillenbourg, P. (1996) Miscommunication in multi-modal collaboration. Paper presented at the American Association for Artificial Intelligence (AAAI) Conference.

VanLehn, K., Jones, R. M., & Chi, M. T. H. (1992). A model of the self-explanation effect. *The Journal of the Learning Sciences*, 2 (1), 1-59.

Vidal, J.M., & Durfee, E.H. (1995). Recursive agent modeling using limited rationality. In (Proceedings ICMAS-95, pp. 376-383).

Webb, N.M (1989) Peer interaction and learning in small groups. *International journal of Educational research*, 13 (1), 21-39.

Weiß, G. (1995). Distributed reinforcement learning. *Robotics and Autonomous Systems*, 15, 135-142.

Weiß, G. (1996). Adaptation and learning in multi-agent systems: Some remarks and a bibliography. In (Weiß & Sen, 1996, pp. 1-21).

Weiß, G. (Ed.) (1997). *Distributed artificial intelligence meets machine learning*. Lecture Notes in Artificial Intelligence, Vol. 1221. Springer-Verlag.

Weiß, G., & Sen, S. (Eds.) (1996). *Adaption and learning in multi-agent systems*. Lecture Notes in Artificial Intelligence, Vol. 1042. Springer-Verlag.

Wooldridge, M., & Jennings, N.R. (Eds.) (1995a). *Intelligent agents*. Lecture Notes in Artificial Intelligence, Vol. 890. Springer-Verlag.

Wooldridge, M., & Jennings, N.R. (1995b). Agent theories, architectures, and languages: A survey. In (Wooldridge & Jennings, 1995a, pp. 1-21).

Wooldridge, M., Müller, J., & Tambe, M. (Eds.) (1996). *Intelligent agents II*. Lecture Notes in Artificial Intelligence, Vol. 1037. Springer-Verlag.