Investigating Reinforcement Learning in Multiagent Coalition Formation

Xin Li and Leen-Kiat Soh

Department of Computer Science and Engineering
University of Nebraska-Lincoln
115 Ferguson Hall, Lincoln, NE 66588-0115
{xinli, lksoh}@cse.unl.edu

Abstract

In this paper we investigate the use of reinforcement learning to address the multiagent coalition formation problem in dynamic, uncertain, real-time, and noisy environments. To adapt to the complex environmental factors, we equip each agent with the case-based reinforcement learning ability which is the integration of case-based reasoning and reinforcement learning. The agent can use case-based reasoning to derive a coalition formation plan in a real-time manner based on the past experience, and then instantiate the plan adapting to the dynamic and uncertain environment with the reinforcement learning on coalition formation experience. In this paper we focus on describing multiple aspects of the application of reinforcement learning in multiagent coalition formation. We classify two types of reinforcement learning: case-oriented reinforcement learning and peer-related reinforcement learning, corresponding to strategic, off-line learning scenario and tactical, online learning scenario respectively. An agent might learn about the others' joint or individual behavior during coalition formation, as a result, we identify them as joint-behavior reinforcement learning and individual-behavior reinforcement learning. We embed the learning approach in a multi-phase coalition formation model and have implemented the approach.

Introduction

Coalition formation in multiagent systems is a process where agents form coalitions and work together to solve a joint problem via coordinating their actions within each coalition (Sandholm 1999, Shehory and Kraus 1998). It is useful as it may increase the ability of agents to execute tasks and improves their payoffs. In general, each autonomous agent is incapable of performing specific global tasks all by itself. So some agents may form coalitions to allocate tasks among them to achieve the global goals.

However, in complex real-world environments, an agent only has incomplete even inaccurate information about the dynamically changing world and the occurrence of events may require the agents to react in a real-time manner. When designing multiagent systems in dynamic real-world environments, it is impossible to foresee all the potential situations an agent may encounter and specify an agent behavior optimally in advance. To increase problem-solving coherence and improve the total performance of the system as a whole, agents should be equipped with learning abilities so that they can learn from their own behaviors as well as their interaction patterns and adapt to the environment (Sycara 1998).

In recent years, reinforcement learning of agents' behaviors has attracted more attention of the community of multiagent systems because of its adaptability to dynamic environments. It has been applied to multiagent problems such as the robotic soccer (e.g., Salustowicz, Wiering, and Schmidhuber 1998), the predator-prey pursuit game (e.g., Kohri, Matsubayashi, and Tokoro 1997), and the prisoner's dilemma game (e.g., Sandholm and Crites 1995).

We have implemented a case-based reinforcement learning (CBRL) approach to multiagent coalition formation problem in dynamic, uncertain, real-time, and noisy environments. In such an environment, (1) each agent only has a partial view of the dynamically changing environment and the uncertain behavior of other agents, (2) the initial states that prompt the decision making process may change while the decision making process is still going on, (3) actions performed are not guaranteed to result in expected outcomes, (4) the coalitions need to be formed in real-time manner, and (5) the accurate event sensing or peer interaction is not guaranteed due to the noise during the agent perception. As a result, a coalition-initiating agent cannot exactly know which peer agents are *able*, or are *willing* to join the coalition to perform a global task. It also cannot exactly expect the computational and communication cost of coalition formation process or coalition's execution outcome.

Our CBRL approach integrates case-based reasoning (CBR) and reinforcement learning (RL) to utilize the agent's past coalition formation experience on the current problem and reinforce the utility of its experience with the current coalition formation outcome. Specifically, we apply case-based reasoning to store and reuse previous coalition formation strategies, and its role in coalition formation is to provide a basis for suitability study of a coalition formation plan given a particular task. We apply reinforcement learning to evaluate and score each plan based on the outcomes, and continuously learn about other agents' behaviors in joining coalitions. Its role in coalition formation is to continuously increase the likelihood of a good plan being selected in the next coalition formation tasks and identify high-utility peer agents as coalition candidates.

Here we investigate the multiple aspects of reinforcement learning application in our CBRL approach: (1) *case-oriented* reinforcement learning vs. *peer-related* reinforcement learning, (2) *strategic* reinforcement learning vs. *tactical* reinforcement learning, (3) *off-line* reinforcement learning vs. *online* reinforcement learning, and (4) *joint-behavior* reinforcement learning vs. *individual-behavior* reinforcement learning. The case-oriented reinforcement learning is to learn about the utilities of cases in casebase while the peer-related reinforcement learning is to learn about other agents' behaviors. They correspond to the strategic, off-line reinforcement learning scenario and the tactical, on-line scenario respectively since the case-oriented reinforcement learning provides a strategic learning approach to facilitate the planning of coalition formation strategies and it occurs outside of the coalition formation process while the peer-related reinforcement learning provides a tactical learning approach to learn how to instantiate the planned strategy and it occurs during the coalition formation process. In addition, we classify the peer-related reinforcement learning into joint-behavior and individual-behavior reinforcement learning. The former is to learn about a peer agent's social characteristics in joining coalitions, e.g., helpfulness of the peer to the agent, while the latter is to learn about a peer agent's personal characteristics, e.g., the availability degree of a specific capability which indicates whether the peer possesses the desired capability to perform a task.

We embed the case-based reinforcement learning in a multi-phase coalition formation model. The model consists of three phases: coalition planning, coalition instantiation, and coalition evaluation. Reinforcement learning is employed in coalition instantiation and coalition evaluation.

Background and Related Work

Reinforcement learning is the process of learning to behave optimally with respect to some scalar feedback value over a period of time (Sen and Weiss 1999). It can be regarded as a memoryless learning technique, in which an agent chooses an action only based on the last observation. In the standard reinforcement learning model, on each step of interaction the agent receives as input some indication of the current state of the environment; the agent then chooses an action to generate as output. The action changes the state of the environment, and the value of this state transition is communicated to the agent through a scalar reinforcement signal. The agent should choose actions that tend to increase the long-run sum of values of the reinforcement signal. It can learn to do this over time by systematic trial and error (Kaelbling, Littman, and Moore 1996).

With the basics of multiagent characteristics and properties defined in recent years, multiagent learning has become an important research issue (Excelente-Toledo and Jennings 2002, Sen and Weiss 1999, Stone and Veloso 2000). Compared with single-agent systems, the multi-agent systems are more complex partially because of the dynamic inter-agent interactions. Agents possibly need to learn from their previous behaviors and other agents' behaviors to decide on the next actions (Alonso et al. 2001). There are generally three aspects: (1) an agent learns about the other agents and their environments by observation in order to predict their behaviors or to produce a model of them (e.g., Hu and Wellman 1998, Nagayuki, Ishii, and Doya 2000); (2) agents learn how to coordinate or cooperate to achieve common goals (e.g., Haynes and Sen 1996, Tan 1993); and (3) an agent meta-learns what particular coordination mechanisms to use (e.g., Prasad and Lesser 1997, Soh and Li 2003, Soh and Tsatsoulis 2001, Sugawara and Lesser 1998). In these areas, few address reinforcement learning to form coalitions among agents. In (Soh and Li 2003, Soh and Tsatsoulis 2001), the learning is about how to negotiate between two agents, at a lower level com-pared to our proposed approach in this paper.

In traditional coalition formation, a rational agent can solve the combinatorial problem optimally without paying a penalty for deliberation. In (Sandholm and Lesser 1995), a bounded rationality is evident in which agents are guided by performance profiles and computational costs in their coalition formation processes. Similarly, our agents are aware of their communication and computational costs as well as time constraint in the coalition formation process. Agents learn about peer agents' characteristics to select low-cost but high-utility coalition candidates and thus reduce the communication and computational cost during coalition formation.

Multiagent Coalition Formation

The application context of our reinforcement learning is multiagent coalition formation. We have designed a Multi-Phase Coalition Formation (MPCF) model to address factors such as uncertainty, noise, real-time and dynamic issues. The model consists of three phases: coalition planning, coalition instantiation and coalition evaluation, as depicted in Figure 1. In *coalition planning*, the coalition-initiating agent derives a coalition formation plan. In *coalition instantiation*, the agent carries out the planned formation strategy, identifying and negotiating with other agents fitting the specifications of the plan. In *coalition evaluation*, the agent evaluates the coalition formation process, the formed coalition (if a coalition is successfully formed), and the coalition execution outcome (if the coalition is executed eventually) to determine the utility of the plan.

In coalition planning, the coalition-initiating agent derives a specific coalition formation plan for the current problem. A coalition formation plan specifies the number of coalition candidates, the number of expected coalition members, the time allocated for coalition instantiation, the allocation algorithm, and the number of messages recommended.

The coalition instantiation phase implements the coalition formation plan to form a coalition. At first, the coalition-initiating agent normalizes the task—dividing the task into separate execution units as different negotiation issues, computing the potential utilities of its peers, and ranking the peers based on their potential utilities. Then the agent concurrently negotiates with each selected peer agent on the set of subtasks in an attempt to form the intended coalition. Each negotiation is argumentative where the initiating agent attempts to persuade the responding agent to perform a task or provide a resource by providing support or evidence for its request (Soh and Tsatsoulis 2001).

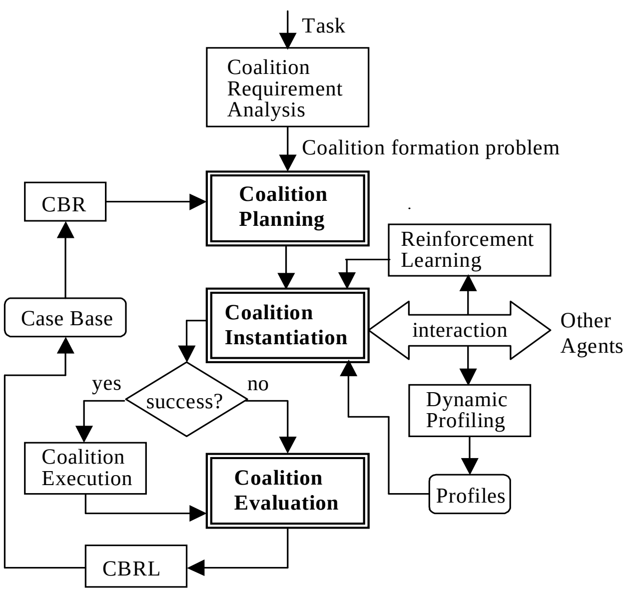

Figure 1. Learning-based MPCF model

The coalition evaluation phase provides the basis for an agent to improve its coalition formation plans. This phase evaluates both the coalition instantiation process (in terms of time spent, number of messages used, number of peers approached, etc.) and the execution outcomes of the subtasks agreed upon in the coalition (in terms of the number of subtasks performed by highly-capable peers, etc.). In general, a good plan is one that uses little computational and communication resources with successful instantiations and subsequent executions.

Reinforcement Learning

Our CBRL design, as shown in Figure 2, is aimed at (1) identifying the situation where a plan was successful and reinforcing that situation-plan pair in a case, and (2) learning about peer agents' behavior in joining coalitions and identifying peer agents of high potential utilities to the current coalition formation. We identify the application of reinforcement learning in our coalition formation model as *case-oriented* reinforcement learning (CRL) and *peer-related* reinforcement learning (PRL). The case-oriented reinforcement learning is applied specific to (1) and the peer-related reinforcement learning specific to (2).

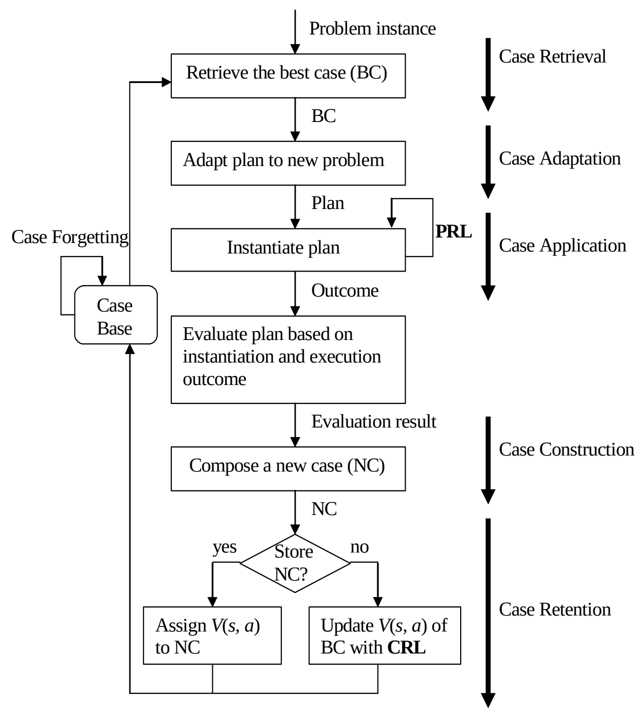

Figure 2. Case-Based Reinforcement Learning. CRL = Case-oriented Reinforcement Learning, and PRL = Peer-related Reinforcement Learning.

A coalition formation case in the casebase consists of a problem description, a solution, an outcome, and its utility. The *problem* description consists of an agent's external and internal environments and the task description. The *solution* part gives a coalition formation plan for the task to conduct the actual coalition formation process. The *outcome* part indicates the subtask allocation result among agents at the end of the coalition formation process, subtasks' execution results in the case that subtasks are eventually executed by coalition members, and the evaluation values to the actual coalition formation process. The *utility* part indicates the quality of the case, specifically, the quality of the solution to solving the current coalition formation problem.

The case-oriented reinforcement learning is to (a) learn the utility of a case based on its coalition formation plan and how well it was applied, and (b) retrieve high-utility cases more often. The peer-related reinforcement learning is to (a) learn the potential utility of each peer agent based on its coalition formation behavior, and (b) approach high-utility peers more often.

In our definition of the reinforcement learning, we follow the multiagent reinforcement learning model (Sen and Weiss 1999) where agents are given a description of the current *state* and have to choose the next *action* from a set of possible actions so as to maximize a scalar reinforcement received after each action. The traditional reinforcement learning can be modeled by a finite-state Markov decision process (MDP) that can be represented by a 4-tuple $<S, A, P, r>$ where S is a set of states, A is a set of actions, P: $S \times S \times A \mapsto [0, 1]$ gives the probability of moving from state s_1 to s_2 on performing action a, and r: $S \times A \mapsto \Re$ is a scalar reward function. However, the reinforcement learning in a dynamic multiagent environment of multiple learning agents cannot be formulated exactly as an MDP as the state transition probability of the environment changes with time due to the uncertain behavior of the other agents. We modified the traditional learning model to adapt to the complex environment.

Case-Oriented Reinforcement Learning

Case-based reasoning is to reuse the past solution in current similar problems. Due to the dynamic and uncertain environment, the past good plan may not be a good solution for the current problem; as a result, the outcome of the plan instantiation may be not good. This can be reflected in the case's utility for the future case reference in planning. We employ the case-oriented reinforcement learning to learn and reinforce the cases' utilities.

After coalition instantiation and coalition execution, the agent evaluates the coalition formation process and the outcome. Coupling the evaluation and the problem description, the agent matches the new case to its casebase. If the new case differs significantly from the existing cases, the agent learns the case to increase the case space. Otherwise, the agent updates the original best case's utility using the evaluation result to reinforce the case. On one hand, this is because the casebase also should keep bad experience as lessons. On the other hand, the dynamic and uncertain environment may make a plan with bad outcome able to produce good outcomes for other problems in the future.

During case retention, the case-oriented reinforcement learning updates the utility of the best case retrieved, the solution of which was used in a coalition formation process. The reinforcement learning algorithm is:

$$V_{t+1}(s,a) \leftarrow (1-\alpha) * V_t(s,a) + \alpha * R_{t+1}$$

where the state s corresponds to the current coalition formation problem; the action a corresponds to taking a coalition formation plan which is adapted from the plan in the best case; $V_t(s,a)$ is the old utility value of the best case while $V_{t+1}(s,a)$ is the new utility value of best case; α is the learning rate ($0 \le \alpha \le 1$); R_{t+1} is the performance evaluation result on the actual instantiation process (i.e., the reward) at time t+1. This reward includes the quality of the coalition formation process and the quality of the coalition. The learning result will be used in the agent's coalition case selection action.

Peer-Related Reinforcement Learning

At the end of coalition planning, the coalition formation plan for the current problem has been decided. The coalition-initiating agent can know which type of peer agents should be approached for the coalition formation. For example, an emergent task needs high-promptness peer agents to reduce coalition instantiation time. In coalition instantiation, the agent will select peer agents as coalition candidates according to the planned candidate type, and concurrently negotiate with each of them on the set of subtasks in an attempt to form the intended coalition. However, an agent only has incomplete information about the environment and other agents, so the coalition-initiating agent needs to learn about the dynamic and uncertain behavior of peer agents in past coalition formation activities to perceive their characteristics and potential utilities as coalition candidates. We employ peer-related reinforcement learning to learn the characteristics of each peer agent and the corresponding potential utility through interactions between agents.

The potential utility of a peer agent as a coalition candidate perceived by an agent is based on the cooperation relationship between the peer and the agent (Soh and Tsatsoulis 2001), and the peer's coalition-derived behaviors, negotiation-derived behaviors, and estimated capabilities. They are recorded in the neighborhood profile. The parameters profiled include: (1) the *helpfulness* of the peer to the agent indicating the satisfaction degree of requests to the peer, (2) the *helpfulness* of the agent to the peer indicating the satisfaction degree of requests from the peer, (3) the *reliance* of the agent on the peer in terms of the ratio of sending requests to the peer among all peers, (4) the *reliance* of the peer on the agent in terms of the ratio of receiving requests from the peer among all peers, (5) the peer's *tardiness* degree indicating the communication delay between the agent and the peer, in terms of the message round-trip time (RTT) between agents, (6) the peer's *hesitation* degree indicating how readily the peer is to agree to a request, in terms of the number of evidence messages the agent needs to provide to persuade the peer, (7) the *availability* degree of capability indicating whether the peer possesses the desired capability to solve task, (8) the *reliability* degree in coalition formation activities based on the standard deviations of the peer's behaviors, and so forth. These parameters reflect varieties of characteristics of the peer agent.

After each interaction (negotiation), the agent A_i updates the potential utility (for future coalition formation activity) of its peer A_j 's k_{th} characteristic $C_{A_j}^k$ in the following manner:

$$PU_{A_i,C_{A_j}^k}(s,a,t+1) \leftarrow (1-\beta) * PU_{A_i,C_{A_j}^k}(s,a,t) + \beta * C_{A_j}^k(A_i,t+1)$$

where the state *s* corresponds to the current coalition formation problem; the action *a* corresponds to coalition candidate selection; $PU_{A_i,C_{A_j}^k}(s,a,t)$ is the old potential utility of $C_{A_j}^k$ and $PU_{A_i,C_{A_j}^k}(s,a,t+1)$ is the updated one; β is the learning rate ($0 \le \beta \le 1$); and $C_{A_j}^k(A_i,t+1)$ is the peer A_j 's k_{th} characteristic as measured by A_i. The learning result will be used in the agent's coalition candidate selection action. The potential utility of peer A_j is the weighted sum of the potential utilities of its characteristics. It is computed when agent A_i selects coalition candidates. The weight values adapt to the candidate type requirement. With this reinforcement, an agent prefers peer agents that have been helpful and *coalition-worthy*.

The peer-related reinforcement learning in our coalition formation model can be identified as *joint-behavior* reinforcement learning and *individual-behavior* reinforcement learning further according to the peer agent's different characteristics revealed through the interactions. The joint-behavior reinforcement learning is to learn about a peer agent's social characteristics in joining coalitions such as helpfulness degree, reliance degree, reliability degree, and so forth. The individual-behavior reinforcement learning is to learn about a peer agent's inherent characteristics such as tardiness degree, hesitation degree, availability degree of capability.

We distinguish the joint-behavior reinforcement learning and individual-behavior reinforcement learning to identify the different roles of different types of characteristics of peer agents in coalition formation. In principle, the joint-behavior reinforcement learning is enough for a coalition-initiating agent to identify the potential utility of a peer agent as coalition candidate based on the past cooperation experience. In the dynamic and uncertain environment, however, peer agents' inherent characteristics may change as the time progresses. It will influence the peer's participation to coalition formation activities. So we also employ individual-behavior reinforcement learning to address the environmental factors.

Case-Oriented RL vs. Peer-Related RL

The case-oriented reinforcement learning occurs at the coalition evaluation phase and the learning result is applied at the coalition planning phase. We apply it into the coalition formation model as a strategic learning approach to facilitate the planning of a coalition, adapting to the real-time and environmental requirements. The peer-related reinforcement learning occurs at the coalition instantiation phase and the learning result is also applied at the coalition instantiation phase. We apply it into the coalition formation model as a tactical learning approach to address how to instantiate a coalition formation plan, taking into account uncertain and dynamic behaviors of the peer agents. In the dynamic, uncertain, and noisy environment, a good coalition formation plan of an agent is not guaranteed to succeed as planned. Thus, a "conceptually good" plan may not be a "practically good" plan. The continual tactical learning during coalition formation is necessary to address the change of the environment.

Corresponding to the different occurrences, the case-oriented reinforcement learning and the peer-related reinforcement learning are in the off-line reinforcement learning scenario and the online scenario respectively. The case-oriented reinforcement learning can be off-line because it does not directly need agents' interactions. The peer-related reinforcement learning must be online because it directly needs the interactions between agents. We combine off-line learning and online learning into our reinforcement learning to address the real-time, dynamic, uncertain, and noisy environmental factors. Then agents can learn in both scenarios of without interactions and with interactions.

In short, the application of case-oriented reinforcement learning in coalition formation is to meta-learn coalition formation processes as a strategic learning approach to conduct the actual coalition formation process. Its off-line application can reduce the agents' computational costs spent on learning during coalition formation. The application of peer-related reinforcement learning in coalition formation is to dynamically learn other agents' behavior in the course of continual interactions as a tactical approach to learn how to instantiate the planned coalition formation strategy. The computational costs for this online learning scenario are necessary.

Experiments and Results

We have implemented a multiagent system where each agent is capable of performing multiple tasks and has multiple resources. Here we present our experiments and results to evaluate the performance of our reinforcement learning approach. Particularly, our experiments were to investigate the impact of learning on the success rate of coalition formation, and on the quality of the coalition formation process.

We report our experiments on four versions of our multiagent design. The first version was **CRLPRL** in which both case-oriented reinforcement learning and peer-related reinforcement learning are used. The second version, **OnlyPRL**, used only the peer-related reinforcement learning. The third version, **OnlyCRL**, used only case-oriented reinforcement learning. The fourth version, **NoLearning**, did not use any learning at all.

In our experiments, there were 9 agents in the system, $A_1 \sim A_9$, each of which could initiate coalition formation activities for task fulfillment. We randomly simulated a series of 40 tasks (timed to occur at different times) for each agent. Each task consisted of different subtasks and required different resources. All tasks required the agent to initiate coalitions. We ran the experiments about a dozen times to obtain the average values used in the following discussions.

Impact of Learning on Coalition Formation Success Rate

Figure 3 shows the impact of reinforcement learning on an agent's success rate in forming coalitions.

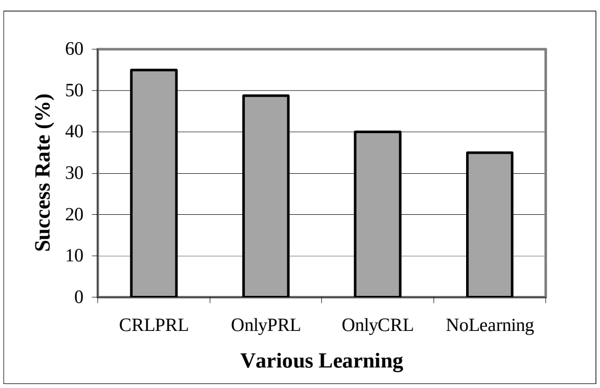

Figure 3. Success rates of coalition formation for different agent versions of learning mechanisms

Note that the success rate of each version of our agent learning design was low (<60%). This was due to the environment the agents operated in. Since the agents were handling their own tasks which overlapped temporally, it was possible for a peer agent not able to entertain or agree to a coalition request. Real-time constraints also played a role. Each task was timed and thus a coalition formation process that had run too long would be terminated. An agent that had been too slow in responding or communication would cause such delays. From Figure 3, we observe the following:

The agent design with both case-oriented reinforcement learning and peer-related reinforcement learning (CRLPRL) outperformed all the other versions in terms of the success rates of coalition formation. This shows that our agents were able to learn to improve their performance in coalition formation.

With learning, the agents were able to form coalitions more successfully. This is based on the observation that NoLearning yielded the lowest success rate at 35%.

Peer-related reinforcement learning (OnlyPRL) yielded better success rate than case-oriented reinforcement learning (OnlyCRL), 48.75% to 40%. This indicates that the peer-related reinforcement learning played a more significant role than the case-oriented reinforcement learning in our environment. In the dynamic, uncertain environment, the tactical, online reinforcement learning on peer agents' behavior played a more significant role on the effectiveness of coalition formation than the strategic, off-line reinforcement learning on the utilities of past coalition formation cases.

Impact of Learning on Coalition Formation Quality

Figure 4 shows the impact of learning on the coalition formation quality, which is an average of how well the coalition formation processes were (e.g., in terms of the number of messages, time spent vs. time expected, etc.) and the quality of the actual coalition formed (e.g., whether an expert peer contributed to the coalition, etc.).

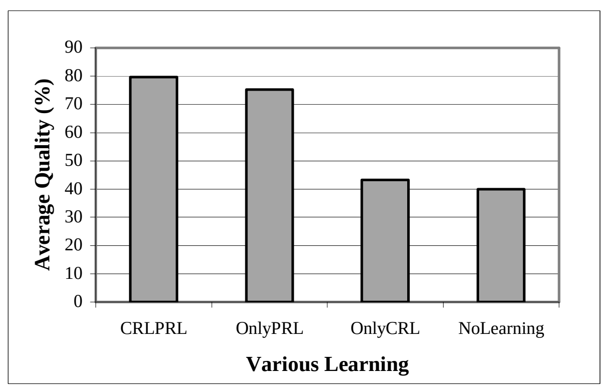

Figure 4. Average coalition formation qualities for different agent versions of learning mechanisms

In general, we observe similar patterns as in Figure 3. CRLPRL outperformed all others with an average quality of 80%. NoLearning and OnlyCRL were far behind, hovering around 40%. This clearly showed that even when there was a good plan (derived from case-based reasoning), the actual coalition formation process or the coalition quality might not be better than that without a good plan (OnlyPRL, for example) in our environment. In the dynamic, uncertain environment, the tactical, online reinforcement learning on peer agents' behavior played a more significant role on the efficiency of coalition formation than the strategic, off-line reinforcement learning on the utilities of past coalition formation cases.

Conclusions

We have described the use of reinforcement learning in a multi-phase coalition formation model. We investigated multiple aspects of the reinforcement learning application in multiagent coalition formation. We have conducted several preliminary experiments and the results have been promising in proving the feasibility of reinforcement learning. With reinforcement learning, our agents form coalitions more effectively and efficiently. Our future work will focus on designing further experiments to test the impact of reinforcement learning on the effectiveness and efficiency of coalition formation based on different degrees of heterogeneity in the agents' characteristics.

References

Alonso, E., D'Inverno, M., Kudenko, D., Luck, M., and Noble, J. 2001. Learning in Multi-Agent Systems. *The Knowledge Engineering Review* 16(3):277–284.

Excelente-Toledo, C. B., and Jennings, N. R. 2002. Learning to Select a Coordination Mechanism. In *Proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'2002)*, 1106-1113, Bologna, Italy.

Haynes, T., and Sen, S. 1996. Learning Cases to Resolve Conflicts and Improve Group Behavior. *International Journal of Human-Computer Studies* 48(1):31-49.

Hu, J., and Wellman, M. P. 1998. Online Learning about Other Agents in a Dynamic Multiagent System. In *Proceedings of the Second International Conference on Autonomous Agents*, 239-246, Minneapolis, MN.

Kaelbling, L.P., Littman, M.L., and Moore, A. W. 1996. Reinforcement Learning: A Survey. *Journal of Artificial Intelligence Research* 4:237-285.

Kohri, T., Matsubayashi, K., and Tokoro, M. 1997. An Adaptive Architecture for Modular Q-Learning. In *Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence (IJCAI'1997),* 820-825, Nagoya, Japan.

Nagayuki, Y., Ishii, S., and Doya, K. 2000. Multi-Agent Reinforcement Learning: An Approach Based on the Other Agent's Internal Model. In *Proceedings of the Fourth International Conference on Multi-Agent Systems*, 215-221, Kyoto, Japan.

Prasad, M. V. N., and Lesser, V. R. 1997. The Use of Meta-Level Information in Learning Situation-Specific Coordination. In *Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence (IJCAI'1997)*, 640-646, Nagoya, Japan.

Salustowicz, R. P., Wiering, M. A., and Schmidhuber, J. 1998. Learning Team Strategies: Soccer Case Studies. *Machine Learning* 33(2/3):263-282.

Sandholm, T. W. 1999. Distributed Rational Decision Making. In G. Weiss (Ed.), *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence* 241-250, the MIT Press.

Sandholm, T. W., and Crites, R. H. 1995. Multiagent Reinforcement Learning in the Iterated Prisoner's Dilemmas. *Biosystems* 37:147-166.

Sandholm, T. W., and Lesser, V. R. 1995. Coalition Formation amongst Bounded Rational Agents. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI'1995),* 662-669, Montreal, Canada.

Sen, S., and Weiss, G. 1999. Learning in Multiagent Systems. In Weiss, G. (ed.), *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence* 259-298, the MIT Press.

Shehory, O., and Kraus, S. 1998. Methods for Task Allocation via Agent Coalition Formation. Artificial Intelligence 101:165-200.

Soh, L.-K. and Li, X. 2003. An Integrated Multi-Level Learning Approach to Multiagent Coalition Formation. In *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence (IJCAI'2003),* 619-624, Acapulco, Mexico.

Soh, L.-K., and Tsatsoulis, C. 2001. Reflective Negotiating Agents for Real-Time Multisensor Target Tracking. In *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence (IJCAI'2001),* 1121-1127, Seattle, WA.

Stone, P., and Veloso, M. 2000. Multiagent Systems: A Survey from a Machine Learning Perspective. *Autonomous Robotics* 8(3):345-383.

Sugawara, T., and Lesser, V. 1998. Learning to Improve Coordinated Actions in Cooperative Distributed Problem-Solving Environments. *Machine Learning* 33(2/3):129-153.

Sycara, K. P. 1998. Multiagent Systems. *AI Magazine* 19(2):79-92.

Tan, M. 1993. Multi-Agent Reinforcement Learning: Independent vs. Cooperative Agents. In *Proceedings of the Tenth International Conference on Machine Learning (ICML'1993),* 330-337, Amherst, MA.