Concept Learning Games: The Game of Query and Response (short paper)

Nima Mirbakhsh, Arman Didandeh
Faculty of Computer Sciences
Institute for Advanced Studies in Basic Sciences (IASBS)
Zanjan, Iran
{*nmirbakhsh, a_didandeh*}*@iasbs.ac.ir*

Mohsen Afsharchi
Department of Electrical and Computer Engineering
University of Zanjan
Zanjan, Iran
afsharchim@znu.ac.ir

***Abstract*—This article deals with the issue of concept learning and tries to have a game theoretic view over the process of cooperative concept learning among agents in a multi-agent system, in which an extreme sense of competition has arisen. This gives birth to a new realm labeled as "Learning Games". We study the cooperative view and give a novel idea to use in competitive environments based on the solution concepts of game theory and an innovative idea of concept scores over an ontology, which is used as our knowledge representation tool for agents. A case study comes at the end and covers the FAQs of this method over a scientific world topic of *AIDS* remedy project.**

***Keywords*-Ontology, Game Theory, Multi-Agent System (MAS), Concept Learning**

I. Introduction

Concept learning mostly looks at the problem of multiple agents working together, trying to learn a concept from or teach it to other agents [11], [10], [4], [3]. Concept learning has always been studied in a cooperative environment of multiple agents. But what if an agent works on the same or similar matters of discussion with some other independent agents in a competitive environment?

An agent could use the points of views from the peer agents of the MAS but should also think about the heavy atmosphere of self-interestedness. It is clear that the learner agent would not like to disregard the chance of taking advantage of the knowledge of the teacher agents. It could not also expect the peer agents to share all their knowledge while any piece of information is worth a lot to any agent. But is it all competition in this environment? Is it a really strict competitive game?

Taking a more optimistic look at the problem, we repeat the above sentence: "any piece of information is worth a lot to any agent". As [4] has indicated, the teacher agent's responses for the learner agent's queries have a vital role in the learner agent's ontology creation and reformation. It is also mentioned in [3] that the queries that the learner agent may ask could be worthful to the teacher agent. So in a competitive MAS, agents should think out of a policy (or better to say, a strategy) in order to lose as less information and gain as more knowledge as they can.

As it is clear, these ideas are very similar to the terminology of game theory, in which cooperation is well studied and competition is known as defect [5]. So we tend to study the problem of MAS concept learning under a game theoretic view.

A method that we wish to take advantage of is the innovative idea of [2], in which they put scores on concepts in an ontology. Although [2] has used this idea mostly to capture the user preferences of a recommendation system, their innovation can model the prior and posterior belief of agents in a MAS regarding to a recently-arrived object. This object is the same as the new piece of information we have in concept learning. This can help agents to evaluate the piece of information they want to share.

Another idea that helps us here is the preference parameter, which models to what extent an agent is willing to share information with its peers. This parameter actually comes from [1] which has studied it under the iterated prisoner's dilemma problem[5].

The general idea of concept learning games is studied in the big context of concept learning, in which the "concept learning and reformation" is explored both from an object-based and also feature-based view in a MAS. This is important while we assume a feature-based ontology in this article, as was mentioned in our previous work[3]. It is also vital to study the usage of query and response for concept learning among agents in a MAS. This is the same infrastructure used in this article as well.

The remaining of this article goes as follows: Section II will name the basic definitions. Section III includes a brief literature review. Section IV reflects the problem statement. Section V studies the game theoretic view of MAS concept learning approach, which includes investigating the importance of lost information versus the gained knowledge for peer agents, and also making the optimized decision based on game theory strategies. Section VII concludes the article.

II. Basic Definitions

While this is a short paper version of this article, the basic definition section is not presented fully here. Instead, we just name the needed definitions with their references for the eager reader to go deep for them. The reader of this article may need the formal definitions of agents, ontologies

978-0-7695-4191-4/10 $26.00 © 2010 IEEE
DOI 10.1109/WI-IAT.2010.161

Table I
THE GENERAL PAYOFF MATRIX FOR A GAME OF PRISONER'S DILEMMA

Agent j/Agent i	defects	cooperates
defects	(a, a)	(b, c)
cooperates	(c, b)	(d, d)

and concepts, which are available both in [3] and in [4].
In this article, we are dealing with a concept learning problem. So from now on, Ag_L will stand for the learner agent and Ag_T is any peer agent of Ag_L which has the teacher agent role. Remember that while each agent is actually playing a game in this article, the words player and agent are assumed interchangeable.

III. LITERATURE REVIEW

For the same reason as in the previous section, the literature review has been pruned from this article (and will be available in the full paper version).

A. Game Theory

As we have discussed before, our approach will need a game theoretic view. For a comprehensive insight on game theory and the related attempts that have been made in this area (near to our concern), some names and references are necessary. [5], [6], [7] could give a bright insight on game theory, esp. for reasoning in MAS over utilities and preferences. [5] also discusses the cooperative and competitive views on interactive agents using a payoff matrix, and also a solution concept named Nash equilibrium. There are also some popular possible interaction scenarios, defined on the payoff matrix. The prisoner's dilemma is the one, attracting many game theorists and social scientists. It is "a non-cooperative and non-zero game, which is played between two agents" [1]. Prisoner's dilemma is a symmetric game, having a payoff matrix of the form in Table I. It has a Nash equilibrium when both players decide to defect. Because of some properties, prisoner's dilemma hinders the possible cooperation. Changing this atmosphere leads to the realm of coalition formation and cooperative game theory [8].

Considering all these issues, almost any game theorist and MAS researcher has come to the point to believe in a *mixed strategy* instead of a certain and rigid decision process. "A mixed strategy allows you to choose between possible choices by introducing randomness into selection" [5]. But answering the question of if a game has a Nash equilibrium with a mixed strategy is a computationally complex problem. Introducing uncertainty about the knowledge of other players (agents) was discussed as Bayesian coalition games in [9]. However the question of "*how could we model uncertainty about our own knowledge of the environment?*" still remains unanswered.

B. Ontological a-priori score

As it has been discussed in [2], the a-priori score of a concept c, $APS(c)$ models the expected preference of each concept for an average user. Following a probabilistic distribution study, [2] define the a-priori score of a concept c with n_c descendants (as the expected value of that distribution) as below:

$$APS(c) = \frac{1}{n_c + 2} \qquad (1)$$

C. Preference in iterated prisoners dilemma

As the authors of [1] have mentioned, they have introduced a preference parameter in the payoff matrix to model the degree to which any player prefers being self-interested or benevolent (κ here). If $0 \leq \kappa \leq 1/2$, the player's preference is to be egoist/self-interested so it will choose to defeat in a game. Also when $1/2 \leq \kappa \leq 1$, the player would desire more to cooperate while it is more altruist and benevolent. For a symmetric payoff matrix like in the prisoner's dilemma we regard to a matrix A (the single-player payoff matrix) like:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

So, the single-player payoff matrix modeling the preference parameter will look like (as in [1]):

$$A^\kappa = \begin{pmatrix} a & b + \kappa(c - b) \\ c + \kappa(b - c) & d \end{pmatrix} \qquad (2)$$

We will use the results of using this preference parameter in an iterated prisoner's dilemma in our approach to decide whether to cooperate or to defect.

D. Concept learning in a MAS

This subsection is completely removed in this short-paper version of our article, and could be found fully in [13], and also in [3] and [4], as it is in our concern. Also [10], [11], [12] has done a lot in this area.

IV. PROBLEM STATEMENT: MAS CONCEPT LEARNING AND GAME THEORY

Without loss of generality and for the simplicity of problem statement, we consider only two above-mentioned agents Ag_L and Ag_T and break our multi-agent interactions into agent-agent mutual interactions. These two independent agents act towards each other in a competitive (however not strictly competitive) environment in which any bit of information from the other peer agents is worth a fortune. For example, consider some scientific teams of individual universities working on a similar project of *AIDS* remedy. Each one of them is granted from a different powerful medicinal company. If any of the scientific teams succeeds, the respective company's stock could raise extremely high, resulting in a great amount of fund for the winner team. So a severe atmosphere of competition exists among scientific teams. However this sense of competition is way too different from the competition in zero-sum interactions[5], while here the problem is not a strict competition interaction.

Besides, it is normal for universities to share their data among each other in order to gain more optimality and reach sooner to the results. Indeed cooperation among universities is well defined all over the world. In this situation, how should the teams behave against each other? Asking this question game theoretically, should they *cooperate* or *defect*? As you see, this seems to be very similar to the prisoner's dilemma, but is there any dominant strategy for this problem?

While both of these strategies seem tempting, another important matter to be considered is that the process could be -and in most situations is indeed- iterated. So it is better to take into account the *shadow of the future* as [5] entitles about a variant of the prisoner's dilemma. If we regard this shadow of iteration, cooperation may come more attractive to teams. But to what extent should the teams accept this cooperation, when they know that real world interaction games are not going to be played forever? Better to say, how should the agents take advantage of cooperation to gain more payoffs but also preserve themselves versus *the sucker's payoff*? This and the iterative view gives rise to a *mixed strategy* which game theorists mostly insist on.

The authors wish to emphasize again that almost any piece of information is really valuable in this context for agents. This means that from Ag_T's point of view, it is really important that what Ag_L is asking. Vice versa, it is also momentous for Ag_L what Ag_T is sending back as its idealogy, because Ag_L can take advantage of Ag_T's progress as well.

More formal to say, the query that Ag_L is sending to Ag_T shows Ag_L's headway in the competition. The response that Ag_T forwards back to Ag_L is also modeling the advancement of Ag_T. As it is obvious, a cycle of mutual decision and action is born here. This is exactly the origin of the iteration and the mixed strategy mentioned above.

So the big question here is what policy should agents take in a concept learning problem, when dealing with peers in a competitive MAS?

V. Game theoretic view of MAS concept learning

Now that the dealing question is clear, it would be better to formally model the problem in order to step forward to the solution. We propose the following process model to study each part of our proposed approach separately:

1. Ag_L and Ag_T's decision to defect or cooperate, called the *competitive MAS interaction processes*;
2. Ag_L's query generation and Ag_T's response generation processes, labeled as the *Information Vector generation*;
3. Ag_L and Ag_T's learning phase from each other's information vector, known as the *learning process*;

It is important to remind that all the knowledge transactions of a MAS is represented as concepts locating in the nodes of an ontology.

A. Competitive MAS interaction processes

Ag_L's query generation process depends not only on its prior knowledge on the query object, as well as Ag_T's response generation process. It is also vital for both agents in the competition to care about the sucker's payoff. To be more precise, they should care about losing the less information in order to gain as more knowledge as possible. This would be much more caring in an uncertain situation like a MAS environment, in which agents have partial access to the environment knowledge and no agent is sure of what it knows for best.

Considering these points, we aim here to look closer to the optimal reasoning of agents in the competition to better produce queries and responses, aspiring to carry out the goals described before. In this path, we try to take advantage of our previous attempts on concept learning and principally the *feature-based representation of concepts* [3]. We also utilize the innovative ideas of [2] and [1].

Actually the preprocess of mind for agents, both in generating queries and responses are somehow similar. Both processes have two main steps. First, the agent, whether Ag_L or Ag_T, will have to detect and identify the query object of concern. Then they both should decide whether to cooperate or to defect. Cooperating and defecting will be clear in the information vector generation phase, but for now, just know that it means for agents to decide on the extent of the information they wish to give out in the competitive atmosphere of MAS. The agent will then exploit the actual environment information in hand to generate an information vector, i.e. a query vector, or a response vector correspondingly. The reason we see the query and the response as vectors is that our representation of environment information is based on features, which is well represented as a vector containing n-tuples of features and feature-values.

1) Concept identification using the ontology: Ag_L detects some object o in the environment and memorizes it using the feature-based representation as a feature vector. This object is the origin of Ag_L's query. So first of all, Ag_L should deliberate o and think out to which concept of any concept c_i in its ontology, o belongs, or even is closest. This is simply done using our proposed method in [3], as is done as below:

$$\begin{aligned} k &= \arg\max P(c_i|o) \\ &= \arg\max P(c_i) \prod P(f_i = v_{ij}|c_i) \quad (3) \end{aligned}$$

Two main considerations are needed here. First is that any concept c_i is identified with a set of features in the ontology of each agent. These are exactly the same feature set F_{c_i} we use in Formula (3) to recognize $P(f_i = v_{ij}|c_i)$ values.

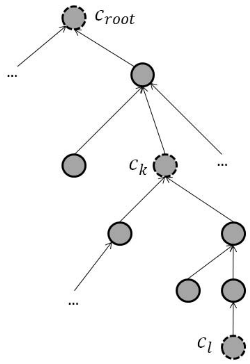

Figure 1. A sample ontology showing concepts c_{root}, c_k and c_l

Secondly, the same actions are done by Ag_T, but only Ag_T uses the query vector from Ag_L to decide on.

2) Deciding whether to cooperate or defect: After the agents (both Ag_L and Ag_T) have detected the feature vector of o and have identified $c_k(c_{k_L}$ and c_{k_T} respectively) as the concept best demonstrating it, they have to decide on their subsequent action which is known as the "D/C decision" standing for the *Defect/Cooperation decision problem*. Here we use the idea of the preference parameter [1], labeled as κ here. To bring this idea into our context, the innovation of [2] will help.

Actually in this article the value for the preference parameter κ, for each agent is not a behavioral property, but is calculated with respect to the implicit and subjective view of any agent. This is the view about a certain concept through the world around for that agent, i.e. to which extent the identified concept c_k is important for each agent.

The process of calculating κ_L and κ_T takes advantage of the ontological knowledge of the agents, known as a-priori score [2] of the concept c_k and its related nodes in the ontology. If we denote the root of an ontology tree with c_{root} and the lowest (most concrete) concept in a path underneath c_k with c_l (as in Figure 1), then the value for κ is calculated as:

$$\kappa = 1 - \frac{\ln(APS(c_k)) - \ln(APS(c_{root}))}{\ln(APS(c_l)) - \ln(APS(c_{root}))} \quad (4)$$

Practically, this value is calculated as κ_T and κ_L (with regard to c_{k_L} and c_{k_T}) and shows the value of abstraction of the detected and identified concept in the Ag_L and Ag_T's ontologies, each. The value κ is always between 0 and 1. The more κ is close to 1, the more abstract the detected concept c_k is. Vice versa, if κ moves towards 0, then c_k is getting more concrete. So in a path from the root of an ontology downwards, the value of κ decreases as the level of abstraction of the concepts in the path decreases too.

Now with respect to the value of κ any agent has calculated, and also according to the valuable work of [1], Ag_L and Ag_T will use their results to decide.

[1] deduces that in an iterated prisoner's dilemma with preference parameter existing. So conclusions could be obtained, based on the values for κ and also the single player payoff matrix (Table I). These conclusions result in some dominant strategies in some cases and also some equilibrium mixed strategies in other cases, helping each agent to decide on their D/C decision.

B. Information Vector (IV) generation

After an agent decides its manner against its peer agent, it should generate the query/response vector. Here we bring the scenarios of cooperation and defect for both learner and teacher agents.

For Ag_L ,it is clear that if the decision is to cooperate, it should send feature vector of o completely as it can percept. For example if o is defined by a vector of $([f_1 = v_1], ..., [f_7 = v_7])$ (features are sorted in order of *information gain* values), cooperation means the query vector below:

$$q = ([f_1 = v_1], ..., [f_9 = v_9])$$

However there are some clear scenarios of defect to name. Ag_L could eliminate some features from the tuple, in order to make the detection phase harder to achieve for Ag_T, i.e. by removing the feature with most information gain value. Thus the query vector is as follows:

$$q = ([f_2 = v_2], ..., [f_9 = v_9])$$

Ag_L could also remove accidentally some features from the feature vector (a generalization of the above scenario). This means that:

$$q = ([f_1 = v_1], ..., [f_4 = v_4], [f_6 = v_6], ..., [f_9 = v_9])$$

Another approach is to let some unrelated features into the query vector to mislead Ag_T's detection. These features can be caught simply from the siblings of c_{k_L} that are not in common with c_{k_L}'s feature set $F_{c_{k_L}}$.

$$q = ([f_1 = v_1], ..., [f_9 = v_9]) \cup ([f_{10} = v_{10}], [f_{11} = v_{11}])$$
$$f_{10}, f_{11} \in F^{sib}_{c_{k_L}} \setminus F_{c_{k_L}}$$

Accordingly for Ag_T, cooperation means letting the exact knowledge exploited from the query vector to be obtained by Ag_L. This exact knowledge is gained during the detection phase as stated before.

On the other hand, defect may be modeled a bit simpler for Ag_T than of Ag_L's. Ag_T could simply choose a superconcept or a subconcept of the c_k it has detected, or even one sibling of c_k.

C. Learning Processes

Consider the situation our agents are now in. Ag_L has decided on the D/C problem, has calculated κ_L, and then has generated an information vector, which we know as the query vector. Then it has sent its information vector to Ag_T. Ag_T has received Ag_L's query and has created its response as another information vector based on κ_T. Then it has sent the response vector back to Ag_L. Now that the process of multi-agent interaction has finished, we shall take a look at how agents can learn from the opponent's information vectors in order to improve their knowledge and reform their ontology. Take into account that we consider Ag_T to always have some knowledge about the query as a teacher (we omit the teachers that have no useful knowledge about the queries from the peer agents). While our approach is similar for both agents Ag_L and Ag_T, we discuss the learning process in two distinct units. We distinguish the situations that Ag_L has a minimum knowledge about o or not when it has identified c_k as the concept representing o. This is modeled with a threshold thr that $P(c_k|o)$ is compared to. If $P(c_k|o) \geq thr$, then we call Ag_L to have a minimum knowledge about o, so it will reform its ontology based on this definite concept c_k. If not, then Ag_L should reform its knowledge-base in another way. The generally agreed upon method in this circumstance is that Ag_L should add the exact concept Ag_T has sent as response into its ontology. Addition of a new concept has been discussed thoroughly in [13]. We believe that Ag_L can add c_{k_T} below c_{k_L}, while c_{k_L} has been identified to be the most similar concept in $\mathcal{O}_L$ to represent o. The new concept node is in fact a set of multiple features f_i and their domains D_{f_i}. These domains are updateable in subsequent transactions over any object detection. To put probabilities on these features and their domains, solely for a new concept, Ag_L puts a uniform distribution over the domain of each feature f_i, for all $f_i \in response_T$.

However if Ag_L has at least a bit of knowledge about o (if $P(c_k|o) \geq thr$), then an incremental approach is needed to reform the concept c_{k_L} in the ontology. Concept reformation has been introduced in [3] and means updating the probabilities for each pair of features and their values, representing any certain concept. To do such updating, we use an update equation similar to the one we used in [3] with a little modification. This equation is as follows:

$$P(f_i = v_{ij}|c_k) = P(f_i = v_{ij}|c_k) \ + $$
$$sgn(P(c_k|o) - thr) \times \alpha \times P(c_k|o) \qquad (5)$$

While these values should represent the probabilities of features f_i taking the value v_{ij}, we revise them through a normalizing phase after the update phase above:

$$For\ all\ f_i \in c_k$$

$$For\ all\ v_{ij^*} \in D_{f_i}$$

$$P(f_i = v_{ij^*}|c_k) = \frac{P(f_i = v_{ij^*}|c_k)}{\sum P(f_i = v_{ij}|c_k)}$$

The same will work for Ag_T using the information vector from the opponent (the query vector).

VI. Conclusion

The problem of MAS concept learning under a game theoretic view is studied and a novel approach is proposed. The approach is based on the process of determining the value of the information pieces which agents want to share. In addition to this evaluation, the preference modeling of an agent also proposed which modeled the willing of the agent to share information with its peers.

References

[1] W. Xianjia, L. Weibing, *Preference and Evolution in the Iterated Prisoner's Dilemma*, Acta Mathematica Scientia 2009,29B(2):456464

[2] V. S. Zuber, B. Faltings, *Inferring User's Preferences using Ontologies*, AAAI, 2006.

[3] N. Mirbakhsh, A. Didandeh and M. Afsharchi, *Incremental Non-Unanimous Concept Reformation Through Queried Object Classification*, Proc. WI/IAT'09, 2009.

[4] M. Afsharchi, B.H. Far, J. Denzinger, *Ontology-Guided Learning to Improve Communication between Groups of Agents*, Proc. AAMAS 2006, Hakodate, 2006, pp. 923–930.

[5] M. Wooldrige, *An Introduction to MultiAgent System*, 2nd ed., Wiley, 2009.

[6] K. Binmore, *Fun and Games: A Text on Game Theory*, D.C.Heath and Company: Lexington, MA, 1992.

[7] S. Parsons, M. Wooldridge, *Game Theory and Decision Theory in Multi-Agent Systems*, Kluwer Academic Publishers, 2000.

[8] G. Chalkiadakis, E. Elkind and N. R. Jennings, *Simple Coalitional Games with Beliefs*, Proc. IJCAI'09, 2009.

[9] S. Ieong, Y. Shoham, *Bayesian Coalitional Games*, AAAI, 2008.

[10] A.B. Williams, *Learning to Share Meaning in a Multi Agent System*, JAAMAS 8(2), 2004, pp. 165–193.

[11] D. H. Fisher, *Knowledge Acquisition Via Incremental Conceptual Clustering*, Kluwer Academic Publishers, Boston, Machine Learning 2: 139-172, 1987.

[12] S. Sen, P.P. Kar, *Sharing a concept*, AAAI Tech Report SS-02-02, Stanford, 2002.

[13] M. Afsharchi: *Ontology Guided Collaborative Concept Learning in Multi-Agent Systems*, PhD Dissertation, 2007, Faculty of Electrical and Computer Engineering, University of Calgary, Calgary, Alberta, Canada.

[14] J. A. Hendler: *National Cancer Institute's Ontology of Cancer*, DAML, DARPA, 2003, http://www.mindswap.org/2003/CancerOntology/nciOncology.owl.

[15] Protege Project, *http://protege.stanford.edu/*, October 2006.