Exploiting Simple Corporate Memory in Iterative Coalition Games

Carlos Mérida-Campos [1] and Steven Willmott[1]

[1] Llenguatges i Sistemes Informàtics,
Universitat Politecnica de Catalunya,
Modul C5 / C6 Campus Nord, UPC
Barcelona, Spain, E-08034
{dmerida, steve}@lsi.upc.es

Primary Contact: Carlos Mérida-Campos
Postal Address: As Above
Tel: +34 93 401 56 46
Fax: +34 93 401 70 14

Abstract. . Amongst the challenging problems that must be addressed in order to create increasingly automated electronic commerce systems are those which involve forming coalitions of agents to exploit a particular market opportunity. Furthermore economic systems are normally continuous dynamic systems – generating many instances of the same or similar problems (the regular calls for tender, regular emergence of new markets etc.).The work described in this paper explores how simple forms of memory can be exploited by agents over time to guide decision making in iterative sequences of coalition formation problems – enabling them to build up social knowledge in order to improve their own utility and the ability of the population to produce increasingly well suited coalitions for a simple call-for-tender economy.

1 Introduction

Among the challenging problems which must be tackled in order to realise some types of Agent Mediated Electronic Systems is the well known coalition-formation problem [Kraus97] – given the emergence of a business opportunity (such as a new market, a government call for tender, a contracting opportunity or a product niche) which companies or individuals could/should work together to best fill this niche – generating profit for themselves and utility for the market environment as a whole?

This problem generally does not just occur *once* in a given environment, it seems likely that many systems will be characterised by populations of agents representing different corporations *regularly bidding/competing for new continuously emerging opportunities.* As in human mediated systems it seems highly likely therefore, that alongside individual considerations for solving a particular coalition formation task, (for a particular call for tender or market for example) social structures within the population (such as who knows who, knowledge of past performance, existing obligations and so forth) which emerge over time will play a major role.

The aim of this paper is to take a simple model of a an iterative commercial world where a population of agents regularly need to form consortia to address new market opportunities and demonstrate simple decision making strategies based on learning between episodes can be exploited to make coalition choices.

Specifically, the paper defines a simple call-for-tender economy based on the iterative coalition worlds described in [Merida04] and explores how agents can exploit two types of information to assist themselves over time:

- Information about past performance – a simple form of society wide reputation.
- Information on previous shared mutual success or failure – a simple form of memory.

Next section defines the world and models applied, Section 3 the experiments Experimental setup and results are provided in Section 4. Sections 5 -7 add analysis and discussion of implications / properties of the model and strategies as well as how they relates to existing work in coalition formation, game theory and a number of other areas.

2 Problem Definition and Models

Taking the Iterative Coalition Formation problem defined in [Merida04], a coalition formation problem ([Sandholm99, Klusch02]) can be defined at its simplest as:

- Given a population **P** of agents and a list of tasks or goals **T**.
- Select subgroups of agents **S1**, **S2**, **S3**, … of **P** to address each of the tasks in **T**.

A variety of problems then address the properties of the subgroups (stability, maximum social welfare, pareto efficiency, etc.), their behaviour (how any payoff is split or how the coalition is maintained to complete the task) and how the coalitions can come into being (the amount of information available, whether agents are cooperative, whether agents can be part of multiple coalitions and so forth). In particular the *coalition formation problem* deals with the process of finding a set of combinations of agents which best solve given tasks in a given problem episode T (on set of tasks) for a given population of agents. From now on defined as a single *coalition game*.

Iterated coalition formation extends this notion by drawing the literature of *iterated games* used in game theory to lead to an *iterative sequence of coalition games*:

- Given a population **P** of agents and an iterated sequence of lists of tasks/goals $\mathbf{T_1}$, $\mathbf{T_2}$, $\mathbf{T_3}$, ….
- Let each task $\mathbf{T_i}$ correspond to a single coalition game $\mathbf{G_i}$.
- Select for that $\mathbf{G_i}$ subgroups of agents $\mathbf{S_{i1}}$, $\mathbf{S_{i2}}$, $\mathbf{S_{i3}}$, … of **P** to address each of the tasks in $\mathbf{T_i}$.

Intuitively this means that a population of agents persists over time to experience a series of coalition games – one after the other.[1] Applying such a system to the analogy of an agent mediated electronic economy:

- Agents represent companies active in a particular domain.
- Public bodies or other corporations issue calls for tender which require one or more companies to resolve – the tasks in the environment.
- Agents can work together in consortia (the subgroups in the environment) to bid for the contract to address these tasks.
- Tasks may be static (similar tasks occurring regularly) or dynamic (changing over time).

Another formulation could map Tasks to the creation of products to market niches (e.g. yearly demand for new fashion items). Other analogies could be for agents representing freelance contractors (combining to bid for small tenders), agents representing software developers (combining to develop products for emerging market niches) or large corporations combining to create new consumer devices for new market niches.

In many of these cases it is highly unlikely that the whole coalition formation and interaction process would be managed automatically by agents (since there are many

1 An obvious generalisation of this are continuous coalition formation problems where problems may occur at any time, have durations and may overlap in time – hence removing the step wise / discrete nature of the iterative model.

factors involved), however the system described here outlines the basic form of the problem and provides a framework for what might be at least partial decision support systems.

3 Game Setup

The world described in the previous section allows us to create a coalition game environment in which agents apply a range of strategies. The environment could be almost arbitrarily complex depending on the precise rules of the world, the nature of the tasks / Agent skills as well as the information / actions available to agents. Further complexity factors include whether the population of agents is allowed to change over time, whether the skills of individuals could change and what restrictions there are on coalitions (such as their size and payoff division within coalitions).

Due to this complexity, the world explored here is simplified to keep both experiment execution and analysis feasible. In particular:

- Experiments are a series of 500 individual coalition formation episodes – each characterized by a single call-for-tender that agents may form coalitions to bid for.
- Only one winning coalition is selected each around according to the function given in Box 1. (Intuitively this function picks the best coalition which meets all criteria OR if there is no such coalition the best coalition overall.)
- Only the winning coalition is awarded with a fixed sum payoff each round and this is assumed split evenly between the members of the winning coalition.
- Agents have no knowledge of each others skills, neither is there a centralized matchmaking system which helps cluster agents based on their skills – skills are only seen/evaluated by the tendering agency at bid evaluation time and not shared.
- Agents only have knowledge of past successes and failures (their own and those shared with others in previous coalitions) and of which agents in the population win at each turn (publicly announced information).
- The number of agents in the population is fixed (agents cannot join or leave the system between rounds) and skills are assigned randomly at the beginning of the iterative sequence – staying fixed across all games.
- The maximum number of agents in a coalition is set at 6 and the number of agents in the world at 102 (a multiple of 6 for convenience) in the experiments carried out here.
- In some experiments the task issued each round may be fixed (randomly generated initially) and in some it may change over time (see Section 4.3).

In this context in each round agents find themselves in a particular coalition (in the first round this is a unitary coalition with only themselves) and can exercise only one choice: whether to Stay in the current coalition for the next game, or leave and join another.

- A task **t** is **fulfilled** by a coalition **C** if the value for all the skills in **t**, is less than or equal to the maximum value among the members of **C** for the same skill.
- For a coalition **C** which fulfils **t** the **surplus** of **C** is the sum of all the skills in the task with value>0 of the difference among the maximum skill value of an agent in the coalition and the required skill value of the task
- For a coalition which does not fulfil **t**, the **deficit** of **C** is defined as the subtraction of de difference among the maximum skill value in the coalition and the skill value for all the skills required in the task.
- The **winner** is chosen as: A) if there is a coalition that fulfils the task is the coalition with maximum surplus, or B) if there is no coalition that fulfils the task is the coalition with minimum deficit

Box 1: *Winner determination function per game.*

Agents which leave coalitions in any given round are randomly ordered and one after the other allowed to select which existing coalition they would like to join (excluding coalitions which already have (6 members) or (with a random probability) whether to form a new coalition. Hence at the end of every round agents which leave are again clustered into a new set of coalitions.

Although this generates a relatively simple world (bounding coalition size bounds complexity for example [Fiaschi04] and agents have seemingly simple choices), the system still exhibits significant complexity: coalitions are partially instantiated at each turn (not allowing full reorganization), knowledge builds up in a complex manner (depending on which agents have been thrown together with others) and the number of possible coalitions is very large even in a small system (for 102 agents with size 6 maximum coalition size the number of combinations is approximately 1.3 billion – for size *m* in a population of *n* agents the value is a binomial number *[n,m]*).

3.1 Agent strategies

In this world, agents therefore need to apply strategies to decide when the stay or leave their current coalition. In this paper we consider just two simple strategies:

- Local Memory Agents (LMA): Agents keep track of which agents they have previously worked with in coalition and apply a Reinforcement Learning (RL) mechanism (see section 3.2) to raise assign/track positive or negative feeling (called *affinity* from now on) for other agents based on whether previous collaboration lead to a positive or negative result.
- Global Memory Agents (GMA): The agent is aware of who is playing the game, and uses a similar RL mechanism increase/decrease the affinity for those agents known (publicly announced) as being in the winning coalition in each round.

Hence in both cases agents track previous results in the world – however in the former they remember their own interactions with others and in the later they remember global indicators of success (independent of whether or not they themselves were involved).

For both strategies, the decision of staying or leaving the coalition is taken in the basis of the “affinities” observed in the following way:

$$Leave(X,Y,\Delta) = (Y - X) > \Delta$$

Formula 1: *Function to decide whether to stay or leave the current coalition*

Where *X* stands for the sum of the affinity towards the members of the agent’s current coalition and *Y* stands for the average affinity that the agent could have in case she decides to leave the current coalition. Y is calculated as the mean of the affinities that belong to the players outside the agent’s coalition multiplied by the maximum coalition size.[2] Delta is a factor that models the degree of willingness of moving to possible profitable coalitions. A delta equal to 0 will mean that the agent will stay in the current coalition unless the possibilities outside of it look better than in the current coalition, a positive delta means that the agent will stay even though there may be better perspectives outside (up to a certain threshold). Negative delta means that the agent is willing to leave the current coalition even if the expectations outside are worst conditions observed outside her current coalition.

3.2 Reinforcement Learning

Reinforcement Learning is used both to determine whether to stay or leave and as a mechanism for deciding where to go in the case of deciding to leave. Each agent maintains a data structure that stores information about the rest of agents in the game (reputation information for GMA, and affinity for LMA). Initially all the agents in the structure have an initial score *Init*. When the agent decides to reward or punish another agent a fixed amount *FixInc* and *FixDec* is added or subtracted respectively. The sum of the scores of agents in the structure is a constant, so each time a set of agents are rewarded, the amount to concede to them is subtracted evenly from the rest of agents. In the same way, when an agent/s is/are punished, the amount of penalty is shared from the rest of agents. The values that scores of agents can have are bounded from *MaxBound* to *MinBound*.

4 Experiments

The following sections describe a set of experiments designed to better help understand the environment, in particular- given limited information and strategies available:

---

2 Note that both X and Y are approximation since A) some members of the current coalition may leave and new individuals join and B) given that if it leaves the agent has some choice of where to move to (depending on its random place in the order of choices) the final affinity with the new coalition may be slightly higher than Y.

- Which strategies are dominant? If any?
- How do different strategies interact?
- Which strategies lead to better global solutions and/or faster convergence to good solutions in the overall population?

Experiments are initiated by generating a population of agents with a random distribution of skills and a task with random corresponding requirements. Each agent in the environment is assigned one of the basic strategies described previously (LMA, GMA or random as a control) as well as a DELTA threshold value. Experiment runs cover 500 coalition episodes – after each of which a winning coalition is designated, agents update their learned data and make new coalition changes. In most experiments the task remains fixed, experiments where the task changes are covered in Section 4.3.

The following parameters are fixed here: 1) Number of agents in the game to 102. (divisible number by the maximum coalition size). 2) maximum coalition size to 6 members. 3) Highly irregular distribution uniformity of the skills in the population and tasks (skill points far from being evenly distributed among the skills) 4) *N1* is set with value 100 and *N2* is set with value 300. 5) The number of different skills *Sk* is set to 10. In the reinforcement learning mechanism 6) the initial value *Init* to 0, and 7) the bounds *MaxBound* and *MinBound* to 1 and -1 respectively, finally 7) the DELTA values used parameter in the RL mechanism for both types of agents were -0.5, -0.2 , 0, 0.2 and 0.5. Those different setups of delta are noted in the experiments as the suffixes --, -, 0, + and ++ respectively.

Each combination of populations / setup was tested 20 times with different random populations. For the last set of experiments (Dynamic Environment) the task was varied over time regularly each 10 or each 50 rounds.

4.1 Pure Populations

The first set of experiments carried out was to assess the performance each strategy by itself. Each population was run with 20 experiments and the behaviour observed. Almost all populations showed a significant turnover in winning coalitions over time with a steady increase in the quality of solutions. This is shown in Figure 1 which plots the sums of the distance from optimal (Smith's Alpha value [Smith62]) for each population type.

In the graph lower means better as coalitions evolved in the population that got very close to the optimum coalitions possible in the populations (calculated separately using a brute method). The results show:

- A significant advantage for LMA Populations over GMA populations.
- A small but significant difference between the performances of LMA by using different DELTAS in that whilst LMA zero performs best, LMA+ and LMA++ performs very nearly as well but LMA- and LMA-- perform worse than the three of them: creating a sequence 0, +, ++, -, --.

- The reverse is true for GMA agents that also perform best around zero, but GMA- outperforms GMA+ (with GMA+, GMA++ performing very poorly).
- GMA+ and GMA++ population performing worst of all (even worst than random)
- Random choice populations performing poorly - showing that all learning based strategies (except GMA+ and GMA++) produced some improvement.

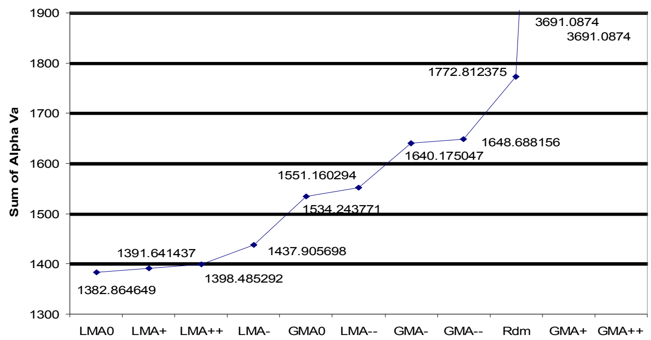

Figure 1: Results for the sum of Smith's Alpha values [Smith62] at the end of each experiment for each one of the pure configurations tested. (Note that the scores for GMA+ are off the scale – see explanatory text for analysis.)

While the comparative values vary with the parameters applied, the results show a relatively stable relationship across variations in thresholds, delta and other parameters.

As later experiments show, the reasons for the relative performance difference between of LMA over GMA is twofold: the type of information which is being learned (and its indicator for success) and the amount/nature of agent mobility between coalitions which each strategy generates. Although global information on winners might be expected to be useful information it appears that the local but pairwise information learned by LMA agents is a stronger indicator for success.

Secondly, in GMA experiments, the decisions agents make to move are based only on their *perception of the probability of encountering (more successful) members* of the population rather than any particular judgment on their current coalition partners – leading to a lack of motivation to leave the current coalition. This effect is best exemplified in the behavior of GMA+, GMA++ populations in which agents quickly form an initial winning (but suboptimal) coalition – all members of which receive global recognition – but which is subsequently never challenged because agents (winners or losers) are never able to reach their decision threshold (0.2 or 0.5) to move elsewhere. Affinity is concentrated in a small number of individuals and averages out with a uniform distribution of negative affinity across the remainder of the population. The fact that agents in GMA+, GMA++ populations don't move is the reason that performance is even worse than a Random Strategy population, as the value of the initial coalition is significantly smaller than the value of a bigger random coalition. Moreover, Random strategy has the advantage of being able to discover optimal coalitions by chance (although they will not persist over time).

As affinity for members in the winning coalition tends to the maximum, negative affinity elsewhere flattens off – never creating a differential large enough to cause mobility: essentially freezing the population in place. At GMA0, GMA- and GMA-- however mobility is assured since members of continually loosing coalitions will always perceive more opportunities elsewhere – thus generating new sets of potential winning coalitions as in LMA. The twin factors of the type of information learned and mobility also suggest why for LMA 0 outperforms +/- but + and ++ outperforms – and – –, and for GMA – outperforms +. These suggest that there may be an optimal value for LMA somewhere between 0 and + DELTA and for GMA somewhere between 0 and – DELTA (or several optima):

Figure 2 illustrates a fragment of a sequence of games where a coalition of score -87 that have been together during more than 200 games and winning many of them (all scores of -87 correspond to the same set of agents in this case) is finally broken by a sustained sequence of better outperforming coalitions Finally, after being together 230 games, its members no longer win regularly and eventually disband as another stronger coalition establishes itself.

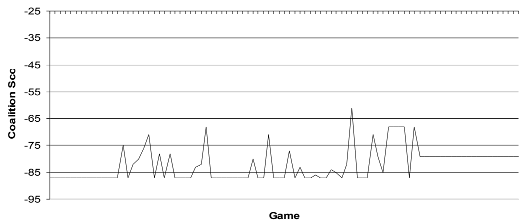

Figure 2: Sequence of experiment of LMA0 pure population that shows an episode of one coalition staying together, winning regularly and eventually being superseded.

The tradeoff between exploration and exploitation is reflected in different ways by the two strategies:

- For LMA, agents learn indicators widely across the population as they move (since they store information about every previous partner) – diffusing learner information across the population: leading to situations where strong coalitions are difficult to establish (as shown below) and some resistance is beneficially to temper exploration.
- For GMA information is concentrated on previous winners – leaving much of the population relatively anonymous and biasing the system towards stability of existing winning coalitions: indicating that more exploration is needed to find strong combinations of agents.

4.2 Mixed Populations

A logical step onwards from pure strategy populations is study of the interactions between different strategies within the same game. Figure 3 shows a summary of the convergence to optimal results for combinations of 50% of one strategy with 50% of another.[3]

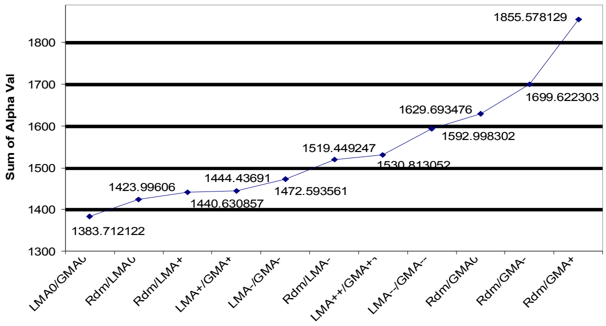

Figure 3: Results for the sum of Smith's Alpha values [Smith62] at the end of each experiment for each one of the mixed configurations tested

From observing the results of convergence we can conclude that the properties observed in pure populations are maintained in the hybrid populations. However, there are a number of interesting effects:

- A 50:50 mix of LMA0:GMA0 performs almost identically to a pure population of LMA0 agents (the best of the pure populations).
- In general for all LMA and GMA strategies, mixing with random strategies has very little impact on the final solution quality – coalition quality achieved is in general only 5-10% below what a pure population may have achieved.
- Mixing with other strategies such as LMA allows even very poorly performing GMA+ and ++ strategies to perform reasonably well.

In each of these cases it is important to consider the nature of the interaction between agents in the environment – in general both winning and loosing coalitions will contain agents of both types. Hence exploratory strategies such as random shake up static strategies such as GMA+ and allow them to function – while retaining some of their strong properties (exploiting knowledge of winners). The strong performance of LMA0:GMA0 is also particularly interesting since it suggests that GMA0 agents are either positively influenced by the dynamics generated by LMA0 or GMA0 decision making creates similar conditions as would be found with all LMA0 agents (note that this mix performs better than Random:LMA0 – indicating that GMA0 agents are having a positive effect).

Furthering this comparison, Figure 3 shows that depending on the experiment either GMA or LMA agents might get the majority of the payoff over 500 runs (and many

3 The two populations have exactly the same skills setup, hence only differing in their strategy.

runs are similar). Although on average LMA populations gain 37% more profit than GMA agents when competing together, this is not a very significant advantage, and suggests that both strategy types play a role in the good performance of this mix.

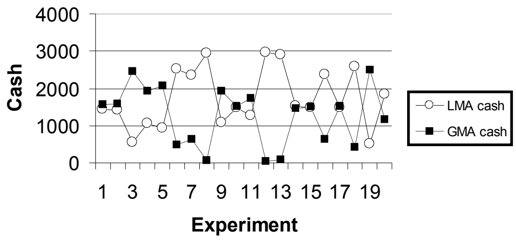

Figure 3: Payoff in each experiment for each Agent population having mixed LMA0 and GMA0 populations (numbers 1-20 indicate different experiment runs).

4.3 Dynamically changing tasks

All experiments discussed up until this point used a the same task in each round of any given 500 game run – enabling agents to learn specifically which combinations work well for the task. With the task varying over time however, system dynamics also change. Figure 4 show relative performance of GMA0 and LMA0 in environments where the task is adapted (see Section 3) either every 10 rounds (fast change).

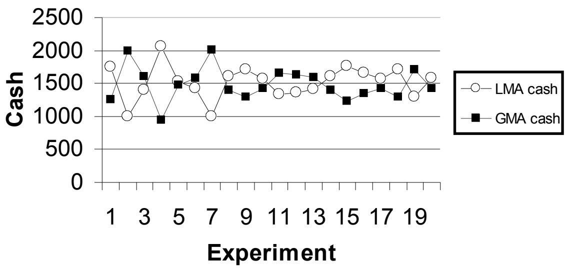

Figure 4: Payoff in each experiment for each Agent population having mixed LMA0 and GMA0 populations and a Dynamic environment where each 10 games the task changes

The results clearly show that A) as the speed of change increases the difference between the two strategies decreases and B) in the fast changing scenario GMA even begins to outperform LMA by a slight margin (although not statistically significant). (Note that quality of solutions was also reduced but this cannot be easily measured since the optimal coalitions change with each task - although performance is poorer however both strategies still outperform random by a considerable margin (LMA gets 51% and GMA 20% more profit on average than Random)

Since these measures of performance are highly dependent on specific parameters chosen it would be unwise to draw general conclusions. However the changes due to dynamism do indicate:

- As dynamism increases LMA knowledge of pairwise relationships is more quickly eroded (made less useful) than GMA information about individuals.
- Potentially at some point LMA knowledge may become less valuable than GMA knowledge.
- There is still value in learning event for short periods of stability such as 10 games (relative to ignoring contextual information).

An interesting deeper point here would be if it was possible to determine whether or not at high dynamicity agents were able to learn some fundamental compatibility traits which reflected the ability of certain subsets of agents to perform well across ranges of tasks – rather than for just one.

5 Contributions and Relationship to Existing Work

Coalition formation problems for economic environments have been tackled basically from game theoretic perspective (pioneering works are [Farrell88],[Greenberg94]) leading to the development of a wide range of techniques. Most of these techniques however focus solving once-off coalition formation problems in which a particular set of agents must be matched against a specific task to find one (possibly optimal) solution.[4] Conversely in much of the research carried out in Agent Mediated Electronic markets usually assumes Agents have much less knowledge about other participants but that systems may endure over long periods of time – providing agents with many opportunities to participate (see [Preist02] and [Faratin00] for example).

In this context, the work presented here explores how simple learning mechanisms can be used over time to build up background knowledge in an environment and subsequently exploit it to make coalition decisions. The long term aim is to use this and mechanisms of social context together with standard coalition formation problems that tackle single instances – so that agents could exploit social relationships / knowledge and task/skill specific knowledge.

Although there is previous work which touches on related issues including coalition formation and learning [Soh03], work on coalition formation through motivation and trust [Griffiths03], and dynamic coalition formation [Klusch02, Soh02], the research closest to that presented here are iterative coalition formation mechanisms (such as

4 This problem is known to be NP-complete in most general cases as it corresponds to a multi-dimensional set-covering problem with the number of states growing with the number of agents, the number of tasks, the number of agents allowed in each coalition and the dimensionality of agent and task characteristics.

[Konishi03]) and work by Shehory, Kraus and others on coalition formation in systems where little information is known about agent skills (such as [Kraus03]).

Still even in these cases work focuses primarily on single coalition formation problems and not related social context. The aim with the work presented here and a similar analysis on adaptation of game theoretic techniques such as tit-for-tat [Merida04] is to begin looking at this wider social context that arises in iterated coalition environments – in particular:

- How social structures accumulate or change over time
- How social structures help or hinder both decision making (by for example reducing the search space) and improve/decrease solution efficiency
- How decision making based on a particular coalition formation problem instance (bidding for a particular tender) and the interpretation of pre-existing social context (reputation, memory, existing agreements, partial coalitions and so forth) interact over time.

The factors are likely to play a significant role in the type of open dynamic trading environments envisaged in many Agent Mediated Electronic Commerce Scenarios and may provide a useful bridge to more abstract work on social agent systems such as work by Castelfranchi, Sichman, Conte and others on Social Power [Sichman94].

6 Discussion points

Returning to the more limited scope of applying simple learning techniques, as described in section 3.2, the world model, game rules and strategies described here are very basic and leave open a range of interesting questions:

- The results show a distinction between different dominant factors in learned knowledge depending on task dynamism – however it is not clear whether these are in fact redundant against one another or conflicting.
- The dominance of information on pair wise relationships (LMA) at low rates of change is not especially surprising, however this raises the question of whether ternary (or larger n-ary) relationships could also be exploited.[5] This would presumably eventually lead to diminishing returns and a storage-space/complexity v's gain tradeoff.
- Since learning about success/failure as used here is used as a surrogate to working with known skill profiles of agents (which are assumed by many coalition formation techniques) an interesting question is to what extent the learned information mirrors skill information and whether in fact it may be possible to infer skill information over time.

---

5 An example rule would be multiplying affinities together for sets of agents which had already been in coalitions together rather than adding as is currently done.

Application of standard reinforcement learning techniques such as [Bush55] and or adaptive changes in learning strategies (forgetting, etc.) would potentially also provide significant improvements in performance. More general topics of interest include:

- Techniques such as the learning outlined here or the game theoretic approach are approximations for actual effectiveness of any particular coalition for the Tasks arising – raising the question of whether strict bounds can be calculated for convergence given particular world characteristics.
- Agents in current systems carry only knowledge between games, however an important class of extensions would be those in which agents gained advantage or influence in proportion to their success – modelling a successful company's opportunity to invest more heavily than its competitors. This would significantly change the dynamics of the world – potentially swinging influence back to individual rather than group success.
- The current setup assumes uniform distribution of payoff within a coalition and no knowledge of agent skills. Relaxing these conditions would open the way for agents to apply a range of existing coalition formation strategies within each coalition game – combining this with learned knowledge from games over time.
- A further generalisation of the game to continuous task arrival would provide an extra dimension of interest with agents needing to integrate considerations of opportunity cost involving.

Each of these additional dimensions of complexity suggests that in most realistic systems; strongly suggesting that in systems of bounded rational agents exploitation of social structures such power relations, reputation, social norms or other features is likely to be critical in reducing complexity explosion in decision making.

7 Conclusions

As Electronic Commerce systems evolve it seems likely not only that they will become more automated but they will also increasingly take on a similar continuous/always on/iterative nature of today's human mediated commercial systems - bringing together concerns of individual behavior, group behavior in a commercial context and durable social structures/knowledge.

The work presented in this paper includes a simple framework for describing iterative economic systems in which agents must regularly perform tasks in teams or coalitions and explores simple learning mechanisms which can be used to guide decision making. The experiments demonstrate:

- How simple corporate memories can be used to track global success metrics and memory of previous joint success over time.
- How agents can use this accumulated information to both increase their own utility and improve the quality of solutions generated by a given population.

- How different types of knowledge (global versus local, unitary v's binary) affects outcomes.

The long term aim of the work presented here is tackle some of the issues raised in sections 2 and 5 and analyze the interplay between decision making based on a particular coalition formation problem instance (bidding for a particular tender) and the interpretation of pre-existing social context (reputation, memory, existing agreements, partial coalitions and so forth) in any decision.

References

[Axelrod81] Axelrod R., Hamilton W. "The Evolution of Cooperation" Science, 211:1390-1396, 1981.

[Bush55] Bush R., Mostseller F. "Stochastic Models for Learning" Wiley: New York. 1955

[Faratin00] Faratin P., Sierra C and Jennings N. "Using similarity Criteria to make negotiation trade-offs" (2000) . In Proceedings of ICMAS'00

[Farrell88] Farrell J., Scotchmer S. "Partnerships". Quarterly Journal of Economics, 103(2), 279-97, 1988.

[Fiaschi04] Fiaschi D., Pacini P.M. "Coalition Formation with Boundedly Rational Agents" in The Complex Dynamics of Economic Interaction. Essays in Economics and Econophysics. Springer, Berlin. 2004.

[Greenberg94] Greenberg J. "Coalition Structures". Handbook of Game Theory. Vol. 2 chap 37, pp. 1305-37, 1994

[Griffiths03] Griffiths N., Luck M. "Coalition Formation through Motivation and Trust" in Proceedings of AAMAS 2003

[Kaelbling96] Kaelbing L.P., Littman M.L., Moore A.W. Reinforcement Learning: a survey. Journal of Artificial Intelligence Research 4:237,285. 1996

[Klusch02] Klusch M., Gerber A. "Dynamic Coalition Formation among Rational Agents".IEEE Intelligent Systems, 17(3):42-47, 2002

[Konishi03] Konishi H., Ray D. "Coalition formation as a dynamic process" (2003). Journal of Economic Theory 110, 2003.

[Kraus97] Kraus. "Negotiation and Cooperation in Multi-Agent Environments" (1997) Journal of Artificial Intelligence. Vol 94: 79-98.

[Kraus03] S. Kraus, O. Shehory, G. Tasse, "Coalition Formation with Uncertain Heterogenous Information", Proccedings AAMAS'03, 2003.

[Lerman03] Lerman K., Galstyan A., "Agent Memory and Adaptation in Multi-Agent Systems" (July 2003). Proc. AAMAS'03

[Lerman00] Lerman K. Shehory O. "Coalition Formation for Large-Scale Electronic Markets" Proc. ICMAS 2000.

[Merida04] Citation for AAMAS paper (forthcoming). [*The submission version of the paper is available online at* lsi.upc.es/~dmerida/aamas/paper.pdf *for the purposes of this review if desired*]

[Preist02] Preist Ch., Mérida C. "Automated Negotiation in Many-to-Many Markets for Imperfectly Substitutable Goods" (2002) Proc. AMEC'02

[Sandholm99] Sandholm T., "Distributed Rational Decision Making" in Multiagent Systems: A Modern Introduction to Distributed Artificial Intelligence, MIT Press, Cambridge, Mass. 1999.

[Sichman94] Jaime Sichman, Rosaria Conte, Yves Demazeau, Cristiano Castelfranchi "A Social Reasoning Mechanism Based On Dependence Networks", In proceedings of the 11th European Conference on Aritificial Intleligence (ECAI'94), 1994.

[Smith62] Smith V. "An Experimental Study of Competitive Market Behaviour". Journal of Political Economics, 70, 111-137, 1962.

[Soh02] Soh L., Tsatsoulis C., "Real-Time Satisficing Multiagent Coalition Formation" Proc. AAAI Workshop Coalition Formation in Dynamic Multiagent Environments, AAAI Press, Menslo Park, Calif. 2002.

[Soh03] Soh L. Li X. "An Integrated Multilevel Learning Approach to Multiagent Coalition Formation" (2003) In proceedings of IJCAI'03