Increasing Realism in Coalition Formation

Michael van de Vijsel and John Anderson
Department of Computer Science, University of Manitoba, Winnipeg, Canada R3T2N2
mikev,andersj@cs.umanitoba.ca

Abstract

Much of the existing research on teams in multi-agent systems focusses on improving the performance of predefined groups of agents. Comparatively little research has been done on the coalition formation *process – how to create teams of self-interested agents. Furthermore, many of the existing coalition formation approaches make restrictive assumptions about either the agent mode or the domain in which the agents interact, that limit the applicability of these approaches to more realistic scenarios. This paper presents a new coalition formation approach that avoids making these assumptions, and evaluates this model against a baseline approach from the perspectives of system throughput and coalition stability.*

1 Introduction

Most work in multi-agent systems involves the improvement of the performance of predefined groups. How those agents came to be in a group is often ignored. This area, commonly known as coalition formation, is an important one both from the standpoint of studying how more effective groups can evolve (and ultimately supporting applications such as electronic marketplaces), as well as for studying the underlying cooperative mechanisms that give rise to groups. Most work on coalition formation approaches has used agent models and domains that include significant restrictions on the capabilities of the agents and their interactions with each other. For example, existing research often assumes that all members of a coalition are equal and bring the same amount of value to the group (e.g. [3, 9]). Agents are also often only allowed to belong to one group at a time (e.g. [2, 5]). Domains often restrict agents to a single goal or type of task (e.g. [2, 7]), or to a set of tasks of different types where there is no conflict between the tasks (e.g. [4]), eliminating the decision about the order to perform the tasks.

We are interested in creating improved approaches to coalition formation that are more reflective of the characteristics of real-world groups. In particular, we are focussing on domains where self-interested agents have conflicting goals, are not restricted to a single group, learn about others based on their interactions, and are of disparate value to a group.

This work proposes a new coalition formation approach that encompasses these aspects, providing increased realism over existing methods. The following sections describe the approach, its implementation in a software domain, and an evaluation comparing it to that of Dutta and Sen [4], a recent approach that also attempts to deal with some of the factors that have previously been ignored in coalition formation. Before this we begin with an overview of related research.

2 Related Work

Breban and Vassileva [2] have outlined a coalition formation approach for an information marketplace that is based on long-term coalitions and trust relationships between agents. When vendor and customer agents interact in this approach, they negotiate a price for a piece of information, which is discounted if they share a coalition. If the agents can agree on a price, the interaction is deemed successful, and if they are not, the interaction is classified negatively, allowing agents to maintain a trust level of others and base the likelihood of joining a coalition upon this trust. While their use of trust relationships is a key differentiator from other research, their approach limits agents to a single coalition at a time, and agents are homogeneous and only ever have a single goal to consider.

Anderson et al. [1] have combined reinforcement learning and coalition formation in a robotic soccer domain. In their work, agents on a soccer team possess different abilities – some are of higher skill than others. Agents begin with no innate knowledge of their domain, and they learn the appropriate behaviours by observing the actions of others. Implicit coalitions are formed when individual agents begin excluding others that they perceive to be of lower skill level, deciding to only learn from those agents that are performing at a higher level. This approach uses coalition formation to control the direction of reinforcement learning. However, the coalitions that are formed are implicit and do not manifest themselves in any shared information or group behaviour. Agents in this approach are not self-interested – they share a collective goal and use more of a team-based approach.

Lerman and Shehory [5] have developed an approach to coalition formation for electronic marketplaces, focussing on scalability. Their approach produces coalitions as emergent behaviour from a set of very simple rules at the agent level. Buyer agents in this approach are homogeneous and follow a set of rules local to their environment – there is no learning or tracking of history. Agents join other agents purchasing the same product as they are, and the price the agents pay for their product is determined by the size of the group they are in. Once the price is set and the order is filled, the agent may leave the group in search of a better price if desired. While the results of the experimentation are positive, the simplicity of the agent model makes the approach unsuitable for implementation in more realistic scenarios.

Dutta and Sen [4] have proposed a partnership formation approach that encompasses some of the characteristics of realistic scenarios. In their work, agents are heterogeneous in ability – each agent has expertise in a particular type of task. When performing tasks in their expertise, they can complete them in less time and with higher quality than tasks that are not in their expertise. They can also request aid in completing a task from another agent. The aid is granted if a *cooperation possibility* exists – that is, if the estimated cost of the helping agent to complete the task is less than the cost of the requesting agent to complete the task.

If a cooperation possibility exists, the helping agent will agree to perform the task based on a probabilistic formula that incorporates not only the cost of the task but the opinion of the helping agent about the requesting agent. This formula is described by:

$$Pr(i,k,j) = \frac{1}{1 + exp^{\frac{C_{ij}^k - \beta * C_{avg}^k - OP_i}{\tau}}} \quad (1)$$

where C_{ij}^k is the cost incurred by agent k to complete task j for agent i; β sets the initial cost an agent is willing to incur when a previously unknown agent has requested help; C_{avg}^k is the average cost of all tasks performed by agent k; and OP_i is the balance of past help that agent k currently has with agent i. An additional parameter, τ is used to adjust the shape of the resulting sigmoidal probability curve.

Each agent maintains an opinion about others with whom it has interacted, represented by the balance of savings it has over time with the other agent. If a helping agent completes a task with less cost than the requesting agent could have done, the difference in cost is stored. Over time, this cost balance is updated as agents help one another. An agent that has helped another agent many times will be less likely to grant aid in the future, until some of the favours it has performed are returned. Thus, this probabilistic formula implements the concept of *reciprocity* in the Dutta and Sen approach.

The domain in which this is implemented is an abstract domain where agents perform tasks that require specific expertise. An agent is an expert at only one (of three) task types, and must either suffer through longer task completion times and poorer quality results, or find others with greater expertise to perform the work for them. In Dutta and Sen's approach, agents begin in an exploratory phase where they choose an agent with whom they have not yet interacted, in an attempt to learn something about as many agents as possible. Once the agents have completed a specified number of tasks, the exploratory phase ends and agents begin to select the agent they feel will be the most useful in terms of cost savings.

This approach encompasses some of the characteristics of real-world groups previously outlined. Agents are heterogeneous, for example, and learn about others as they act in the world. However, there is no conflict between goals, because the order of task completion is irrelevant, and agents can complete as many tasks for others as they are given in one turn. Coalitions are also not explicit – each agent simply has its own group that it is likely to cooperate with. Since agents begin with no opinion about others, we can say for the purposes of comparison that an agent's coalition consists of those agents thought of with a positive opinion value. That is, those of whom the agent's opinion has increased since first interacting with them.

3 Experimental Domain

We have illustrated that even the most recent coalition formation approaches that have been proposed are lacking some of the characteristics of realistic scenarios outlined earlier. Thus, we propose a new coalition formation approach and agent model, which we will term the vandeVijsel agent model, which encompasses these characteristics. We begin by describing the experimental domain for the evaluation of this approach, in order that the approach can be described with examples.

The domain for experimentation is a package-delivery domain based on that described by Sen and Dutta [6]. That domain consisted of a central package depot, where agents received packages that were be delivered to a location on one of a number of radial fins extending outwards from the depot. An agent could carry one of its own packages at a time, along but not across fins, and could also carry an additional package to an address on the same fin for another agent.

In order to make this domain more realistic, we have extended it in several important ways. First, packages are picked up from and delivered to depots scattered randomly across the environment (a less restricted grid). Agents are assigned multiple packages upon arrival at a depot, in order that they might have multiple, potentially conflicting goals (delivery addresses in opposing directions). Packages are assigned an initial payoff value, calculated as a factor α times the Manhattan distance between the starting and destination points of the package. This payment is reduced by one every time cycle, and thus α must be a value that is proportional to the size of the domain to ensure mainly positive values. In our implementation this is set to 3 times the sum of the length and width of the grid. Further, a penalty is assessed if packages are never delivered (lost).

The implementation of this software domain places agents randomly on the grid and allows them to move, and also supports interactions between agents in the same grid locations, as well as the maintenance of coalitions. The next section describes the vandeVijsel agent model, and describes how these agents form effective coalitions.

4 The vandeVijsel Agent Model

In order to encompass greater realism in our agent model, we have made vandeVijsel agents fallible in terms of memory. A low memory attribute leads to losing packages and forgetting depot locations, affecting individual performance and effectiveness in a coalition. We have also made these agents variable in terms of speed, trust (the tendency to trust others with packages rather than delivering them itself), and honesty (the degree to which these attributes are exaggerated when reported to others, in order to appear more attractive to a potential coalition). Each agent is given a random value between 1 and 10 for each of these values when the agent is created, and these values remain static. Agents have a maximum payload of 10 packages at a time, restricting the degree to which they can help others, and do not have prior knowledge of depot locations. Each of these limitations deals with elements in Section 2 that were noted problems with prior research.

An agent's speed attribute, rather than fixing a uniform rate of speed, deals with the likelihood of movement from one grid unit to another in a given cycle. A random number is generated, and movement is automatically allowed 40% of the time, and disallowed 20% based on this value. The remaining 40% is governed by a second random value from 0–9, and if this value is less than the agent's speed attribute, the agent is allowed to move. This results in best case movement of 80% of the time (speed value of 10), and worst case movement of 44% (speed value of 1). Over time, an agent with a speed value of 10 will move 20% further over a period of time than an agent with a speed value of 5.

The basic action each agent must decide upon each turn from the standpoint of the domain is the direction of travel. As it is a self-interested agent, it chooses this to maximize payoff of the packages it carries. It thus chooses a package to deliver such that the sum of the payoff of this package along with the remaining payoffs of other packages are maximized (for further details, see [8]). The exception to this occurs when the agent is carrying a package for another agent, in which case that this will take priority over the agent's own package. As described in the next subsection, an agent will only agree to deliver a package for another agent if the package in question is within 10 units of the agent's current destination. Thus, it will not leave the area without first delivering the package for the other agent.

Once the agent has made the decision on direction of movement, it moves and then checks if its new location contains a package depot. If so, it adds the depot to its list of encountered depots, delivers any packages addressed to this depot (the payoff of these packages goes to the agents originally assigned the packages), and asks for new packages (if it has room in its payload).

After moving, memory fallibility is handled by generating a random number between 1 and 1000, and if (10 minus the agent's memory attribute) is less than this value, this causes the agent to forget one depot at random. A similar likelihood governs the agent forgetting a package, which will lead to a penalty for the agent (or the agent that trusted this one with the package) when it is not delivered.

4.1 Forming and Maintaining Coalitions

After each agent moves, it is given a list of others (*encountered agents*) that occupy the same location, and the possible interactions that could be performed between each are processed individually. First, the agent determines if it shares a coalition with the encountered agent. If the agents share a coalition, the agent can request aid in delivering packages from the encountered agent. The encountered agent will agree to deliver a package for the asking agent if the package's destination is within 10 units of its current destination and it has room in its payload.

If a coalition is not shared, the agent begins by evaluating the encountered agent for suitability in coalitions of which it is a member. If the sum of the *reported* attributes of the encountered agent is no more than 5 points below the average of a coalition, and no individual attribute is more than 5 points below the average of that attribute in the coalition, the encountered agent is considered suitable. The agent determines a list of suitable coalitions, ordered by average attribute difference, and proposes membership to each in turn. The encountered agent evaluates these via the same criteria but from the opposite viewpoint: it is interested that the coalition not be too low-valued (the same standard of 5 is used), and membership is rejected if the coalition is not suitable to the encountered agent. If no coalition is still shared at this point, the agent can also propose to form a new coalition with the encountered agent, and the evaluation of this is then performed using the reported attributes of each of these agents. If a coalition is shared at any point, the ability to make delivery assistance requests ensues immediately.

Because this relies on reported values, dishonest agents can flatter themselves to gain access to better coalitions. A base attribute inflation factor of $(10 - h)/3$, where h is the agent's honesty value, is employed to create reported attribute values. This factor is adjusted for each agent attribute value individually when the agent is created by adding a random value between 0 and itself, allowing variation between agents of the same honesty and between attributes in the same individual.

Once coalitions exist, they must be maintained. Maintenance involves removing agents that are not living up to the performance expectations of the coalition (largely due to agents having been deceptive about their abilities). The performance of an agent is assessed every cycle after delivering 10 packages for coalition members, allowing only a limited time for exploitation by poorly-performing agents. After movement and encounters are processed, the average payoff is calculated for all members of a coalition, and those falling below the average by more than a specified factor are removed. Through experimentation, this factor has been set at the sum of the X and Y dimensions of the grid. Coalitions cease to exist when they contain only one agent.

5 Dutta/Sen Agent Model

In order to examine the efficacy of this approach, we have adapted the partnership formation approach of Dutta and Sen [4] described in Section 2 to this domain. To allow for the disparate agent expertise central to that approach, each Dutta/Sen agent is given the ability to be an expert in a particular quadrant of the grid. This expertise allows them to move 100% of the time in their expertise quadrant, and 50% of the time when outside their expertise quadrant.

Movement in the Dutta/Sen agent model is the same as in the vandeVijsel agent model, with the exception of the expertise quadrant. Agents receive a list of encountered agents after movement, and the cost of having encountered agents assist with package delivery is estimated. In cases where others can offer a cost savings, assistance is requested and the encountered agent will provide aid with probabilistically based on Equation 1 if it has room for an additional package, duplicating the basic coalition formation approach used in [4]. Recall that part of this equation involves increasing the likelihood of cooperation based on the opinion of another agent, as captured by the cost savings realized (or not) by that agent. These savings are tallied whenever an agent completes delivery of a package for another agent, at which time the estimates of that agent's expertise are also updated.

6 Evaluation

In order to evaluate the performance of the vandeVijsel agent model as compared to the Dutta/Sen model, we ran 10 trials of 50,000 time cycles for each agent type in a software simulation. Each trial consisted of 500 agents operating in a 100x100 grid with 100 package depots. We examined both the throughput and coalition stability of each approach.

Figure 1 shows the aggregate throughput by all 500 agents of each type, averaged over the 10 trials. The vandeVijsel agents achieved a 38% improvement in throughput over the Dutta/Sen agents. The single factor that could exert the greatest influence on this

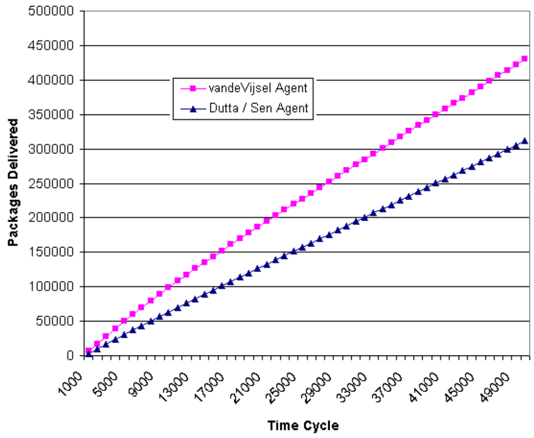

Figure 1: Comparison of overall system throughput for the vandeVijsel agent and Dutta/Sen agent, averaged over 10 trials

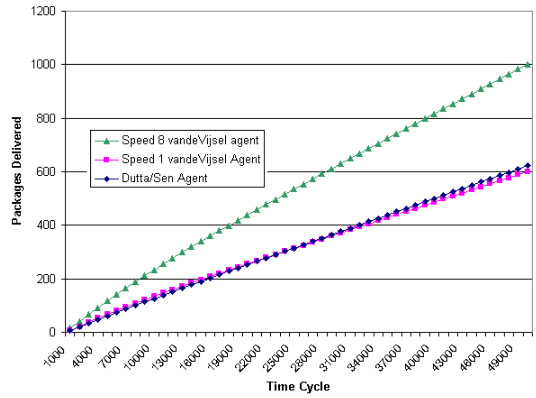

Figure 2: Comparison of actual throughput for vandeVijsel agents with speed values 1 and 8, and the average Dutta/Sen agent

result other than the coalition formation approach is the speed of the agents: if one group of agents are artificially faster than the other without regard to forming coalitions, this comparison would be biased.

In order to produce an speed comparison between these two approaches, we must obtain the effective speed of an average agent of each type. The speed of a vandeVijsel agent is directly related to the agent's speed attribute (resulting in a worst-case movement rate of 44% and best-case movement rate of 80%). The effective speed of a Dutta/Sen agent depends directly on the amount of time the agent spends inside and outside its quadrant of expertise. Tracking this value, we found that on average, a Dutta/Sen agent spends 22402 time cycles (45%) in its expertise quadrant, and 27598 time cycles (55%) outside of its expertise quadrant. Since Dutta/Sen agents move 100% of the time in their expertise quadrant, and 50% outside of this quadrant, this translates to an effective movement rate of 72.5% for a Dutta/Sen agent.

To compare these movement rates, we can translate the percentage movement rate for a Dutta/Sen agent to the speed attribute value required by the equivalent vandeVijsel agent. A vandeVijsel agent is guaranteed to move 40% of the time, and the other 40% is determined by its speed attribute. Thus, to calculate the effective speed attribute given the average rate of movement of a Dutta/Sen agent, we subtract 40% from the Dutta/Sen agent's effective movement rate, and divide the result by 4. We thus find that the movement rate accomplished by a Dutta/Sen agent is equivalent to a vandeVijsel agent with a speed attribute of 8.125. The random generation of attributes in vandeVijsel agents results in an average speed attribute of 5, showing that the Dutta/Sen agents are moving significantly faster, on average, than our agents. In addition, given this effective attribute value, we would expect the throughput of the Dutta/Sen agent to match that of a similarly fast vandeVijsel agent.

In fact, we find quite the opposite. Figure 2, illustrating throughput broken down by speed, shows the average Dutta/Sen agent displaying a throughput similar to that of a vandeVijsel agent with a speed attribute of 1. Not only are our agents attaining a higher throughput, they are doing so while moving more slowly, on average, than the Dutta/Sen agents. There is an additional factor to consider in this that also speaks positively of our approach. vandeVijsel agents have a range of honesty, trust, and memory values: agents can be deceptive regarding their abilities, lack the ability to trust others and gain the benefit of coalitions, lose packages, and forget the locations of depots. Despite these more challenging conditions, the performance of vandeVijsel agents is still higher.

We also compared the two approaches on the basis of coalition stability: the rate of change in the membership of the coalitions over time. In order to make a valid comparison, we must ensure that we are comparing similar coalition concepts between the two models. For the vandeVijsel agent model, coalitions are explicit, so changes are easily tracked. For the Dutta/Sen agent model, the closest concept is the group of agents for which an agent has a positive opin-

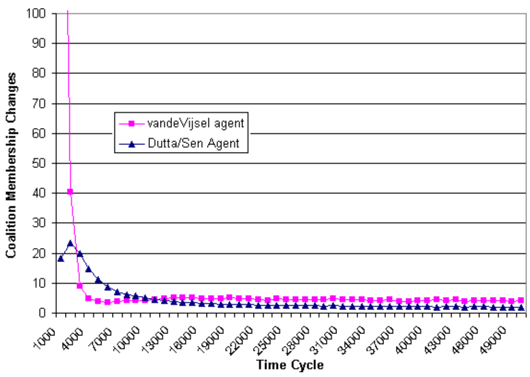

Figure 3: Coalition Stability in each approach, per coalition, averaged over 10 trials

ion, which ultimately controls the likelihood of cooperation in that approach. This results in a single coalition for each Dutta/Sen agent.

Since the opinion stored by a Dutta/Sen agent represents a balance of cost savings, it can oscillate between positive and negative values as two agents establish a reciprocative relationship. Only over time will the cost savings between two agents of complementary expertise increase to the point that they are both positive. For this reason, we examine the number of positive opinions at specific time intervals and evaluate the rate of change using the values at these intervals, rather than an overall change.

Finally, we must also consider the effect of multiple coalitions in the vandeVijsel approach, and the effects of deceptive agents (neither of which are present in the Dutta/Sen approach). Deceptive vandeVijsel agents exaggerate their abilities in order to gain admission to coalitions for exploitation purposes. They do this repeatedly, and are repeatedly removed when coalitions realize that their abilities are not what they claim. This results in a significantly higher number of membership changes that are not possible in the other approach, and so we compare stability on a per coalition basis. Figure 3 illustrates that viewed in this fashion, both models reach stability after approximately 10000 time cycles. The vandeVijsel agent exhibits slightly more changes on average, largely due to the potential for dishonesty among agents.

7 Conclusions and Future Work

In this paper, we have described a new approach to coalition formation, which takes into account a number of important factors that are exhibited by real-world groups. We have also compared an implementation of this approach to that of Dutta and Sen [4], a recent approach that also attempts to take some of these factors into account. This evaluation illustrated that we could obtain significant throughput improvement and similar coalition stability, while operating under fewer assumptions than Dutta and Sen. Further details of these experiments are available in [8]. Future work will entail dealing with other elements of realism, in particular working with agents whose attributes evolve over time, and populations that are not closed.

References

[1] John Anderson, Brian Tanner, and Jacky Baltes. Dynamic coalition formation in robotic soccer. In *Proceedings of the AAAI-04 Workshop on Forming and Maintaining Coalitions and Teams in Adaptive Multiagent Systems*, San Jose, CA, July 2004.

[2] Silvia Breban and Julita Vassileva. Using inter-agent trust relationships for efficient coalition formation. In *Proceedings of the 15th Canadian Conference on AI*, May 2002.

[3] Christopher H. Brooks and Edmund H. Durfee. Congregating and market formation. In *Proceedings of the 1st IJCAA-MAS*, July 2002.

[4] Partha Sarathi Dutta and Sandip Sen. Forming stable partnerships. *Cognitive Systems Research*, 4(3):211–221, 2003.

[5] K. Lerman and O. Shehory. Coalition formation for large-scale electronic markets. In *Proceedings of ICMAS-2000*, pages 216–222, July 2000.

[6] Sandip Sen and Partha Sarathi Dutta. The evolution and stability of cooperative traits. In *Proceedings of the 1st IJCAA-MAS*, 2002.

[7] Maksim Tsvetovat and Katia Sycara. Customer coalitions in the electronic marketplace. In *Proceedings of Agents-2000*, pages 263–264, 2000.

[8] Michael van de Vijsel. Increasing realism in coalition formation for multi-agent systems. Master's thesis, University of Manitoba, Winnipeg, Canada, 2005.

[9] Junichi Yamamoto and Katia Sycara. A stable and efficient buyer coalition formation scheme for e-marketplaces. In *Proceedings of Agents-2001*, May 2001.