Selecting partners

Bikramjit Banerjee and Sandip Sen

Department of Mathematical & Computer Sciences
University of Tulsa
bikram@euler.mcs.utulsa.edu, sandip@kolkata.mcs.utulsa.edu,
URL:http://www.mcs.utulsa.edu/~sandip/

Abstract. The goal of a rational agent is to maximize utility. We consider situations where a rational agent has to choose one of several contenders to enter into a partnership. We assume that the agent has a model of the likelihood of different outcomes and corresponding utilities for each such partnership. Given a fixed, finite number of interactions, the problem is to choose a particular partner to interact with where the goal is to maximize the sum of utilities received from all the interactions. We develop a multinomial distribution based mechanism for partner selection and contrast its performance with other well-known approaches which provide exact solution to this problem for infinite interactions.

1 Introduction

In an open environment, agents have to be extremely cautious about which other agents to interact or partner with. The goal of a self-interested agent would be to interact with or enter into partnership with those agents that will produce maximal local utility for this agent [Tes98]. Obviously, such an agent will also have to achieve its local goals effectively to maximize its utility. For discussions in this paper, however, we will concentrate exclusively on utilities received by interacting with other agents.

We assume that each agent interaction will ultimately generate some utility for each of the interacting agents. From a single agent X's point of view, interaction with a collection of agents can be thought of as entering a partnership (also called coalitions in game theory). We are only interested in the utility received by the agent X by interacting with a coalition. We will not concern ourselves with issues of how the coalition generates the revenue and the process by which the generated revenue is distributed among the partners.

We assume that an agent can get one of several payoffs or utilities for joining a particular coalition, and that there is a static probability distribution that governs which of the payoffs is received at any particular interaction. Our usage of the term "interaction" needs further clarification: by one interaction of an agent X with a coalition Y we refer to the entire process of X joining the coalition, the coalition generating some revenue R, and X receiving some share of that revenue r_X^Y which, as mentioned above, is determined by a probability distribution. This probability distribution of different utilities that X can receive from coalition Y will be referred to as the *payoff-structure of Y from X's viewpoint*, P_X^Y. In a later section, we will discuss a representative scenario that justifies our assumptions.

Some combination of *a priori* domain models, observation or experience-based learning schemes as well as word-of-mouth transmissions can be used to arrive at these payoff-structures. In this paper, however, we will not address the issue of how these payoff-structures are generated. Rather, we will focus on how to select a coalition given the payoff-structures for each of the coalitions an agent can interact with. More precisely, we will consider choosing a single coalition to interact with repeatedly when the total number of interactions are known ahead of time. We believe, and our theoretical as well as experimental results will show, that the particular coalition chosen should vary depending on the set of payoff- structures for all the coalitions under consideration, and the total number of interactions. Our goal in this paper is to design such a coalition or partner selection mechanism that identifies the most beneficial partnership given a finite number of interactions that the selecting agent is going to partake in.

A sample problem: Consider a situation where an agent has to select between two partners for the next n interactions. An interaction with one of these partners, say A, yields returns of 10 and 20 units with probabilities 0.7 and 0.3 respectively. On the other hand, each interaction with the other partner, B, yields payoffs of 1000 and 4 units with probabilities 0.01 and 0.99. The expected utility from B is greater than that from A but if the agent takes into consideration the finite number of interactions, n, that it intends then A may be more attractive than B, since it returns higher rewards in both cases than the one that B returns almost certainly. Our contention is that to make such decisions, an agent must take into account the *number of intended interactions*, and when that is done the consequent decision may (depending on the payoff-structure as in this example) be quite unlike what risk-aversion dictates.

2 Coalition Formation

There has been a significant amount of work in game theory and multiagent systems field in the area of coalition formation. Representative work include:

Searching for optimal coalition structure: How should a group of agents be partitioned into subgroups or coalitions such that the sum of the revenue generated by all such coalitions is maximized [SLA+99]?

Decision mechanisms for forming coalitions: How should agents decide on which coalition to join [Ket94,SL97,SK95,SK96,ZR94]?

Payoff division in a coalition: How should the revenue generated in a coalition be divided among the partners [Ket94,LR57]?

Almost all of this body of research ignores, as we do, the issue of how the coalitions generate their revenues or the nature of problem solving adopted by individual agents after they form a coalition.

We address a slightly different problem of coalition selection under uncertainty. Perhaps Ketchpel's work [Ket94] is most related. That work addresses a process for negotiating who becomes the leader versus the member in a group based on how much one is willing to pay others to join a group. We, on the other hand, do not concern ourselves with our agent X building a coalition. Rather, X is selecting which coalition to join given only summary information of what payoff it can expect from that coalition. The payoff-structure encoding in the form of a probability distribution over possible payoffs for joining a coalition is the summary information on which X must base its decision. Another important distinction and the key focus of this work is the prior determination of how many interactions an agent is going to have.

Our basic hypothesis is that the most beneficial partnership to interact with can change based on how many interactions we will have. For example, consider a simple scenario of choosing from the two following partnerships. One partnership always give a steady return whereas the other offers a small return most of the times and very infrequently returns an astronomical amount. If one is only going to interact for a few times, it might be prudent to interact with the first partnership, but if one is going to have a prolonged sequence of interactions, perhaps it is worthwhile to choose the second partnership expecting that the "jackpot" is more likely to be hit at least once and will more than compensate for the smaller returns in the other interactions. Our goal is to go beyond this informal heuristic and provide a formal and well-founded decision procedure based on probability theory for this partnership selection problem.

3 Payoff-structures of partnerships

We now present a formal representation for *payoff-structure* of a partnership or coalition for an agent X outside the coalition. It can be represented as an n-tuple - $\{(u_1, p_1)(u_2, p_2) \ldots (u_n, p_n)\}$ where p_i stands for the probability that the interaction with this partnership will yield a payoff $u_i, \quad \forall i = 1 \ldots n$. X will have such a tuple for each of the other partnerships it can interact with. We note that n is not a constant but varies from one partnership to another, i.e., the number of possible payoffs depends on the particular partnership X is interacting with.

4 Selecting the potentially most beneficial partnership for limited interactions

We are interested in devising a procedure that, given a pool of possible partnerships and the number of interactions (N) that the agent wants to make, will select a partnership that is most likely to return a maximum total utility over these N interactions. To this end, we make pairwise comparisons among the partnerships and select the most profitable one. Let p_{ij}^N be the probability that partnership i returns greater utility than j. If $p_{ij}^N > p_{ji}^N$ then we regard partnership i to be more beneficial than partnership j. Hence, we define the objective function for selection between agents i and j, as

$$S(i,j) = \begin{cases} i \text{ if } p_{ij}^N > p_{ji}^N \\ j \text{ otherwise} \end{cases}$$

We note here, that $S(i,j)$ is *not transitive* (i.e. if $S(i,j) = i$ and $S(j,k) = j$ does not imply that $S(i,k) = i$). We will illustrate this non-transitivity with an example scenario in a later section.

Our proposed decision mechanism first finds $S(i,j)$ values for all possible partnership pairs (i,j), and chooses the partnership that is returned for the maximum number of times (i.e. the partnership with the maximum number of wins against other partnerships). This mechanism trivially selects a partnership if it wins against all other partnerships. If there is a tie among a subset of these partnerships then we can use arbitration mechanism from the voting theory [Str80], with p_{ij}^N as a relative measure of mandate between i and j, for each pair (i,j) in that subset. For example, assume that there is a tie among partnerships i, j, and k. Then the arbitration mechanism gives $p_{ij}^N + p_{ik}^N$ votes to partnership i, $p_{ji}^N + p_{jk}^N$ votes to partnership j, and $p_{ki}^N + p_{kj}^N$ votes to partnership k. If there is still a tie, we choose one of these partnerships randomly.

Let the payoff-structures of partnerships i and j be given by sequences $\{(u_k^i, p_k^i)\}$ and $\{(u_k^j, p_k^j)\}$, and their lengths be n_i and n_j respectively. We can divide N into n parts, where each part is 0 or a positive integer. The kth part in such a division represent the number of times the kth utility value was returned when interacting with a given partnership. Let any such decomposition be represented as C_n^N. There are $\begin{pmatrix} N+n-1 \\ n-1 \end{pmatrix}$ such distinct decompositions. If we regard such a decomposition of N interactions as actually arising in exchanges with partnership i, then the utility received as a result, is given by

$$U_i^{C_{n_i}^N} = \sum_{k=1}^{n_i} C_{n_i}^N(k) * u_k^i$$

where $C_{n_i}^N(k)$ is the kth component of $C_{n_i}^N$. Then the value of p_{ij}^N is computed according to the rule for conditional probability, as

$$p_{ij}^N = \sum_{C_{n_i}^N} \mathbf{Pr}[C_{n_i}^N].\mathbf{Pr}[U_i^{C_{n_i}^N} > U_j | C_{n_i}^N]$$

where $\mathbf{Pr}[U_i^{C_{n_i}^N} > U_j | C_{n_i}^N]$ (i.e. the probability that i returns greater net-utility, given a particular decomposition of its N interactions into n_i parts), is, in turn, computed as

$$\mathbf{Pr}[U_i^{C_{n_i}^N} > U_j | C_{n_i}^N] = \sum_{C_{n_j}^N} \mathbf{Pr}[C_{n_j}^N].h(C_{n_i}^N, C_{n_j}^N)$$

with $h(C_{n_i}^N, C_{n_j}^N)$ being a boolean function to decide whether or not i returns greater utility, given a pair of decompositions of N interactions into n_i and n_j parts respectively, and it is defined as

$$h(C^N_{n_i}, C^N_{n_j}) = \begin{cases} 1 \text{ if } U_i^{C^N_{n_i}} > U_j^{C^N_{n_j}} \\ 0 \text{ otherwise.} \end{cases}$$

Now, suppose $C^N_{n_i} = \langle x_1, x_2, x_3 \ldots x_{n_i} \rangle$ with $x_1 + x_2 + x_3 + \ldots x_{n_i} = N$. Then we compute $\mathbf{Pr}[C^N_{n_i}]$ as

$$\mathbf{Pr}[C^N_{n_i}] = \frac{N!}{x_1!x_2!x_3!\ldots x_{n_i}!}(p^i_1)^{x_1}(p^i_2)^{x_2}\ldots(p^i_{n_i})^{x_{n_i}}$$

from the multinomial distribution. It can be shown that the computational complexity of this scheme for selecting a partner from among A subjects, is $O(A^2N^{2n})$ where $n = Max_i\{n_i\}$. We also note that in general, $p^N_{ij} + p^N_{ji} \leq 1$ as there may be some $C^N_{n_i}, C^N_{n_j}$ for which $U_i^{C^N_{n_i}} = U_j^{C^N_{n_j}}$.

An example evaluation: Suppose the payoff-structures of three partnerships i, j and k are $\{(1,\ 0.4)\ (10,\ 0.6)\}$, $\{(5,\ 0.5)\ (5,\ 0.5)\}$, and $\{(11,\ 0.4)\ (3,\ 0.6)\}$. Here $n_i = n_j = n_k = 2$. We choose N=1. For each partnership there are 2 decompositions of 1 trial into two parts, viz. $\langle 0, 1\rangle$ and $\langle 1, 0\rangle$. To compute p^N_{ij}, we see that case $\langle 0, 1\rangle$ for i produces benefit = 10. This is greater than the benefits for both cases of j (benefit 5 from both $\langle 0, 1\rangle$ and $\langle 1, 0\rangle$). Hence $h(\langle 0,1\rangle_i, \langle 0,1\rangle_j) = h(\langle 0,1\rangle_i, \langle 1,0\rangle_j) = 1$. Again for case $\langle 1, 0\rangle$ of partnership i the benefit is 1, which is lesser than that of j for both of its cases. That means, $h(\langle 1,0\rangle_i, \langle 0,1\rangle_j) = h(\langle 1,0\rangle_i, \langle 1,0\rangle_j) = 0$.

Hence

$$\begin{aligned} p^N_{ij} &= \mathbf{Pr}[\langle 0,1\rangle_i] * \{\mathbf{Pr}[\langle 0,1\rangle_j] * 1 + \mathbf{Pr}[\langle 1,0\rangle_j] * 1\} \\ &+\mathbf{Pr}[\langle 1,0\rangle_i] * \{\mathbf{Pr}[\langle 0,1\rangle_j] * 0 + \mathbf{Pr}[\langle 1,0\rangle_j] * 0\} \end{aligned}$$

where subscript of any decomposition indicates the respective partnership with which this case arises. Now,

$$\mathbf{Pr}[\langle 0,1\rangle_i] = \frac{1!}{0!1!}(0.4)^0 * (0.6)^1 = 0.6$$

$$\mathbf{Pr}[\langle 1,0\rangle_i] = \frac{1!}{1!0!}(0.4)^1 * (0.6)^0 = 0.4$$

It is also immediately seen that

$$\mathbf{Pr}[\langle 0,1\rangle_j] = \frac{1!}{0!1!}(0.5)^0 * (0.5)^1 = 0.5$$

$$\mathbf{Pr}[\langle 1,0\rangle_j] = \frac{1!}{1!0!}(0.5)^1 * (0.5)^0 = 0.5$$

Consequently, $p_{ij}^N = 0.6 * (0.5 + 0.5) + 0.4 * (0 + 0) = 0.6$. Similarly, we can compute $p_{ji}^N = 0.4$, $p_{jk}^N = 0.6$, $p_{kj}^N = 0.4$, $p_{ik}^N = 0.36$, $p_{ki}^N = 0.64$. As a result,

$$S(i,j) = i,$$
$$S(j,k) = j,$$
$$S(i,k) = k.$$

As outlined earlier, at this situation (a non-transitive case) the arbitration by voting comes into the picture, and i gets 0.96 votes, j gets 1.0 votes, and k gets 1.04 votes. Consequently, k is chosen.

It is worthwhile to explore this example further so that we can appreciate the notion that non-transitivity is the exception rather than the rule. For $N = 2$, we have $p_{ij}^N = 0.84$, $p_{ji}^N = 0.16$, $p_{jk}^N = 0.36$, $p_{kj}^N = 0.64$, $p_{ik}^N = 0.4752$, $p_{ki}^N = 0.5248$. Hence partnership k is the obvious preferred choice, and transitivity holds. Again, for $N = 3$, the direct choice is i.

We have come up with this elaborate procedure after exploring various other alternatives that were computationally simpler, but inadequate nevertheless. The closest of these is an approximation for p_{ij}^N for all pairs (i,j), using Hoeffding Inequality [Vid97]. This approximation is really a crude one, and the problem called for better bounds. The direct application of Chernoff's Theorem [Bil86] provides tighter bounds, but requires "large number of interactions" which defeats the very purpose of the problem. Lastly, the mechanism we have presented, computes p_{ij}^N accurately, using the rules for multinomial distribution of probability values, and conditional probability.

5 Comparative evaluation with a decision mechanism for infinite interactions

To compare our decision mechanism with a standard reference we chose the *Expected Utility Maximization Principle (MEU)* [LR57]. The MEU principle prescribes interacting with partnership i given by

$$i = \arg \max_{j \in Partners} \sum_{k=1}^{n_k} u_k^j * p_k^j,$$

where $Partners$ is the set of partners the agent can interact with.

This principle is guaranteed to maximize the total utility received by the agent if the agent interacts infinitely often. In an open environment, agent relationships are often ephemeral, and *infinite interactions* are impractical. The obvious question is whether our strategy will be able to outperform a MEU choice when the assumption of infinite interactions do not hold. In particular, if we know that an agent is interested in a relationship for a *finite*, short period, a partnership with smaller expected utility may return more net utility than another partnership with smaller expected utility. It would be interesting to evaluate if this conjecture is true and if so, for what range of interactions?

At this point we can also observe that the decision mechanism developed by us have the following properties:

- The strategy reduces to the MEU strategy for infinite interactions.
- It is based solely on the *payoff-structure* of the partnerships and the number of interactions to be performed.

6 Evaluation scenario

Decisions of this kind attain significant proportions in any domain where limited application of acquired knowledge is required with a fair degree of confidence in the outcomes. We can visualize a computational marketplace, where agents distribute computational tasks among service-providers (in return of payoffs, that may be determined by the time taken, quality of the service etc.), through broker agents. Such a broker has estimates of the payoff-structures for various service-providers, and is faced with the problem of choosing the potentially-most-beneficial recipient (the benefit being the portion of the payoff of the recipient, that the broker charges as the intermediary), for a set of tasks (the number of such tasks can be a measure of N). The broker and the recipient thus enter a partnership or collaboration which the broker offers to maintain for sufficient tasks (or for a sufficient period) that he is confident enough, will produce the desired payoff.

The procedure outlined in this paper enables the broker to objectively evaluate the potential of various prospective recipients, as an explicit function of the intended number of interactions, and this is where the procedure gains its capability to suggest an alternative to the MEU-choice.

7 Experimental results

For the purpose of experimentation, we have considered only two partnerships. Of these, the payoff-structure of the MEU-partnership is $\mathcal{M}$, and that of the other partnership (non-MEU) is $\mathcal{NM}$. We can imagine, there are other partnerships in the marketplace, and the agent's choices may vary across several of these partnerships. However, for the purpose of illustration of non-MEU choices against a given MEU-partnership, only one non-MEU partnership is sufficient.

In the experiments, the payoff structure of the $\mathcal{NM}$ partnership is $\{(2, 0.1)$ $(10, 0.3)$ $(12, 0.6)\}$ with an expected utility of 10.4, and we vary the payoff-structure of the MEU-partnership, such that its maximum possible payoff decreases, keeping the expected utility fixed at 10.56. This means that according to the MEU strategy it is preferable to choose the second partnership. Figures 1 and 2 shows the comparative performance of the Non-MEU and MEU partnerships. For a given number of interactions, the actual payoff generation is simulated, and the percentage of cases out of 1000 simulations, in which the Non-MEU partnership yields greater payoff than the other is plotted in each graph. The shaded regions in a plot refer to the values of N for which, given the

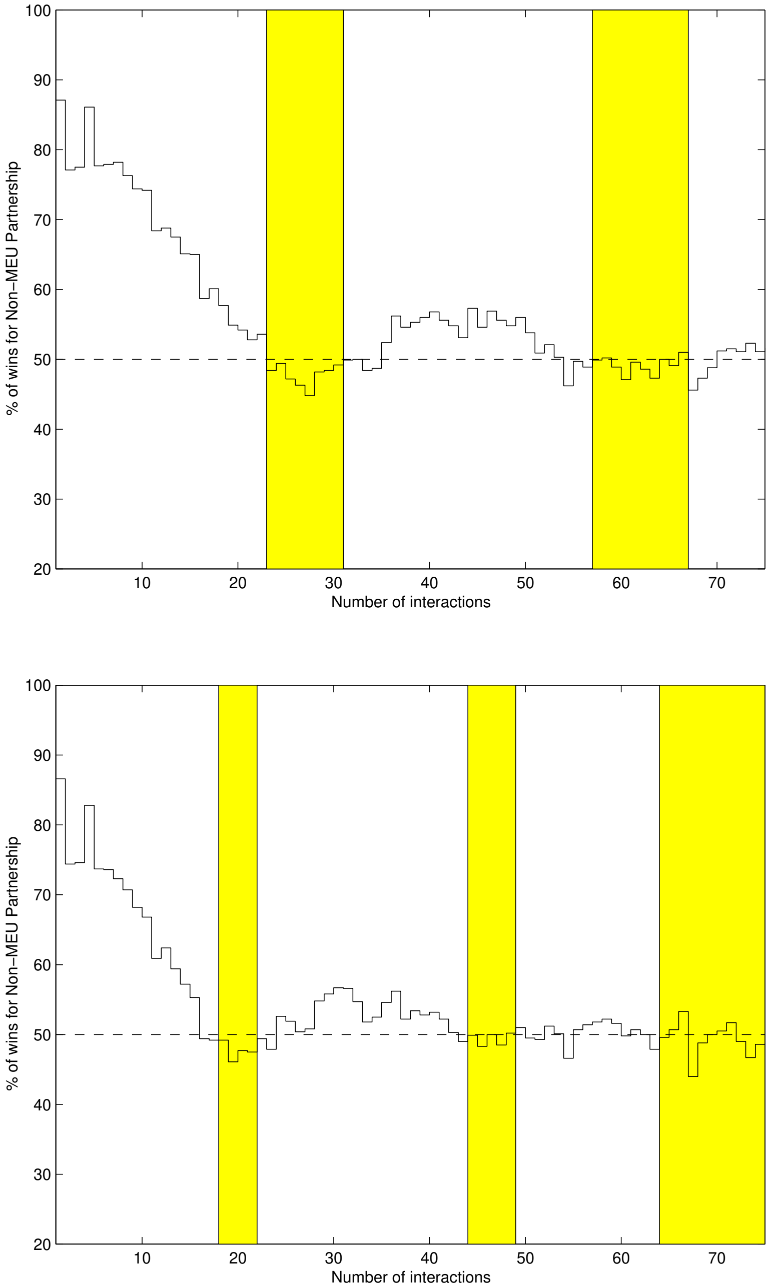


Fig. 1. Simulation of benefit from $\mathcal{NM} = \{(2,\ 0.1)\ (10,\ 0.3)\ (12,\ 0.6)\}$ against $\mathcal{M} = \{(8,\ 0.97)\ (90,\ 0.02)\ (100,\ 0.01)\}$(left) and $\mathcal{M} = \{(8,\ 0.96)\ (70,\ 0.025)\ (75,\ 0.015)\}$(right)

payoff-structure, the function S selects the MEU-partnership. For all other values of N (i.e. the unshaded zones), the choice is the Non-MEU partnership. The values below the dotted-line signify cases where the MEU-partnership outperforms the other in actual simulation. The first 75 interactions have been plotted in each case.

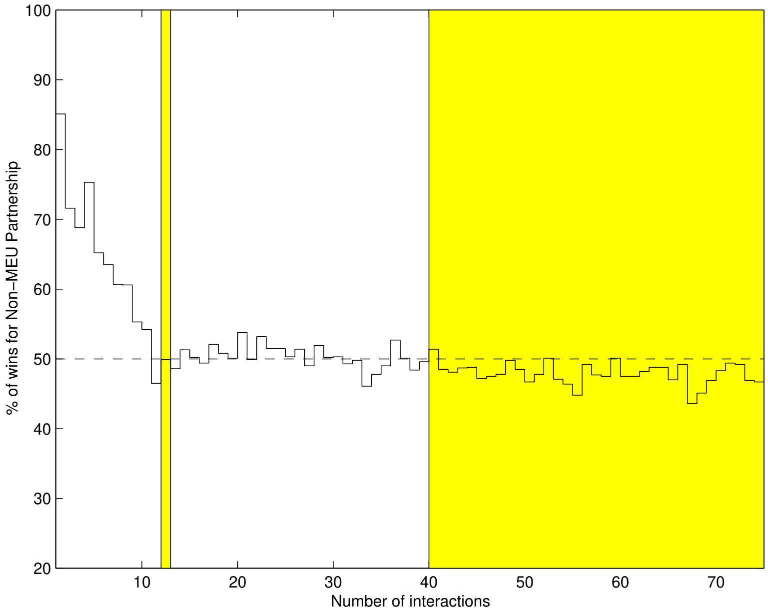
Figures 1 and 2 show close agreement between the procedural choice with the results of the simulations. An interesting observation from these figures is that the choice function S, contrary to intuition, changes its output more than once. The MEU-partnership appears in multiple distinct regions in the figures 1 and 2. Such change in choice depends on the probability distributions over the utilities, the multinomial coefficients, and the decomposition-pattern of the number of interactions. We did not find any regular pattern in these variations to summarize the choice with a simple heuristic. The summary description we can provide is that the "bands of dominance" of the MEU strategy increases in width with more interactions until after a relatively large number of interactions it becomes totally dominant. The last observation is consistent with the fact that for large interactions, the MEU strategy is accurate and sufficient to identify the most desirable partnership. We also note that the sample payoff-structures assumed in the experiments are representative of the type of payoff-structures where the efficacy of our procedure is clearly demonstrable. In other types of payoff-structures the results of our procedure cannot be any worse than the MEU-strategy.

It can be observed that as the highest payoff from the MEU partnership increases (and the corresponding probability of receiving that payoff is decreased to keep the expected utility constant), it takes more interactions before the first onset of dominance of the MEU partnership. This can be explained by the fact that if the expected utility of a partnership is based largely on the superlative payoff from an infrequent event (the jackpot in our earlier example), it would take more interactions to benefit from the occurrence of that infrequent event. For a more limited number of interactions, it is advisable to choose a partnership that returns a consistently high payoff even though its expected utility is less.

Another interesting observation from figure 2(right) is the unusually frequent variation of the procedural-choice in a short range of N. Because the difference in the expected utilities of $\mathcal{M}$ and $\mathcal{NM}$ is already low, the lowering of the highest utility in $\mathcal{M}$ reduces the structural difference between $\mathcal{M}$ and $\mathcal{NM}$. As a result, the choice becomes highly sensitive to N. This sensitivity, in particular, is beyond the scope of MEU-strategy.

8 Conclusion

In this paper, we have considered the problem of an agent deciding on which partnership to interact with given the number of interactions and a model of the environment in the form of payoff structures of each of these partnerships. We have developed a probability-theory based procedure for making the selection



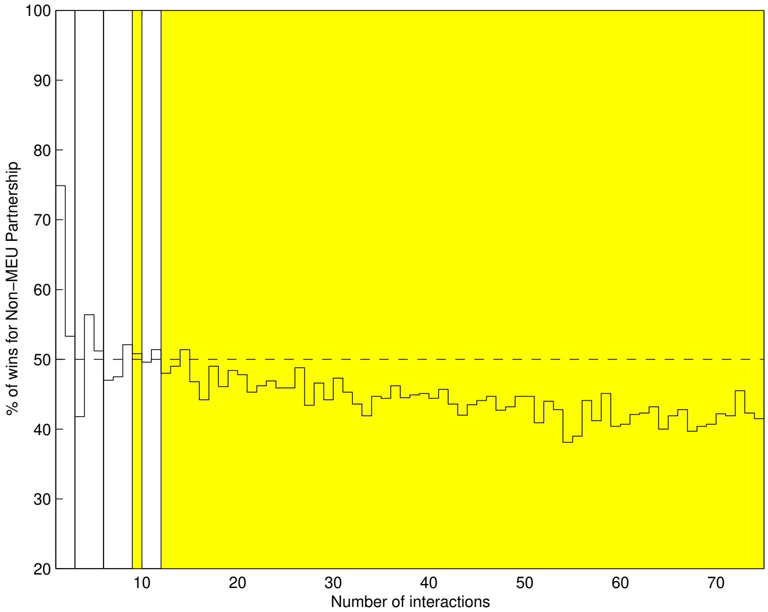


Fig. 2. Simulation of benefit from $\mathcal{NM} = \{(2,\ 0.1)\ (10,\ 0.3)\ (12,\ 0.6)\}$ against $\mathcal{M} = \{(8,\ 0.94)\ (50,\ 0.04)\ (52,\ 0.02)\}$(left) and $\mathcal{M} = \{(8,\ 0.8)\ (20,\ 0.1)\ (21.6,\ 0.1)\}$(right)

and compared its performance with the well known MEU principle which solves this problem exactly when the number of interactions is infinite.

The agreement between the performance predicted by our selection procedure and the simulated values show that the procedure is well adapted for the sensitive differences between the payoff-structures of two coalitions. This procedure effectively captures a broad range of these subtleties and hence, is a more fine-grained measure of the relative efficacy of coalitions, than the MEU-strategy.

A related and practically important problem is to devise a strategy for interaction where the payoff-structure for coalitions is incompletely known. It would be instructive to compare the performance of Bayesian update schemes versus model-free reinforcement learning methods on these problems [Mit97]. This, admittedly more complex, problem is akin to the multi-arm bandit problem [SU95]. A critical issue in this problem is that since learning is on-line, the approximation of the payoff structure from limited sampling has to be accurate to take advantage of any non-MEU strategies. Off-line learning of the payoff structure of partnerships by observing them interacting with other agents can also provide approximate models. In the latter cases, the present work can be readily used.

References

[Bil86] Patrick Billingsley. *Probability and Measure*. John Wiley & Sons, 1986.

[Ket94] Steven Ketchpel. Forming coalitions in the face of uncertain rewards. In *Twelfth National Conference on Artificial Intelligence*, pages 414–419. MIT Press/AAAI Press, 1994.

[LR57] R. Duncan Luce and Howard Raiffa. *Games and Decisions: Introduction and Critical Survey*. Dover, New York, NY, 1957.

[Mit97] T.M. Mitchell. *Machine Learning*. WCB McGraw-Hill, Boston, MA, 1997.

[SK95] Onn Shehory and Sarit Kraus. Task allocation via coalition formation among autonomous agents. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 655–661, August 1995.

[SK96] Onn Shehory and Sarit Kraus. A kernel-oriented model for coalition-formation in general environments: Implementation and results. In *Proceedings of the Thirteent National Conference on Artificial Intelligence*, pages 133–140, August 1996.

[SL97] Tuomas Sandholm and Victor Lesser. Coalitions among computationally bounded agents. *Artificial Intelligence*, 94(1):99–137, 1997.

[SLA+99] Tuomas W. Sandholm, Kate Larson, Martin Andersson, Onn Shehory, and Fernando Tohme. Coalition structure generation with worst case guarantees. *Artificial Intelligence*, 111(1-2):209–238, 1999.

[Str80] Philip D. Jr. Straffin. *Topics in the theory of voting*. The UMAP expository monograph series. Birkhauser, Boston, MA, 1980.

[SU95] Marcos Salganicoff and Lyle H. Ungar. Active exploration and learning in real-valued spaces using multi-armed bandit allocation indices. In *Proceedings of the Twelfth International Conference on Machine Learning*, pages 480–487, 1995.

[Tes98] L. Tesfatsion. *Preferential Partner Selection in Evolutionary Labor Markets: A Study in Agent-Based Computational Economics*. Springer-Verlag, Berlin, 1998.

[Vid97] Mathukumalli Vidyasagar. *A theory of Learning and Generalization*. Springer, 1997.

[ZR94] Gilad Zlotkin and Jeffrey S. Rosenschein. Coalition, cryptography, and stability: Mechanisms for coalition formation in task oriented domains. In *Twelfth National Conference on Artificial Intelligence*, pages 432–437. MIT Press/AAAI Press, 1994.