Abstract—Modern multi-organisational coalitions are capable of bringing diverse sets of capabilities, assets and information sources to bear on complex and dynamic operations. However, the successful completion of these operations places demands on the trust between coalition partners. When it is necessary to rely on other partners, decision-makers must be able to make rapid and effective trust assessments and decisions. We propose a decision-theoretic model which allows controls to be used, as well as trust, to increase confidence in initial interactions. We consider explicit incentives, monitoring and reputation as examples of such controls.

I. INTRODUCTION

Modern coalition operations require multiple organisations to act in a coordinated manner in order to achieve some high level goal. Such operations therefore require trust at multiple levels. In military coalitions for example, one unit trusts that others will be in position at the right time to provide support, while commanders must trust that their subordinates will follow their instructions. In computational multi-agent systems, trust has long been recognised as a vital concept in open multi-agent systems (MAS), where agents may be self-interested, diverse and deceptive [1]. By relying on others, agents place their own interests at risk, which introduces the need for trust [2].

Since the success of a coalition can hinge on its ability to effectively delegate tasks and share information and assets, there exists a need for trust-aware decision support tools which explicitly factor trust into the decision making process. In this paper, we seek to address aspects of the decision-making problem in contexts which are highly dynamic. Such systems are characterised by a high rate of population turnover, where the agent population frequently and rapidly changes, and high degrees of social change, where the social structures among agents frequently change.

It can be difficult to form stable trust relationships using existing techniques (e.g. [3], [4]). In dynamic systems, agents may only participate for a short time before leaving, and so it may be impossible to gather sufficient evidence to evaluate the trustworthiness of partners, or third-party opinion providers. Subsequently, agents may be unable to make effective decisions about whether to interact, and with whom. When agents are sensitive to risk, and can choose not to interact, systems can become ‘paralysed’; without interactions, evidence cannot be gathered, trust cannot be formed, and interactions may not take place.

In order to overcome this problem, organisations may make use of controls [5] which permit interaction when trust is low, providing initial evidence from which to bootstrap trust evaluations. In this paper, we investigate ways in which agents may use controls to mitigate the perceived risk in initial interactions. We consider three kinds of controls which can be used in the cases where trust is insufficient, outlined in the following section.

II. DELEGATION STRATEGIES

Given that we assume a probabilistic trust evaluation model, we take a decision-theoretic approach to the problem of partner selection in situations involving risk. Trustees have a choice of two behaviours (effort levels) which entail different probabilities of task success. The trustor’s goal is to determine the best action to take for the decision. We allow trustors to consider a number of alternative delegation strategies, including abstaining from delegating altogether:

1) Simple delegation: the trustor delegates the task without considering the trustee’s effort choice.

2) Delegate with monitoring: the trustor invokes the trustee to adopt the desired effort level, and pays the monitoring cost to observe which effort levels were selected by trustees, so that appropriate incentives may be offered conditional on observed effort.

3) Delegate without monitoring: the trustor invokes the trustee to adopt the desired effort level using incentives, but does not monitor; incentives are conditional on the task outcome (which is a function of the effort level chosen).

4) Delegate with reputational incentive: trustors produce contracts which attempt to influence the trustee’s effort level choice by including the loss or gain the trustee would suffer from a change in reputation as an explicit incentive.

5) Abstain from delegation: the trustor abstains from delegation entirely, if this is deemed the optimal course.

Delegation proceeds as follows. The trustor $x$ observes the environment and obtains sets of potential candidates $Y$ and recommenders $R$. Next, the trustor evaluates the candidates using his trust evaluation model, producing a set of opinions (we may assume these to be probabilistic estimates of task success). For each candidate, the trustor computes the best contract from each of the alternative delegation strategies (i.e. unmonitored, monitored, etc.). The candidate $y$, whose best contract maximises the trustor’s EU, is selected. The trustor $x$ proposes the contract to $y$, who can then accept or refuse. If $y$ accepts, $x$ observes and subjectively evaluates the outcome.
and may, by monitoring, observe the effort level choice of $y$. The contract is enacted, transferring rewards and penalties between the agents as specified. If $y$ refuses, $x$ can seek another candidate from $Y$. Finally, the trust model is updated to reflect the new observation. We refer readers to the full version of this paper [6] for more details regarding this model.

### III. Evaluation

We evaluate our approach within a simulated multi-agent environment. In each time step, trustees must select the most appropriate partner based on the level of trust, and the expected utility of the most desirable contract with that partner. The model was evaluated under the following conditions:

- **Simple**: only unconditional trust evaluations are used.
- **Unmonitored**: trustee choice is considered, using conditional trust, but monitoring is not allowed.
- **Unmonitored + Monitored**: both unmonitored and monitored strategies are available.
- **Unmonitored + Monitored + RI**: unmonitored, monitored and reputational incentive (RI) strategies are available.

Figure 1 shows the performance of our model in each of the four cases above, measuring the average change in utility experienced by the trustees in the society in each interaction. The Simple strategy obtained an average utility gain of 1.8 over the course of the experiment. In the Unmonitored case, the model performs very poorly, obtaining an average utility of around 0. This is to be expected, as this strategy has no way of obtaining conditional evidence, and so cannot build trust. In the Monitored mode, the model performs much better, rising to around 2.4 by 500 interactions. The Reputational Incentive mode outperforms the others, achieving around 3.6. The results in these cases were found to be statistically significant using ANOVA testing at the $p < 0.001$ level ($F(2,1494) = 11784, p = 2.2 \cdot 10^{-16}$). Post-hoc comparison using a Tukey HSD test indicated that the means in all cases were significantly different from each other. This supports our hypothesis that monitoring and reputational incentives can increase the decision-making performance.

Figure 2 shows the number of trustees monitoring in each interaction step. Monitoring is initially preferred, as agents learn about the conditional behaviours of their partners. Over time, the benefit of monitoring decreases, so agents begin to favour unmonitored delegation forms, using reputational incentive. Monitoring can never be completely avoided, however, as the population of agents is continually changing.

### IV. Conclusions

Trust alone is necessary but not sufficient when making delegation decisions in the presence of uncertainty. In the presence of risk and uncertainty, trust is required to facilitate interactions, but interactions are required to obtain evidence from which trust can be built. In highly dynamic coalitions, building trust is even more difficult. By employing controls in addition to trust, trustees can mitigate some of the perceived risk in their interactions, and be motivated to delegate, providing crucial initial interactions required to bootstrap trust.

### ACKNOWLEDGEMENTS

This is a summary of the IJCAI 2011 paper [6] prepared for the internal International Technology Alliance research project meeting. Please refer to the IJCAI 2011 full version for details.

This research was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the author(s) and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

### REFERENCES