Agent Support for Policy-Driven Mission Planning Under Constraints *

Murat Şensoy and **Daniele Masato** and **Timothy J. Norman**
Martin Kollingbaum and **Chris Burnett**
Computing Science, University of Aberdeen, AB24 3UE, Scotland, UK

Katia Sycara and **Jean Oh**
Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Abstract

Forming *ad-hoc* coalitions between military forces and humanitarian organizations is crucial in mission-critical scenarios. Very often coalition parties need to operate according to planning constraints and regulations, or policies. Therefore, they find themselves not only in need to consider their own goals, but also to support coalition partners to the extent allowed by such regulations. In time-stressed conditions, this is a challenging and cognition-intensive task. In this paper, we present intelligent agents that support human planners and ease their cognitive burden by detecting and giving advice about the violation of policies and constraints. Through a series of experiments conducted with human subjects, we compare and contrast the agents' performance on a number of metrics in three conditions: agent support, transparent policy enforcement, and neither support nor enforcement.

Introduction

International coalitions between military forces and humanitarian organizations are increasingly becoming a necessity to respond effectively to mission-critical tasks, for example peace-keeping operations, relief missions, and so on. Although coalitions are formed so that members can cooperate and benefit from each other's skills and capabilities, they may be formed rapidly and without much previous co-training. In addition, because of their different nature and nationalities, coalition members are required to adhere to policies of differing provenance. Some policies may be private to a specific coalition party, hence they cannot be disclosed to the partners, whereas others may be public. As a consequence, policies may very well conflict and lead to situations where some actions may be forbidden and permitted/obliged at the same time[1] (Elagh 2000). Self-interested behaviour should also be expected from the coalition partners. Indeed, although members need to engage in collaborative planning to successfully complete assigned tasks, they often pursue individual goals that can significantly affect the outcome of the joint effort, for example in terms of cost sustained to complete the mission, or success in the achievement of individual and joint goals. Policies and self-interest place a considerable cognitive burden on human planners and coordinating personnel involved in the decision-making process for the scenarios mentioned above. In this paper, we expand on previous work (Burnett et al. 2008; Kollingbaum et al. 2009; Sycara et al. 2009) on agent support to ease such cognitive burden by conducting additional experiments and presenting a detailed analysis of our results.

The rest of this paper is organised as follows: we first define the concept of policy and the types of policies employed in our experiments; we then present the reasoning mechanism and the design considerations according to which the agents have been implemented; finally, we present the experiment methodology and a detailed analysis of the results.

Policies

Our definition of policy is based on previous work in the area of normative systems and norm-governed agency (Dignum 1999; Kollingbaum 2005; Lopez y Lopez, Luck, and d'Inverno 2004). We specify an obligation, permission or prohibition on a particular action with two conditions – an activation condition and an expiration/fulfilment condition – determining whether a policy is relevant to the human planner. Let us define the set $Expr$ as the set of all possible well-formed formulae consisting of first-order predicates over terms (constants, variables and the operators $\wedge$, $\vee$ and $\neg$). A policy can then be defined as in Definition 1. Policies regulate the actions of coalition partners under specific circumstances. In particular, they specify the obligations that have to be fulfilled, the prohibitions that constrain/forbid particular actions, and the permissions that define the range of actions that are allowed.

Definition 1 A policy, expressing an obligation, permission or prohibition is a tuple $\langle v, \rho, \varphi, a, e \rangle$ where:

- $v \in \{O, P, F\}$ is a label indicating whether this is an obligation, permission or prohibition.
- ρ is a role identifier for the norm addressee.

* **Acknowledgement** 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.

[1] This case is indeed covered by our experimental scenario.

IF:	you have acquired clearance from party B to use route R on day D
AND:	you have intelligence that there is no danger on route R on day D
THEN:	you are permitted to deploy ground vehicles on route R on day D

```
(defrule A-P1 "Party A - Policy 1"
  (action-message
     (sender        PartyB)
     (message-id    BC-A1)
     (route         $?route)
     (day           ?day))
  (intelligence
     (intel-type    nodanger)
     (route         $?route)
     (day           ?day))
 =>
  (assert
     (response
        (modality       permission)
        (response-id    A-P1)
        (route          ?route)
        (day            ?day))))
```

Figure 1: Example policy and equivalent Jess rule

- φ describes the action regulated by this policy.
- $a \in Expr$ is the activation condition.
- $e \in Expr$ is the expiration/fulfilment condition.

The policy is activated for the entities playing role ρ whenever a hold and remains active until e holds. ■

We can categorize policies depending on their scope. Below we list some of these categories:

- **Resource policies:** we assume there is a set of resources available to the coalition and each resource is owned by a coalition partner. Hence, there may be coalition policies that represent coalition-wide resource sharing agreements, and individual policies that represent member-specific resource sharing constraints.
- **Information sharing policies:** these describe what information can, must or must not be shared. Coalition members may want to protect certain pieces of information, so to safeguard their own interests, and such cases would be codified by individual policies. These policies could refer to specific aspects of information such as its source.
- **Action policies:** coalition partners may have policies that oblige, permit or prohibit specific actions, and hence influence the planning activity.
- **Default policies:** these express the upfront "normative" position of a coalition partner (or the coalition itself, if there exists an agreement) and determine what default policy has to be assumed in case no explicit permission or prohibition is defined. The choice is usually between two possibilities: if something is not explicitly allowed, it is prohibited; if something is not explicitly prohibited, it is allowed. *We assume the latter holds in our experiments.*

Agent Support

Policies may impose restrictions on information sharing between coalition partners and on the actions that may be part of a plan. Similarly, planning constraints may force planners to comply with some restrictions. This can have a negative impact on the quality of planning outcomes. The presence of diverse policies among partners may also adversely impact the planning process. Finally, in some situations, planners may not recognize that violations have occurred.

We consider the use of agents to monitor communication and planning activities of coalition partners and reason about possible policy and plan constraint violations. An agent operates in a supportive role to human planners in order to ease their cognitive load during planning. The agent is assigned solely to a specific human planner and operates in a monitoring, controlling and/or advising capacity. Agents can aid planners by reporting policy violations, by informing the planners about the policies that have to be observed in particular situations and advising planners on alternative courses of action in order to act in a policy-abiding manner. Agents also analyse the current plan and inform human planners about the violation of planning constraints. The agent support was designed with four criteria in mind:

1. **Reasoning about policies:** the agent has to correctly assess if a planner complies with policies or violates them.
2. **Reasoning about planning constraints:** the agent has to determine which constraints are violated and report these to the planner.
3. **Level of agent visibility:** the agent design has to be balanced in terms of how the agent makes its presence known to the planner, its degree of pro-activity and reactivity to the user's actions, while minimising the possibility of irritating the planner.
4. **User dependence on the agent:** the agent has to be designed so that its intervention helps the planner learn and navigate its policy/goal space more effectively, rather than blindly relying on the assistance of the agent.

Aiding Strategies

In the development of these agents, we were particularly interested in experimentally comparing different agent aiding strategies. We compared the following two strategies:

- A **critic agent** that detects policy violations of coalition partners in the course of communication activities between them and during their planning. The agent (a) intercepts messages or (b) interrupts the planning of actions that violate policies in order to inform the sender about the set of policies violated. The sender can then decide whether to adhere to such an advice or to overrule the agent recording a reason for the overruling.
- A **censor agent** that interferes with the communication by deleting parts of the exchanged messages (or blocks them completely) that violate policies. In that case, the receiver is informed that a message is either truncated or completely censored.

The difference between the two types of agents is in their policy-related feedback to the human planner and their subsequent interaction. The critic agent, besides reasoning about policies, also monitors plan steps committed by a human planner and reasons about the effect of policies on planned actions. The censor agent, on the other hand, is not concerned with effects of policies on planned actions. It only intercepts and forbids the transmission of messages that contain policy violations.

Reasoning about Policies and Constraints

For agents to become operational, they must have access to plans and to communication activities. We use a traditional forward-chaining mechanism – the expert system shell Jess (Hill 2003) – to implement the policy and constraint reasoning as a set of rules, as illustrated in Figure 1.

Let us assume there are two organisational entities called "Party A" and "Party B", and let us assume we hold information about messages (e.g., a message of type "BC-A1" expressing that Party B granted clearance to use a particular route) and intelligence about safety (e.g., an attribute "inteltype" stating if that intelligence indicates danger). The translation of policies into rules of an expert system shell is then straightforward[2]. According to Definition 1, line 14 defines the type of policy v (permission), line 1 defines the role identifier ρ (A), lines 15–17 define the action φ, and lines 2–10 define the activation condition a (there is no expiration condition for this policy). The fact that Party B has sent a message that grants the use of a specific route on a given day, and the fact that Party A holds intelligence that there is no danger along the route on that day, leads the agent to generate a response expressing the permission for Party A to use the given route on that day. Effectively, this describes the basic reasoning cycle of the agent: (a) detecting the current situation changed by arriving messages expressing commitments for action of the coalition partner or revealed intelligence, as well as new planned actions, (b) reasoning about these changes with the policies encoded as rules, and (c) collecting the responses. In case a policy becomes relevant, a response is recorded and used to indicate a violation or the activation of an obligation for action. As policies expressing obligations may be fulfilled over time, it is also necessary to perform maintenance activities in order to remove those responses that become irrelevant to the current situation.

Human-Agent Experiments

In this section, we shall describe the nature of our experiments and how they were carried out, then we shall present a detailed analysis of our preliminary results.

Experimental Task

The experimental scenario is characterised by the interaction of two organisational entities, each of which has its own goals and is regulated by its own set of policies. Cooperation between the entities is required in order for them to achieve their goals and comply with the set of policies. The two partners are: a *humanitarian relief organization* with the individual goal of rescuing all injured civilians from a hostile region; a *military organization* tasked with the individual goal of defeating all hostile strongholds in the same area. Both parties will incur in costs for achieving their goals and should aim to minimise those as well.

The experimental situation requires two test subjects playing one of the two roles as a coalition partner. The player representing the humanitarian organization is regarded as "Party A", whereas "Party B" represents the military organization. Both players are provided with details about their private *goals*, *resources*, *intelligence*, their *capabilities*, *constraints*, and *policies*. Parties are given different maps (see Figure 2) that outline locations from where injured people have to be evacuated by Party A, or that represent insurgent strongholds that have to be defeated by Party B. These destinations have numerical requirements: for Party A, a specific number of wounded have to be evacuated; for Party B, strongholds have a specific resistance value that has to be overcome by military means.

Players can deploy *resources* (e.g., Jeeps, Helicopters) to fulfill their individual goals. Resources have a finite *capacity* for transporting wounded people or, in case of military hardware, a specific military *strength*. The players have to plan multiple deployments of their finite resources in order to achieve their individual goals. Each deployment incurs specific costs which sum up to the overall cost of a plan. Deployments are taking place along given routes (according to the maps provided) and at a specific time. Because both parties operate in the same area, they have to collaborate so that their plans are complementary. Moreover, there are dependencies between the parties' plans. For example, Party A will need military escort through dangerous areas, hence Party B will have to arrange its own plan so that it can provide such a service as well as achieve its individual goals. Both parties have certain *intelligence* about the tactical situation in the field. Depending on a party's policies, part of this intelligence may be disclosed to the coalition partner.

Each player is governed by a set of policies inspired by international guidelines for humanitarian/military cooperation (Wheller and Harmer 2006). In order to produce plans that honour those policies, communication, collaboration and individual sacrifice of utility are necessary. For example, some policies specify preconditions that constrain the deployment of resources. Choices made by one player may produce a situation that fulfills preconditions of its partner's policies, but a conflict may also occur, i.e., choices made by one player may affect the coalition partner. In order to obey policies, the players must recognise such conflicts and negotiate alternative plans in order to avoid the violation of policies.

Materials and Procedure

In order to test the influence of agents on the collaborative planning process of human teams, we created a software environment that allows human subjects to engage in the kind of collaborative planning described above. Of particular interest is capturing of information exchanged between the partners and how agents can provide advice and feedback about the planning actions to human test subjects.

[2]Some cases, however, require more work, in particular where the maintenance and observation of state over time is required.

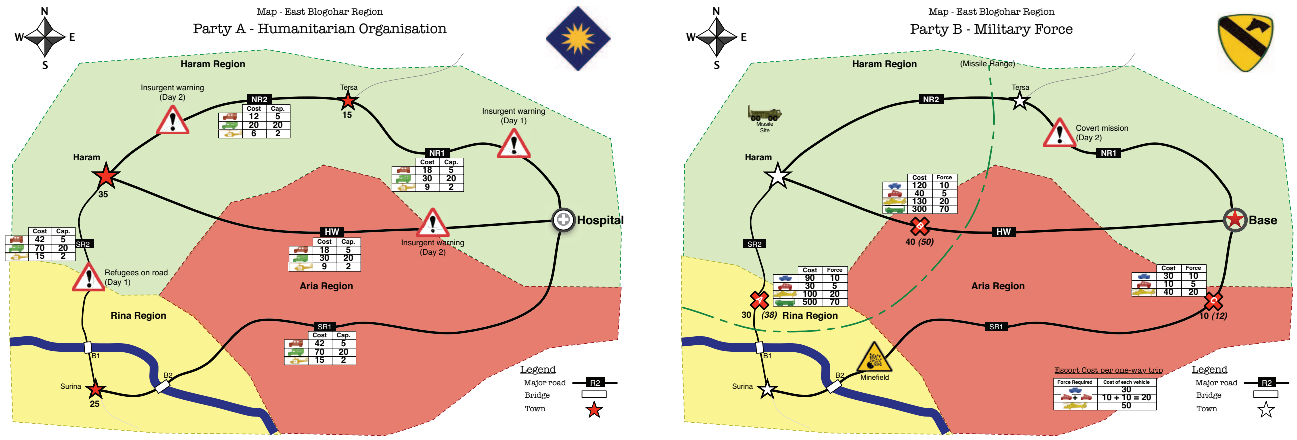

Figure 2: Maps for Party A and Party B

Twelve teams of two paid subjects each were recruited to participate in the study. Each team member played the game on a dedicated PC. Subjects sat in different rooms, so they could not look at each other's screen, and verbal communication was not allowed. The subjects were also forbidden to share note sheets or other such aids – they could only describe their intentions, commitments and planned resource deployments by using a *structured representation of messages* provided by a graphical user interface. Experiments were conducted in the three experimental conditions: the unaided condition (*control*), the condition where the agent acted as a *critic*, and the condition where the agent acted as a *censor*. Participants were given the briefing documents and map for the role they were playing (Party A or Party B) without making them aware of the goals, resources and map of the other party. The briefing materials explained the mission objectives, resources, policies, resource deployment costs, planning constraints (e.g., a Jeep can only carry 5 wounded in each deployment). Each subject was instructed on how to use the interface through a video; the video explained how to send a message to the other party, how to add plan steps, deploy resources and so on. The subjects were also shown what kind of information was displayed in different areas of the interface, and briefed about the functionality of each area. Each subject was then given a practice problem (a simplified scenario based on their role), and allocated 10 minutes to solve the problem collaboratively. After the practice problem, participants were allocated 60 minutes to create a plan for their role in the complete experimental scenario. They were reminded to fulfill their obligations and interact with the other party in ways compatible with their own policies. Any exchange of information about obligations, plans, routes or resources was then recorded for further analysis.

Results

We shall analyse the results of our experiments in three steps. First, we shall demonstrate the effect of agent assistance on the violation of policies. Second, we shall analyse the effect of *critic*, *censor* and *control* conditions on the interactions between the parties. Finally, we shall demonstrate the effect of agent assistance on the achievement of mission objectives.

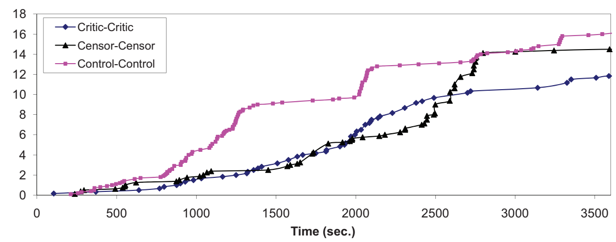

Figure 3: Cumulative no. of *attempts* to violate prohibitions

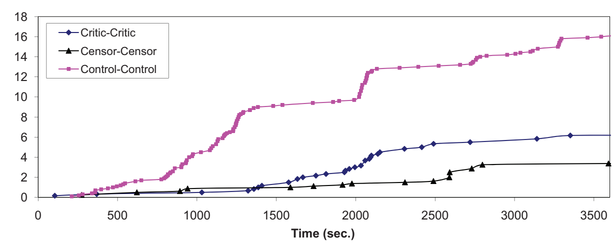

Figure 4: Cumulative no. of prohibition violations

Violation of prohibitions Figure 3 represents the cumulative number of attempts to violate prohibitions over time. In the *control* condition, all user attempts at committing a violation were successful; in the *critic* and *censor* conditions, all attempted violations were intercepted by the agent. In the *critic* condition, users were advised about the attempted violation but could override the agent's warnings, whereas in the *censor* condition, the violation of prohibitions regarding communication was prevented by the agent. For example, Party A had a policy stating: *"You are forbidden to share intelligence from insurgents with Party B"*. Party A also had the following intelligence from insurgents: *"There will be*

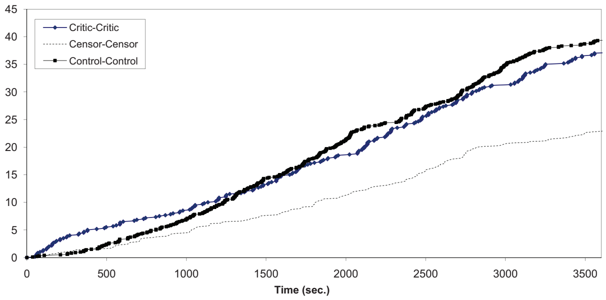

Figure 5: Cumulative no. of attempts to send messages

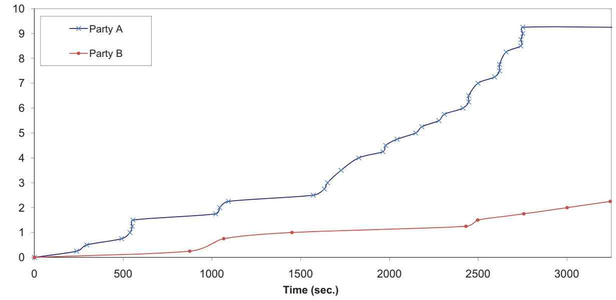

Figure 6: Cumulative no. of messages censored for Parties A & B

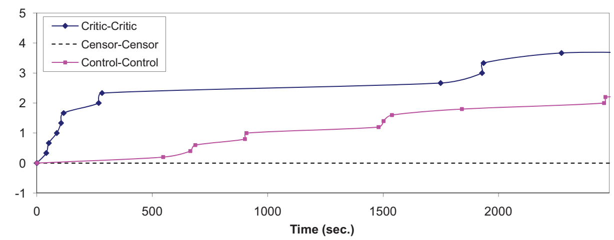

Figure 7: Cumulative no. of intelligence messages sent by Party A

a large number of refugees along road SR2 on Day 1". If Party A tried to send this intelligence, the communication would be transparently prevented by the censor agent.

As shown in Figure 3, the *control* condition, shows the highest number of attempts to violate prohibitions throughout the experiments. Violation attempts in the *critic* and *censor* conditions were almost the same until around the 45^{th} minute of the experiments. However, in the last 15 minutes of the experiments, possibly due to the time-pressure, the number of violation attempts increased in the *censor* condition and approached that of the *control* condition. These results show that, overall, agent support led to a smaller number of attempts to violate prohibitions.

Figure 4 shows the cumulative number of actual prohibition violations over time. These results demonstrate how effective agent interventions were at reducing the number of violations over time; that is, the number of violations in the *control* condition far exceeded that of the *censor* and that of the *critic* conditions. In the *censor* condition, the agent automatically prevented violations if these violations are related to communication policies. On the other hand, in the *critic* condition, the agent just alerts the users about the violation and allows them to decide whether to violate the policy or not. Towards the end of the experiments, the censor agent was twice as effective as the critic agent in preventing violations. This is because the critic agent allowed players to make their own decisions about violations while the censor agent simply enforced the policies.

Communication between parties In this section, we evaluate how agents affected the communication between the parties during the experiments. Figure 5 shows the total number of messages that subjects attempted to send over time. The figure shows that, in the *control* and *critic* conditions, planners attempt to send approximately the same number of messages. This indicates that the critic agent does not hamper dialogue between the parties. In the *censor* condition, instead, the subjects attempted to send significantly fewer messages. This may be considered as the effect of censoring; that is, censor agents block messages between parties if that would result in a policy violation. As a consequence, the parties consider each other less cooperative and they become less willing to communicate.

Figure 6 shows the number of censored messages in the *censor* condition. A considerable number of Party A's messages were censored, compared to those censored for Party B. This is an effect of the scenario; Party A had intelligence from insurgent sources stating the existence of dangers and refugees along specific roads; sharing intelligence from insurgents with Party B was however prohibited by Party A's policies. Nonetheless, Party A tended to share this intelligence with Party B because it needed Party B's support (e.g., to obtain an escort in case of danger), resulting in agent interventions to block those messages.

Figures 7 and 8 show the cumulative number of intelligence messages sent by Parties A and B. The figures show that, in the *censor* condition, Party A could not send any intelligence message to Party B, and all attempts to send such messages are blocked by the agent. By contrast, Party B can send some intelligence messages to Party A.

Planning performance Figure 9 shows how many wounded could be rescued using Party A's current plan as the experiment progressed. This is an estimation calculated using the intelligence and information that Party A had at that time. The figure reveals that significantly more wounded were rescued in the *critic* condition, while the number of rescued for other conditions were close, with the control condition marginally outperforming the censor condition towards the end of the experiments. Note also that better plans (those with a larger number of wounded rescued) were established significantly more rapidly in the *critic* condition than in the other two conditions. The costs of the overall plans at the end of experiments were 680 for the *critic* condition, 430 for the *censor* condition, and 480 for the *control* condition.

Figures 10 and 11 show the size of the current plan and

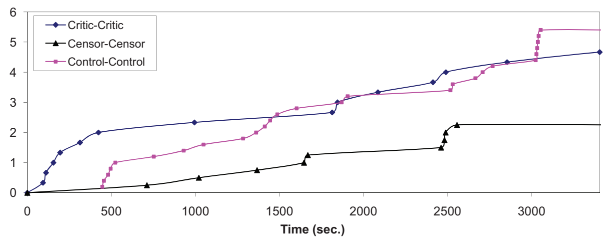

Figure 8: Cumulative no. of intelligence messages sent by Party B

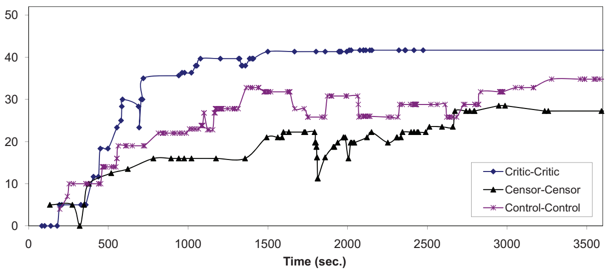

Figure 9: Total no. of wounded rescued by Party A

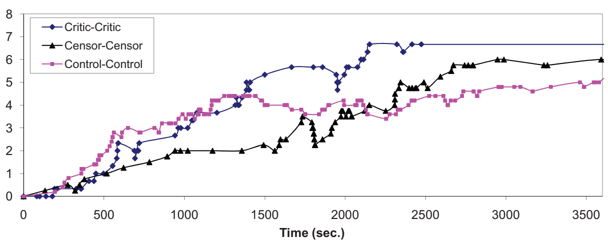

Figure 10: Plan size for Party A

the total number of plan step removals for Party A throughout the experiments. Plan step removals indicate instability in the plans being constructed. These figures indicate that subjects in the *critic* condition produced larger plans with fewer plan step removals, while the plan step removals were more in the *control* and *censor* conditions. Figure 12 shows the total number of plan step removals for Party B. The figure indicates that the *control* condition had the highest number of plan step removals with the total number over the experiments being similar for the *critic* and *censor* conditions. Nonetheless, it is interesting to note that, similarly to Party A, Party B's plans are more stable towards the end of the experiments when the subject is aided by the critic agent.

We also assessed how many insurgents were captured using Party B's current plan at the end of the experiments. On the basis of the information known to Party B at the end of the experiments, Party B was estimated to capture on average 46 insurgents in the *censor* condition, and 35 in the *critic* condition. The actual figure, computed assuming that all the critical information (known and unknown) was available to Party B, was 23 for both conditions. The reason for this is

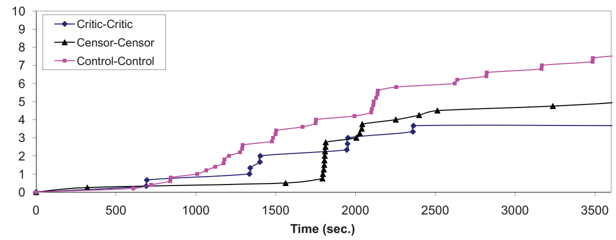

Figure 11: Cumulative no. of plan step removals in Party A's plans

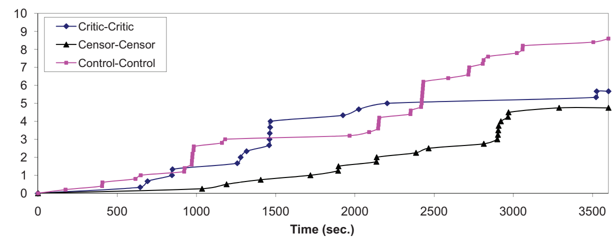

Figure 12: Cumulative no. of plan step removals in Party B's plans

that Party A had intelligence from insurgent sources about dangers which, if known by Party B, would have decreased the estimated number of captured insurgents. Party A, however, was prevented from sharing any intelligence with Party B in the *censor* condition, hence Party B *overestimated* the success of its plan; the actual number of captured insurgents was half the estimate (46 versus 23). On the other hand, in the *critic* condition the agents helped the planners to achieve a more accurate prediction of the actual number of captured insurgents (35 versus 23).

The costs of the overall plans at the end of experiments were 700 for the *censor* condition and 720 for the *critic* condition. This supports the hypothesis that a transparent policy enforcement led Party B to overestimate what it could actually achieve. Party B believes that more insurgents can be captured by incurring the same cost as in the *critic* condition, *underestimating* the actual cost of its plans. The critic agent offered a more accurate cost estimation to the subjects.

The size of plans for Party B in different conditions is shown in Figure 13. The plan size was not significantly different, but slightly larger in the *control* condition. Figure 14 shows how many plan steps related to escorting Party A appeared in Party B's plan. Interestingly, for the *critic* condition, Party B's plans contained around 3.5 steps to escort Party A at the end of experiments, while this number is 1.5 and 0.3 for the *control* and *censor* conditions. This is a further indicator (along with the relative number of messages illustrated in Figure 5) that leads us to conclude that critic agents fostered more cooperative planning.

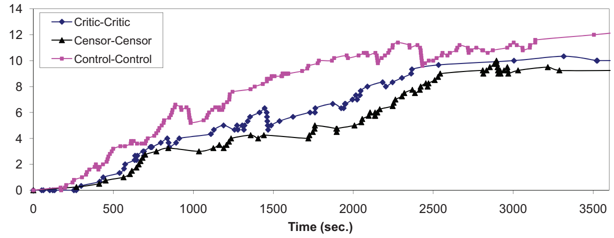

Figure 13: Plan size for Party B

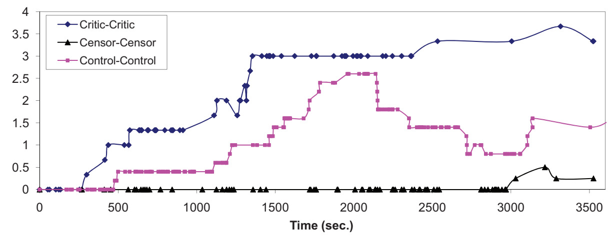

Figure 14: No. of escort plan steps in Party B's plans

Conclusions

In this paper we have presented intelligent agents that support human planners in a joint planning scenario under policy and planning constraints. In our scenario, a humanitarian organisation and a military force had to cooperate to achieve their individual goals within a normative setting. We have run a number of experiments with human subjects in three different conditions: a control condition with no agent support, a critic condition where humans planners were supported by agents, and a censor condition in which agents transparently enforced policies and constraints.

Our results show that critic agents:

- *Fostered collaboration between coalition partners*. Planners established and stabilised plans quickly, and produced higher quality plans with respect to the main mission objectives – rescuing wounded and capturing insurgents.
- *Effectively limited the number of policy violations* attempted and committed by planners, *without hampering the communication* between the partners.
- Enabled planners to *develop a more accurate prediction of the plan outcomes*, both in terms of objective achievement and costs incurred for resource deployment.

Censor agents successfully limited the number of attempted and committed policy violations, yet they introduced several undesirable side effects. First, the communication between parties was negatively affected, producing a significant reduction in collaboration between planners. Planners also developed a misleading view of their plan outcomes; they overestimated the success of their plans or underestimated the costs incurred in their deployments. Finally, planners showed poorer planning performance, struggling to finalise their plans with their partners. Critic agents have therefore been proved to provide effective support to humans when compared to the control and censor conditions.

References

Burnett, C.; Masato, D.; McCallum, M.; Norman, T. J.; Giampapa, J. A.; Kollingbaum, M.; and Sycara, K. 2008. Agent Support for Mission Planning Under Policy Constraints. In *Proc. of the Second Annual Conf. of the Int. Technology Alliance*.

Dignum, F. 1999. Autonomous agents with norms. *Artificial Intelligence and Law* 7(1):69–79.

Elagh, A. 2000. On the Formal Analysis of Normative Conflicts. *Information and Communications Technology Law* 9(3):207–217.

Hill, E. F. 2003. *Jess in Action: Java Rule-Based Systems*. Greenwich, CT, USA: Manning Publications Co.

Kollingbaum, M. J.; Giampapa, J. A.; Sycara, K.; Norman, T. J.; Burnett, C.; and Masato, D. 2009. Automated Aiding Strategies for Decentralized Planning with Interdependent Policies (Short Paper). In *Proc. of Eighth Int. Conf. on Autonomous Agents and Multiagent Systems*.

Kollingbaum, M. 2005. Norm-governed Practical Reasoning Agent, PhD thesis.

Lopez y Lopez, F.; Luck, M.; and d'Inverno, M. 2004. Normative Agent Reasoning in Dynamic Societies. In *Proc. of the Third Int. Conf. on Autonomous Agents and Multiagent Systems*.

Sycara, K.; Norman, T. J.; Giampapa, J. A.; Kollingbaum, M. J.; Burnett, C.; Masato, D.; McCallum, M.; and Strub, M. H. 2009. Agent Support for Policy-Driven Collaborative Mission Planning. *The Computer Journal, Special Issue on Advances in Sensing, Information Processing and Decision Making for Coalition Operations within the US/UK Int. Technology Alliance*.

Wheller, V., and Harmer, A. 2006. Resetting the rules of Engagement. Trends and Issues in Military-Humanitarian Relations. Technical report, HPG Research Report.