© The Author 2005. Published by Oxford University Press on behalf of The British Computer Society. All rights reserved.
For Permissions, please email: journals.permissions@oupjournals.org
doi:10.1093/comjnl/bxh000

Agent Support for Policy-driven Collaborative Mission Planning

Katia Sycara[1], Timothy J. Norman[2], Joseph A. Giampapa[1], Martin J. Kollingbaum[1], Chris Burnett[2], Daniele Masato[2], Mairi McCallum[2] and Michael H. Strub[3]

[1]*Robotics Institute, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, USA*
[2]*Department of Computing Science, University of Aberdeen, Aberdeen AB24 3UE, Scotland, UK*
[3]*USARL Exchange Scientist at DSTL*
Email: katia@cs.cmu.edu

In this paper, we describe, how agents can support collaborative planning within international coalitions, formed in an ad-hoc fashion as a response to military and humanitarian crises. As these coalitions are formed rapidly and without much lead time or co-training, human planners may be required to observe a plethora of policies that direct their planning effort. In a series of experiments, we show how agents can support human planners and ease their cognitive burden by giving advice on the correct use of policies and catch possible violations.

Received 03 October 2008; revised 05 February 2009

1. INTRODUCTION

In order to respond rapidly and effectively to military and humanitarian crises, the formation of international coalitions by different countries and organizations is regarded as a necessity. Coalitions – an organizational form whose members engage in collaborative activities – are typically motivated by the fact that no single nation or humanitarian relief organization has all the necessary capabilities or resources to undertake particular tasks alone. Planning collaborative missions for responding effectively to crisis situations is a complicated task that requires mission planners and coordinating personnel to observe a plethora of rules or constraining policies in their attempt to deploy resources and troops in an efficient manner.

Characteristically, these coalitions are formed rapidly and without much lead time or co-training. This gives rise to a set of problems. First, although coalition partners intend to collaborate for the fulfillment of common goals, they are also assumed to have individual goals whose fulfillment they are pursuing as well. In other words, coalition members are self-interested, although it is assumed that by being part of the coalition, both the individual utility and also the group utility are increased. Second, and partially as a result of the existence of individual goals, there is varying trust among the members of a coalition. Third, coalition members typically operate under different policies that range from security policies to policies about how to conduct their missions (e.g different rules of engagement in military coalitions). Policies from different coalition partners may be in conflict and, therefore, hamper the collaborative planning of missions. With that, coalitions clearly differ from *teams*: in teamwork [1], the assumption is that team members only hold common goals and that they engage in collaborative planning and execution in pursuit of these common goals. Because of the lack of self-interest, as observable in coalitions, issues of trust and policy conflicts do not usually occur in teams.

One of the most crucial challenges for coalitions is how to construct joint plans in the presence of self interest, individual goals and diverse policies, especially in time-stressed situations, where coalition members have only limited co-training for recognizing and resolving their differences. Among others, we are mainly confronted with the following issues:

- in mission-critical scenarios, coalitions consist of members from different countries or organizations, are formed rapidly and without much lead time
- coalition members have to adhere to certain policies (explicit obligations, permissions and prohibitions) that determine what planned actions are actually allowed, forbidden or obliged and what information can be (or must be) disclosed and

communicated to other coalition partners in the course of a collaborative planning activity

One possibility to address these challenges is to create automated agents that assist coalition partners in their policy-driven planning activities. We have implemented an experimental framework that allows us to evaluate the effectiveness of different types of intelligent automated agents aiding the planning process. This framework allows coalition partners to exchange information necessary for creating their individual plans and records the plans themselves. Agents, which operate within this framework, are aware of the policies and reason about the plans established and the communication taking place between the coalition partners.

In this paper, our aim is to study how policies impact on the planning performance of coalition partners and how agents can most effectively aid such a collaborative decision-making process. In particular, we investigate how the presence of agents impacts on policy infringement (especially the reduction of unintended violations), which effect they have on the quality of a plan and how they ease the cognitive overload of human planners in their attempt to take into account high numbers of (often new and unknown) norms and policies during their planning task and their communication with coalition partners. For our investigation, we consider coalitions with a small number of members, each with their own organizational policies, and observe their performance in terms of constructing plans for fulfilling shared and individual goals.

The rest of the paper is organized as follows: in section 2, we describe the reasoning of the agents about policies, their implemention and the software environment, within which they operate and provide feedback to the human planners. Section 3 describes the experimental setup and the results of these experiments. Section 4 provides a discussion of how agents in our experimental setup impacted on the planning performance within a coalition and proposes future directions of research. Section 5 presents related work. Section 6 concludes the paper.

2. AGENT SUPPORT

Policies may impose restrictions on information sharing between coalition partners and on the set of actions that may be part of a plan. This can have a negative impact on the quality of planning outcomes. Additionally, the presence of diverse policies among coalition partners may adversely impact the planning process. Moreover, in some situations, planners may not recognize that violations occurred.

We consider the use of agents to monitor communication and planning activities of coalition partners and reason about possible policy violations. Agents, as we utilize them, do not form coalitions themselves or are part of a virtual organization. Such 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 human planners by reporting policy violations, by informing the human planner about the policies that have to be observed in a particular situations and advising human planners on alternative courses of action in order to act in a policy-abiding manner. The agent support was designed with three criteria in mind:

(i) Reasoning about Policies: The agent has to correctly assess whether a planner complies with policies or violates them
(ii) In terms of the agent presenting itself to a user, its design has to be balanced in terms of the degree of visibility – how the agent makes its presence known to the user, the degree of pro-activity and reactivity to user actions, and minimizing the possibility of irritating the user
(iii) User dependence on the agent: the agent has to be designed so that its intervention helps the user learn and navigate its policy / goal space more effectively, rather than blindly relying on the assistance of the agent.

We provide *each* player with a personal agent. We designed the agent as an unobtrusive monitor of human communication and planning activities. The agent's reaction is based on criteria resulting from multiple information sources and not just a simple and uninformed reaction to user actions. The agent of each party in the collaboration monitors the messages the party wants to send, the messages the party receives and the plan operators he/she proposes. In this way, the agent follows the human lead, it is aware of information that is arriving to the human and is monitoring the human's intentions by observing the formation of their tentative and final plans.

2.1. Aiding Strategies

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

- a *critic* agent that detects policy violations of coalition partners in the course of communication activites 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;
- a *censor* agent, that interferes with the communication by deleting parts of the exchanged messages

(or blocks it completely) that contain policy violations; 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.

2.2. Reasoning about Policies

In order for the agent to work effectively in tandem with the human planner, the agent must intercept any communication and planning action *before* it is actually performed (message received by another coalitions partner, plan action added to plan) in order to provide the human planner with warnings in case violations occur and possibly advise how to remedy such a situation. The agent, therefore, has to maintain a representation of the "social burden" or "normative position" of the human planners [2] – the policies that inform their planning and communication behaviour, which are the obligations, permissions and prohibitions that are currently relevant to them. A monitoring agent reasons about a *potential* normative position of a human planner (and not the actual situation), because a possible future normative situation is anticipated by intercepting actions performed by the human planner. Based on this "outlook" at a potential future normative situation, the agent can reason about the normative consequences of an action and can advise the human planner accordingly.

In our experiments, policies were given to the human planner in a written form. The following policy is a typical example taken from our test scenario (see section 3):

Example 1. "IF you want to deploy an ambulance along route R on day D for a rescue operation, THEN you are obliged to obtain a commitment of escort from your coalition partner"

This obligation will become relevant to the human planner in the course of planning such a rescue operation, if the deployment of this specific resouce is planned. By becoming relevant, this policy adds to the current "social burden" of the human planner – the human planner has to observe this obligation (beside possible other activated norms) and see to it that it is fulfilled. This obligation will be fulfilled when such a commitment of escort is obtained. In that case, we regard the obligation to *expire*. We, therefore need to specify these additional fulfillment or *expiration conditions*. The example above would then be complete by ammending it with the following information:

Example 2. "IF you want to deploy an ambulance along route R on day D for a rescue operation, THEN you are obliged to obtain a commitment of escort from your coalition partner. IF you have aquired a commitment of escort before day D THEN this obligation is fulfilled"

Independent of whether an obligation is fulfilled, it will also be de-activated in case that the activating circumstances no longer hold. In case of the above example, if the human planner decides to discard the planned deployment of an ambulance, this obligation is no longer relevant. The following example shows a prohibition:

Example 3. "IF you know that the route R on day D is dangerous for deployments, THEN you are prohibited from deploying an ambulance along route R on day D for a rescue operation"

This prohibition becomes relevant if there is knowledge about danger on the given route available. As is obvious, this can also be formulated as a permission: "If there is no knowledge of danger ... THEN you are permitted ...". It shows that, in the design of policies, we have to clarify the *default normative position* for a coalition partner, from which point of view the policies are designed. Are we assuming that "everything is permitted that is not explicitly prohibited" or do we take the stance that "everything is prohibited that is not explicitly permitted"? For the design of our policies, we decided that, per default, any plan and communication action is permitted and that we provide explicit prohibitions to disallow such actions. In the same way as obligations, prohibitions must be augmented with conditions that indicate the circumstances under which a violation occurs. Prohibitions are activated under conditions that describe violating circumstances – with an activation of a prohibition, its violation is indicated, whereas in its deactivated state, it is regarded as not being violated.

We describe here the normative position of the human planner, which is monitored by the agent, as the set of currently *activated* policies. Policies are *active*, if they are relevant to the human planner under current circumstances and have to be observed. The set Ω comprises the currently activated policies:

Definition 2.1. *The set Ω comprises the currently active policies, containing the permissions given, the obligations that must be fulfilled, and the prohibitions that are potentially under threat of violation.*

If Ω contains activated prohibitions then the agent may inform the human planner what violations of policies could *potentially* occur. By providing information about those violations and active obligations to the

human planner, the monitoring agent may be able to motivate her to correct her behavior.

With respect to an implementation of such an agent, with each occurrence of either a communication or planning action, we regard the set Ω being discarded, the activations of all policies checked afresh and a new set Ω' created. Ω' represents the *potential* normative position of the human planner that would obtain, if the intercepted actions take place. If the coalition finishes its collaborative planning activity, the set Ω^{final}, maintained for an individual coalition partner by its monitoring agent, can have the following states:

(a) Ω^{final} is empty or contains only permissions – the human planner has a clean record with all obligations fulfilled and no prohibitions violated, or

(b) Ω^{final} still contains obligations and/or prohibitions – this indicates, that those obligations were not fulfilled and the violation of prohibitions persisted beyond the planning session.

At that point in time, Ω^{final}, represents the *actual* normative state of the human planner.

Our representation of policies follows our earler work [3]. We specify an obligation, permission or prohibition on a particular action with two condition – an activation and an expiration/fulfillment condition – determining whether a policy is relevant to the human planner. If we define the set *Expr* as the set of all possible well-formed formulae comprising first-order predicates over terms (constants, variables and the operators $\wedge$, $\vee$ and $\neg$, then a policy can be defined in the following way:

DEFINITION 2.2. *A policy, expressing an obligation, permission, prohibition is a tuple* $\langle \nu, \rho, \varphi, a, e \rangle$*, where*

- $\nu \in \{\mathcal{O}, \mathcal{P}, \mathcal{F}\}$ *is a label indicating whether this is an obligation, permission or prohibition*
- ρ *is a role identifier for a norm addressee*
- φ *describes the action regulated by this policy*
- $a \in Expr$ *is the activation condition*
- $e \in Expr$ *is the expiration condition*

This definition displays in a simple fashion the elements that characterize an implementation of our policies – they are ascribed to a specific role (in our experiments, we have the roles "Party A" as the humanitarian organization and "Party B" as the military organization) and are activated/de-activated under certain conditions. The policies themselves exist in two forms, (a) formulated in simple "IF ... THEN ..."-statements that are given to human planners, and (b) implemented as a set of rules, expressing their activation/de-activation, in order to allow agents a processing of these policies and the reasoning about their current activation state.

2.3. Agent Implementation

For agents to become operational, they must have access to plans and to communication activities. We use a traditional forward-chaining mechanism (expert system shell Jess [4]) for implementing the policy reasoning for an agent as a set of rules. A policy will experience activations and de-activations under specific circumstances. In order to correctly implement activation and de-activation of policies, each policy is expressed by a set of rules and data structures recording such an activation state. The agent operates in a fixed monitoring cycle:

(i) detect the current situation changed by arriving messages expressing the coalition partners' commitments for action or revealed intelligence, as well as new planned actions,
(ii) reason about these changes (activations/de-activations of policies), and
(iii) create the new set of activated policies.

The agent has to intercept both communication and planning actions in order to update an internal representation of the normative situation at hand. The encoding of policies in the Jess language occurs as a set of rules to correctly represent the activation state of a policy.

In stage (i) of the monitoring cycle of the agent, messages received or intended to be sent, as well as plan steps intended to be added to the current plan, will be recorded by the agent. In stage (ii), the agent will reason about this recorded information with rules such as displayed in figure 1. In figure 1, we provide an example of how essential parts for maintaining the activation state of a policy are implemented in Jess. In this example, one of the human planners is confronted with the following policy:

EXAMPLE 4. "If: you wish to perform ground or air operations along route R on day D, THEN you are obliged to obtain a clearance from Party B for route R on day D"

As shown in figure 1, the agent will recognize the intention of the human planner for adding such a particular plan step with rule *A* and record a "policy-activation". This activation of a policy will be recognized in stage (iii) of the monitoring cycle of the agent and regarded as being an element of the set of activations Ω. Rules *B* and *C* maintain de-activation and re-activation (explicit re-activation must be considered due to the properties of how rules fire in Jess). Rule *B* will remove the "policy-activation" record. Such a removal indicates two different cases, either the *fulfillment* of the obligation or just a simple de-activation. An obligation is fulfilled, if the human planner received a commitment of support by a coalition partner in the next monitoring cycle, after the activation was recorded and observed in stage (iii) of the previous monitoring cycle (becoming an element of Ω). The fulfillment will be recorded and the "policy-activation" record removed. If, on the other hand, the human planner already received a commitment for

Rule A

```
(defrule AP-2
    "If:  you wish to perform ground or air operations along route R on day D,
     THEN you are obliged to obtain a clearance from Party B for route R on day D
    "
    (PlanStep (plan_step_ref ?plan_step_ref)(route $?route)(day ?day)(state "added"))
    =>
    (assert
        (policy-activation
              (modality  "obligation")(policy_id  "AP-2")(route ?route)(day ?day)
                  (plan_step_ref ?plan_step_ref)(activation-state "active")))
)
```

Rule B

```
(defrule AP-2-retract-on-commit
    ?f <-(policy-activation
             (modality  "obligation")(policy_id  "AP-2")(route $?route)(day ?day)
             (activation-state "active"))
    (received-COMMIT (msg_ref ?msg_ref))
    (Message         (msg_ref ?msg_ref)(route $?route)(day?day))
    =>
    (retract ?f)
)
```

Rule C

```
(defrule AP-2-reactivate-planstep-on-withdraw
    ?f<-(PlanStep          (plan_step_ref ?plan_step_ref)(route $?route)(day ?day))
        (received-WITHDRAW (msg_ref ?msg_ref1))
        (Message           (msg_ref ?msg_ref1)(route $?route)(day ?day))
     =>
    (assert
        (policy-activation
              (modality  "obligation")(policy_id  "AP-2")(route ?route)(day ?day)
                  (plan_step_ref ?plan_step_ref)(activation-state "active"))))
)
```

Rule D

```
(defrule planstep-discarded
    (PlanStep          (plan_step_ref ?plan_step_ref) (state "discarded"))
    ?f<-(policy-activation (policy_id ?policy_id)(plan_step_ref ?plan_step_ref))
    =>
    (retract ?f)
)
```

FIGURE 1. Implementation of Activation/De-activation Rules of a Policy in Jess

support before the mentioned planning activity, rule B will counteract rule A immediately and remove the "policy-activation" record. With that, no activation of this obligation will be recorded in stage (iii) of the monitoring cycle. Rule C covers the case of the coalition partner withdrawing its commitment – the obligation will become relevent to the human planner again. Finally, rule D will remove a "policy-activation" record, if the human planner decides to remove/discard the intended plan step.

These four rules represent an implementation for this particular policy. They embody principles of implementation, but the actual approach taken has to be decided for each policy individually (rules A, B, C are policy-specific, whereas rule D is a general-purpose maintenance rule for plan steps).

2.4. The Experimental Environment

As pointed out before, the interaction between human planner and agent depends on the aiding strategy employed. It influences how the agent deals with the information collected in stage (iii) of its monitoring cycle. In *censor* mode, the agent will silently record policy violations/fulfillments and manipulate the communication between coalition partners in a way that eliminates/censors parts within messages (or complete messages) that violate policies. With that, we recreate traditional censoring of written communication. In *critic* mode, the agent operates in a verbose fashion. It informs the human planner about the policies currently active as collected in stage (iii) and may also advise the agent how to resolve violations.

We built a software environment that allows human planners to communicate during planning, create a plan and, if tests are performed in the *critic* condition, interact with an agent acting as the *critic*. We strongly simplified the communication between the human planners – they interact in writing and exchange messages in a pre-specified form. This allowed us to keep the reasoning of the agent about communication activities simple and focus on the main purpose of our experiments – to compare how different agent aiding strategies impact on the quality of the plans produced by the coalition partners. Such a simplification has a certain influence on the design of the user interface of our software environment – human planners must be given the means to create their messages according to the given format.

Figure 2 shows the experiment interface as it is used to test agents aiding in the *critic* condition. It contains areas that allow the agent to provide feedback about potential violations of prohibitions or which obligations are currently active. For this, the agent produces output in specific "'Advice" and "Reminders" areas (see Figure 2). The user can either accept the advice given by the *critic* agent, for example, by discarding a plan step or not sending a message, or the user can go ahead with its planning and communication actions despite the warnings of the agent. To allow the user to reject or "override" the agent's advice and warnings about impending violations, an "override" button is part of the interface.

The interface also contains a communication and a planning area. The communication area consists of two parts, the area where the human planner can assemble messages in the given pre-specified form by choosing values from pull-down menues and, in addition, a normal "chat" interface, where they may type any free-form text. Communication via this "chat" interface is, currently, not analyzed by the agent (and planners are discouraged to use it). Finally, in the planning area, human planners can assemble their plan.

In the *control* and *censor* conditions, the interface does not contain the "Advice" and "Reminders" areas for the agent. The *censor* agent intercepts any message that contains policy violations and prohibits the message from being sent and informs the recipient that the sender's message contained a policy violation and was suppressed. As we see in the section of experimental results, this behavior of the censor agent has some interesting consequences in terms of the performance of the planning partners.

The experimental environment is extensible in the sense that additional chat partner can join, but the communication and planning menus within the GUI are specific to the role a participant takes on. Therefore, role-specific information has to be imported into these menus.

3. EXPERIMENTS

We chose a rescue mission as our example scenario and performed a set of experiments to investigate the effectiveness of agents supporting a collaborative planning effort in the context of this scenario.

In this scenario, we assume that there are two parties that form a coalition, a humanitarian relief organization with the individual goal of rescuing injured civilians from a potentially hostile region, and a military organization that has to coordinate its military objectives with the evacuation activities. In the experimental setup, the humanitarian organization is regarded as "Party A" and the military organisation as "Party B".The goal of this coalition is to find a joint plan for rescuing as many injured people from a dangerous region to a hospital in the shortest possible time. The optimal situation for Party A would be to provide medical attention and evacuation as soon as possible. For this, party A may need support from party B, for example, an escort through a dangerous region. Party B, on the other hand, has military objectives that, potentially, may be in conflict with the support given to party A.

We assume that both parties have a set of resources such as ambulances, field hospitals/paramedic units, helicopters, Jeeps, etc.. These resources (e.g. ambulances) have a finite capacity for transporting wounded people or, in case of military hardware, a specific military *strength*. Both parties have to plan multiple deployments of their finite resources in order to achieve their individual goals – rescue all wounded (Party A) or defeat all insurgent strongholds (Party B). Each deployment incurs specific costs, whose sum determines the overall cost of a plan. Each party has the goal to minimize costs. During their planning activity, the coalition partners will allocate these resources to be used in planned actions. We assume that party A and party B have a small set of capabilities they may plan to utilize in pursuing a mission. Party A can either evacuate wounded people, taking round-trips to their location or dispatch a paramedic unit to provide medical care at their location directly. Party B may either support party A by providing escort through dangerous terrain or pursue own military goals by attacking enemy strongholds. These destinations have numerical requirements – in case of Party A, a specific number of wounded has to be evacuated, in case of Party B, insurgent strongholds have a specific resistance value that has to be overcome by military means. Deployments are taking place along given routes (according to the map of Party A as depicted in figure 3) and at a specific time. As both parties operate in the same area to pursue their individual goals, they have to collaborate so that their plans are complementary. Moreover, there are dependencies between the parties – for example, Party A will need military escort through dangerous areas. Party B has 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 [5]. These policies are private knowledge to each party. In order to produce plans that honor those policies, communication, collaboration and individual sacrifice of utility may be necessary.

Figure 4 demonstrates a kind of conflict between specific parameterizations (in this example, the route and day of deployments) of capabilities, where one player's actions can cause the other's to become prohibited. For example, a particular road could be

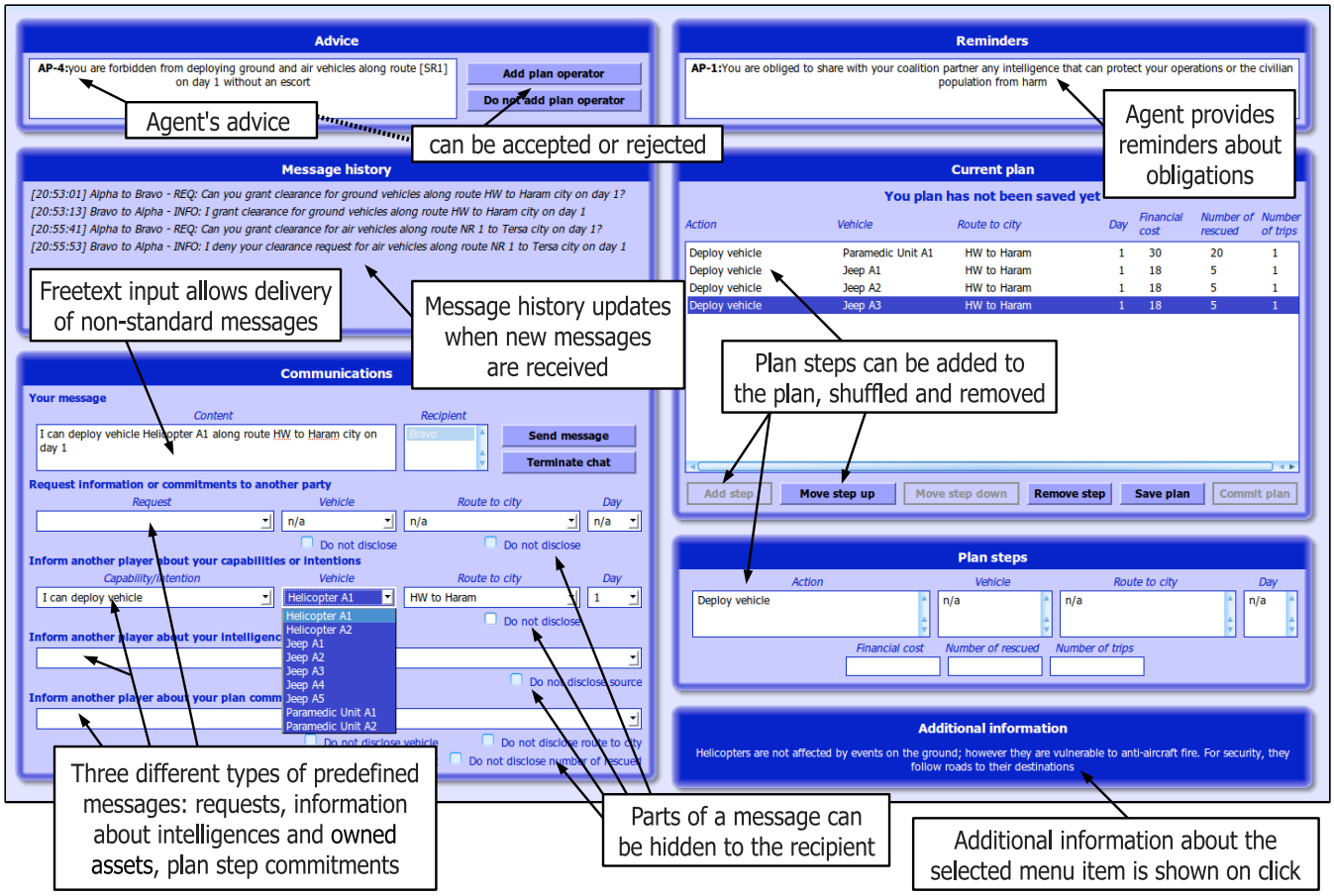


FIGURE 2. The Communication & Planning Interface in the Critic Condition.

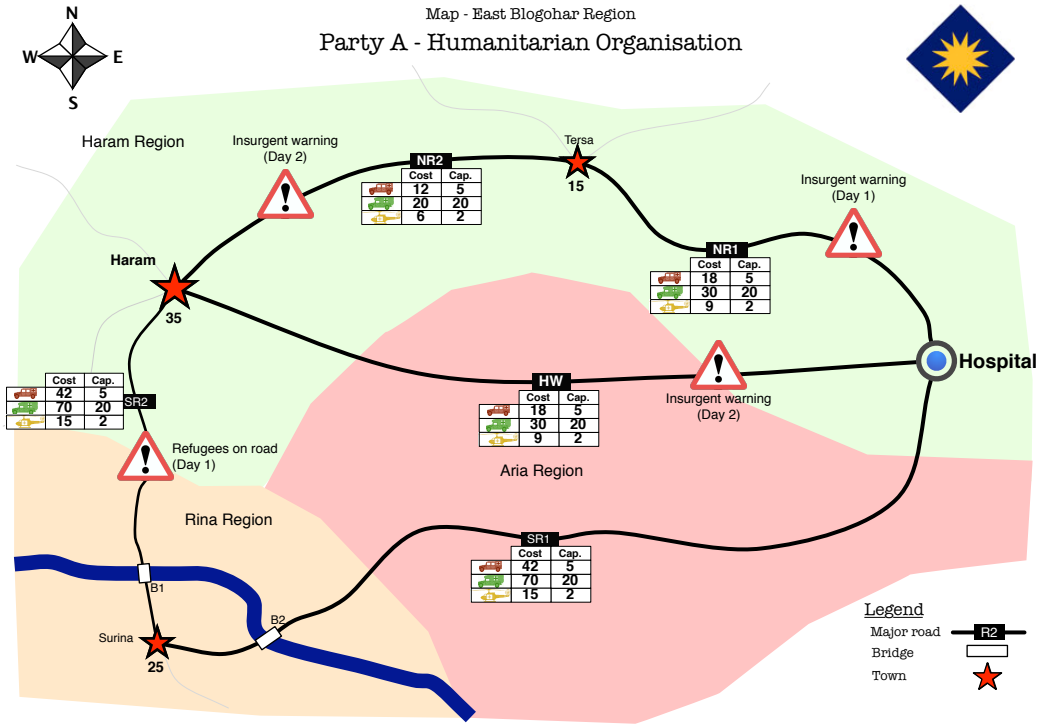


FIGURE 3. Map for Party A

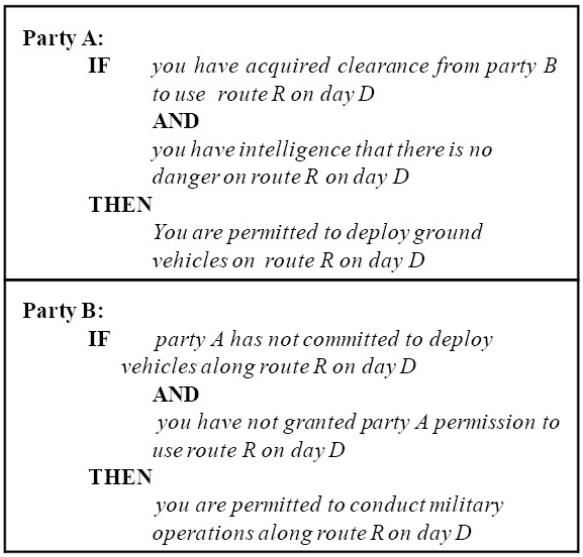


FIGURE 4. Example Policies

an attractive choice for both players. However, a state in which both players are using this road at the same time would result in policy violations for one (or possibly both) of the players. To avoid this, players must negotiate and compromise. However, there may be information sharing policies that can complicate negotiation by prohibiting players from revealing certain information. For example, Party A has such a policy forbidding her from revealing intelligence if this intelligence comes from insurgent sources.

3.1. Materials and Procedures

Thirty teams of two paid subjects each were recruited to participate in the study. The test population was homogeneous in terms of planning experience. These teams were tested in their collaborative planning effort in one of three conditions, the unaided condition (control), the condition where the agent acted as a "critic", and the condition where the agent acted as a "censor", resulting in ten teams operating in each of the three conditions. The test subjects were forbidden to share computer screens, note sheets or other such aids and worked isolated from each other. They could only describe their intentions, commitments and planned resource deployments by using either a structured representation of messages or a free-form text chat box of the experiment software environment (see figure 2). The test subjects were given written documents as well as shown a video briefing them about the impending task explaining the mission objectives, resources, policies, resource deployment costs and planning constraints (e.g. a jeep can take only 5 wounded in each deployment). In a first step, a team performed a practice problem as a warmup in order to become familiar with the planning process. In particular, the practice problem of party A was: "Plan the lowest cost emergency medical evacuation to the village of Tersa on Day 1. Be sure to do so in a way that is compatible with your policies. What is the total cost of your operation?" As the second step, the team then performed the complete planning problem in one of the described experimental conditions. The total allotted time to finish the whole experiment including reading the briefing, video viewing and performing the practice problem, was 2 hours.

3.2. Results

The purpose of the experiment was to examine team planning performance under the three different experimental conditions. We investigated a variety of measures of performance with respect to the *outcome of collaboration*, for each of the parties: (a) fulfillment of mission goals, (b) cost of final mission plan, (c) compliance with policies (number of policy violations and number of satisfied obligations), and (d) timeliness of mission fulfillment.

Notice that the planning task is challenging for the following reasons. First, the subjects are given only a limited amount of time to complete the task. Therefore, they are under considerable pressure to finish the task within the allotted time. Second, the subjects are asked to coordinate with another party that has its own mission goals. Therefore, although there are dependencies in the plans of the two parties, the parties have to operate with the cognitive load of coordination, with incomplete information and with another party with whom they have not worked together before. Moreover, in the case of Party B, fulfilling obligations to provide escort for the evacuation of the wounded, possibly leads to additional cost and delay in fulfilling her own mission goals (defeating the insurgent strongholds). Third, the subjects are asked to plan under policy constraints that include permissions, prohibitions and obligations as well as plan constraints (e.g. vehicle capacity constraints, fulfillment of preconditions of plan steps) to optimize performance. Fourth, optimized performance is multi-dimensional including the degree of goal fulfillment, e.g. number of evacuated and treated injured patients, the cost of resource deployment and timeliness.

We hypothesized that since policies impose restrictions on the planning activity, the planning performance as measured by degree of goal fulfillment or cost considerations would be worse for conditions that encourage policy compliance (critic) or enforce policies (censor).

The violations of Figure 5 were determined by reviewing the transcripts of the parties' dialoges, the final recorded state of the entered plan steps, and by reviewing the log files of the agents that were monitoring the humans' dialogues and their planning actions[4]. Agents were monitoring the users' actions for policy violations in all three conditions, even if there

[4]SE stands for Standard Error.

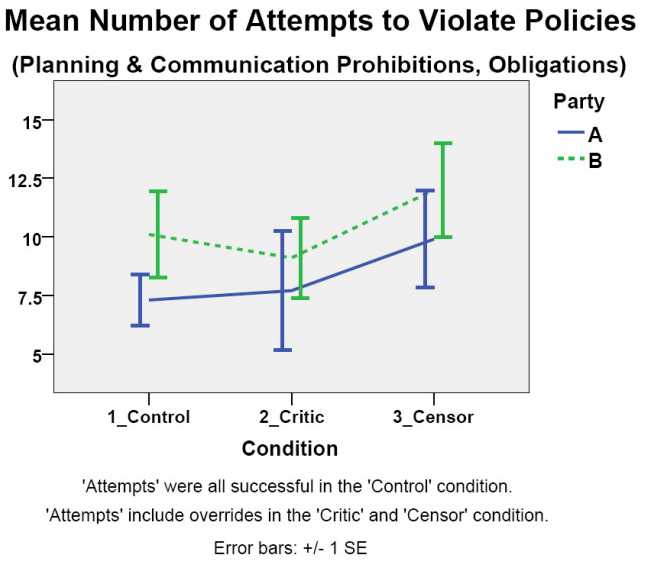


FIGURE 5. User-Committed Policy Violations

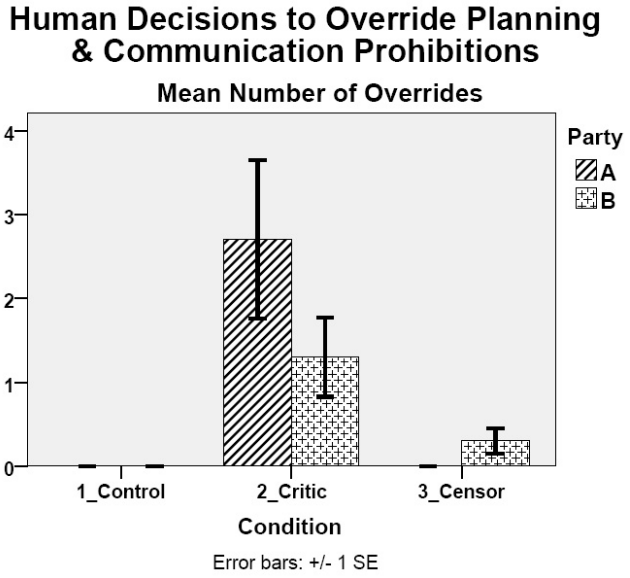


FIGURE 6. User Override of Agent Enforcement

was no agent output in the control condition. The policy violations are of three types:

planning prohibitions e.g. *You are forbidden from deploying paramedic units into the Rina region.*,

communication prohibitions e.g. *If your intelligence is from insurgent sources, then you are forbidden from sharing it with Party B.*, and

obligations that remained unfulfilled at the end of the session; e.g. *If you wish to perform ground or air operations along route R on day D, then you are obliged to obtain a clearance from Party B for route R on Day D.*

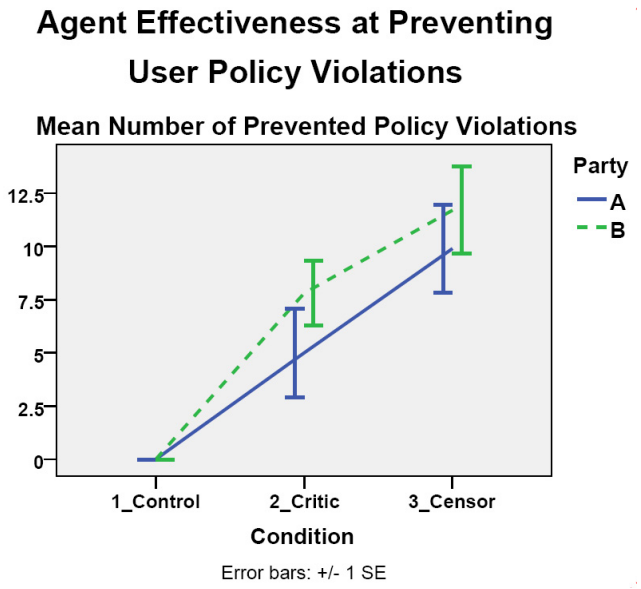


FIGURE 7. The mean number of times per user, where an agent prevented policy violation.

Figure 5 illustrates the number of *attempted*[5] policy violations: in the control condition, all user attempts at committing a violation were successful; in the agent-aided conditions, critic and censor, all attempted violations were intercepted by the agent. In the critic condition, users were allowed to override the agent's warnings and commit the violations; in the censor condition, the policy violating communication was suppressed, but some users circumvented this by communicating through the chat window (see Figure 6).[6]

No significant differences in attempted policy violations were found between conditions. This indicates that the agent assistance did not induce human users to attempt more policy violations. There was slightly more variance among subjects in the censor condition, possibly due to uncertainties arising from suppressed communications. In the censor condition, the range of strategies varied from making fewer policy violations, to attempting to repeat the censored message multiple times, to taking an exhaustive approach, in which the user, whose message was being censored, would attempt all possible actions until one is no longer censored, thus performing an exhaustive though naïve generate-and-test approach. Although Party B has a slightly larger "policy burden", in that they must plan for themselves as well as with Party A, this difference did not affect the number of violations.

[5]Here, *attempted* does not imply any intentionality on the part of the user, i.e. the policy violation could be accidential.

[6]Planning prohibitions were tallied for subjects in all conditions. Since the censor agent did not censor a subject's private plan steps without there being a corresponding communication, there could not be any user overrides of planning prohibitions to include in Figure 6.

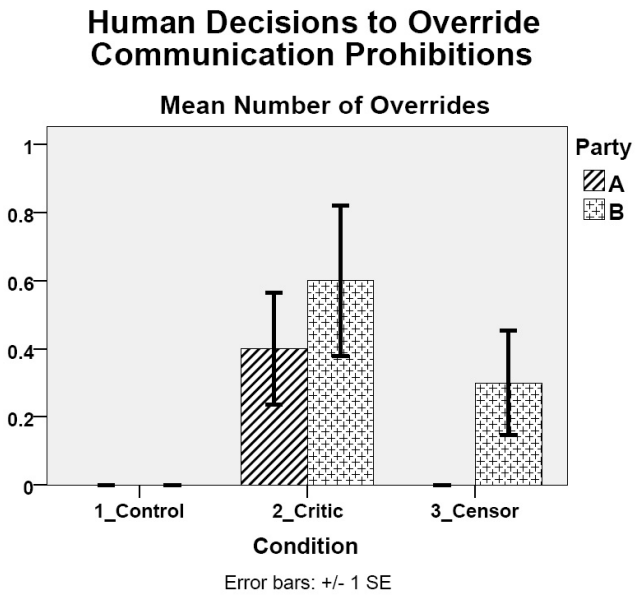


FIGURE 8. User Overrides of Agent Prevention of Communication Policy Violations

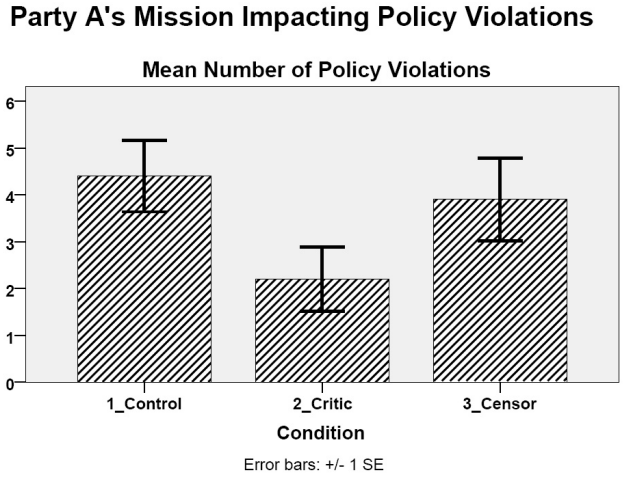


FIGURE 9. Mission Impacting Policy Violations

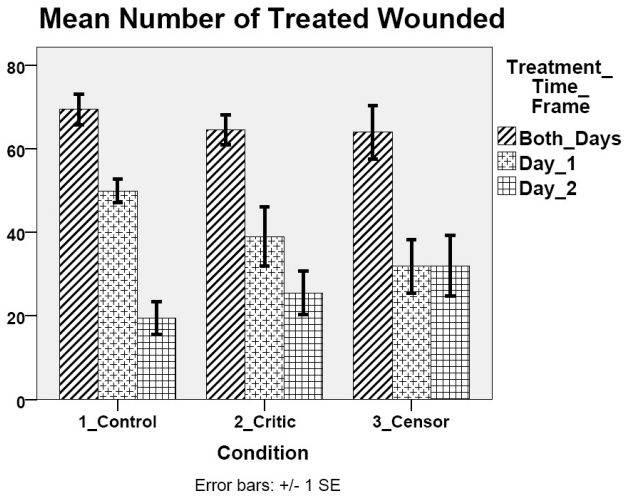


FIGURE 10. Mean Number of Wounded Evacuated by Party A

Figure 7 illustrates the effectiveness of agent interventions at preventing policy violations. That is, it shows the number of times that an agent detected a violation, acting either as a critic or censor, and the user complied with the agent's feedback. All three types of policy violations are represented. ANOVAs found differences for both parties across the three conditions: (1) both ($F_{2,57} = 23.13, p < 0.001$), (2) Party A ($F_{2,27} = 8.58, p = 0.001$) and (3) Party B ($F_{2,27} = 15.35, p < 0.001$). Post-hoc tests revealed consistent differences ($p < 0.001$) for all three pairwise comparisons: control vs. critic, control vs. censor, and critic vs. censor.

Figure 8 reports the mean number of overrides of communication prohibitions. Communication policies are ones where the parties must communicate through sending messages to each other, for example Party A requesting an escort from Party B or Party B granting an escort to Party A.

Additionally, we analyzed the data of policy violations for policies that were mission impacting.

Mission impacting policies are the ones that impact actual deployment of resources to attain mission goals, for example deployment of helicopters or jeeps to evacuate wounded or deployment of field paramedic units to treat wounded on site. An example of a mission impacting policy for Party A is, "If you do not have a commitment of escort on route R on day D from Party B, then you are forbidden from deploying ground and air vehicles along route R on day D". This policy clearly impacts the deployment of resources, and breaking this policy may endanger the safety of the wounded. On the other hand, a policy such as, "If your intelligence is from insurgent sources, you are forbidden from sharing it with Party B", does not directly affect A's or B's resource deployment or attainment of mission objectives, although it may affect the relationship of Party A with the insurgents (the insurgents may consider violation of this policy as a breach of promise of neutrality on the part of Party A). In Figure 9, although not reaching statistical significance ($F_{2,27} = 2.18, p = 0.133$), we see that the subjects in the critic condition made the fewest number of mission impacting policy violations (MIPVs), and was the only group to have any members (4 out of 10) with zero such errors. We should also note that the censor agent would not provide feedback on policy violations that were introduced by the addition of a plan step, so it is not surprising that the control and censor conditions have more closely-related means.

Figure 10 presents the mean number of wounded treated in the plans created by Party A in the three

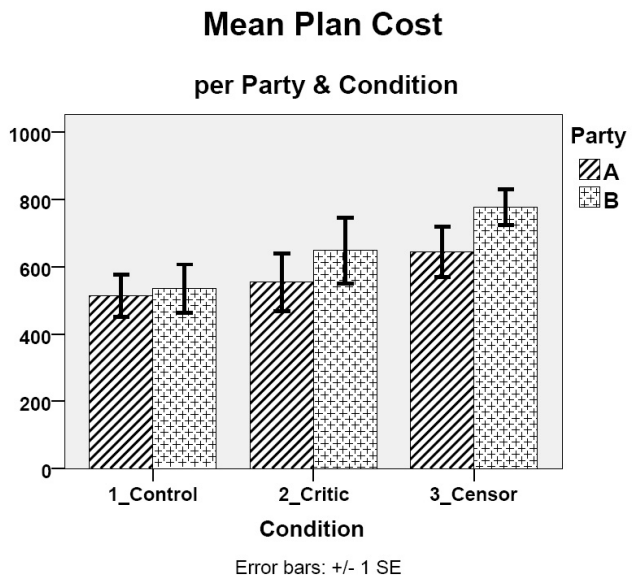


FIGURE 11. Mean Plan Costs for Party A and Party B

conditions.[7] Evacuating and treating the wounded was the mission objective for Party A, explicitly given in A's briefing materials. Since this was the mission objective of A, we expect that most of party A's efforts would be in fulfilling this objective by treating as many wounded as possible. The figure shows the mean number of wounded that were treated each day and also the mean number of wounded treated over the two days of the plan. A weak difference was found for day 1 ($F_{2,27} = 2.50, p = 0.101$), with a clearer distinction between the control and censor conditions in post hoc tests ($p = 0.035$). In the censor condition, as in the control condition, the maximum number of wounded treated on the first day was 60. The only difference between the two conditions was that the subjects in the censor condition were not allowed to violate any policies. The reason the human subjects in the censor condition did not have a perfect score of treating all 75 wounded on day 1 is because they had overestimated the capacity of a paramedic unit to treat wounded, not due to a policy violation. Otherwise, they had secured all the necessary clearances and escorts. The top performer in the critic condition, and the only subject across the three conditions to treat all 75 wounded on day 1 without violating any policies, did so with the critic agent providing advice 13 times in the course of the session.

Figure 11 shows the mean plan cost in each condition for the two parties. The cost of each plan step was determined by multiplying the cost of a resource deployment by the number of times the resource must traverse a route. Thus, for example, the cost of a deployed paramedic unit would be multiplied once, since paramedic units do not return to the hospital, but the cost of deploying a jeep would be multiplied by two for every round-trip between destination and hospital it is scheduled to make. Although the subjects were required to calculate the costs of their plans, many subjects made errors in their calculations, hence this statistic is based on the corrected plan cost.

Users in the aided conditions generated plans with marginally higher costs ($F_{2,57} = 3.12, p = 0.052$) with the greatest difference occurring between the control and censor groups ($p = 0.016$).

Figure 11 shows that for both Party A and Party B, plans in the control condition had the least cost, followed by plans in the critic condition. Plans in the censor condition had the highest cost. The results of Figure 10 and Figure 11 support our experimental hypothesis for both parties, namely performance as measured by cost was higher in the control condition than in the aided conditions.

4. DISCUSSION

The results reported in Section 3.2 have shown that agents can have a positive impact on the enforcement of policies. We saw that in the unaided case, individuals would make on average from $7 - 10$ policy violations, with all individuals making at least three policy violations, per session. We also saw that the rate of individuals overriding agent advice dropped to 3 for party A and 1 for party B, with many individuals not making any policy violations at all.

This article reports experiments with two types of aiding strategies: a critic and a censor. Neither form of assistance prompted the human subjects to attempt more or fewer policy violations, as evidenced by the lack of statistical significance in differences shown in Figure 5. There was a minority of human subjects, however, that adjusted the ways in which they used the agents as a result of the type of agent intervention. For example, in the censor condition, some users would try to exhaustively generate-and-test communications for granting clearances or committing to escorts against the censorship of the agent.

There is a slight degradation of performance between subjects in the control and in the critic conditions (figures 10 and 11). We hypothesize that this behavior is due to the critic agent focusing the user's attention on avoiding policy violations rather than on the objectives of their task.

The following characteristics were observed about the censor condition.

- Of the two agents, critic and censor, the censor agent was the most effective at preventing policy violations.
- The censor agent was unable to provide feedback on mission impacting policy violations (MIPVs) that were introduced as plan steps, so its performance

[7] Party B's metric of number of strongholds defeated was not discriminating for evaluating subject performance.

in reducing MIPVs cannot be distinguishable from the control condition.
- The Party A subjects in the censor condition were most distracted from their mission objective of treating as many wounded as possible on day 1. We hypothesize that the lack of direct feedback to the user committing the violation may be the cause.
- Similarly for both parties, the plan costs were greatest in the censor condition. We hypothesize that the lack of direct feedback to the user committing the violations may cause confusion, distracting the user from being mindful of their plan costs.

We also collected data on the efficiency of the planning process. In particular, we found that the number of messages exchanged by the subjects was approximately equal in the control and censor conditions and slightly higher in the critic condition. We think this may be due to the fact that the human subjects, alerted by the critic about their policy violations, were more careful to try to rephrase their messages to avoid policy violations.

For future work, we plan to investigate the reasons why users may override an agent's advice.

5. RELATED WORK

Policy-based reasoning (e.g. Grid computing [6, 7], KAoS [8], Ponder2 [9, 10]) is used to regulate access to shared resources or to enforce certain properties (quality of access / service / data etc.) [11]. The resources that the policies and norms protect are electronic documents [8, 9] or access to computational resources [7, 6]. A classic assumption is that what is not explicitly permitted is prohibited (e.g. granting access rights) and the concept of an obligation is not present. Recent work [12, 13] introduces richer concepts for describing policies (e.g. "obligation policies"). In the situation of sharing computational resources, the tacit assumption is that all interested parties are cooperative, and therefore the best solution is to optimize a fair allocation of the shared or contended resource among all individuals that wish to use it. All individual agents agree to this assumption by agreeing to be part of a virtual community of potential users of the contended resource, and agree to abide by a fair use policy for that resource. Sanctions are an effective means of enforcing policy compliance by disobedient agents.

Our concept of a policy is inspired by previous work [3] and is aligned with research into normative systems, in particular work on norm-governed agency [14, 15, 16] and Electronic Institutions [17, 18] with [19] discussing how human norms can be represented in a form that allows agents within e-institutions to process them. In this paper, we expand on work presented in [20] and describe a specific application of normative agents, where agents do not form virtual organisations, but observe the behaviour of human planners within a coalition. In this setting, the agent is focussed on understanding current knowledge held by the human planners – what intelligence they hold about the current state of the world, what commitments they received and made, as well as what requests they expressed and what their current plan is. It is important to note that in our scenario, there is no consensus on a shared set of policies nor is it the goal to arrive at such an agreement. The contended resources that are desired by more than one stakeholder in our scenario are the pathways to the cities, either by land or air, and, in the case of party A, some of the military hardware to get there. It is infeasible to model all the possible users of the resources as belonging to the same virtual organization. Even as there is some degree of cooperation, it is not possible to form shared goals among all the stakeholders. For example, party B shares an interest with party A to achieve safe passage, whereas party A may have secret communications with other stakeholders to achieve the same goal. The norms serve as self-interested general rules of safety rather than guidelines for cooperative and fair sharing.

By using a rule-based language for encoding policies and, consequently, a rule engine such as the Java Expert System Shell (Jess) [4] for processing these specifications, we followed a similar implementation path as described in [21], in particular implementing the reasoning about norms in Jess.

6. CONCLUSION

In this paper, we discussed difficulties in establishing joint plans within coalitions in the face of individual goals and diverse policies of the coalition members. For a planner, it would be easiest to operate without any restrictions, constraints or regulations on the operations that may be added to a plan. In a social context and, in particular, a diverse one such as coalitions of independent partners, this is not possible. As we showed, detailed and, sometimes, even conflicting policies have to be dealt with in practice, when coalitions try to engage in collaborative planning. With such policies and restrictions in place, planning becomes more complicated. We therefore advocate agent support for policy-based planning activities within coalitions. In this paper, we demonstrated how agents can be integrated into the dialogical process of human planners establishing a collaborative plan. We described two agent-based strategies for assisting the collaborative planning process: (a) a "critic" that provides active feedback about the fulfillment of policies and (b) a "censor" agent that silently manipulates the interaction between human planners so that their interaction and information exchanged takes place according to given policies. We have outlined an experimental framework that allows us to evaluate the

effects of these strategies in the context of a military-humanitarian scenario and presented data that shows the impact of agent support on the planning outcomes.

ACKNOWLEDGEMENT

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

[1] Cohen, P. R. and Levesque, H. J. (1991) Teamwork. *Noûs*, **25**, 487–512. Special Issue on Cognitive Science and Artifical Intelligence.

[2] Sergot, M. (2001) A Computational Theory of Normative Positions. *ACM Transactions in Computational Logic*, **2**, 581–622.

[3] Kollingbaum, M. (2005) Norm-governed Practical Reasoning Agents. PhD thesis University of Aberdeen.

[4] Friedman-Hill, E. (2003) *Jess in Action*. Manning.

[5] Wheller, V. and Harmer, A. (2006) Resetting the rules of Engagement. Trends and Issues in Military-HumanitarianRelations. Technical report. HPG Research Report.

[6] Boella, G. and van der Torre, L. (2006) Security policies for sharing knowledge in virtual communities. *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, **36**, 439–450.

[7] Pearlman, L., Welch, V., Foster, I., Kesselman, C., and Tuecke, S. (2002) A community authorization for group collaboration. *Proceedings of the Third International Workshop on Policies for Distributed Systems and Networks (POLICY'02)* 10pp. IEEE Computer Society. 0-7695-1611-4/02.

[8] Suri, N., Bradshaw, J., Burstein, M., Uszok, A., Benyo, B., Breedy, M., Carvalho, M., Diller, D., Groth, P., Jeffers, R., Johnson, M., Kulkarni, S., and Lott, J. (2003) Toward DAML-based policy enforcement for semantic data transformation and filtering in multi-agent systems. *Proceedings of the Third Joint Conference on Autonomous Agents and Multi Agent Systems (AAMAS03)*, 14-18 July.

[9] Russello, G., Dong, C., and Dulay, N. (2007) Authorisation and Conflict Resolution for Hierarchical Domains. *IEEE Workshop on Policies for Distributed Systems and Networks*.

[10] Charalambides, M., Flegkas, P., Pavlou, G., Bandara, A., Lupu, E., Russo, A., Dulay, N., Sloman, M., and Rubio-Loyola, J. (2005) Policy Conflict Analysis for Quality of Service Management. *6th IEEE Workshop on Policies for Distributed Systems and Networks (Policy 2005)*.

[11] Boella, G., Sauro, L., and van der Torre, L. (2004) Social viewpoints on multiagent systems. *AAMAS '04: Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems*, Washington, DC, USA, pp. 1358–1359. IEEE Computer Society.

[12] Zhao, H., Lobo, J., and Bellovin, S. M. (2008) An Algebra for Integration and Analysis of Ponder2 Policies. *2008 IEEE Workshop on Policies for Distributed Systems and Networks POLICY 2008*.

[13] Kagal, L. and Finin, T. (2007) Modeling Conversation Policies using Permissions and Obligations. *Journal of Autonomous Agents and Multi-Agent Systems*, April, pp. 197–206. Springer-Verlag.

[14] Dignum, F. (1999) Autonomous Agents with Norms. *Artificial Intelligence and Law*, **7**, 69–79.

[15] Lopez y Lopez, F., Luck, M., and dÍnverno, M. (2004) Normative Agent Reasoning in Dynamic Societies. *Third International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 2 (AAMAS2004)*.

[16] Oren, N., Panagiotidi, S., Vazquez-Salceda, J., Modgil, S., Luck, M., and Miles, S. (2008) Towards a Formalisation of Electronic Contracting Environments. *Proceedings of Coordination, Organization, Institutions and Norms in Agent Systems, the International Workshop at AAAI 2008*.

[17] Rodríguez-Aguilar, J.-A. (2001) On the Design and Construction of Agent-mediated Electronic Institutions. PhD thesis Institut d'Investigació en Intel·ligència Artificial (IIIA), Consejo Superior de Investigaciones Científicas (CSIC), Campus UAB, Bellaterra, Spain Spain.

[18] Vasconcelos, W. (2004) Norm Verification and Analysis of Electronic Institutions. *Lecture Notes in Artificial Intelligence*. Springer-Verlag.

[19] Vazquez-Salceda, J., Aldewereld, H., Grossi, D., and Dignum, F. (2008) From Human Regulations to Regulated Software Agents' Behaviour: connecting the abstract declarative norms with the concrete operational implementation. *Artificial Intelligence and Law*, **16**.

[20] Burnett, C., Masato, D., McCallum, M., Norman, T. J., Giampapa, J., Kollingbaum, M. J., and Sycara, K. (2008) Agent Support for Mission Planning Under Policy Constraints. *Proceedings of the Second Annual Conference of the International Technology Alliance*, pp. 100–107.

[21] García-Camino, A., Noriega, P., and Rodríguez-Aguilar, J. A. (2005) Implementing Norms in Electronic Institutions. *Procs. 4th AAMAS*.