Policy-driven Planning in Coalitions - a Case Study

Martin J. Kollingbaum\textsuperscript{[1,\ast]}, Joseph A. Giampapa\textsuperscript{[1,\dagger]}, Katia Sycara\textsuperscript{[1,\ddagger]}, and Timothy J. Norman\textsuperscript{[2,\diamond]}

\textsuperscript{1}Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213, USA  
\{\textsuperscript{\ast}mkolling,\textsuperscript{\dagger}garof,\textsuperscript{\ddagger}katia\}@cs.cmu.edu  
\textsuperscript{2}Dept. of Computing Science, Univ. of Aberdeen, Aberdeen AB24 3UE, UK  
\textsuperscript{\diamond}t.j.norman@abdn.ac.uk

Abstract. A collaborative planning effort between partners that form coalitions may be complicated by policies that regulate what actions they may deploy in their plans and, in particular, what information they are allowed to exchange during the planning process. We are interested in situations where coalitions have to be formed ad-hoc without much co-training. For this, we investigate how agents can support human planners in producing good plans while observing the normative standards that regulate their planning and communication behavior. Based on an implementation of such norm-processing agents, we conducted a set of experiments, where human test subjects were conducting collaborative planning tasks under the guidance of these agents. A summary of experimental results is provided in the paper.

1 Introduction

Constructing joint plans within coalitions in time-stressed situations poses a particular challenge, especially in the face of individual goals, self interest and with coalition members having only limited co-training for recognizing and resolving their differences. During planning, coalition partners may also have to take into account specific policies that describe what their obligations, permissions and prohibitions are in terms of the actions comprising a plan and the communication necessary to coordinate planning activities with coalition partners. With such policies in place, human planners are under pressure to produce high-quality plans, while adhering to all their obligations and prohibitions as described by their policies. In this paper, we describe how agents can assist human planners in a monitoring and advisory capacity in their policy-driven (norm-driven) planning efforts. The scope of this work is to investigate how policies (or norms) influence collaborative planning and whether agents can ease the cognitive burden for human planners to create high-quality shared plans in the face of such policies.

In the following, we describe the development of agents that observe both communication and planning activities and provide feedback on how these actions might impact on a human planner’s normative situation. We conducted a set of experiments to investigate the effectiveness and impact of such agents in
a collaborative planning scenario and provide insights about the results of these experiments.

2 Agents monitoring Policies

We consider the use of agents to monitor communication and planning activities of coalition partners and reason about possible policy violations. As the human planners are involved in a distributed planning problem, they may not be aware of potential differences and conflicts between the different policies held by the coalition partners. This may adversely impact on the planning process due to the time and effort it may take to identify and reconcile policy conflicts. Also, situations may occur that call for certain policies to be violated in order to produce a viable plan. We base our notion of policies on normative concepts, in particular, these are:

– the obligations that must be fulfilled,
– the prohibitions that constrain/forbid particular actions, and
– the permissions that define the range of actions that are allowed.

We regard policies to be relevant (or active) during a planning activity under specific conditions only. Due to this conditional nature, a human planner may not recognize that policies are in conflict with those of coalition members, that violations occurred or that policies are relevant to the present circumstances. Based on what planning actions are taken or what is communicated, this normative position may change – for example, an obligation may become fulfilled or a specific planned action violates a prohibition. These changes have to be observed and remembered by the human planner (which is a cognitively demanding task) and make collaborative planning under such norms/policies a complicated task. Agents can provide assistance to a human planner by detecting and advising a planner when policies become active, when policy violations occur and may also propose courses of actions that may resolve such conflicts.

Agents, as we utilize them, do not form coalitions themselves or are part of a virtual organization. These agents operate in a supportive role to a human planner in order to ease the cognitive burden on human planners during a collaborative planning effort. The agent is assigned solely to a specific human planner and operates in a monitoring, controlling and/or advising capacity. 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 in order to provide the human planner with warnings in case violations were to occur and possibly advise how to rectify such a situation. The agent, therefore, has to maintain a representation of the normative position of the human planner, which is a potential one (and not the actual situation). Based on this “outlook” at a potential future normative situation, the agent can reason about appropriate responses to that.

In the experiments performed, we were interested in how an agent can support a collaborative planning activity. In particular, we investigated two supporting strategies or “aiding conditions” for an agent:
– in the critique condition, the agent detects policy violations that are incurred by the human planners in their communication and planning behavior. In case of a violation, the agent either (a) intercepts the sending of a message or (b) interrupts the planning of actions in order to inform the human planner about the set of policies that become activated due to these intended and violating actions – the planner can then decide whether to adhere to such an advice or to ignore the agent and intentionally violate a policy;
– in the censor condition, the agent still monitors the activities of the human planner, but silently interferes with the communication by deleting offending parts of the exchanged messages (or blocking them completely) in order to avert 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 int only intercepts and forbids the transmission of messages that contain policy violations. The critique agent does not force the user to a particular action, it merely provides advice and suggestions, which the user can accept or reject. Both conditions are compared to a third control condition (the unaided condition), where the test subjects did not have any agent support.

2.1 Modelling Policies

Policies are given to the human planner in a verbalized form. The following example is taken from our example domain outlined in subsequent sections:

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 policy will become relevant to the human planner in the course of planning such a rescue operation, if the deployment of this specific resource is intended. In becoming relevant, it adds to the current “social burden” of the human planner – it 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 have expired. 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 acquired a commitment of escort along route R 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 to deploy 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: the point of view from which 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 only in cases where this does not obtain (explicit permissions may be included to specify particular exceptions to a given prohibition, e.g. “You are ONLY permitted, IF ...”). In the same way as obligations, prohibitions must be augmented with conditions that indicate the circumstances under which a violation occurs. Due to the specific way the agent processes norms (as described later), prohibitions have an activation condition that describes the (set of) violating circumstances – with an activation of a prohibition, its violation is indicated, whereas in its deactivated state, it is regarded as not violated.

In accordance with definitions used in previous work [1], we describe here the normative position of the human planner, which is monitored by the agent (and, therefore, is also the normative position of the agent), as the set \( \Omega \) of currently activated policies:

**Definition 1.** The set \( \Omega \) comprises the currently instantiated policies, containing the permissions given, the obligations that must be fulfilled, and the prohibitions that are potentially under threat of violation.

Note, that the agent intercepts the actions of the monitored human planner, before these actions actually take place. This allows the agent to assess the normative consequences of an action, before harm is done and can inform the human planner accordingly. If \( \Omega \) contains activated prohibitions then the agent signals potential violations and not actual ones. By providing information about those violations and active obligations to the human planners, the monitoring agent may be able to motivate them to correct their behavior.

With respect to an implementation of such an agent, with each occurrence of either a communication or planning action, we regard the set \( \Omega \) being discarded, the activations of all policies checked afresh and a new set \( \Omega' \) created. \( \Omega' \) 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 $\Omega^{\text{final}}$, maintained for an individual coalition partner by its monitoring agent, can have the following states:

(a) $\Omega^{\text{final}}$ is empty or contains only permissions – the human planner has a clean record with all obligations fulfilled and no prohibitions violated, or
(b) $\Omega^{\text{final}}$ 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, $\Omega^{\text{final}}$, represents the actual normative state of the human planner.

Our representation of policies follows our earlier work [2]. 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 $\text{Expr}$ as the set of all possible well-formed formulae comprising first-order predicates over terms (constants, variables and the operators $\land$, $\lor$ and $\neg$), then a policy can be defined in the following way:

Definition 2. A policy, expressing an obligation, permission, prohibition is a tuple $\langle \nu, \rho, \varphi, a, e \rangle$, where

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

3 Planning and Communication Environment

For agents to become operational, they must have access to plans in development and communication activities. We use a traditional forward-chaining mechanism (expert system shell Jess [3]) to implement the policy reasoning mechanism for an agent. According to our model, a policy will experience activations and de-activations under specific circumstances. In order to correctly implement their
activation and de-activation, each policy is expressed by a set of rules and data structures recording such an activation state.

As we noted before, the agent operates in a fixed monitoring cycle:

(a) detect the current situation changed by arriving messages expressing the coalition partners’ commitments for action or revealed intelligence, as well as new planned actions,
(b) reason about these changes, and
(c) 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.

3.1 Conversations during Planning

In terms of communication between human planners, we strongly simplified and restricted the way conversations within a coalition takes place. In order for agents to easily monitor communication and reason about the messages exchanged, human planners converse in writing, using a specific set of message types.

In this conversation, a planning party may request another coalition member to commit to a specific action or to provide particular information. The planning party itself may also be the target of such a request. On the other hand, planning parties may pro-actively offer information or commit to specific actions as they develop their own plan. Finally, a planner may have to change its plan and, therefore, withdraw commitments or withdraw requests or offers. By identifying these general types of conversations, we can establish a set of message types:

<table>
<thead>
<tr>
<th>Performative</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>REQUEST</td>
<td>commitment, information</td>
</tr>
<tr>
<td>OFFER</td>
<td>commitment</td>
</tr>
<tr>
<td>INFORM</td>
<td>information</td>
</tr>
<tr>
<td>ACCEPT</td>
<td>commitment</td>
</tr>
<tr>
<td>PROVIDE</td>
<td>commitment, information</td>
</tr>
<tr>
<td>DENY</td>
<td>commitment, information</td>
</tr>
<tr>
<td>WITHDRAW</td>
<td>commitment, request for information</td>
</tr>
</tbody>
</table>

Table 1. Message Types

Messages according to these types are used in conversations that follow particular transitions as shown in figure 1. Dialogs for requesting a commitment or pro-actively offering it are shown. A REQUEST (for commitment) has to be answered with either a PROVIDE (a commitment) or a DENY. In the proactive case, an OFFER can be answered with either an ACCEPT or a DENY. Requests, commitments, offers and the acceptance of offers can be withdrawn in a separate WITHDRAW conversation. The state transitions in figure 1 show
annotations such as A:RCx etc. These indicate that, for example, party A sends a REQUEST for a commitment (the x is a placeholder for particular domain-specific information).

Fig. 1. Request and Offer Conversations

Not shown is INFORM, as it is a special case. It allows to disclose particular information proactively such as the set of plan steps in the planner’s own plan or confidential information, without requests from another party or expecting a reply.

For planning, a human planner is provided with a set of domain-specific actions and basic manipulation operations to assemble a plan.

4 Scenario

We chose rescue missions as an 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, rescue helicopters, Jeeps, etc. 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 its own military goals by attacking enemy strongholds.

In support of the communication necessary during planning, a set of domain-specific speech acts is provided according to the types of performatives outlined above. Previously, we determined that there is a need for the performatives REQUEST, OFFER, INFORM, PROVIDE, ACCEPT, DENY and WITHDRAW to enable regimented conversations between human planners. We also saw that these conversations are subject to the exchange of either commitments or specific information necessary to the decision process of whether a specific action should be planned or not. The kind of commitments or information exchanged with these messages is domain-specific. In our case, we assume that party A and B exchange the following commitments:

- guarantee safety of a road (B to A)
- provide escort (B to A)
- evacuate (A to B)
- dispatch (A to B)

For example, party A can issue a REQUEST to party B asking for an escort. Party B would then either provide this commitment or deny the request. Information about the following aspects may be exchanged as well:

- intelligence
- intelligence source
- plan step
- specifics (request/provide more details regarding what road/resource/day)

With that, the parties are enabled to disclose intelligence or the source of intelligence, or to communicate planned activities and whether a road is safe for evacuating wounded etc.
5 Experiments

The human planner interacts with the agent via the application as depicted in figure 2 – the interface is shown in a mode where agents operate in the “critique” condition. The purpose of this interface is to provide the user with the possibility of sending messages to coalition partners, maintain a plan and interact with an agent. If the agent operates in the “critique” condition, it will report back on the results of its monitoring the communication and planning activities of the human planner. The following information is provided to the user in the top section of this interface (characterizing a potential normative situation):

- reminders about the current set of active obligations (right upper window)
- the set of potential violations of prohibitions (left upper window) – the user can either ignore this warning (intentional violation of a norm) or accept it

If the user accepts the warning of a policy violation, the offending action – sending a message or adding a particular plan step will be aborted. If the user ignores the warning these actions would go ahead.

In addition, the user manages its communication by assembling messages via pull-down menus (left part of the interface) or uses a chat window to directly converse with a coalition partner (which is not monitored by the agent), as well as its plan in the right part of the interface.
5.1 Experimental Task

The experiments conducted are characterized by the interaction of two human test subjects acting as planners in the role of the humanitarian organization (Party A) and the military organization (Party B). Both test subjects are provided with details about their private goals, resources, intelligence, their capabilities, and policies. Both parties are given different maps that outline locations or destinations, from where – in the case of Party A – injured people must be evacuated, or that represent insurgent strongholds that must be defeated by Party B. These destinations have numerical requirements – in case of Party A, a specific number of wounded to be evacuated, in case of Party B, insurgent strongholds have a specific resistance value that must be overcome by military means and incur costs.

Thirty teams, each comprising two paid subjects, were recruited to participate in the study. 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. 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 warm up 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.

6 Results

The results of these experiments have shown that agents can have a positive impact on the enforcement of policies. We saw that in the unaided experiment condition (no agent monitoring and feedback), individuals would make on average from 7 – 10 policy violations, with all individuals making at least three policy violations, per session. With reference to Figure 3, we saw that the rate of individuals deliberately violating policies dropped to a median average of one violation per session, 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 4. 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.

The following characteristics were observed about the critic condition.
– 60% of the subjects in both Parties A and B felt that it was necessary to override the critic agent in order to complete their plans (Figure 3).

– only one subject out of thirty actually reached the mission objective of treating 100% of the wounded on the first day of the mission. The subject did so without violating any policies, but with the assistance of 13 interventions from the critic agent.

– there is a slight degradation of performance between subjects in the control and in the critic conditions. 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 mean plan cost is slightly higher in the critic condition than in the control condition. When considering, however, that 4 out of 10 subjects in the critic condition did not violate any mission-impacting policies, it is possible that Party A’s plan costs are better approximations of the true plan costs.

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. Only 3 out of 20 individuals actually circumvented the censorship of the censor agent, each one committing only 1 violation.

– The censor agent was unable to provide feedback on mission impacting policy violations (MIPVs) that were introduced as plan steps, so its performance in reducing MIPVs cannot be distinguished 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.

Results from these experiments provide us with feedback that informs the design of future policy-enforcing agents. For example, the censor agent was clearly the undisputed best enforcer of policy. Its enforcement reliability was countered by a significant increase of plan costs and distraction from the mission objectives for Party A. The critic agent immediately flagged policy violations, so it was possible for at least one subject in all three conditions to achieve the perfect plan (e.g. treating all the wounded on day 1), as well as enabling others to have zero mission impacting policy violations. A possible next agent to test would be a critiquing censor agent that: is capable of critiquing plan steps, provides direct feedback to the user that it is censoring, and does so without allowing the censored subject to override it.
7 Related Work

Policies have been used in disparate fields, ranging from security models of programming languages to the management of resources in distributed IT systems [4]. In this context, policies are usually regarded as “permissions” that allow the performance of specific actions such as access to data or the use of network resources. This view has its limitations: it is assumed that what is not explicitly permitted is prohibited and the concept of an obligation is not present traditionally. Recent work [5, 6] introduces richer concepts for describing policies (e.g. “obligation policies”).

Our concept of a policy is strongly aligned with research into normative systems, in particular work on norm-governed agency, virtual organizations [7–10] and Electronic Institutions [11–14]. In this paper, we expand on work presented in [15] 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. With such monitoring agents in place, we share similar concerns as those posed in the context of Electronic Institutions, as outlined, for example, in [13, 14]. In a very similar fashion, we must monitor the communication behavior of coalition partners, assess the eligibility of the messages exchanged and reason about the current normative state. By using a rule-based language for encoding policies and, consequently, a rule engine such as the Jess Expert System Shell for processing these specifications, we followed a very similar implementation path as described in [13, 12], in particular implementing the reasoning about norms in Jess [16].

8 Conclusion

In this paper, we discussed difficulties in establishing joint plans within coalitions in the face of self interest, 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 and optimal plans may be hard to achieve. 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 results.

9 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


