# Arguments over Co-operative Plans


Rolando Medellin-Gasque[1], Katie Atkinson[1], Peter McBurney[2], and Trevor Bench-Capon[1]

[1] University of Liverpool, Department of Computer Science, Liverpool UK
{medellin,katie,tbc}@liverpool.ac.uk
[2] King's College London, Department of Informatics, London UK
peter.mcburney@kcl.ac.uk



**Abstract.** Autonomous planning agents that share a common goal should be able to propose, justify and share information about plans. To reach an agreement on the best plan, strategies for persuasion and negotiation can be used by agents in order to share their beliefs about the world and resolve conflicts between the agents. We present an argumentation scheme and associated critical questions to create and justify plan proposals where plans are combinations of actions requiring several agents for their execution. An analysis of different ways in which actions can combine is presented and then associated with the argumentation scheme and the critical questions. We believe these elements are necessary to enable agents to engage in rational debate over co-operative plan proposals.



Key words: plan proposal, argumentation schemes, critical questions, co-operation.


## 1 Introduction

Planning in Artificial Intelligence is concerned with the automatic synthesis of action strategies from a description of actions, sensors and goals [11]. The planning literature has been focusing in recent years on overcoming strong assumptions about plan generation. The complexity of distributed systems restricts the application of single-agent planning strategies to distributed problems usually because a local agent view is not sufficient. A common assumption in AI planning is that the planner has accurate and complete knowledge of the world and the capabilities of other agents. We want to overcome this assumption and provide strategies where agents with different views of the world are able to propose and justify complex plans. Our goal is to provide autonomous agents with tools to justify plans in terms of acceptable arguments and enable them to critique and defend plans to refine them and choose between them. The dialogue is intended to support planning tasks such as plan modification, choosing the best plan and even establishing coordination strategies for the execution of a plan.

In this paper we present an argumentation scheme to propose and justify plans based on the argumentation scheme for action proposals of Atkinson *et al.* in [3]. We extend the concept of action used in [3] with action-elements taken from the PDDL 2.1 Planning Specification [3] presented in [10] such as time constraints and invariant conditions.

[3] Planning Domain Description Language (PDDL) is an attempt to standardize planning domain and problem description languages developed for the International Planning Competitions.

Thus, this work extends the action proposal model of [3] to more complex types of action-proposals involving several durative actions performed by several agents. An analysis over different ways to combine actions to form plans is also presented in order to create more specific critical questions. The analysis is based on interval algebra proposed by Allen in [1]. Allen's interval relations define the basic relations between time intervals. We relate time intervals with the action duration in order to define ways in which actions may be combined in a co-operative plan. Furthermore, we present critical questions grouped in 6 categories that address specific elements of the plan-proposal. As a basis to formalize our argumentation scheme we will use a formal model developed in [17], an Action-based Alternating Transition System (AATS). This transition system defines actions that may be performed by agents through the states from which these could be performed and the states that will result, with a particular focus on the simultaneous action of a group of agents. This makes AATSs especially suitable for situations where co-operation is important.

The paper is structured as follows: Section 2 presents the action representation and proposal including action combinations. Sub-section 2.3 introduces the AATS notation to formalize the action proposal in sub-section 2.4. In section 3 we present the plan proposal as an argumentation scheme of AATS models together with critical questions associated to the extended action and the ways in which actions can combine. In section 4 we develop an example using the elements presented in this paper. Section 5 comments on related work and finally, in section 6, we conclude the paper and discuss future research work.

## 2 Action Representation and proposal

### 2.1 Action representation

Actions usually are represented as operations an agent is able to perform from a state where some preconditions hold. We want to extend this action definition and incorporate elements useful when representing and reasoning about temporal plans. We use actions as presented in the PDDL 2.1 specification [10] which have elements to express temporal domain descriptions over plans. In the PDDL 2.1 specification, instead of having an action with preconditions and effects, actions are represented as durative actions with elements to express more precisely temporal conditions and effects. The durative action representation is as follows: initially, the action can start at a point in time when a set of preconditions hold, once the action start, " start effects" become true. Action has a duration and "invariant conditions" (distinct from preconditions) and are accessible though the duration of the action. Actions are not black-boxes and access to start effects is available during the performance of the action. The end of the action is given by "termination conditions" where upon, end effects become true [4]. The planning community is still developing ways to create planners that handle temporally extended actions. Our intention in presenting this durative action representation is to consider all the elements

[4] This model still represents a simplified model of time; durative actions could be extended to allow effects to be asserted at arbitrary points during the interval of execution, or to be a function of duration ("until" actions).

needed by agents to engage in argumentative dialogues over co-operative plans. In the next sub-section we will define different ways in which actions may combine to form a plan.

### 2.2 Action combinations

We now define the way in which actions can be combined to form plans. By action combinations we mean the different ways in which atomic actions could be combined in a plan in terms of concurrency, repetition and temporal aspects. Even if there is just one action there could be variants such as its periodicity or whether the action execution is optional. Two or more actions could be defined in a plan as a sequence (as in classical AI planning) or as a set of actions with no particular order (partial-order planning) that could, but need not, overlap. We want to cover both cases and others focusing on aspects such as the order of the actions and their periodicity.

The analysis is based on the interval algebra proposed by Allen in [1]. Interval algebra is based on the 13 possible primitive relationships (6 of which are inverses) between two time intervals (Figure 1). We apply a similar model to combinations of actions. Most of the interest in Allen's representation for time intervals comes from a mechanism by which the time relationships between the pairs of intervals can be propagated through the collection of all intervals. The notion of disjunction of interval relationships can be used to declare multiple paths and interactions. This idea gives us reason to think this analysis could be extended for larger plans. We add to the Allen list cases focusing on specific properties such as the periodicity, optionality and interleaving of actions. The 14 cases are presented in the following list, for arbitrary actions $alpha$ and $beta$:

AC1.- Action $\alpha$ occurs exactly *k* times, where *k* is a non-negative integer ($\alpha(k)$).
AC2.- Optionally, action $\alpha$ occurs exactly *k* times ($\alpha(k, o)$).
AC3.- Action $\alpha$ occurs from 1 to *k* times, where $k > 1$ ($\alpha((1 - k)$).
AC4.- Optionally, action $\alpha$ occurs from 1 to *k* times ($\alpha(1 - k, o)$).
AC5.- Action $\alpha$ precedes $\beta$ ($precedes(\alpha, \beta)$) (Figure 1a).
AC6.- Action $\alpha$ meets $\beta$ ($meets(\alpha, \beta)$) (Figure 1b).
AC7.- Action $\alpha$ overlaps $\beta$ ($overlaps(\alpha, \beta)$) (Figure 1c).
AC8.- Action $\alpha$ starts $\beta$. ($starts(\alpha, \beta)$) (Figure 1d).
AC9.- Action $\alpha$ is entirely in action $\beta$ ($entirely(\alpha, \beta)$ (Figure 1e).
AC10.- Action $\alpha$ finishes $\beta$. ($finishes(\alpha, \beta)$ (Figure 1f).
AC11.- Action $\alpha$ equals $\beta$. ($equals(\alpha, \beta)$ (Figure 1g).
AC12.- Action $\alpha$ or action $\beta$ but not both ($\alpha|\beta$))
AC13.- Both actions interleaving concurrently (overlapping) over periods of time until completion of both ($iC(\alpha, \beta)$.
AC14.- Both actions executed not concurrently over periods of time until completion of both. $i(\alpha, \beta)$.

The purpose of this analysis is to cover all of the ways in which actions may be combined in a plan with questions that match the specific action combinations. This analysis covers cases where plans are formed of one or two atomic actions. Perhaps,

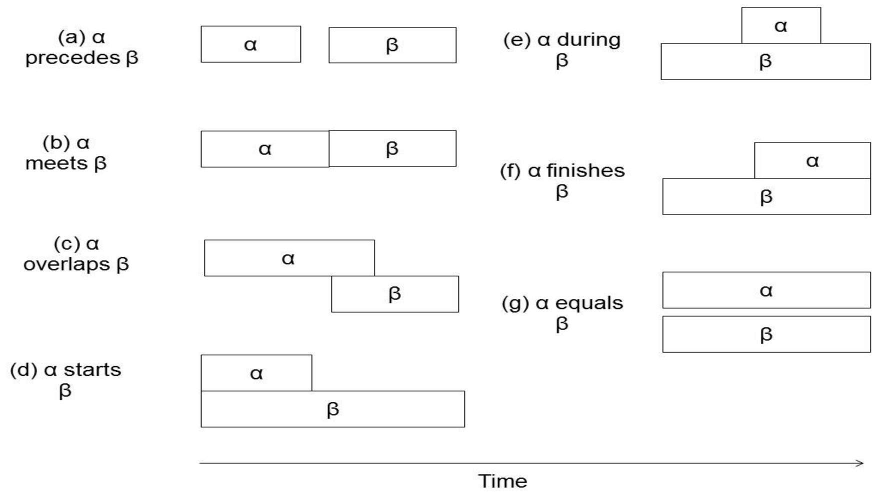


**Fig. 1.** Allen's possible primitive time relationships between two intervals labelled $\alpha$ and $\beta$. Time is represented by the horizontal axis. (a) to (f) have inverses.

plans comprising one or two actions seem too simple for the purposes of creating real-world plans, but nevertheless, we want to identify basic cases in which actions may be combined before extending this to larger plans.

### 2.3 Action-based Alternating Transition Systems

We use Action-based Alternating Transition Systems (AATS) as introduced in [17] as a basis for our formalism to represent action and plan proposals. AATS models define joint-actions that may be performed by agents in a state and the effects of these actions. In particular, an AATS model defines semantic structures useful to represent joint-actions for multiple agents, their preconditions and the states that will result from the transition. An *AATS* is an (n+7)-tuple of the form:

$$S = \langle Q, q_0, A_g, Ac_1, ..., Ac_n, \rho, \tau, \Phi, \pi \rangle$$

where:

- $Q$ is a finite non-empty set of states;
- $q_0 \in Q$ is the initial state;
- $A_g = \{1, ..., n\}$ is a finite non-empty set of agents;
- $Ac_i$ is a finite, non-empty set of actions, for each $i \in Ag$, where $Ac_i \cap Ac_j = \emptyset$ for all $i \neq j \in Ag$;
  Now we can say that a joint action $j_{Ag}$ for the set of agents $Ag$ is a tuple $(\alpha_i, .., \alpha_n)$ where for each $\alpha_j (j \leq n)$ there is some $i \in Ag$ such that $\alpha_j \in Ac_i$. We denote the set of all joint-actions $J_{AG}$. Given an element $j$ of $J_{AG}$ and an agent $i \in Ag$, $i's$ action in $j$ is denoted by $j_i$ .

- $\rho : Ac_{Ag} \rightarrow 2^Q$ is an action precondition function, which for each action $\alpha \in Ac_{Ag}$ defines the set of states $\rho(\alpha)$ from which $\alpha$ may be executed;
- $\tau : Q \times J_{Ag} \rightarrow Q$ is a partial system transition function, which defines the state $\tau(q, j)$ that would result by the performance of $j$ from state $q$, note that, as this function is partial, not all joint actions are possible in all states (cf. the pre-condition function above);
- $\Phi$ is a finite, non-empty set of atomic propositions; and
- $\pi : Q \rightarrow 2^{\Phi}$ is an interpretation function, which gives the set of primitive propositions satisfied in each state: if $p \in \pi(q)$, then this means that the propositional variable $p$ is satisfied (equivalently, true) in state $q$.

In [2] Atkinson and Bench-Capon extended this transition system to enable representation of a theory of practical reasoning related to arguments about action through which values [5] were added to the system. The extensions are:

- $Av_i$, is a finite, non-empty set of values $Av_i \subseteq V$, for each $i \in Ag$.
- $\delta$: $Q \times Q \times Av_{Ag} \rightarrow \{+, -, =\}$ is a valuation function which defines the status (promoted(+), demoted(-) or neutral (=)) of a value $v_u \in Av_{Ag}$ ascribed by the agent to the transition between two states: $\delta(q_x, q_y, v_u)$ labels the transition between $q_x$ and $q_y$ with one of $\{+, -, =\}$ with respect to the value $v_u \in Av_{Ag}$.

### 2.4 Proposals for Action

Argumentation schemes are stereotypical patterns of defeasible reasoning used in everyday conversational argumentation. In an argumentation scheme, arguments are presented as general inference rules where under a given set of premises a conclusion can be presumptively drawn [19]. Artificial Intelligence has become increasingly interested in argumentation schemes due to their potential for making significant improvements in the reasoning capabilities of artificial agents [7] and for automation of agent interactions. In [20], Walton explains: "*...arguments need to be examined within the context of an ongoing investigation in dialogue in which questions are being asked and answered*". Critical questions are a way to examine the acceptability of arguments. Depending on the nature of the critical question, they can be used to critique several aspects of the argument. Usually, critical questions provide pointers which would make the argumentation scheme inapplicable or could lead to a valid way to attack the argument, either defeating the argument on one of its premises or on its presumptive conclusion.

The action proposal presented in [3] is as follows: In the current circumstances $R$, we should perform action $A$ to achieve new circumstances $S$ which will realize some goal $G$ which will promote some value $v$. Furthermore, in [2] the authors re-stated the argumentation scheme in terms of the extended AATS. Figure 2 presents an action as in the PDDL 2.1 specification (presented in section 2.1). So, we can extend the action proposal from [3] with elements from the PDDL 2.1 specification. The extended action proposal and AATS representation are presented in Table 1.

For the purpose of this paper, time is discrete and actions take a single time step, thus we will not represent durative actions elements from section 2.1 in the plan proposal in the next section. Nevertheless, in the critical questions' section, time elements

[5] Our use of the term values follows [4] where values are qualitative social interests of agents.

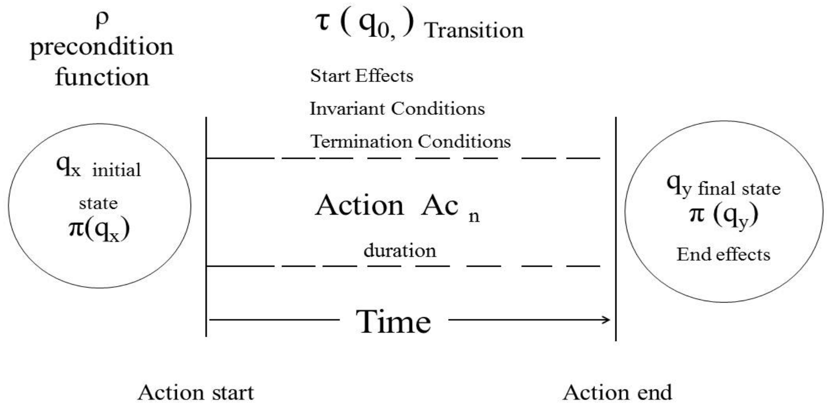


**Fig. 2.** Action Proposal Representation.

**Table 1.** Argumentation scheme for actions

| **Action Proposal** | **as an AATS model** $AS2$ |
|---|---|
| In the current circumstances $R$ | In the initial state $q_0 = q_x \in Q$ |
| we should perform action $A$ at time $t$ with duration $d$ | agent $i \in Ag$ should participate in joint action $j_n \in J_{Ag}$ |
| to achieve start effects from point $t$ | where $j_{n^i} = \alpha_i$ |
| given invariant conditions | such that $\tau(q_x, j_n)$ is $q_y$ |
| action finishing by termination conditions | such that $p_a \in \pi(q_y)$ and $p_a \notin \pi(q_x)$ |
| to achieve new circumstances $S$ | or $p_a \notin \pi(q_y)$ and $p_a \in \pi(q_x)$ |
| which will realize some goal $G$ | such that for some $v_u \in Av_i$ , $\delta(q_x, q_y, v_u)$ is +. |
| which will promote some value $v$ | |

are considered. Future work will be focused on representing durative actions within the action and plan proposal and the representation of action elements such as the propositions satisfied during the transition, which do not arise in [3].

## 3 Plan Proposal and Critical Questions

We now present our argumentation scheme in terms of the action elements presented above. Our plan proposal $ASP$ is as follows: Given a social context $X$ [6] in the current circumstances $q_0$ holding preconditions $\pi(q_0)$, plan $PL$ should be performed to achieve

[6] The "social context" was an extension to the argumentation scheme presented in [5] where agents use a social structure to issue valid commands between them.

new circumstances $q_x$ , that will hold postconditions $\pi(q_x)$ which will realize the plan-goal $G$ which will promote value(s) $V_G$.

The valid instantiation of the scheme pre-supposes the existence of a regulatory environment or a social context X in which the proponent has some rights to engage in a dialogue with the co-operating agent. Current circumstances are represented by an initial state $q_0$. The agent acting as the proponent propose plan $PL$ as a finite set of linked action-combinations. The plan leads to a state in which propositions $\pi(q_x)$ and the plan-goal $G$ is achieved (where $G$ is an assignment of truth values to a set of propositions $p \subseteq \Phi$) and a non-empty set of values associated with the plan is promoted. Table 2 presents the plan proposal as an AATS model.

**Table 2.** Plan Proposal $ASP$

| **Plan Proposal** | **as an AATS model** |
|---|---|
| Given a social context X,<br>in the current circumstances $q_x$<br>holding preconditions $\pi(q_x)$<br>plan $PL$ should be performed<br>to achieve new circumstances $q_y$<br>that will hold postconditions $\pi(q_y)$<br>which will realize the plan-goal $G$<br>which will promote value(s) $V_G$. | Given social context $\Delta$ ,<br>In the initial state $q_0$ = $q_x \in Q$, where $\pi(q_0)$,<br>agents $i, j \in Ag$ should execute plan $PL$,<br>where $PL$ is a finite set of joint-actions $j_n$<br>such that $PL = \{j_0, .., j_n\}$<br>and $\{j_0, .., j_n\} \in J_{Ag}$ and $j_n = \{\alpha_i, .., \alpha_j\}$<br>with transition given by $\tau(q_x, PL)$ is $q_y$ ,<br>where $\tau(q_0, \{j_1, .., j_n\}) = \tau(\tau(q_0, j_1), (j_2, ..j_n))$<br>and $\tau(q_x, \{\}) = q_x$<br>such that $p_a \in \pi(q_x)$ and $p_a \notin \pi(q_y)$ where $G = p$<br>and ($V_G \subseteq V$ such that $v_1 \in V_G$<br>iff $\delta(q_x, q_y, v_1)$ is $+$)<br>and $V_G \neq \emptyset$ |

### 3.1 Critical Questions for plan proposals

A benefit of having critical questions associated with an argument scheme is that the questions enable dialogue participants to identify points of challenge in a debate or locate premises in an instantiation of the argument scheme that can be recognized as questionable. Most of the critical questions are created from argumentation scheme elements and represent a valid way to challenge proposals that could identify sources of disagreement about a particular element of the argumentation scheme. A question can be seen as a weak form of attack on a particular element of the argument scheme given different beliefs about the world of the agent posing the question. Critical questions then could be used to create Dialogue Games for agents where the participants put forward arguments instantiating the argumentation scheme and opponents to the argument challenge it through objections based on critical questions. Argumentation-based dialogues are used to formalize dialogues between autonomous agents based on theories of argument exchange. In [18] a classification is given based on the role the question plays in the context of the argumentation scheme. A question could be used to: criticize a scheme premise, point to exceptional situations in which the scheme should not be used, set conditions for the proper use of the scheme, or point to other arguments that might be used to attack the scheme. Furthermore, questions could argue for an incompatible conclusion like : Are there (better) reasons not to do plan A or plan B? We classify our set of critical questions into 6 layers (also presented in Figure 3).

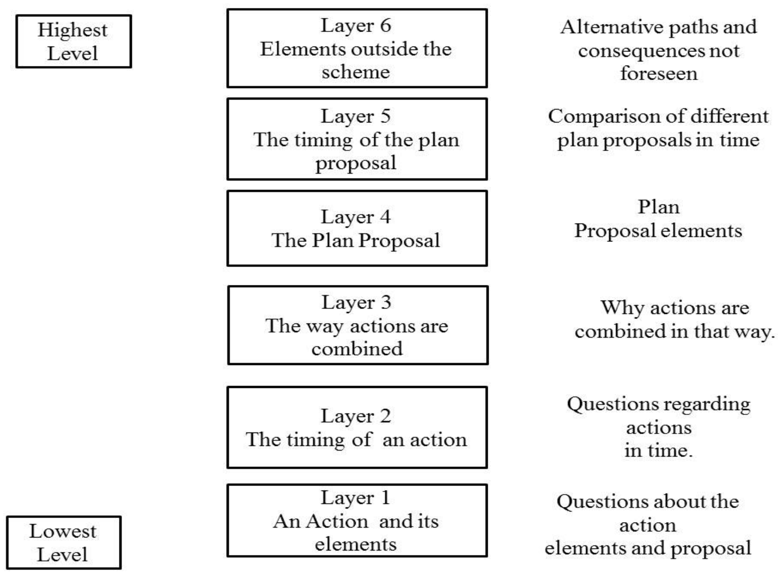


**Fig. 3.** Critical Question Layers.

- Layer 1.- An action and its elements (Lowest level).
- Layer 2.- The timing of a particular action.
- Layer 3.- The way actions are combined.
- Layer 4.- The plan proposal overall.
- Layer 5.- The timing of the plan proposal.
- Layer 6.- Elements outside the scheme (alternative paths or consequences not foreseen) (Highest Level).

The layers are derived from the different categories of critical questions that relate to the different elements of the argumentation scheme. Each layer groups questions according to the level of detail on which they focus. At the plan proposal level, for example, the critical questions are all those which are independent of the way in which actions are composed inside the plan i.e. the way in which actions are combined. This classification allows us then, to consider questions at each layer separately. Furthermore, this classification gives us elements to create a strategy to select critical questions in a dialogue. Having the critical questions classified an agent could pick a layer and narrow the scope of available questions. An agent then could focus on a specific level of the proposal *e.g.* either the plan proposal or specific sequences of actions. A strategy like this involves a dialogue-protocol where rules to issue such questions are specified. It could be that the answer to a critical question in one layer imposes constraints within another layer, so this may affect the optimal ordering in which the layers are addressed. Appropriate participant strategies, and their possible relationships with the dialogue protocol, are left for future work.

Our set of critical questions is based on the set developed for action proposals in command dialogues presented in [5]. The complete list of 66 critical questions necessary to comprehensively question all relevant aspects of the plan proposals is presented in [14]. We believe this analysis enables plan proposals to be questioned in a comprehensive way in order to be justified. We present here example questions for each layer.

**Layer 1. An action $\alpha$ and its elements (9 questions)**.
- Is the action $\alpha$ possible?
- Are the action invariants conditions as stated by proponent?
- Are the termination conditions as described as possible?
**Layer 2. The timing of an action (10 questions)**.
- What is the earliest time the action $\alpha$ can start?
- Is the action $\alpha$ possible to finish at the specified time?
- What is the earliest time the action $\alpha$ can end?
**Layer 3. The way actions are combined (7 questions).**
- (For sequential actions) Can the order of the actions be changed?
- (For concurrent actions) Is there a conflict in any of the invariants conditions of the actions?
**Layer 4. The plan proposal (18 questions)**.
- Are the current circumstances $R$ as stated by proponent?
- Is the value $v$ indeed a legitimate value?
- Can the desired goal $G$ be realized?
**Layer 5. The timing of the plan proposal (11 questions)**.
- Is the starting point for the plan $PL$ fixed? If not, what is the range allowed?
- Is the plan $PL$ possible at the specified time?
**Layer 6. Elements outside the scheme (11 questions)**.
- Does performing the plan $PL$ have a side effect which demotes some other value $v$?
- Is there an alternative plan $PL$ to realize the same goal $G$?

## 4 Example

To illustrate our approach we will use our argumentation scheme in the context of agents representing organizations in a conflict zone. The example was first introduced in [8] and also used in [16] to illustrate a similar problem regarding planning and dialogues for autonomous agents. The situation is the following: two agents, one representing an Non-Governmental Organization ($NGO$) and one representing a peace keeping force ($KF$), are working in a conflict zone.

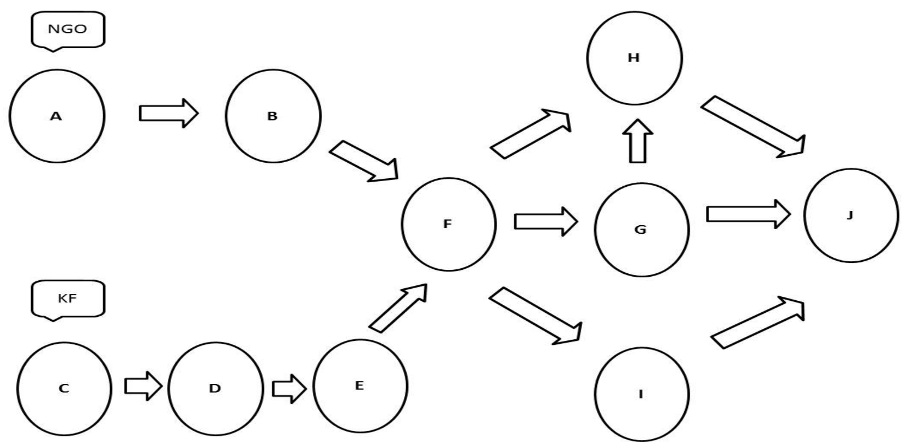


**Fig. 4.** Example NGO.

The initial conditions are: Agent $NGO$ is based at zone A and agent $KF$ is based at zone C. The joint-goal is that agent $NGO$ reaches zone J safely to help the villagers there. A initial sub-goal is to meet in zone F. The values involved are: $v_1$ representing humanitarian help and $v_2$ representing $NGO$ security. The restrictions are: $NGO$ can traverse the routes (A,B),(B,F),(F,H),(I,J) independently, but for all the other routes it needs to be accompanied by $KF$. $KF$ can traverse any route. At any time, some disruption may flare up at zone G. If this happens, only the $KF$ agent has the surveillance data to know this is happening, and must go to zone G to suppress the disturbance. Furthermore, $NGO$ cannot traverse the routes where zone G is involved if there is a conflict. Finally agent $NGO$ is able to see all the zones and routes only when in zone F. The routes between zones are shown as arcs in Figure 4. The list of possible actions and joint-actions is presented in Table 3.

**Table 3.** Actions and joint-actions.

| **Actions** | **Joint-actions** |
|---|---|
| $\alpha_0 = move_{NGO}(X, Y)$ | $j_0 = (idle_{NGO}, idle_{KF})$ |
| $\alpha_1 = idle_{NGO}(X)$ | $j_1 = (idle_{NGO}, control_{KF})$ |
| $\alpha_2 = move_{KF}(X)$ | $j_2 = (idle_{NGO}, move_{KF})$ |
| $\alpha_3 = move_{KF}(X, Y)$ | $j_3 = (move_{NGO}, idle_{KF})$ |
| $\alpha_4 = control_{KF}(X)$ | $j_4 = (move_{NGO}, control_{KF})$ |
| | $j_5 = (move_{NGO}, move_{KF})$ |

Our strategy to coordinate the agents is based on a persuasion dialogue where the $NGO$ agent propose a plan and engages in a dialogue where $KF$ need to accept all the actions in the plan to execute it. Another strategy could involve agents creating a plan from the top following a deliberation dialogue. As mentioned in section 3.1, a dialogue-protocol for engaging in such dialogues is left for future work.

We present now a possible dialogue between 2 agents based on the scenario presented above (possible plans $PL_1, PL_2, PL_3$ are presented in table 4).

1. $NGO$ presents a proposal with plan $PL_1$ to reach zone F and promote $v_1$ humanitarian help (the generation and agreement on sub-goals is out of the scope of this paper).
2. $KF$ present a question: Does performing the plan $PL_1$ have a side effect which demotes the value $v_2$ $NGO$ security?
3. $NGO$ provides justification asserting none of the effects demote $v_2$.
(The value of having a set of critical questions related to a plan proposal is that an agent has several options to questions and/or attack the elements presented. In this example we only present one question for space reasons. We believe the exchange of arguments using critical questions in a dialogue allows agents to choose the best possible plan).
4. $KF$ accepts plan $PL_1$. (From this point we assume the plan $PL_1$ is executed and agents have a new view of the world. Agents are now in zone F).
5. $NGO$ presents plan $PL_2$ to reach zone J (goal) and promote values $v_1$ and $v_2$ (Different sets of values involved may lead to different plans).
6. $KF$ rejects $PL_2$ and provides evidence that shows demotion of value $v_2$ (Agent $KF$ detects a conflict in zone G).
7. $NGO$ accepts $KF$ rejection.
8. $KF$ presents plan $PL_3$ to reach zone J.
9. $NGO$ rejects joint-action $j_5 = (move_{NGO}(H, J), (sodfe), control_{KF}(G))$ [7] referring to a constraint in $NGO$ domain not evident to $KF$.(Agent cannot travel alone on route (H-J)). The plan is then partially accepted by $NGO$. From here agents exchange arguments in the action level, assuming the first action of $PL_3$ : $j_5 = (move_{NGO}(F, H), (e), move_{KF}(F, G))$ was accepted.
10. $KF$ proposes a modification to plan $PL_3$. We assume the modification is based on a replanning process based on new information. The new sequence proposed is $(j_1, j_2, j_5)$.
11. $NGO$ accepts modification to plan $PL_3$. and dialogue-goal is reached.

A detailed dialogue in AATS terms is presented in table 5. This detailed example-dialogue represents joint-states with a sub-index $(q_0, .., q_{18})$. Each joint-state represent the state of agent $NGO$, the state of agent $KF$ and the conflict status. For example, the initial state $q_0$ is given by the function $\pi(q_0) = \{In(A)_{NGO}, In(C)_{KF}, conflict(0)\}$.

## 5 Related Work

Our approach is influenced by work on argumentation for practical reasoning [2] and dialogues about plans [6, 15, 16]. Regarding dialogues and plans, Tang, Norman and

[7] Action sequences are specified using the action combination analysis of section 2.2. $(sodfe)(j_0, j_1)$ means: joint-action $j_0$ Starts, Overlaps, During, Finishes or Equals joint-action $j_1$.

**Table 4.** Plans.

| Time | Plan |
|---|---|
| | **Plan $PL_1$ to reach zone F** |
| 1 | $j_5 = equals(move_{NGO}(A, B), move_{KF}(C, D))$ |
| 2 | $j_5 = equals(move_{NGO}(B, F), move_{KF}(D, E))$ |
| 3 | $j_2 = equals(idle_{NGO}(F), move_{KF}(E, F))$ |
| | **Plan $PL_2$ to reach zone J** |
| 4 | $j_5 = equals(move_{NGO}(F, G), move_{KF}(F, G))$ |
| 5 | $j_5 = equals(move_{NGO}(G, J), move_{KF}(G, J))$ |
| | **Plan $PL_3$ to reach zone J** |
| 4 | $j_5 = equals(move_{NGO}(F, H), move_{KF}(F, G))$ |
| 5 | $j_4 = sodfe(move_{NGO}(H, J), control_{KF}(G))$ |
| | **Modified plan $PL_3$ to reach zone J** |
| 4 | $j_5 = equals(move_{NGO}(F, H), move_{KF}(F, G))$ |
| 5 | $j_1 = sodfe(idle_{NGO}(H), control_{KF}(G))$ |
| 6 | $j_2 = sodfe(idle_{NGO}(H), move_{KF}(G, H))$ |
| 7 | $j_5 = sodfe(move_{NGO}(H, J), move_{KF}(H, J))$ |

Parsons in [16] establish a model for individual and joint agents' actions suitable for describing the behaviour of a multi-agent team, including communication actions. Tang *et al's* work has been focused on setting a basis for implementing multi-agent planning dialogues based on argumentation that take into account the communication needs for the plan to be executed successfully. The model uses policies to generate plans and the communication needs are embedded in the policy algorithm generation. From the work of Tang *et al.* we are particularly interested in the techniques used to combine planning and dialogue models using policies. In our approach agents propose and justify previously created plans and then engage in a dialogue to justify the actions and possibly modify the plan. The approach in Tang *et al.* embeds the communication policy in the planning algorithm.

In [6] Belesiotis, Rovatsos and Rahwan develop an argumentation mechanism for reconciling conflicts between agents over plan proposals. The authors extend a protocol where argument-moves enable discussion about planning steps in iterated dispute dialogues as presented in [9]. The authors then introduce a logic for arguments about plans based on the situation calculus. From this approach we are interested in their protocol based on iterated disputes. We plan to modify the approach extending the way a plan proposal makes use of critical questions in the dispute tree.

Another approach is presented in [15]. Onaindia *et al.* present the problem of solving cooperative distributed planning tasks through an argumentation-based model. The model allows agents to exchange partial solutions, express opinions on the adequacy of candidate solutions and adapt their own proposals for the benefit of the overall task. The argumentation-based model is designed in terms of argumentation schemes and critical questions whose interpretation is given through the semantic structure of a partial order planning paradigm. The approach assumes a lack of uncertainty and deterministic planning actions, thus, focuses only on questions concerned with the choice of actions.

**Table 5.** Possible Dialogue for the $NGO$ example.

| | Agent | Locution | Variables | Comments |
|---|---|---|---|---|
| 1 | $NGO$ | ProposePlan($PL_1$) | $(q_0, PL_1)$ is $q_{11}$ | Initial state $NGO$ in zone A, $KF$ in zone C. |
| 2 | | | $v_1+$ | Final state $NGO$ and $KF$ in zone F. |
| 3 | | | | Promotion of value "humanitarian aid". |
| 4 | $KF$ | Question() | $\delta(q_0, q_1, v_2)$ is $-$ | Is there a side effect that demotes $v_2$? |
| 5 | $NGO$ | Provide() | $\delta(q_0, q_1, v_2)$ is $=$ | Value $v_2$ is not demoted. |
| 6 | $KF$ | AcceptProposal($PL_1$) | | Proposal accepted for plan $PL_1$ |
| 7 | $NGO$ | ProposePlan($PL_2$) | $(q_{11}, PL_2)$ is $q_{16}$ | Initial state $NGO$ and $KF$ in zone F. |
| 8 | | | $G \in \pi(q_{16})$ | Final state $NGO$ and $KF$ in zone J |
| 9 | | | $v_1+$ and $v_2+$ | goal reached in $q_{16}$, values promoted |
| 10 | $KF$ | RejectProposal($PL_2$) | $PL_2$ | |
| 11 | | Provide() | $\delta(q_{11}, q_{16}, v_2)$ is $-$ | Demotion of $v_2$ |
| 12 | | Provide() | $conflict \in \pi(q_2)$ | |
| 13 | $NGO$ | AcceptRejection() | | |
| 14 | $KF$ | ProposePlan($PL_3$) | $(q_{11}, PL_3)$ is $q_{18}$ | $NGO$ and $KF$ in zone F. |
| 15 | | | $v_1+$ and $v_2+$ | $NGO$ and $KF$ in zone J. |
| 16 | $NGO$ | AcceptAction() | $j_5$ | |
| 17 | | RejectAction() | $j_4$ | |
| 18 | | Provide() | $(q_7, j_4) \notin Q$ | $NGO$ cannot travel alone proposed route. |
| 19 | $KF$ | ProposeActions() | $j_1, j_2, j_5$ | Goal reached in $q_{17}$ |
| 20 | $NGO$ | AcceptActions() | $j_1, j_2, j_5$ | |

The argumentation scheme, based on the scheme for action proposal from [3] is of the form:

In the current circumstances and considering the current base plan $\Pi_i$, agent $ag_i$ should perform the refinement step $\Pi^{'}$, which will result in a new partial plan $\Pi_j$, which will realise some sub-goals *G*, which will promote some values *V*.

Our work is very similar in approach to this work in the sense that plans should be entities treated at a detailed level when arguing about them. We go further and consider plan proposals in more detail referring to action elements and combination of actions. Furthermore, our argumentation scheme is related to a more comprehensive set of critical questions, giving an agent more options to critique and enhance a proposal. We believe these elements allow an agent to question and/or attack the argument in a more targeted fashion, facilitating the modification of more types of plans and faster identification of differences between participants.

## 6 Conclusion

Our research aims at contributing to solving problems related to multi-agent planning, where agents need to agree on plans given different views of the world and of other agents' capabilities. We believe our main contribution in this paper is that we have articulated a novel list of critical questions related to an argumentation scheme for plan proposals as different combination of actions including temporal aspects. The critical questions address each element of a proposed plan and so they are comprehensive with respect to the representation we have chosen for plan proposals. We believe every component and every interaction of components in our representation of a proposal for plan is subject to a possible critical question

The importance of this work is that it enables a proposal for plan execution to be considered rationally and automatically by software agents engaged in deliberation over the plan of action. The critical questions enable the proposed plan to be questioned/challenged in a comprehensive and organized manner, and to be clarified or defended in response, as appropriate. Indeed, it is possible to use the critical questions as the basis for an agent dialogue game protocol in which one participating agent may propose, and then clarify or defend a plan of action, while other agents question or challenge this proposal. For example, Atkinson and colleagues in [3] develop such a dialogue protocol for proposals for single actions. Whether the proposed plan of action survives such questions and attacks in the dialogue will depend upon the facts about the world underlying the proposal, and the ability of the proponent agent to defend his proposal from attack. Consequently, the acceptability or otherwise of the proposed plan will depend upon the outcome of the multi-agent dialogue based upon the critical questions, and vice-versa.

The multi-agent dialogue is a form of game-theoretic semantics for the statement of a plan of action in the same way as Hintikka's game-theoretic semantics for first-order logic [12] interprets well-formed formulae involving existential and universal quantifiers as equivalent to two-party games between a proponent and an opponent of some proposition. Our approach will interpret proposals for plans in terms of dialogue games between agents defending and attacking the proposal. Our work in this paper can therefore be seen as part of a larger effort to develop computational semantics for plans of actions between interacting software agents [13].

Future work includes analysis on how to represent formally action elements not yet accounted for in the formalization, such as the duration of actions and action invariants. One limitation in this work is that we only considered plans comprising two actions, effectively a plan for each agent; how to decompose these plans into a number of actions and the issues that arise from the interaction of their components is something we will consider in the next phase of our research. To support this theory we will also implement a prototype where agents use a protocol that allows them to engage in dialogues about plan proposals in a single solution.

## References


1. J. F. Allen. Towards a general theory of action and time. *Artificial Intelligence*, 23(2):123–154, 1984.
2. K. Atkinson and T. Bench-Capon. Practical reasoning as presumptive argumentation using action based alternating transition systems. *Artificial Intelligence*, 171(10-15):855–874, 2007.
3. K. Atkinson, T. Bench-Capon, and P. McBurney. A dialogue game protocol for multi-agent argument over proposals for action. *Autonomous Agents and Multi-Agent Systems*, 11(2):153–171, 2005.
4. K. Atkinson, T. Bench-Capon, and P. McBurney. Computational representation of practical argument. *Synthese*, 152(2):157–206, 2006.

5. K. Atkinson, R. Girle, P. McBurney, and S. Parsons. Command Dialogues. In I. Rahwan and P. Moraitis, editors, *Argumentation in Multi-Agent Systems*, Fifth International Workshop, pages 93–106, Berlin, Heidelberg, 2009. Springer-Verlag.
6. A. Belesiotis, M. Rovatsos, and I. Rahwan. A generative dialogue system for arguing about plans in situation calculus. In P. McBurney, I. Rahwan, S. Parsons, and N. Maudet, editors, *Argumentation in Multi-Agent Systems*, volume 6057 of *Lecture Notes in Computer Science*, pages 23–41. Springer Berlin / Heidelberg, 2010.
7. F. Bex, H. Prakken, C. Reed, and Douglas Walton. Towards a formal account of reasoning about evidence: argumentation schemes and generalisations. *Artificial Intelligence Law*, 11(2-3):125–165, 2003.
8. C. Burnett, D. Masato, M. Mccallum, T. J. Norman, J. Giampapa, M. J. Kollingbaum, and K. Sycara. Agent support for mission planning under policy constraints. In Proceedings of the Second Annual Conference of the ITA, Imperial College, 2008.
9. P. E. Dunne and T. Bench-Capon. Two party immediate response disputes: properties and efficiency. *Artificial Intelligence*, 149(2):221–250, 2003.
10. M. Fox and D. Long. PDDL2.1: An extension to PDDL for Expressing Temporal Planning Domains. *Journal Artificial Intelligence Research (JAIR)*, 20:61–124, 2003.
11. H. Geffner. Perspectives on Artificial Intelligence Planning. In *Eighteenth National Conference on Artificial intelligence*, pages 1013–1023, Menlo Park, CA, USA, 2002. American Association for Artificial Intelligence.
12. J. Hintikka. *The Game of Language: Studies in Game-Theoretical Semantics and Its Applications*, volume 22 of *Synthese Language Library*. D. Reidel, Dordrecht, The Netherlands, 1983.
13. P. McBurney and S. Parsons. Dialogue games for agent argumentation. In I. Rahwan and G. Simari, editors, *Argumentation in Artificial Intelligence*, chapter 13, pages 261–280. Springer, Berlin, Germany, 2009.
14. R. Medellin-Gasque, K. Atkinson, P. McBurney, and T. Bench-Capon. Critical Questions for Plan Proposals. Technical Report ULCS-11-003, Department of Computer Science, University of Liverpool, UK, March 2011.
15. E. Onaindía, O. Sapena, and A. Torreño. Cooperative Distributed Planning through Argumentation. In *International Journal of Artificial Intelligence*, volume 4, pages 118–136. CESER Publications, 2010.
16. Y. Tang, T. J. Norman, and S. Parsons. A model for integrating dialogue and the execution of joint plans. In *AAMAS '09: Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems*, pages 883–890, Richland, SC, 2009. International Foundation for Autonomous Agents and Multiagent Systems.
17. W. van der Hoek, M. Roberts, and M. Wooldridge. Social laws in alternating time: Effectiveness, feasibility, and synthesis. *Synthese*, 156:1–19, 2007.
18. B. Verheij. Dialectical argumentation with argumentation schemes: an approach to legal logic. *Artificial Intelligence and Law*, 11(2-3):167–195, 2003.
19. D. N. Walton. *Argumentation Schemes for Presumptive Reasoning*. Lawrence Erlbaum Associates, Mahwah, NJ, USA, 1996.
20. D. N. Walton. Justification of argumentation schemes. *Australasian Journal of Logic*, 3, 2005.