Group Intention is Social Choice with Commitment

Guido Boella[1], Gabriella Pigozzi[2], Marija Slavkovik[2], and Leendert van der Torre[2]

[1] guido@di.unito.it
[2] {gabriella.pigozzi, marija.slavkovik, leon.vandertorre}@uni.lu

Abstract. A collaborative group is commonly defined as a set of agents, which share information and coordinate activities while working towards a common goal. How do groups decide which are their common goals and what to intend? According to the much cited theory of Cohen and Levesque, an agent intends g if it has chosen to pursue goal g and it has committed itself to making g happen. Following the same line of reasoning, a group intention should be a collectively chosen goal with commitment. The literature often considers a collective goal to be one of those individual goals that are shared by all members. This approach presumes that a group goal is also an individual one and that the agents can act as a group if they share the beliefs relevant to this goal. This is not necessarily the case. We construct an abstract framework for groups in which common goals are determined by social choice. Our framework uses judgment aggregation to choose a group goal and a multi-modal multi-agent logic to define commitment and revision strategies for the group intentions.

1 Introduction

An agent acts according to its beliefs and intentions. According to what does a group act? We would expect that in order for groups to act, jointly or by coordinating their activities, they need to establish what to believe i.e., form epistemic attitudes, and what to aim for, i.e., to form motivational attitudes. In existing frameworks for collaborative activities [9, 12, 16] and group decision-making protocols [10], the formation of group attitudes is defined only for a specific type of groups. These groups consist of agents that engage in pursuing a group goal only when the members have the same beliefs regarding this goal, or when they are successful in reaching an agreement on a given set of beliefs.

Consider a group of robots in charge of office building maintenance. One candidate group goal for them is to clean the meeting room. The decision rule to adopt the goal is that the room needs to be cleaned when the following conditions (reasons) are met: the floors are dirty or the garbage bin is full, and people do not occupy the room. To decide whether to pursue this goal, the robots need to decide if the reasons to adopt the goal are true. The robots cannot check whether the room is occupied or what state it is in. Hence, to estimate the state of affairs, the robots need to rely on their individual beliefs, which may diverge.

Assume there are three robots in the group. One robot believes that the room is occupied and thus, according to it, the group should not adopt the goal. According to the other two robots, the group should adopt the goal. One robot believes that the garbage bin is full and the floors clean and the other that the floors are dirty. The question is how to should one aggregate the beliefs of the robots in this case? The majority of the robots would estimate that the goal should be adopted. However the group is not univocal

and as a result, the goal would not be chosen for a group goal when the robots reason according to [9, 12, 16]. A general method for forming group attitudes needs to specify how group attitudes are formed also when agents have disagreeing beliefs or limited persuasive abilities. A framework that allows for such general methods is still lacking.

Consider now that the group adopts the goal, but before pursuing it, the robots learn that there is a seminar scheduled in the meeting room. The group has to de-commit from their intention to clean the room. However, to be able to do so, the group has to have a commitment strategy that allows de-committing upon change in the reasons for the goal adoption. Furthermore, the group needs to be able to reconsider its reasons for goals, and goals, after de-committing. If they simply drop the goal, without "remembering" why, they would not be able to re-deliberate and commit again once the seminar is over.

An intelligent agent reacts to the changes in its environment and a group should be capable of doing the same. When new information becomes available the group faces a choice: to remain committed to its group attitudes or to reconsider them. The most sensitive commitment strategy proposed by Rao and Georgeff [21] allows for a group reaction only when the goal is accomplished or impossible. However, the existence of a goal is intertwined with the existence of some beliefs [2]. Consequently, there is a need for a commitment strategy that reacts to new information regarding those beliefs on which the goal hinges. Furthermore, groups need to know not only when to de-commit from their goals but also how to reconsider their goal-related beliefs.

The research question we address in this paper is:

How can groups choose and reconsider their goals?

A good methodology for collectively choosing and reconsidering goals is one that can be used by any group of agents regardless of the homogeneity of the individual beliefs of its members and their persuasion abilities. The relation between individual goals and beliefs can be specified and analyzed in modal agent logics like BDI_{LTL} [22]. The challenge in group goal generation is to incorporate the aggregation of individual attitudes into collective ones, as studied in merging, judgment aggregation and social choice [3, 11, 13, 14, 17]. However, using this approach is not straightforward. The main difficulty lies in the inability to use judgment aggregation directly in a BDI_{LTL} framework. Properties of judgment aggregation and modal agent logic are of different kinds, as they were initially developed for different purposes.

Our research question thus breaks down into the following sub-questions:

1. *How to aggregate individual (epistemic and motivational) attitudes?*
2. *What are the desirable properties for this aggregation?*

A good methodology for collectively choosing and reconsidering goals is also a methodology that is dynamic enough to allow for the group to change its epistemic and motivational attitudes in light of new information. Having such a methodology increases the autonomy and reactivity of groups. Hence we need to answer the following sub-question as well:

3. *Which commitment and reconsideration strategies should be available for groups?*

We thus focus on finding the following solutions:

Formal framework. We extend a multi-agent modal language to be able to represent judgment aggregation in it.

Choice of aggregation. Judgment aggregation is an abstract framework that allows for various desirable properties to be specified for the aggregation procedure. The task is to determine which aggregation properties are necessary and desirable for aggregating individual beliefs and goals.

Commitment and reconsideration strategies. Within our formal framework, we define *when* to de-commit from intentions and *how* to change them.

Multiple goals. Since a group can have more than one goal, we need to model the effect that the commitment to (and reconsideration of) one goal has over the commitment to (and reconsideration of) other goals.

We make the following assumptions. The group has a fixed membership – agents do not join or depart from the group. The group has a set of candidate group goals and an order of priory over these goals. The decision rules for each candidate goal are available to the group. How the decision rules are learned is outside of the scope of this paper. Here we do not consider how plans are generated, executed and revised once the group goals are selected or reconsidered, nor we consider whether the group goals are executed via individual or joint actions.

The layout of the paper is as follows. In Section 2 we introduce judgment aggregation. In Section 3 we extend a multi-modal agent formalism to capture the aggregation of individual attitudes. Sections 4 and 5 respectively study the commitment and reconsideration strategies. Related work, conclusions and outlines for future work are in Section 6.

2 From individual attitudes to group goals

Let us consider again the example of the robot cleaning crew from Section 1.

Example 1. Let $C = \{w_1, w_2, w_3\}$ be a crew of cleaning robots. We denote the group goal to clean the meeting room with g_1, and the reasons to adopt this goal with: there are no people in the room (p_1), the room is dirty (p_2), the garbage bin in it is full (p_3). The individual beliefs of the robots on whether g_1 should be the group goal are justified by individual beliefs on p_1, p_2, p_3 using the decision rule $(p_1 \wedge (p_2 \vee p_3)) \leftrightarrow g_1$.

A group of agents could collectively decide to adopt or reject a goal by voting. However, the goals of the agents are not independent from their beliefs [2], which we express using decision rules. When the decision is whether to adopt a goal or not, we also need to explain why this goal should (not) be adopted. Having reasons for (not) adopting a goal enables the agents to re-considerer this goal in light of new information. Judgment aggregation deals with the problem of reaching decisions for a set of logically dependent issues, by aggregating individual opinions on these issues.

2.1 Judgment aggregation preliminaries

A general overview of judgment aggregation is given in [17]. Here we present the terminology and definitions of judgment aggregation we use in our framework.

Consider a logic $\mathcal{L}$ with entailment operator $\models$. In a judgment aggregation framework, an agenda $\mathcal{A} \subseteq \mathcal{L}$ is the pre-defined set of issues, on which every agent casts her judgments. E.g., the agenda of Example 1 is $\mathcal{A} = \{p_1, p_2, p_3, g_1\}$. Consider a valuation function v such that a judgment "yes" on issue a is a valuation $v(a) = 1$, while a judgment "no" on the same issue is a valuation $v(a) = 0$. A profile is the set of all judgments assigned, on the agenda issues, by the decision-making agents.

Definition 1 (Profile). *Let $\mathcal{N} = \{1, 2, \ldots\}$ be a set of agent names, $\mathcal{A} \subseteq \mathcal{L}$ an agenda and $\overline{\mathcal{A}} = A \cup \{\neg a \mid a \in \mathcal{A}\}$. A profile π is the set of judgments for the agenda items, submitted by the agents in the group: $\pi \subseteq \mathcal{N} \times \overline{\mathcal{A}}$. We define two operators:*
The judgment set for agent i is $\pi \Rightarrow i = \{a \mid (i, a) \in \pi\}$.
The set of all the agents who judged "yes" to $a \in \overline{A}$ is $\pi \Downarrow a = \{i \mid (i, a) \in \pi\}$.

Definition 1 extends the definition of a profile given [17] with the operators $\Rightarrow$ and $\Downarrow$. We introduce these operators to ease the explanation of the various aggregation properties we present later on. To get a better intuitive grasp on these operators, the reader should envision the profile as a two-dimensional object with the agenda items identifying the columns and the agents identifying the rows:

$$\pi = \begin{array}{c} \\ w_1 \\ w_2 \\ w_3 \end{array} \begin{array}{cccc} p_1 p_2 & p_3 & g_1 \\ \left\{\begin{array}{cccc} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{array}\right\} \end{array}$$

π is a possible profile for Example 1. We identify $\pi \Rightarrow w_2$ as the row labeled w_2 and $\pi \Downarrow p_2$ as the 1 entries in the column labeled p_2, which identify the agents who casted judgement "yes" on p_2.

In judgment aggregation, the judgments over $\mathcal{A}$, both individual and aggregates, are constrained by decision rules $\mathcal{R} \subseteq \mathcal{L}$. The set $\mathcal{R}$ contains only formulas r such that all the non-logical symbols of $\mathcal{A}$ occur in $r \in \mathcal{R}$. For instance, in Example 1, the decision rule is: $(p_1 \wedge (p_2 \vee p_3)) \leftrightarrow g_1$ and all agents respect it. In general, each agent could follow a different decision rule and yet another decision rule can be imposed for the group. In judgment aggregation, the agents are allowed to submit only those judgment sets which are consistent with $r \in \mathcal{R}$. Often, the agents are also required to cast judgments on all the agenda issues. We construct Definition 2 to formalize what it means for a judgment set to be admissible.

Definition 2 (Admissible profiles). *A judgment set $\pi \Rightarrow i$ is complete if and only if for every $a \in \mathcal{A}$, either $(i, a) \in \pi$ or $(i, \neg a) \in \pi$. A judgment set $\pi \Rightarrow i$ is consistent with $r \in \mathcal{R}$ if and only if $\{r\} \cup (\pi \Rightarrow i) \not\models \bot$. A judgment set $\pi \Rightarrow i$ is admissible if it is consistent with the given $r \in \mathcal{R}$ and complete for $\mathcal{A}$. The set of all admissible judgment sets for a given $\mathcal{A}$ and r is denoted by $\mathcal{W}$.*
A profile π is admissible if $\pi \Rightarrow i$ is admissible for all $i \in \mathcal{N}$. The set of all admissible profiles for agents $\mathcal{N}$ is denoted by Π.

We denote $con(r, \varphi) = 1$ if φ is consistent with r and $con(r, \varphi) = 0$ otherwise. For the profile in Figure 1, $con((p_1 \wedge (p_2 \vee p_3)) \leftrightarrow g_1, \pi \Rightarrow i) = 1$ for every $i \in C$.

We now present the definition of a judgment aggregation function, as given in [17].

Definition 3 (**Judgment aggregation function**). *Given an agenda $\mathcal{A}$ and agents $\mathcal{N}$, a judgment aggregation (JA) function is $f : 2^{\mathcal{N} \times \overline{\mathcal{A}}} \mapsto 2^{\overline{\mathcal{A}}}$.*

We refer to $f(\pi)$ as the collective judgment set for $\mathcal{N}$. We denote the result of $f(\pi)$ with $\perp$ when the JA function produces an inadmissible judgment set. If $f(\pi) \neq \perp$ then we call $f(\pi)$ a decision and denote it by $\mathcal{D}_\pi$. Figure 1 illustrates an example of a judgment aggregation function and profile, for which the collective judgment set is $\perp$ because $con((p_1 \wedge (p_2 \vee p_3)) \leftrightarrow g_1, \{p_1, p_2, p_3, \neg g_1\}) = 0$.

The JA function is a useful abstraction, because many properties of judgment aggregation can be defined in terms of it. It then can be studied which properties can be accepted together (avoiding impossibility results). Given a JA function f, we describe the most common properties in the JA literature.

Universal domain. f satisfies universal domain if and only if its domain is Π.

Anonymity. Given a profile $\pi \in \Pi$, let $\hat{\pi} = \{\pi \Rightarrow 1, \ldots, \pi \Rightarrow n\}$, be the multiset of all the individual judgment sets in π. Two profiles $\pi, \pi' \in \Pi$ are permutations of each other if and only if $\hat{\pi} = \hat{\pi}'$. f satisfies anonymity if and only if $f(\pi) = f(\pi')$ for all permutation π and π'.

Unanimity on $a \in \overline{\mathcal{A}}$. f satisfies unanimity on $a \in \overline{\mathcal{A}}$ if and only if for every profile $\pi \in \Pi$ it holds: if for all $i \in \mathcal{N}$, $(i, a) \in \pi$, then $a \in f(\pi)$.

Collective rationality. f satisfies collective rationality if and only if for all $\pi \in \Pi$, $con(r, f(\pi)) = 1$ for a given $r \in \mathcal{R}$, and either $a \in f(\pi)$ or $\neg a \in f(\pi)$ for every $a \in \mathcal{A}$.

Constant. f is constant when there exists $\varphi \in 2^{\overline{\mathcal{A}}}$ such that for every $\pi \in \Pi$, $f(\pi) = \varphi$.

Independence. Let $\Phi = \{\pi \Downarrow a \mid a \in \mathcal{A}\}$ for every $\pi \in \Pi$. Let $f_1, \ldots, f_m$ be functions defined as $f_j : \mathcal{A} \times \Phi \mapsto \{0, 1\}$. The JA function f satisfies independence if and only if for all $\pi \in \Pi$, there exist functions f_i such that for all $\varphi \in f(\pi)$ it holds that $\varphi = \{a \mid a \in \mathcal{A}, f_j(a, \pi \Downarrow a) = 1\} \cup \{\neg a \mid a \in \mathcal{A}, f_j(a, \pi \Downarrow a) = 0\}$.

The best known example of $f_j : \mathcal{A} \times \Phi \mapsto \{0, 1\}$ is the simple majority voting f_m which counts how many agents expressed judgment "yes" on agenda item a. The function f_m returns a if that number of agents is greater than $\lceil \frac{n}{2} \rceil$ and $\neg a$ otherwise.
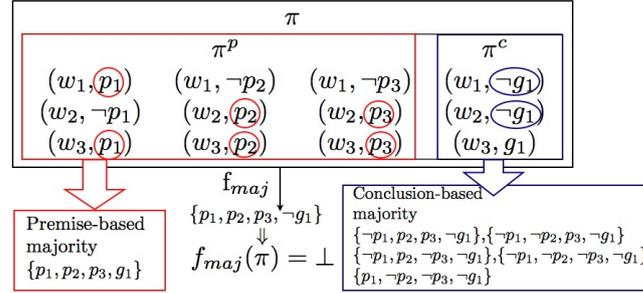
Fig. 1. A profile, issue-wise majority aggregation, premise-based and conclusion-based majority.

The issue-wise majority function f_{maj} is defined as

$$f_{maj} = \left\{ \begin{array}{l} \varphi = \{f_m(\pi \Downarrow a) \mid a \in \mathcal{A}\} \text{ if } \varphi \text{ is complete and consistent} \\ \perp \text{ otherwise} \end{array} \right\}$$

The JA function f_{maj} satisfies *universal domain*, *anonymity*, *unanimity*, *completeness*, and *independence* but it does not satisfy *collective rationality*, as it can be seen on Figure 1. All judgement aggregation functions which satisfy *universal domain*, *anonymity*, *independence* and *collective rationality* are *constant* [17]. The most debated [3, 17] is *independence*. The reason why it is convenient to have independence is because it is a necessary condition to guarantee the non-manipulability of f [8]. An aggregation function is non-manipulable, if no agent can obtain his sincere judgment set φ selected as the collective judgment set by submitting another judgment set φ'.

Two aggregation procedures that violate *independence* but guarantee *universal domain*, *anonymity* and *collective rationality* have been proposed in the literature: *premise-based* and *conclusion-based* procedures. A distinguishing feature of judgment aggregation with respect to social choice theory is the distinction between premises and conclusions. The agenda is a union of two disjoint sets: the premise set ($\mathcal{A}^p$), and the conclusion set ($\mathcal{A}^c$). $\mathcal{A} = \mathcal{A}^p \cup \mathcal{A}^c$, $\mathcal{A}^p \cap \mathcal{A}^c = \varnothing$. In Example 1, $\mathcal{A}^p = \{p_1, p_2, p_3\}$ and $\mathcal{A}^c = \{g_2\}$. We give a general definition on when a JA function is premise- or conclusion-based in Definition 4. The literature on judgment aggregation [17] defines premise- and conclusion-based procedures only in terms of issue-wise majority.

Definition 4. *Let* $\pi^p = \{(i, a) \mid (i, a) \in \pi, a \in \mathcal{A}^p\}$, $\pi^c = \{(i, a) \mid (i, a) \in \pi, a \in \mathcal{A}^c\}$ *and let the premise and conclusion aggregation functions be defined as* $f^p : \Pi \mapsto 2^{\mathcal{A}^p}/\varnothing$ *and* $f^c : \Pi \mapsto 2^{\mathcal{A}^c}/\varnothing$ *correspondingly. The JA function* f *is premise-based if and only if there exists a* f^p *such that* $f^p(\pi^p) \subseteq f(\pi)$ *and conclusion-based if and only if there exists a* f^c *such that* $f^c(\pi^c) \subseteq f(\pi)$, *for all* $\pi \in \Pi$.

An example of premise and conclusion-based aggregations is given in Figure 1. Intuitively, a JA function is premise-based when the collective judgment set on the premises is obtained as a result of aggregating only the premise profile π^p. The decisions $f(\pi)$ are those $\varphi \in \mathcal{W}$ for which $f^p(\pi^p) \subset \varphi$ (or alternatively $f^c(\pi^c) \subset \varphi$ for conclusion-based aggregations). However, it is possible that, depending on the decision rule, there are more than one $\varphi \in \mathcal{W}$ that satisfy the condition $f^p(\pi^p) \subset \varphi$ (or alternatively $f^c(\pi^c) \subset \varphi$). This is what happens for the conclusion-based procedure we illustrate in Figure 1.

2.2 Judgment aggregation for BDI agents

In this section first we set the problem of finding collective decisions for group goals in the context of judgment aggregation. We then argue that the aggregation function used for this problem should satisfy: *collective rationality*, *universal domain*, *unanimity* and select a unique $\varphi \in \mathcal{W}$. For a democratic group, in which all agents have equal say on what the group attitudes should be, *anonymity* should be satisfied. For a group in which the agents have different levels of expertise, *anonymity* can be omitted.

We use Gg to denote that "g is a group goal". A set of the candidate group goals is the set $\mathcal{G} = \{Gg \mid g \in \mathcal{L}\}$ which contains all the goals which the group considers to adopt. For each goal $Gg \in \mathcal{G}$, the group has at its disposal decision rules $\mathcal{R}_g \subseteq \mathcal{R}$ and an agenda $\mathcal{A}_g$ composed of the goal g (conclusion), and all the reasons (premises) relevant for g, which were given by the decision rule. Each agent submits her judgments on the agenda thus generating a profile π_g, such that $con(\mathcal{R}_g, \pi_g \Rightarrow i) = 1$ for all $i \in \mathcal{N}$.

The decision rules contain all the constraints which the individual and collective judgment sets should satisfy. These constraints contain three types of information: rules describing how the goal depends on the reasons (justification rules $\mathcal{R}_g^{just}$), rules describing the constraints of the world inhabited by the agents (domain knowledge $\mathcal{R}_g^{DK}$) and rules that describe how g interacts with other candidate goals of the group (coordination rules $\mathcal{R}_g^{coord}$). Hence, the decision rule for a group goal g is $\mathcal{R}_g = \mathcal{R}_g^{just} \cup \mathcal{R}_g^{DK} \cup \mathcal{R}_g^{coord}$.

Example 2 (Example 1 revisited). Consider the cleaning crew from Example 1. $\mathcal{R}_{g_1}^{just}$ is $(p_1 \wedge (p_2 \vee p_3)) \leftrightarrow Gg_1$ and $\mathcal{A}_{g_1} = \{p_1, p_2, p_3, Gg_1\}$. Suppose that the crew has the following candidate group goals as well: *place the furniture in its designated location* (g_2) and collect recyclables from garbage bin (g_3). The agendas are $\mathcal{A}_{g_2} = \{p_4, p_5, p_6, p_7, Gg_2\}$, $\mathcal{A}_{g_3} = \{p_3, p_8, p_9, Gg_3\}$. The justification rules are $\mathcal{R}_{g_2}^{just} \equiv (p_4 \wedge p_5 \wedge (p_6 \vee p_7)) \leftrightarrow Gg_2$ and $\mathcal{R}_{g_3}^{just} \equiv (p_8 \wedge p_9 \wedge p_3) \leftrightarrow Gg_3$. The formulas $p_4 - p_9$ are: the furniture is out of place (p_4), the designated location for the furniture is empty (p_5), the furniture has wheels (p_6) , the furniture has handles (p_7), the agents can get revenue for recyclables (p_8), there is a container for the recyclables (p_9).
An example of a domain knowledge could be $\mathcal{R}_{g_2}^{DK} \equiv \neg p_4 \rightarrow \neg p_5$, since it can not happen that the designated location for the furniture is empty while the furniture is not out of place. Group goal Gg_3 can be pursued at the same time as Gg_1, however, Gg_2 can only be pursued alone. Thus the coordination rule for all three goals is
$\mathcal{R}_{g_1}^{coord} = \mathcal{R}_{g_2}^{coord} = \mathcal{R}_{g_3}^{coord} \equiv ((Gg_2 \wedge \neg(Gg_1 \vee Gg_3)) \vee \neg Gg_2)$.

We want the justifications for a goal to explain when a goal should be adopted/refuted. Having collective justifications for a goal enables the agents to, when the world changes, consider adopting a goal that has been rejected previously. To this end, we take into consideration justification rules which are of form $Gg \leftrightarrow \Gamma$, where $\Gamma \in \mathcal{L}$ such that all the non-logical symbols of Γ occur in $\mathcal{A}_g^p$ and $\{Gg\} = \mathcal{A}_g^c$.

We now continue to discuss the desirable properties for the aggregation of individual beliefs and goals. We need a JA function that satisfies *universal domain* to be able to aggregate all admissible profiles. We can use only JA functions that satisfy *collective rationality*. If the collective set is not complete, for example, if it contains only a collective judgment on the goal, then we do not know why the goal was (not) adopted and consequently we would not know when to revise it. For example, the cleaning crew decides for the goal g_3 (to collect recyclables), without having the reasons like p_9 (a container where to put them). If the world changes and $\neg p_9$ holds, the robots will continue to collect recyclables.

The aggregation of all admissible profiles should produce a set consistent with the decision rule because otherwise we would not be generating the group goals and justifications for them. For the same reason we need an aggregation method that selects a unique $\varphi \in \mathcal{W}$.

To guarantee that $con(\mathcal{R}_g, f(\pi_g)) = 1$, we need to choose between conclusion-based and premise-based procedures. At first glance, the premise-based procedure seems an obvious choice since it will produce complete collective judgment sets under our decision rules. However, upon closer inspection, this procedure has notable drawbacks.

As it is observable from the profile in Figure 1; a premise-based procedure may lead

the group to adopt a conclusion that the majority (or even the unanimity) of the group does not endorse. In our case, the conclusion is the goal and a premise-based aggregation may establish a group goal which none of the agents is interested in pursuing. To avoid this we need to aggregate using a conclusion-based procedure. In particular we want a conclusion-based procedure that has the property of unanimity on Gg.

Given that our decision rules are of the form $g \leftrightarrow \Gamma$, there exist profiles for which a conclusion-based procedure will not produce complete collective set of judgments. However, the conclusion-based aggregation can be supplemented with an additional procedure that completes the produced set of judgments when necessary. Such aggregation procedure is the complete conclusion-based procedure (CCBP) developed in [20]. This CCBP satisfies *universal domain* and is *collectively rational*. However, it does not always produce a unique decision. CCBP produces a unique collective judgment for the goal, but it can generate more than one set of justifications for it. This is an undesirable, but not an unmanageable property. To deal with ties, as it is the practice with voting procedures, the group can either determine a default set of justifications for adopting/rejecting each candidate goal, or it can appoint one member of the group as tiebreaker. Tie-breaking problems in judgment aggregation are the focus of our ongoing research.

The CCBP from [20] also satisfies *anonimity*. Whether this is a desirable property for a group of artificial agents depends entirely on whether the group is democratic or the opinions of some agents are more important. CCBP can be adjusted to take into consideration different weights on different agents' judgment sets.

3 Formal framework for modeling group attitudes

In this section we introduce the language of modal multi-agent logic in which we represent individual and collective mental attitudes. We then combine the methodology of judgment aggregation with this representation language and show how collective attitudes are generated. To model commitment to a group goal and reconsideration of a group goal we use temporal logic.

3.1 Modal multi-agent logic

Just like modal agent logic is concerned mainly with the relation between the individual goals and beliefs over time, modal multi-agent logic is concerned with the relation between group goals and beliefs over time. We use modal multi-agent logic to represent: the agenda, individual judgments, collective judgments and new information that may prompt goal revision. In line with judgment aggregation proper, we do not use the formal language to represent the judgment aggregation function, but only the arguments of this function and the results from it. We assume that there is a service, available to the agents, that elicits the individual judgments, performs the aggregation and makes the aggregation results available, to the members and for plan-generation.

Agenda issues in judgment aggregation are usually represented by propositional formulas. This is not a viable option in our case. First, we want to represent the difference between a goal and the supporting reasons by means of representing the distinction between conclusions and premises explicitly in the logic. Second, the logic should represent the distinction between individual and collective judgments. We distinguish

conclusions from premises by using a single K modal operator Gg for representing that "g is a group goal".

The obvious choice for modeling the judgment "true" on agenda issue a, of an agent i, is the modal operator belief $B_i a$ (correspondingly $B_i \neg a$ for a judgment "false" on a). However, we find that beliefs are ill suited for modeling collective judgments of agents. While a belief $B_i a$ is an exclusively private mental state, judgments are public and contributed for the decision-making purposes of the group. A judgment is thus closer to a public commitment than to a private belief. Hence, we model judgments by using the modal operator of acceptance $A_S a$ [19]. $A_S a$ denotes: agents in S accept a. The operator $A_S a$ allows us to represent both individual judgments, $S = \{i\}$, for $i \in \mathcal{N}$ and collective judgments with $S = \mathcal{N}$. We define the group acceptance $A_{\mathcal{N}} a$ to be the result of applying judgment aggregation to the individual acceptances. We present the formal definitions of profile and judgment aggregation function in our logic in Definition 6. The modal operator A_S we use is inspired by the modal operator of the *acceptance logic* of [19]. The details on the relation between acceptance logic and our acceptance operator are given in the Appendix.

We represent the new information that becomes available to the agents with a normal modal K operator E. Ea denotes: "it is observed that a is true".

Lastly, to model how the group attitudes evolve with reconsideration we need a temporal logic. By using LTL we do not need to distinguish between path formulas and state formulas, but we are able to quantify over traces. Just as in BDI_{LTL}, where $B\square a$ denotes that a is believed to be necessary, we use $E\square \neg a$ to denote that a is observed to be impossible.

We now give the syntax of our modal multi-agent logic AGE_{LTL}.

Definition 5 (Syntax). *Let Agt be a non-empty set of agents, with $S \subseteq Agt$, and L_P be a set of atomic propositions. The admissible formulae of AGE_{LTL} are formulae ψ_0, ψ_1 and ψ_2 of languages $\mathcal{L}_{prop}$, $\mathcal{L}_G$ and $\mathcal{L}_{AE_{LTL}}$ correspondingly:*
$\psi_0 ::= p \mid (\psi_0 \wedge \psi_0) \mid \neg\psi_0$
$\psi_1 ::= \psi_0 \mid G\psi_0$
$\psi_2 ::= \psi_0 \mid A_S\psi_1 \mid E\psi_2 \mid \boldsymbol{X}\psi_2 \mid (\psi_2 \boldsymbol{U} \psi_2)$
where p ranges over L_P and S over 2^{Agt}. Moreover, $\Diamond\phi \equiv \top\boldsymbol{U}\phi$, $\square\phi \equiv \neg\Diamond\neg\phi$, and $\phi\boldsymbol{R}\phi' \equiv \neg(\neg\phi\boldsymbol{U}\neg\phi')$.

We can now adjust the definition for a judgment aggregation function. We represent individual judgments with $A_{\{i\}}a$ with intuitive reading "agent i judges a as true" and $A_{\{i\}}\neg a$ with reading "agent i judges a as false".

Definition 6 (JA in AGE_{LTL}). *Let $\mathcal{N} = \{1, 2, \ldots\}$ be a set of agent names and $\mathcal{G} \subseteq \mathcal{L}_G/\mathcal{L}_{prop}$ a set of candidate goals.*
An agenda $\mathcal{A}_g \subseteq \mathcal{L}_G$ for goal $Gg \in \mathcal{G}$ is a set of formulas such that $\mathcal{A}_g = \overline{\mathcal{A}_g^p} \cup \overline{\mathcal{A}_g^c}$, with $\overline{\mathcal{A}_g^p} = \mathcal{A}_g^p \cup \{\neg a \mid a \in \mathcal{A}_g^p\}$, $\mathcal{A}_g^p \subseteq \mathcal{L}_{prop}$ and $\overline{\mathcal{A}_g^c} = \{Gg, \neg Gg\}$.
A profile of judgments is the set $\overline{\pi} = \{A_{\{i\}}a \mid i \in \mathcal{N}, a \in \mathcal{A}_g\}$.
$\overline{\pi} \Rightarrow i = \{a \mid A_{\{i\}}a \in \overline{\pi}\}$ is the judgment set of agent i.
$\overline{\pi} \Downarrow a = \{i \mid A_{\{i\}}a \in \overline{\pi}\}$ is the set of all the agents that accepted a.
Given a set of decision rules $\mathcal{R}_g \subset \mathcal{L}_{prop}$, a profile is acceptable if and only if, for all $i \in \mathcal{N}$ and all $a \in \mathcal{A}$, $con(\mathcal{R}_g, \pi \Rightarrow i) = 1$ and either $A_{\{i\}}a$ or $A_{\{i\}}\neg a$. The set of all

acceptable profiles for $\mathcal{N}$ and $\mathcal{A}_g$, given $\mathcal{R}_g$ is $\overline{\Pi}$.

The decision for a profile $\overline{\pi}$, $\mathcal{D}_{\overline{\pi}} = \{A_{\mathcal{N}} a \mid a \in f^a(\overline{\pi}), f^a : \overline{\Pi} \mapsto 2^{\mathcal{A}_g}\}$.

For instance, the profile in Figure 1, is $\overline{\pi}_{g_1} = \{A_{\{w_1\}} p_1, \neg A_{\{w_1\}} p_2, \neg A_{\{w_1\}} p_3,$ $A_{\{w_1\}} \neg Gg_1, A_{\{w_2\}} \neg p_1, \ldots, A_{\{w_3\}} Gg_1\}$. The decision using premise-based majority would be $\{A_C p_1, A_C p_2, A_C p_3, A_C Gg_1\}$.

AGE_{LTL} can be used to model that an agent does not have a judgment on an agenda issue, but we will work with the assumption that either $A_{\{i\}}$ or $A_{\{i\}} \neg a$ for all agents and agenda issues.

AGE_{LTL} has Kripke semantics. As in Schild [22], a Kripke structure is defined as a tuple $\mathcal{M} = \langle W, \mathcal{R}, \mathcal{G}, \mathcal{E}, \mathcal{A}, L \rangle$. W is a set of possible situations, and $\mathcal{R}$ is the temporal relation over situations $\mathcal{R} \subseteq W \times W$. $\mathcal{G}$ is the goal relation over situations $\mathcal{G} \subseteq W \times W$, while $\mathcal{E}$ is the observation relation over situations $\mathcal{E} \subseteq W \times W$. Let $\Delta = 2^{\mathcal{N}} \times Inst$. $\mathcal{A} : \Delta \mapsto W \times W$ maps every $S \in \Delta$ to a relation $\mathcal{A}_S$ between possible situations. L is a truth assignment to the primitive propositions of L_P for each situation $w \in W$, i.e., $L(w) : Prop \mapsto \{true, false\}$.

Temporal formulas are validated in the standard manner [15]. Normal modal formulas $G\psi$ and $E\psi$ have standard semantics, see for example [4]. Acceptance formulas $A_S \psi$ are validated according to the semantics of acceptance logic presented in [19]. The axiomatization of AGE_{LTL} is given in the Appendix. Note that AGE_{LTL} is a fusion of the decidable logics: LTL, acceptance logics and two K modal logics.

3.2 Generation of group goals

The mental state of the group is determined by the mental states of the members and the choice of judgment aggregation function. We represent the mental state of the group by a set χ of AGE_{LTL} formulas. The set χ contains: the set of all candidate goals for the group $\mathcal{G} \subseteq \mathcal{L}_G / \mathcal{L}_{prop}$ and, for each $Gg \in \mathcal{G}$, the corresponding decision rules $\mathcal{R}_g$, as well as the individual and collective acceptances made in the group regarding agenda $\mathcal{A}_g$. The set χ is common knowledge for the group members. An agent uses χ when it acts as a group member and its own beliefs and goals when it acts as an individual.

To deal with multiple, possibly mutually inconsistent goals, the group has a priority order $\succsim_x$ over the group goals $\mathcal{G} \subset \chi$. To avoid overburdening the language with a $\succsim_x$ operator, we incorporate the priority order within the decision rules $\mathcal{R}_{g_i}^{just} \equiv \Gamma_i \leftrightarrow Gg_i$. We want the decision rules to capture that if Gg_i is not consistent (according to the coordination rules) with some higher priority goals $Gg_1, \ldots, Gg_m$, then the group can accept Gg_i if and only if none of $Gg_1, \ldots, Gg_m$ is accepted. Hence, we replace the justification rule $\mathcal{R}_{g_i}^{just} \in \chi$ with $\mathcal{R}_{g_i}^{pjust} \equiv (\Gamma_i \wedge \bigwedge_j^m (A_{\mathcal{N}} \neg Gg_j)) \leftrightarrow Gg_i$, where $Gg_j \in \mathcal{G}$, $Gg_j \succsim_x Gg_i$ and $Gg_i \wedge Gg_j \wedge \mathcal{R}_{g_i}^{coord} \models \bot$.

Example 3. Consider the goals and rules of the robot crew C from Example 2. Assume the crew has been given the priority order $Gg_1 >_\chi Gg_2 >_\chi Gg_3$. χ contains: $\mathcal{G} = \{Gg_1, Gg_2, Gg_3\}$, one background knowledge rule, one coordination rule, three justification rules, out of which two are new priority modified rules:
$\{\mathcal{G}, \neg p_4 \rightarrow \neg p_5, (Gg_2 \wedge \neg(Gg_1 \vee Gg_3)) \vee \neg Gg_2, Gg_1 \leftrightarrow (p_1 \wedge (p_2 \vee p_3)),$
$Gg_2 \leftrightarrow (p_4 \wedge p_5 \wedge (p_6 \vee p_7) \wedge A_C \neg Gg_1), Gg_3 \leftrightarrow (p_8 \wedge p_9 \wedge p_3 \wedge (A_C \neg Gg_2)\}$.

The agents give their judgments on one agenda after another starting with the agenda for the highest priority candidate goal. Once the profile $\overline{\pi}$ and the decision $\mathcal{D}_{\overline{\pi}}$ for a goal g are obtained, they are added to χ. To avoid the situation in which the group casts judgments on an issue that has already been decided, we need to remove decided issues from $\mathcal{A}_g$ before eliciting the profile for this agenda.

The group goals are generated by executing **GenerateGoals**$(\chi, \mathcal{N})$.

```
function GenerateGoals(χ, S):
for each Ggᵢ ∈ G s.t. [∀Ggⱼ ∈ G: (Ggⱼ ≳ Ggᵢ) ⇒ (A_N Ggⱼ ∈ χ or A_{C:χ}¬Ggⱼ ∈ χ)]
```
$\quad\quad \{ \ B := (\{a \mid A_{\mathcal{N}} a \in \chi\} \cup \{\neg a \mid A_{\mathcal{N}} \neg a \in \chi\}) \cap \mathcal{A}_{g_i};$
$\quad\quad\quad$ *comment // B is the set of already collectively accpted issues from $\mathcal{A}_{g_i}$*
$\quad\quad\quad \mathcal{A}_{g_i}^* := \mathcal{A}_{g_i}/B;$
$\quad\quad\quad \overline{\pi}_{g_i} := elicit(S, \mathcal{A}_{g_i}^*, \chi);$
$\quad\quad\quad \chi := \chi \cup \overline{\pi}_{g_i} \cup f^a(\overline{\pi}_{g_i}); \ \ \}$
```
return χ.
```

$elicit$ requests the agents to submit complete judgment sets for $\overline{\pi}_{g_i} \subset \chi$. We require that $elicit$ is such that for all returned $\overline{\pi}$ it holds $con(\chi, \overline{\pi} \Rightarrow i) = 1$ for all $i \in \mathcal{N}$ and that $con(\chi, f^a(\overline{\pi})) = 1$. When a higher priority goal Gg_i is accepted by the group, a lower priority incompatible goal Gg_j cannot be adopted regardless of the judgments on the issues in $\mathcal{A}_{g_j}$. Nevertheless, Although $elicit$ will provide individual judgments for the agenda $\mathcal{A}_{g_j}$. If the acceptance of Gg_i is reconsidered, we can obtain a new decision on Gg_j because the profile for Gg_j is available.

Example 4. Consider the χ for robots given in Example 3. The following calls to $elicit$ are made in the given order. First, $\overline{\pi}_{g_1} = elicit(\mathcal{N}, \mathcal{A}_{g_1}^*, \chi)$ with the $GenerateGoals(\chi) = \chi' = \chi \cup \overline{\pi}_{g_1} \cup f^a(\overline{\pi}_{g_1})$. Second, $\overline{\pi}_{g_2} = elicit(\mathcal{N}, \mathcal{A}_{g_2}^*, \chi')$, with $GenerateGoals(\chi') = \chi'' = \chi' \cup \overline{\pi}_{g_2} \cup f^a(\overline{\pi}(g_2))$. Last, $\overline{\pi}_{g_3} = elicit(\mathcal{N}, \mathcal{A}_{g_3}^*, \chi'')$, with $GenerateGoals(\chi'') = \chi''' = \chi'' \cup \overline{\pi}_{g_3} \cup f^a(\overline{\pi}_{g_3})$. Since there is no overlapping between agendas $\mathcal{A}_{g_2}$ and $\mathcal{A}_{g_1}$, $\mathcal{A}_{g_1}^* \equiv \mathcal{A}_{g_1}$ and $\mathcal{A}_{g_2}^* \equiv \mathcal{A}_{g_2}$. However, since $\mathcal{A}_{g_2} \cap \mathcal{A}_{g_3} = p_3$, then $\mathcal{A}_{g_3}^* = \{p_8, p_9, Gg_3\}$.

Proposition 1. *$GenerateGoals$ terminates if and only if $elicit$ terminates and does not violate the candidate goal preference order.*

The proof is straightforward.

4   Commitment strategies

The group can choose to reconsider the decision (acceptance or rejection) on a group goal in presence of new information. Whether the group chooses to reconsider depends on how committed it is to bring the goal about. According to Cohen and Levesque [5], an agent intends g if it has chosen to pursue goal g and it committed itself to making g happen. Following the same line of reasoning, we define group intention to be $I_{\mathcal{N}} g \equiv A_{\mathcal{N}} Gg$ and read it as "the agents $\mathcal{N}$ intend g". We defined collective acceptance as resulting from judgment aggregation, which is a social choice method. Thus, in our framework, group intention is social choice with commitment. The level of commitment of a group to a collective acceptance depends on the choice of commitment strategy.

These are the three main commitment strategies (introduced by Rao and Georgeff [21]):

Blind commitment: $I_i g \rightarrow (I_i g \mathbf{U} B_i g)$

Single-minded commitment: $I_i g \rightarrow (I_i g \mathbf{U} (B_i g \vee B_i \square \neg g))$

Open-minded commitment: $I_i g \rightarrow (I_i g \mathbf{U} (B_i g \vee \neg G_i g))$

These commitment strategies only consider the relation between beliefs regarding g and Gg. Instead, a commitment to a goal can now be reconsidered upon new information on either one of the agenda issues in $\mathcal{A}_g$ and also upon new information on a higher priority goal.

The strength of our framework is exhibited in its ability to describe the groups' commitment not only to its decision to adopt a goal, but also to its decision to reject a goal. Namely, if the agents decided $I_{\mathcal{N}} g_i$ and $A_{\mathcal{N}} \neg g_j$ are committed to both $I_{\mathcal{N}} g_i$ and $A_{\mathcal{N}} \neg g_j$. Commitment to reject g allows for g to be reconsidered and eventually adopted if the state of the world changes.

Let $\mathcal{N}$ be a set of agents with a set of candidate goals $\mathcal{G}$. Let $Gg_i, Gg_j \in \mathcal{G}$ have agendas $\mathcal{A}_{g_i}, \mathcal{A}_{g_j}$. We use $p \in \mathcal{A}_{g_i}^p$ and $q_i \in \mathcal{A}_{g_i}^c$, $q_j \in \mathcal{A}_{g_j}^c$. The profiles and decisions are $\overline{\pi}_{g_i}$ and $f^a(\overline{\pi}_{g_i})$; $Gg_j > Gg_i$, and Gg_j cannot be pursued at the same time as Gg_i.

We use the formulas $(\alpha_1) - (\alpha_7)$ to refine the blind, single-minded and open-minded commitment. Instead of the "until", we use the temporal operator *release*: $\psi \mathbf{R} \phi \equiv \neg(\neg\psi \mathbf{U} \neg\phi)$, meaning that ϕ has to be true until and including the point where ψ first becomes true; if ψ never becomes true, ϕ must remain true forever. Unlike the *until* operator, the *release* operator does not guarantee that right hand-side formula will ever become true, which in our case translates to the fact that an agent could be forever committed to a goal.

(α_1) $Eg_i \mathbf{R} I_{\mathcal{N}} g_i$

(α_2) $\bot \mathbf{R} A_{\mathcal{N}} \neg Gg_i$

(α_3) $(E\square\neg g_i \vee Eg_i) \mathbf{R} A_{\mathcal{N}} q_i$

(α_4) $A_{\mathcal{N}} \neg q_j \mathbf{R} A_{\mathcal{N}} q_i$

(α_5) $A_{\mathcal{N}} p \rightarrow (E\neg p \mathbf{R} A_{\mathcal{N}} q_i)$

Blind commitment: $\alpha_1 \wedge \alpha_2$.

Only the observation that the goal is achieved (Eg_i) can release the intention to achieve the goal $I_{\mathcal{N}} g_i$. If the goal is never achieved, the group will always be committed to it. If a goal is not accepted, then the agents will not reconsider accepting it.

Single-minded commitment: α_3.

Only new information on the goal (either that the goal is achieved or had become impossible) can release the decision of the group to adopt /reject the goal. Hence, new information is only regarded if it concerns the conclusion, while information on the remaining agenda items is ignored.

Extended single-minded commitment: $\alpha_3 \wedge \alpha_4$.

Not only new information on g_i, but also the collective acceptance to adopt a more important incompatible goal g_j can release the intention of the group to achieve g_i. Similarly, if g_i is not accepted, the non-acceptance can be revised, not only if g_j is observed to be impossible or achieved, but also when the commitment to pursue g_j is dropped (for whatever reason).

Open-minded commitment: $\alpha_3 \wedge \alpha_5$.

A group will maintain its collective acceptances to adopt/reject a goal as long as the new information regarding all collectively accepted agenda items is consistent with $f^a(\overline{\pi}_{g_i})$.

Extended open-minded commitment: $\alpha_3 \wedge \alpha_4 \wedge \alpha_5$.

Extending on the single-minded commitment, a change in intention to pursue a higher priority goal Gg_j can also release the acceptance of the group on Gg_i.

Once an intention is dropped, a group may need to reconsider its collective acceptances. This may cause for the dropped goal to be re-affirmed, but a reconsideration process will be invoked nevertheless.

5 Reconsideration strategies

In Section 3.2 we defined the mental state of the group χ. We can now define what it means for a group to be *coherent*.

Definition 7 (Group coherence). *Given a Kripke structure $\mathcal{M}$ and situations $s \in W$, a group of $\mathcal{N}$ agents is coherent if the following conditions are met:*

(ρ_1): $\mathcal{M} \models \neg(A_S a \wedge A_S \neg a)$ for any $S \subseteq \mathcal{N}$ and any $a \in \mathcal{A}_g$.

(ρ_2): If $\mathcal{M}, s \models \chi$ then $\chi \not\models \bot$.

(ρ_3): $\mathcal{M}, s \models \bigwedge \mathcal{G} \rightarrow \neg\square\neg g$ for all $Gg \in \mathcal{G}$.

(ρ_4): Let $Gg \in \mathcal{G}$ and $\mathcal{G}' = \mathcal{G}/\{Gg\}$, then $\mathcal{M} \models (\bigwedge \mathcal{G} \wedge E\square\neg g) \rightarrow \boldsymbol{X}(\neg Gg)$.

(ρ_5): Let $p \in \mathcal{A}_g^p$ and $q \in \{Gg, \neg Gg\}$. $Ep \wedge (Ep \, \boldsymbol{R} \, A_{\mathcal{N}} q) \rightarrow \boldsymbol{X} A_{\mathcal{N}} p$

The first condition ensures that no contradictory judgments are given. The second condition ensures that the mental state of the group is logically consistent in all situations. The third and fourth conditions ensure that impossible goals cannot be part of the set of candidate goals and if g is observed to be impossible in situation s, then it will be removed from $\mathcal{G}$ in the next situation. ρ_5 enforces the acceptance of the new information on the group level, when the commitment strategy so allows – after a is observed and that lead the group to de-commit from g, the group necessarily accepts a.

A coherent group accepts the observed new information on a premise. This may cause the collective acceptances to be inconsistent with the justification rules. Consequently, the decisions and/or the profiles in χ need to be changed in to ensure that ρ_1 and ρ_2 are satisfied. If, however $\square\neg g$ or g is observed, the group reconsiders χ by removing Gg from $\mathcal{G}$. In this case, the decisions and profiles are not changed.

For simplicity, at present we work with a world in which the agents' knowledge can only increase, namely the observed information is not a fluent. A few more conditions need to be added to the definition of group coherence, for our model to be able to be applicable to fluents. E.g., we need to define which observation is accepted when two subsequent contradictory observations happen.

For the group to be coherent at all situations, the acceptances regarding the group goals need to be reconsidered after de-commitment. Let $\mathcal{D}_g \subset \chi$ contain the group acceptances for a goal g, while $\overline{\pi}_g \subset \chi$ contain the profile for g. There are two basic ways in which a collective judgment set can be reconsidered. The first way is to elicit a new profile for g and apply judgment aggregation to it to obtain the reconsidered $\mathcal{D}_g^*$. The second is to reconsider only $\mathcal{D}_g$ without re-eliciting individual judgments. The first approach requires communication among agents. The second approach can be done

by each agent reconsidering χ by herself. We identify three reconsideration strategies available to the agents. The strategies are ordered from the least to the most demanding in terms of agent communication.

Decision reconsideration ($\mathcal{D}$-r). Assume that Ep, $p \in \mathcal{A}_g^p$, $q \in \{Gg, \neg Gg\}$ and the group de-commited from $A_{\mathcal{N}}q$. The reconsidered decision $\mathcal{D}_g^*$ is such that p is accepted, i.e., $A_{\mathcal{N}}p \in \mathcal{D}_g^*$, and the entire decision is consistent with the justification rules, namely $con(\mathcal{R}_g^{pjust}, \mathcal{D}_g^* \Rightarrow \mathcal{N}) = 1$. If the $\mathcal{D}$-r specifies an unique $\mathcal{D}_g^*$, for any observed information and any $\mathcal{D}_g$, then χ can be reconsidered without any communication among the agents. Given the form of $\mathcal{R}_g^{pjust}$ (see Section 3.2), this will always be the case.

However $\mathcal{D}$-r is not always an option when the de-commitment occurred due to a change in collective acceptance of a higher priority goal g'. Let $q' \in \{Gg', \neg Gg'\}$. Let the new acceptance be $A_{\mathcal{N}}\neg q'$. $\mathcal{D}$-r is possible if and only if $\mathcal{D}_g^* = \mathcal{D}_g$ and $con(\mathcal{R}_g^{pjust}, \mathcal{D}_g \cup \{A_{\mathcal{N}}\neg q'\}) = 1$. Recall that $A_{\mathcal{N}}q'$ was not in $\mathcal{A}_g$ and as such the acceptance of q' or $\neg q'$ is never in the decision for $\overline{\pi}_g$.

Partial reconsideration of the profile (Partial $\overline{\pi}$-r). Assume that Ea, $a \in \mathcal{A}_g$, $Gg \in \mathcal{G}$. Not only the group, but also the individual agents need to accept a. The *Partial $\overline{\pi}$-r* asks for new individual judgments be elicited. This is done to ensure the logical consistency of the individual judgment sets with the observations. New judgments are only elicited from the agents i which $A_{\{i\}}\neg a$.

Let $W \subseteq \mathcal{N}$ be the subset of agents i s.t. $A_{\{i\}}\neg a \in \chi$. Agents i are s.t. $A_{\{i\}}a \in \chi$ when the observation is $E\neg a$. Let $\overline{\pi}_g^W \subseteq \overline{\pi}_g$ be the set of all acceptances made by agents in W. We construct $\chi' = \chi/\overline{\pi}_g^W$. The new profile and decision are obtained by executing $GenerateGoals\,(\chi', W)$.

Example 5. Consider Example 4. For $\overline{\pi}_{g_1}, \overline{\pi}_{g_2}$ and $\overline{\pi}_{g_3}$ of the robot crew C, the decisions $\mathcal{D}_{g_1} = \{A_C p_1, A_C \neg p_2, A_C p_3, A_C Gg_1\}, \mathcal{D}_{g_2} = \{A_C p_4, A_C p_5, A_C p_6, A_C p_7, A_C \neg Gg_2\}$ and $\mathcal{D}_{g_3} = \{A_C p_8, A_C p_9, A_C Gg_3\}$ are made. Assume the group de-commits on Gg_1 because of $E\neg p_2$. If the group is committed to Gg_3, the commitment on Gg_3 will not allow for $A_{\mathcal{N}}p_3$ to be modified when reconsidering Gg_1. Since $A_{\mathcal{N}}p_3$ exists in χ', p_3 will be excluded from the (new) agenda for g_1, although it was originally in it. *elicit* calls only on the agents in W to complete $\overline{\pi}_{g_1} \in \chi'$ with their judgment sets.

Full profile reconsideration ($\overline{\pi}$-r). The full profile reconsideration is the same with the partial reconsiderations in all respects except one – now $W = \mathcal{N}$. Namely, within the full profile revision strategy, each agent is asked to revise his judgment set by accepting the new information, regardless whether he had already accepted the information or not.

5.1 Combining revision and commitment strategies

Unlike the Rao and Georgeff commitment strategies [21], in our framework the commitment strategies are not axioms of the logic. We require that the commitment strategy is valid in all the models of the group and not in all the models of AGE_{LTL}. This allows the group to define different commitment strategies and different revision strategies for different goals. It might even choose to revise differently depending on which information triggered the revision. Choosing different revision strategies for each goal, or each

type of new information, should not undermine the coherence of the group record χ. The conditions of group coherence of the group ensures that after every reconsideration χ must remain consistent. However, some combinations of commitment strategies can lead to incoherence of χ.

Example 6. Consider the profiles and decisions in Example 5 . Assume that initially the group chose open-minded commitment for $I_C g_1$ and blind commitment for $I_C g_3$, with goal open-minded commitment for $A_C \neg G g_2$. If $E g_1$ and thus $I_C g_1$ is dropped, then the extended open-minded commitment would allow $A_C \neg G g_2$ to be reconsidered and eventually $I_C g_2$ established. However, since the group is blindly committed to $I_C g_3$, this change will not cause reconsideration and as a result both $I_C g_2$ and $I_C g_3$ will be in χ thus making χ incoherent.

Problems arise when $sub(\mathcal{R}_{g_i}^{pjust}) \cap sub(\mathcal{R}_{g_j}^{pjust}) \neq \varnothing$, where $sub(\mathcal{R}_{g}^{pjust})$ denotes the set of atomic sub-formulas of g $(Gg_i, Gg_j \in \mathcal{G})$. Proposition 2 summarizes under which conditions these problems are avoided.

Proposition 2. *Let α' and α'' be the commitment strategies selected for g_i and g_j correspondingly. $\chi \cup \alpha' \cup \alpha'' \neg \not\models \bot$ (in all situations):*
a) if $\phi \in sub(\mathcal{R}_{g_i}^{pjust}) \cap sub(\mathcal{R}_{g_j}^{pjust})$ and $p \in \mathcal{A}_{g_i} \cap \mathcal{A}_{g_j}$, then α_5 is either in both α' and α'' or in none;
b) if Gg_i is more important than Gg_j and G_j and G_i cannot be accepted at the same time, then $\alpha_4 \in \alpha''$.

Proof. The proof is straightforward. Namely, if the change on acceptance of Gg_i causes the decision on Gg_j to induce group incoherence, we are able to de-comit from Gg_j. If we were not able to de-comit on Gg_j group coherence is blocked. If the change on collective acceptance on Gg_i is caused by an observation on a premise $p \in \mathcal{A}_{g_i} \cap \mathcal{A}_{g_j}$ then condition a) ensures that the commitment to the collective acceptance regarding Gg_j does not block group coherence. If the change on collective acceptance on Gg_j is caused by a change in commitment to a higher priority goal the condition b) ensures that a commitment regarding Gg_j does not block group coherence. Condition b) allows only "goal sensitive" commitments to be selected for lower level goals.

6   Conclusions

We constructed a group decision-making framework by combining judgment aggregation and multi-agent modal logic. We identified the desirable judgment aggregation properties for aggregation in collaborative groups. Our multi-agent modal logic AGE_{LTL} is an extension of BDI_{LTL} with modal operators for representing individual and collective acceptances and observations of new information. We extend the commitment strategies of Rao and Georgeff [21] to increase the reactivity of the group to new information. Having a group goal Gg in our framework does not imply that the members individually have the goal Gg and groups can have different levels of commitment to different goals.

Our framework is intended for groups that engage in joint activity. Our framework is applicable when it is not possible to assume that the agents persuade each other on a single position and goal, but it is necessary anyway that the group presents itself as

a single whole from the point of view of beliefs and goals, and above all as a rational entity that has goals justified by the beliefs it holds, and it is able to revise these goals under the light of new information. This requirement was held by Tuomela [24] and adopted in agent theory by Boella and van der Torre [1] and Lorini [18]. There are many situations where the proposal of the paper can be applied. For example in an opensource project, where several people have to discuss online to agree on which is their position on issues (e.g. which algorithm is better for a certain task) and which is their goal (e.g. delivering a new realize on which date).

Work on collaborative group activities [9, 12, 16] and group decision-making protocols [10] focus on how to define collective intentionality and how to distribute the collective intentions over the agents. We define group intention to be the collective acceptance of a group goal and focus on defining commitment strategies for the collective acceptances. An advantage of our framework is its ability to allow groups to commit to a decision to reject a goal, thus having the option to reconsider rejected goals. Furthermore, we do not only show when to reconsider, but also how, by defining reconsideration strategies. Table 1 summarizes our commitment and reconsideration strategies.

Commitment to	Release on					Change	How		
$A_{\mathcal{N}}(\neg)Gg$	$\Box\neg g$	g	Gg_j	$\mathcal{A}_g^p$		χ	$\circledast\mathcal{D}(g)$	$\circledast\overline{\pi}_g$	JA
Blind	✓								
Single-minded	✓	✓				$\mathcal{D}$-r	✓		
Extended	✓	✓	✓			Partial $\overline{\pi}$-r		✓	✓
Open-minded	✓	✓		✓		Full $\overline{\pi}$-r		✓	✓
Extended	✓	✓	✓	✓					

Table 1. $Gg_j > Gg$ and can not be pursued at the same time with Gg. $\circledast\mathcal{D}(g)$ denotes: collective attitudes for g are reconsidered. $\circledast\,\overline{\pi}_g$ denotes: the profile (all or some parts of it) is re-elicited.

Icard et al. [23] consider the joint revision of individual attitudes, with the revision of beliefs triggering intention revision. However, they do not allow for the revision of intentions to cause a revision of beliefs. We focus on joint reconsideration of group attitudes and we allow for both the change in epistemic and in motivational attitudes can be a cause for reconsideration.

In our framework, the new information is simultaneously available to the entire group. In the future we intend to explore the case when only some members of the group observe the new information. The only assumptions we make regarding the connectivity of the members is that they are able to communicate their acceptances and receive the aggregation result. The problem of elicitation and communication complexity in voting is a nontrivial one [6, 7] and we intend to study these properties within our framework.

In the work we presented, we do not consider how individual acceptances are formed. We can take that $B_i\phi \rightarrow A_{\{i\}}\phi$, but this need not be the case. We can define dishonest agents as those for which $B_i\phi \rightarrow A_{\{i\}}\phi$ does not hold. In this case, the agent might declare $A_{\{i\}}\phi$ while it does not believe ϕ. The question is whether there are scenarios in which incentives for doing so arise. Furthermore, given that the some reconsideration strategies call for re-elicitation of judgments, can an agent have the incentive to deliberately give judgments that would lead to sooner re-elicitation? We intend to devote more attention to answering these questions as well as studying the manipulability properties of our decision-making framework.

References

1. G. Boella and L. van der Torre. The ontological properties of social roles in multi-agent systems: Definitional dependence, powers and roles playing roles. *Artificial Intelligence and Law Journal (AILaw)*, 15(3):201–221, 2007.

2. C. Castelfranchi and F. Paglieri. The role of beliefs in goal dynamics: Prolegomena to a constructive theory of intentions. *Synthese*, 155:237–263, 2007.

3. B. Chapman. Rational aggregation. *Politics, Philosophy and Economics*, 1(3):337–354, 2002.

4. B. F. Chellas. *Modal Logic: An Introduction*. Cambridge University Press, Cambridge, 1980.

5. P. R. Cohen and H. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42(2-3):213–261, 1990.

6. V. Conitzer and T. Sandholm. Vote elicitation: Complexity and strategy-proofness. In *AAAI/IAAI*, pages 392–397, 2002.

7. V. Conitzer and T. Sandholm. Communication complexity of common voting rules. In *ACM Conference on Electronic Commerce*, pages 78–87, 2005.

8. F. Dietrich and C. List. Strategy-proof judgment aggregation. STICERD - Political Economy and Public Policy Paper Series 09, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE, Aug 2005.

9. B. Dunin-Keplicz and R. Verbrugge. Collective intentions. *Fundam. Inf.*, 51(3):271–295, 2002.

10. B. Grosz and L. Hunsberger. The dynamics of intention in collaborative activity. *Cognitive Systems Research*, 7(2-3):259–272, 2007.

11. S. Hartmann, G. Pigozzi, and J. Sprenger. Reliable methods of judgment aggregation. *Journal of Logic and Computation*, forthcoming.

12. N. R. Jennings. Controlling cooperative problem solving in industrial multi-agent systems using joint intentions. *Artif. Intell.*, 75(2):195–240, 1995.

13. S. Konieczny and R. Pino-Pérez. Merging with integrity constraints. *Lecture Notes in Computer Science*, 1638/1999:233–244, 1999.

14. L. Kornhauser and L. Sager. The one and the many: Adjudication in collegial courts. *California Law Review*, 81:1–51, 1993.

15. F. Kröger. *Temporal Logic of programs*. Springer, Berlin, 1987.

16. H. J. Levesque, P. R. Cohen, and J. H. T. Nunes. On acting together. In *Proceedings of the Eighth National Conference on Artificial Intelligence*, pages 94–99, 1990.

17. C. List and C. Puppe. Judgment aggregation: A survey. In P. Anand, C. Puppe, and P. Pattanaik, editors, *Oxford Handbook of Rational and Social Choice*. Oxford, 2009.

18. E. Lorini and D. Longin. A logical account of institutions: From acceptances to norms via legislators. In *KR*, pages 38–48, 2008.

19. E. Lorini, D. Longin, B. Gaudou, and A. Herzig. The logic of acceptance. *Journal of Logic and Computation*, 19(6):901–940, 2009.

20. G. Pigozzi, M. Slavkovik, and L. van der Torre. A complete conclusion-based procedure for judgment aggregation. In *First International Conference on Algorithmic Decision Theory Proceedings*, pages 1–13, 2009.

21. A. S. Rao and M. P. Georgeff. Intentions and rational commitment. In *In First Pacific Rim Conference on Artificial Intelligence (PRICAI-90) Proceedings*, pages 94–99, 1993.

22. K. Schild. On the relationship between bdi logics and standard logics of concurrency. *Autonomous Agents and Multi-Agent Systems*, 3(3):259–283, 2000.

23. T.Icard, E.Pacuit, and Y. Shoham. A dynamic logic of belief and intention. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Twelfth International Conference (KR-10)*, page forthcoming, 2010.

24. R. Tuomela and K. Miller. Groups beliefs. *Synthese*, 91:285–318, 1992.

Appendix – Relations between AGE_{LTL} and acceptance logic and axiomatization of AGE_{LTL}

Here we elaborate in more detail on the fusion logic AGE_{LTL} we use. The modal operator $A_S\phi$ we use is equivalent to the modal operator $A_{S:x}\phi$ of the *acceptance logic* of [19] with one syntactic and one semantic exception.

The operator $A_{S:x}\phi$ uses x ranging over a set of labels to describe the context under which the acceptance is made. In our case the context is that of the group and since we deal with only one group, we have no use of these labels. The context labels play no role in the semantics of the acceptance logic formulas.

On the semantic level, the axioms for $A_S\phi$ are all the axioms of $A_{S:x}\phi$ except two: the axiom inclusion ($Inc.$) and the axiom unanimity ($Un.$). Dropping ($Un.$) and ($Inc.$) does not affect the decidability of the logic of acceptance. ($Un.$) (not to be confused with unanimity in judgment aggregation) states that if $A_{\mathcal{N}:x}\phi$, then $\forall i \in \mathcal{N}$, $A_{\{i\}:x}\phi$. In our case, it is the aggregation of individual acceptances that determines the collective acceptance. Since we use the acceptances to model judgments, we do not want an axiom that states that the individual judgments mirror the collective judgments. The agents use the collective acceptance when functioning as a group and their private beliefs when acting as individuals. In our framework we do not model the private mental states, but only individual acceptances which are "declared" to all the agents in the group.

($Inc.$) states that if a the group C accepts φ, so will any subgroup $B \subset C$. In our case, the judgment aggregation over the profile containing only the judgment sets of B can produce a different collective judgment set than the profile containing all the judgment sets of C.

The axiomatization of the AGE_{LTL} logic is thus:

(**ProTau**) All principles of propositional calculus
(**LTLTau**) All axioms and derivation rules of LTL
(**K-G**) $G(\phi \to \psi) \to (G\phi \to G\psi)$
(**K-E**) $E(\phi \to \psi) \to (E\phi \to E\psi)$
(**K-A**) $A_S(\phi \to \psi) \to (A_S\phi \to A_S\psi)$
(**PAccess**) $A_S\phi \to A_M A_S\phi$ if $M \subseteq S$
(**NAccess**) $\neg A_S\phi \to A_M \neg A_S\phi$ if $M \subseteq S$
(**Mon**) $\neg A_S\bot \to \neg A_M\bot$ if $M \subseteq S$
(**MP**)   From $\vdash \phi$ and $\vdash (\phi \to \psi)$ infer $\vdash \psi$
(**Nec-A**) From $\vdash \phi$ infer $\vdash A_S\phi$
(**Nec-G**) From $\vdash \phi$ infer $\vdash G\phi$
(**Nec-E**) From $\vdash \phi$ infer $\vdash E\phi$

Given $\mathcal{M} = \langle W, \mathcal{R}, \mathcal{G}, \mathcal{E}, \mathcal{A}, L \rangle$ and $s \in W$, the truth conditions for the formulas of AGE_{LTL} (in a situation s) are:

- $\mathcal{M}, s \not\models \bot$;
- $\mathcal{M}, s \models p$ if and only if $p \in L(p)$;
- $\mathcal{M}, s \models \neg\phi$ if and only if $\mathcal{M}, s \not\models \neg\phi$;
- $\mathcal{M}, s \models \phi \wedge \psi$ if and only if $\mathcal{M}, s \models \phi$ and $\mathcal{M}, s \models \psi$;
- $\mathcal{M}, s \models A_{\mathcal{N}}\phi$ if and only if $\mathcal{M}, s' \models \phi$ for all $(s, s') \in \mathcal{A}$;

- $\mathcal{M}, s \models G\phi$ if and only if $\mathcal{M}, s' \models \phi$ for all $(s, s') \in \mathcal{G}$;
- $\mathcal{M}, s \models E\phi$ if and only if $\mathcal{M}, s' \models \phi$ for all $(s, s') \in \mathcal{E}$;
- $\mathcal{M}, s \models \mathbf{X}\phi$ if and only if $\mathcal{M}, s' \models \phi$ for the s', $(s, s') \in \mathcal{R}$
- $\mathcal{M}, s \models \phi\mathbf{U}\psi$ if and only if $\mathcal{M}, s \models \phi$; $\mathcal{M}, s^i \models \phi$ for all s^i, $i \in \{1, 2, \ldots, k\}$ such that $\{(s, s^1), (s^1, s^2), \ldots (s^{k-1}, s^k)\} \in \mathcal{R}$ and for s^{k+1} such that $(s^k, s^{k+1}) \in \mathcal{R}$ it holds $\mathcal{M}, s^{k+1} \not\models \phi$ and $\mathcal{M}, s^{k+1} \models \psi$.

A formula ϕ is true in an AGE_{LTL} model $\mathcal{M}$ if and only if $\mathcal{M}, s \models \phi$ for every situation $s \in W$. The formula ϕ is valid (noted $\models_{AGE_{LTL}}$) if and only if ϕ is true in all AGE_{LTL} models. The formula ϕ is AGE_{LTL}-satisfiable if and only if the formula $\neg\varphi$ is not AGE_{LTL} valid.